

UNIVERSIDADE DE BRASÍLIA
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO

Combinações de Modelos de Previsão
da Estrutura a Termo da Taxa de Juros:
Aplicações ao Caso Brasileiro

Rafael Cavalcanti de Araújo

Orientador: Daniel Oliveira Cajueiro

Trabalho apresentado como requisito à conclusão do curso de mestrado do Programa de Pós-Graduação em Economia da Universidade de Brasília.

Brasília
Janeiro
2011

À minha amada esposa: Cyntia

Agradecimentos

A Deus pelo dom da vida.

A Cyntia, pelo amor, abdicção, compreensão e por ser a mãe de Arthur.

Aos meus pais, Eduardo e Nazaré, pelos valores e dedicação. Aos meus irmãos, Bruno e Ricardo, pelo apoio nos momentos importantes.

Ao meu orientador, Daniel, pelo incentivo, presteza e tempestividade.

Sumário

1	Introdução	10
2	Revisão Teórica Geral	14
2.1	Modelos para a ETTJ	14
2.2	Combinação de Previsões	21
3	Modelos da ETTJ considerados	28
3.1	Modelos de Previsão da ETTJ	28
3.1.1	Passeio Aleatório	28
3.1.2	Autorregressivo Univariado	29
3.1.3	Vetor Autorregressivo	30
3.1.4	Nelson-Siegel Dinâmico	32
3.1.5	FSN-ECM	34
4	Métodos de Combinação de Previsões	39
4.1	Combinações Simples	40
4.2	Erro Quadrático Médio Invertido	41
4.3	Mínimos Quadrados Ponderados	42
4.4	Combinações via Modelo de Conjunto de Confiança	43
5	Descrição dos Dados	46

6	Resultados	51
6.1	Procedimentos de Previsão	51
6.2	Procedimentos de Avaliação dos Resultados	51
6.3	Resultados dos Modelos Individuais	53
6.4	Resultados dos Métodos de Combinação	61
7	Conclusão	66
8	Referências Bibliográficas	69
9	Apêndice	80
9.1	Figuras	80

Lista de Figuras

1	Evolução das Taxas de <i>swap</i> DI-Pré	48
2	Traço do Erro Quadrático Acumulado (inclui período de treinamento)	60
3	Traço do Erro Quadrático Acumulado, horizonte de um mês	64
4	Traço do Erro Quadrático Acumulado, horizonte de três meses	64
5	Traço do Erro Quadrático Acumulado, horizonte de seis meses	65
6	Traço do Erro Quadrático Acumulado, horizonte de 12 meses	65
7	Taxas observadas e previstas de <i>swap</i> de 3 meses, horizonte de um mês	81
8	Taxas observadas e previstas de <i>swap</i> de 12 meses, horizonte de um mês	81
9	Taxas observadas e previstas de <i>swap</i> de 18 meses, horizonte de um mês	82
10	Taxas observadas e previstas de <i>swap</i> de 30 meses, horizonte de um mês	82
11	Taxas observadas e previstas de <i>swap</i> de 3 meses, horizonte de três meses	83
12	Taxas observadas e previstas de <i>swap</i> de 12 meses, horizonte de três meses	83
13	Taxas observadas e previstas de <i>swap</i> de 18 meses, horizonte de três meses	84

14	Taxas observadas e previstas de <i>swap</i> de 30 meses, horizonte de três meses	84
15	Taxas observadas e previstas de <i>swap</i> de 3 meses, horizonte de seis meses	85
16	Taxas observadas e previstas de <i>swap</i> de 12 meses, horizonte de seis meses	85
17	Taxas observadas e previstas de <i>swap</i> de 18 meses, horizonte de seis meses	86
18	Taxas observadas e previstas de <i>swap</i> de 30 meses, horizonte de seis meses	86
19	Taxas observadas e previstas de <i>swap</i> de 3 meses, horizonte de 12 meses	87
20	Taxas observadas e previstas de <i>swap</i> de 12 meses, horizonte de 12 meses	87
21	Taxas observadas e previstas de <i>swap</i> de 18 meses, horizonte de 12 meses	88
22	Taxas observadas e previstas de <i>swap</i> de 30 meses, horizonte de 12 meses	88

Lista de Tabelas

1	Estatísticas Descritivas das Taxas de <i>swap</i> DI-Pré	50
2	(Traço) Raiz do Erro Quadrático Médio Relativo, horizonte de um mês	55
3	(Traço) Raiz do Erro Quadrático Médio Relativo, horizonte de três meses	56
4	(Traço) Raiz do Erro Quadrático Médio Relativo, horizonte de seis meses	57
5	(Traço) Raiz do Erro Quadrático Médio Relativo, horizonte 12 meses	58

Resumo

Questões como quebras estruturais e problemas de especificação tornam difícil a possibilidade de existência de um modelo de previsão da estrutura a termo da taxa de juros que domine os demais competidores. Essa dissertação tem como objetivo principal identificar a existência de métodos de combinação que permitam obter previsões da curva de juros superiores às dos modelos individuais para o caso brasileiro. A análise dos resultados dos modelos individuais confirma não ser possível encontrar um modelo que consistentemente produza menor erro de previsão. Adicionalmente, os desempenhos relativos podem variar ao longo do tempo. Os problemas de utilização de modelos individuais podem ser reduzidos pelo intermédio de métodos de combinação de previsões. Os resultados alcançados para o período analisado permitem atestar ganhos de previsão consistentes ao longo tempo. Em particular, quanto mais longo o horizonte considerado, maior a contribuição das combinações.

Palavras-chave: Estrutura a Termo da Taxa de Juros, Modelos de Previsão, Combinação de Previsão, Modelo de Conjunto de Confiança.

Classificação JEL: G1, E43, E47, C5.

Abstract

Issues like structural breaks and misspecification make it difficult to find any term structure of interest rates model that dominates all competitors. This dissertation aims to identify the existence of combination methods that produce superior forecast results compared to individual models applied to Brazilian data. Empirical results confirm not being possible to determine a individual model capable of consistently produce superior forecasts. Furthermore, the relative performance of these models can change over time. The problems of using individual models can be reduced by applying forecast combination schemes. The results for the period considered show consistent forecast gains of combination techniques over time. In particular, the higher the forecast horizon, the higher the contribution of forecast combination.

Keywords: Term Structure of Interest Rates, Forecasting Models, Forecast Combination, Model Confidence Set.

JEL Classification: G1, E43, E47, C5.

1 Introdução

A estrutura a termo da taxa de juros (ETTJ) é uma função que associa uma única taxa de juros para cada maturidade de um mesmo tipo de ativo de renda fixa. A disposição gráfica da ETTJ em um determinado ponto no tempo é conhecida como curva de juros¹. As diferentes maturidades da ETTJ são estreitamente relacionadas, visto que as taxas dos vencimentos mais longos representam a média ajustada ao risco da trajetória esperada da taxa de curto prazo.

A previsão da ETTJ de forma adequada é relevante para inúmeras áreas. Para os bancos centrais, a ETTJ ajuda a explicar o mecanismo pelo qual a política monetária afeta as decisões dos agentes econômicos. Apesar das ações de política monetária impactarem diretamente a taxa de juros de curto prazo, são as taxas com maturidades mais longas que desempenham papel fundamental nas decisões dos agentes econômicos. Uma vez que a ETTJ reflete as expectativas dos agentes com relação à trajetória futura da taxa de juro de curto prazo, a mesma embute informação acerca das decisões correntes e futuras de política monetária. Logo, a maneira pela qual os bancos centrais podem afetar as taxas de juros de prazos mais longos é influenciando as expectativas dos agentes com respeito às decisões futuras de política monetária.

Para os tesouros dos países, a ETTJ sinaliza as taxas pelas quais a administração

¹Apesar dessa diferença conceitual, os termos ETTJ e curva de juros são tratados como sinônimos ao longo desse estudo.

da dívida pública deve pagar para que o mercado adquira os seus títulos prefixados². No caso das instituições financeiras, a curva de juros fornece os fatores de descontos necessários para calcular o valor de mercado dos diversos ativos existentes em seus balanços. Em particular, a área de gestão de carteira utiliza modelos de curva de juros para fins de alocação de recursos enquanto a área de risco tenta replicar os movimentos da ETTJ para gerenciar os limites de exposição ao risco dos ativos de renda fixa de diferentes maturidades.

Para as firmas, a ETTJ afeta as decisões de investimento ao refletir os custos de oportunidade utilizados na análise dos seus projetos a serem implementados. Para as famílias, a curva de juros influencia as decisões de consumo e poupança ao representar os preços relativos pelos quais os indivíduos devem pagar (ou receber) para antecipar (ou postergar) seus planos de consumo. Para a área de conjuntura econômica, a ETTJ é de grande interesse, pois é considerada um importante indicador antecedente da atividade econômica. Além destes, diversos outros agentes econômicos também fazem uso da curva de juros nas suas atividades, como exemplo: seguradoras e entidades de previdência.

Assim, dada sua relevância, trabalhos nas áreas de finanças e econometria tentam modelar e prever a dinâmica da taxa de juros. Em particular, foram desenvolvidos vários modelos para tentar prever a evolução da ETTJ. Esses modelos podem ser

²Títulos cuja rentabilidade é definida no momento da aquisição. Já a rentabilidade dos títulos pós-fixados é determinada ao longo da existência do título e é vinculada à variação do seu indexador.

agrupados em quatro principais áreas: modelos lineares, modelos afins, modelos de fatores e modelos de regressão da taxa *forward*.

A implementação dos modelos de previsão da ETTJ no Brasil é recente e os resultados obtidos até o momento não permitem apontar para um único modelo capaz de produzir previsões superiores aos demais³. Tal fato ocorre devido diferentes modelos apresentarem desempenhos de previsão superiores aos seus concorrentes, a depender do horizonte de tempo à frente, da maturidade e do período considerados.

Uma vez que os diferentes modelos de previsão podem apresentar algum tipo de viés de especificação (Timmerman (2006)), o presente trabalho se justifica pela necessidade de averiguar se a utilização de diferentes técnicas de combinação contribui para a melhoria das previsões da ETTJ, aplicada ao mercado brasileiro.

É nesse esforço de prever o comportamento da curva de juros brasileira que este trabalho se insere. O objetivo principal é verificar a existência de técnicas de combinação⁴ que permitam obter previsões das taxas de juros superiores às dos modelos individuais para os dados no Brasil. Mais especificamente, o problema aqui colocado é verificar se a utilização de métodos de combinação de modelos individuais produz previsões superiores às obtidas pelos modelos individuais. Em caso afirmativo, serão analisados os esquemas de combinação mais adequados para as diferentes situações.

³A mesma constatação se aplica para o mercado americano.

⁴Ao longo do trabalho, os termos técnicas, métodos e esquemas são tratados como sinônimos.

O presente trabalho, além da introdução, contém os seguintes capítulos:

- O segundo capítulo apresenta uma revisão bibliográfica dos modelos de previsão da curva de juros e dos esquemas de combinação de previsões.
- O terceiro capítulo expõe os modelos de previsão individuais utilizados para a curva de juros, usando os dados brasileiros.
- O quarto capítulo apresenta os diferentes métodos de combinação aplicados no trabalho.
- O quinto capítulo descreve os dados utilizados para representar a ETTJ brasileira.
- O sexto capítulo avalia o desempenho fora da mostra dos modelos individuais e dos esquemas de combinação aplicados. A análise será efetuada por meio do uso de estatísticas, de teste de significância e de análise gráfica.
- O sétimo capítulo encerra o trabalho, relacionando as conclusões obtidas, as limitações do trabalho e propondo questões para pesquisas futuras.

2 Revisão Teórica Geral

2.1 Modelos para a ETTJ

Por muito tempo a literatura da taxa de juros se concentrou no debate da Hipótese da Expectativa da ETTJ, postulada por Fisher (1896). Nesta, as taxas de juros sobre títulos de diferentes maturidades são primariamente relacionadas com as expectativas dos agentes dos mercados sobre a evolução da taxa de juros no futuro. Sobre esta hipótese, além de outras adicionais (como a existência de prêmio de risco⁵ nulo ou constante), as taxas da curva de juros podem fornecer previsões ótimas para a evolução da taxa à vista, no sentido de que nenhum modelo alternativo produziria previsão superior.

Contudo, inúmeros estudos empíricos⁶ para os dados americanos rejeitaram a Hipótese da Expectativa, ao constatar que os prêmios de risco oscilam ao longo do tempo.

⁵Excesso de retorno exigido por um investidor avesso ao risco para aplicar recursos em títulos de maiores maturidades ao invés de investir em aplicações com retorno atrelado à taxa livre de risco (taxa de um dia estabelecida pelos bancos centrais), dado que existe incerteza sobre a evolução do cenário econômico, em particular, sobre a taxa de juro real e inflação. Nesse sentido, mesmo que o investidor tenha eventuais ganhos em relação ao principal no título, o mesmo está sujeito a perdas em relação ao custo de oportunidade - aplicação atrelada à taxa livre de risco. Em alguns casos, também há compensação em razão de problemas de liquidez.

⁶Ver, por exemplo, em Della Corte *et al.*, (2006).

Posteriormente, a literatura da ETTJ concentrou esforços no campo teórico. Os primeiros trabalhos focaram o conceito de equilíbrio, no qual um pequeno conjunto de variáveis de estado⁷ descrevem a ETTJ. Estes tipos de modelos procuram explicar a dinâmica da taxa de juros de curto prazo, com parâmetros estimados por intermédio de métodos estatísticos. O modelo pioneiro de Vasicek (1977) foi seguido por outros, tais como Cox *et al.* (1985), Brennan e Schwartz (1979) e Longstaff e Schwartz (1992).

Os modelos de equilíbrio usam tipicamente a classe dos modelos afins da ETTJ⁸. Nesta abordagem, as taxas de juros de diversas maturidades são funções afins (constante mais linear) de um pequeno número de fatores latentes ou não observados⁹. Estes fatores, por sua vez, podem ser extraídos de um painel de taxas de juros ao longo do tempo e por meio das maturidades. Além disso, os modelos afins impõem restrições em equações transversais que excluem oportunidades de arbitragem. Duffie e Kan (1996) generalizaram a literatura e Dai e Singleton (2000) caracterizaram o conjunto de modelos admissíveis e identificáveis.

Em seguida, uma outra classe de modelos, denominada livre de arbitragem¹⁰, foi introduzida. O foco era a boa aderência das taxas observadas em um determinado ponto no tempo, de modo a não permitir oportunidade de arbitragem, qualidade es-

⁷Componentes fundamentais do processo da taxa de juros.

⁸Para uma abordagem dos modelos afins, ver Campbell *et al.*, (1997) e Piazzesi (2003).

⁹Singleton (1980) é a primeira aplicação da análise dos fatores latentes para a curva de juros.

¹⁰*Arbitrage-free*, em inglês.

sencial para o apereçamento dos derivativos de taxas de juros¹¹. Assim, os parâmetros desta classe de modelos são calibrados de modo a coincidir o preço do ativo subjacente do modelo com o seu preço de mercado, usado como dado de entrada. Uma vez que estes modelos não focaram a dimensão de séries temporais, as propriedades dinâmicas não são descritas. Consequentemente, não apresentam utilidade para previsão. Entre os modelos da classe de livre de arbitragem destacam-se: Ho-Lee (1986), Hull e White (1990), Black *et al.* (1990) e Heath *et al.* (1992).

Como os modelos afins possuem propriedades dinâmicas, Duffee (2002) testou a eficácia de suas previsões. O autor verificou que estes modelos produzem resultados inferiores aos do modelo de passeio aleatório (PA), usados na literatura como referencial (*benchmark*) para testar a eficácia de previsão dos demais modelos.

O baixo desempenho de previsão dos modelos afins estimulou Diebold e Li (2006) a adotar uma perspectiva de previsão fora da amostra de forma explícita. No trabalho, os autores assumem a abordagem de componentes exponenciais de Nelson e Siegel (1987) e a adaptam a um modelo dinâmico com três parâmetros, conhecido como DNS (Dynamic Nelson-Siegel, em inglês). Estes parâmetros, por sua vez, foram interpretados como fatores latentes da curva de juros¹². Os resultados obtidos do modelo foram promissores, apresentando previsões superiores ao passeio

¹¹Para uma descrição detalhada dos modelos livres de arbitragem, ver Fabozzi *et al.* (2002).

¹²Originalmente descritos por Litterman e Scheinkman (1991) como: nível, inclinação e curvatura.

aleatório para os dados americanos, especialmente em horizontes mais longos. A grande repercussão dos resultados levaram o DNS a ser frequentemente usado como referencial de comparação dos modelos posteriormente desenvolvidos.

Após o trabalho de Diebold e Li (2006), vários autores desenvolveram pesquisas em modelos de previsão da ETTJ baseados no modelo de Nelson-Siegel. De Pooter (2007) estendeu o modelo dinâmico de Nelson-Siegel e constatou que a introdução de novos fatores (de forma similar ao feito em Svenson (1995)) contribui para o incremento da previsão fora da amostra.

Koopman *et al.* (2007) introduzem parâmetros variantes no tempo nos modelos de Nelson-Siegel. Diebold *et al.* (2008) estendem o modelo para um contexto global, modelando um grande conjunto de curva de juros de diferentes países em uma abordagem que permite tanto fatores específicos de um país como fatores globais.

Outra vertente de modelos de previsão é baseada na interação entre a macroeconomia e a curva de juros. Uma abordagem usualmente adotada é estender os modelos de fatores latentes existentes, adicionando fatores macroeconômicos observáveis e estudando a interação entre eles. Neste sentido, Ang e Piazzesi (2003) desenvolveram um modelo afim de três fatores latentes, acrescidos de variáveis macroeconômicas. Monch (2008) analisa o conteúdo preditivo dos componentes principais da curva de juros dentro de um modelo de fatores dinâmicos envolvendo dados macroeconômicos de taxas de juros. Ludvigson e Ng (2009) usam um modelo dinâmico de fatores que sumarizam as informações extraídas de um grande número de séries

macroecômicas¹³ Eles mostram que os fundamentos macroeconômicos contribuem de forma importante para a previsão do prêmio de risco dos títulos americanos¹⁴.

Diebold *et al.* (2006) estenderam a abordagem de Nelson e Siegel (1987) para incluir fatores macroeconômicos observáveis¹⁵. De Pooter *et al.* (2010), também partindo do modelo Nelson-Siegel, atestam ganhos preditivos que podem ser obtidos quando variáveis macroeconômicas são incluídas.

Outros autores também usaram modelos com dados macroecômicos para a curva de juros, porém, sem ter como base a abordagem dos modelos de fatores. Evans e Marshall (1998) extraem informações de pontos múltiplos da curva de juros usando modelos de vetor autorregressivo (VAR) estruturais. Já Hordal *et al.* (2006) utilizam um modelo macroestrutural onde uma pequena economia é combinada com um modelo de taxa de juros que impõe ausência de arbitragem.

Apesar de boas propriedades estatísticas, os modelos baseados na abordagem de Nelson e Siegel (1987) não impõem a restrição de ausência de oportunidade de

¹³Como aponta Stock *et al.* (2006), apesar de terem objetivo comum (resumir as principais fontes de variação e covariação entre as variáveis e reduzir o número de variáveis a serem analisadas), a análise de componentes principais e a análise de fatores assumem diferentes hipóteses. Na análise de fatores, as covariâncias entre as variáveis devem-se à presença de um pequeno número de fatores latentes. Por outro lado, na análise de componentes principais não se assume a causalidade das covariâncias entre as variáveis.

¹⁴Importante destacar que a consideração de informações macroeconômicas para a previsão da ETTJ está além do escopo do trabalho.

¹⁵Inflação, capacidade instalada da indústria e taxas do *federal funds*.

arbitragem, conceito fundamental na literatura devotada à análise de apreçamento de títulos. Esta condição implica que ETTJ não deve permitir lucros livre de risco em títulos de diferentes maturidades. Ang e Piazzesi (2003) verificam que a imposição de não arbitragem aumenta a eficácia da previsão para o caso dos modelos afins.

No intuito de conciliar o modelo Nelson-Siegel com a ausência de arbitragem, Christensen *et al.* (2009a) derivaram um modelo afim que mantém a estrutura de fatores da abordagem de Nelson e Siegel (1987), o que deu origem a uma nova classe de modelos afins da ETTJ, conhecidos como AFDNS (Arbitrage-Free Dynamic Nelson-Siegel, em inglês). Christensen *et al.* (2009b) também desenvolveram generalizações do modelo Nelson Siegel, com a adição de um fator de curvatura (como em Svenson (1995)) e um fator de inclinação, além de impor a ausência de arbitragem.

Entretanto, o incremento da previsão advindo da imposição de ausência de arbitragem não é consenso na literatura. Duffe (2009) argumenta que esta restrição é irrelevante para a melhoria da previsão, servindo apenas para fornecer um mecanismo que especifica a forma funcional da compensação do risco, a qual, por sua vez, impõe restrições sobre o comportamento dinâmico das taxas.

Bowsher e Meeks (2008) não adotam a abordagem de Nelson-Siegel. Em vez disso, introduzem um modelo de cinco fatores, conhecido como FSN-ECM, no qual *splines* são usadas para modelar a curva de juros. Além disso, os autores argumentam que os nós dessas *splines* atuam como fatores. Embora esta abordagem permita

uma curva de juros mais flexível, a intuição dos fatores é perdida.

Para o caso brasileiro, Lima *et al.* (2006) estimaram um modelo de vetor autorregressivo e um de vetor de correção de erros (VEC) para prever taxas de juros de longo prazo. Apesar de úteis em construir cenários qualitativos para a curva de juros, seus resultados empíricos não fornecem boas previsões em termos quantitativos. Vereda *et al.* (2008), a partir da abordagem de Evans e Marshall (1998), propõem novas abordagens VAR e concluem que a incorporação de variáveis macroeconômicas contribui para melhorar desempenho da previsão de longo prazo para a taxa de juros.

Vicente e Tabak (2007) aplicam os modelos afim e o DNS. Seus resultados sugerem que as previsões obtidas deste último são superiores ao primeiro, em horizontes mais longos. Para o horizonte de previsão um mês à frente, nenhum dos modelos supera o passeio aleatório. Assim como De Pooter (2007) para os dados americanos, Almeida *et al.* (2007) estendem a abordagem de Nelson-Siegel à ETTJ no Brasil ao introduzir um segundo fator de curvatura, encontrando ganhos preditivos em relação ao DNS. Por sua vez, Almeida *et al.* (2007) analisam diferentes critérios de escolhas das cargas dos fatores (*factor loadings*) e concluem que os valores dos parâmetros do DNS devem ser adaptados a depender da maturidade e do horizonte de previsão considerados.

Cajueiro *et al.* (2009) aplicam os modelos DNS e FSN-ECM para os horizontes de um, três, seis e doze meses à frente. Comparando os resultados dos dois modelos

com o passeio aleatório, os autores verificam que enquanto o modelo de FSN-ECM apresenta resultados preditivos superiores para horizontes de tempo de curto prazo, o modelo de DNS produz melhores previsões para horizontes mais longos¹⁶.

Leite *et al.* (2009) desenvolvem um modelo de previsão para a curva de juros brasileira baseado em informações macroeconômicas, expectativa de inflação coletada pelo Banco Central do Brasil junto às instituições financeiras¹⁷ e o prêmio de risco das taxas a *forward*. Utilizando um horizonte de tempo seis meses à frente, encontram resultados superiores aos modelos dinâmicos de Nelson-Siegel e ao passeio aleatório.

Por fim, Guillen e Tabak (2008) examinam a Hipótese de Expectativa para ETTJ brasileira. Seus resultados sugerem a rejeição desta hipótese ao identificar prêmios de risco variantes ao longo do tempo. Especialmente, caracterizaram o prêmio de risco da curva de juros brasileira como dependente das condições de liquidez internacional e de fatores domésticos.

2.2 Combinação de Previsões

A combinação de previsões corresponde à combinação de duas ou mais previsões

¹⁶Assim como em Cajueiro *et al.* (2009), curto prazo é definido nesse trabalho como horizontes de previsão de 1 e 3 meses à frente, enquanto longo prazo corresponde aos horizontes de 6 e 12 meses à frente. Adicionalmente, consideram-se as seguintes definições das maturidades: curtas (até 3 meses), médias (de 6 a 18 meses) e longas (acima de 18 meses).

¹⁷Relatório Focus.

individuais obtidas de um painel de previsões, de forma a produzir uma única previsão. Bates e Granger (1969), com seu trabalho inicial, explicitaram os benefícios das combinações de modelos quando um quantitativo de previsões é limitado. Além disso, mostraram que as combinações das previsões de diferentes modelos geralmente superam as previsões individuais.

Na literatura de taxa de juros, a combinação de previsões baseia-se na constatação empírica de que diferentes modelos apresentam bom desempenho de previsão a depender do período, horizonte de previsão e maturidade analisados.

Diebold e Pauly (1987) focam as quebras estruturais¹⁸ para defender a utilização de pesos variantes no tempo ao apontar argumentos para subotimilidade de pesos fixos no tempo: diferentes velocidades de aprendizagem dos modelos de previsão podem tornar um modelo particular melhor em relação aos demais ao longo do tempo; a configuração de um modelo pode torná-lo melhor para previsão em algumas situações do que em outras; alguns modelos podem ser mais vulneráveis às mudanças macroeconômicas; e a não linearidade na estrutura macroeconômica leva a erros de variância de previsão não constantes, o que conduz à necessidade da adoção de pesos não fixos.

Diebold e Lopez (1996) partem do argumento do elevado custo de combinar conjuntos de informações em um único modelo para justificar o emprego de combinação

¹⁸Por quebras estruturais entendem-se mudanças permanentes no vetor de parâmetros de um modelo.

de previsões. Classificam os métodos de combinação em dois grupos: variância-covariância e baseados em regressões. No segundo grupo, os autores relacionam as seguintes variações: combinação de pesos variantes no tempo, contração Bayesiana de combinação de pesos, regressões de combinações dinâmicas e regressões de combinações não lineares.

Hendry e Clements (2001) defendem que a combinação adiciona valor à previsão quando os modelos individuais apresentam diferentes problemas de especificação e o processo gerador de dados é sujeito a mudanças determinísticas intermitentes. Afirmam que na ocorrência destas situações é provável que a média domine os métodos de pesos estimados. Os autores ainda propõem métodos alternativos de combinação de previsões: correção de viés e correção de interceptos.

Timmerman (2006) defende que, diante de uma perspectiva teórica, a combinação de previsões oferece ganhos de diversificação, o que torna atrativo combinar previsões individuais. O autor cita alguns argumentos para justificar a combinação de previsões. O primeiro é a diversificação de carteira, cuja premissa é que dada a presença de informação privada, os conjuntos de informações das previsões individuais geralmente não são observados pelo usuário de previsão. Neste caso, não é possível agregar os conjuntos de informações e construir um grande modelo que abranja cada modelo de previsão.

O segundo é que previsões individuais podem ser afetadas por quebras estruturais de formas diferentes. O autor argumenta que, em razão da dificuldade de serem

detectadas quebras estruturais de forma tempestiva, é plausível que, na média, a combinação de previsões de modelos com diferentes graus de adaptabilidade produza previsões superiores às dos modelos individuais.

O terceiro é que as previsões dos modelos podem estar sujeitas a problemas de viés de especificação desconhecido. Logo, a combinação de previsões ao longo de diferentes modelos pode ser vista como uma maneira de tornar a previsão mais robusta contra problemas de viés de especificação.

O quarto é que as previsões podem ser baseadas em diferentes funções de perda. Assim, caso existam previsões individuais com funções de perda com alguma simetria, é possível que se encontre combinações destas previsões melhores que as individuais.

Stock e Watson (2006) abordam casos em que o número de previsões a serem combinadas é grande. Nestas situações, destacam as seguintes vantagens em combinar previsões de diferentes modelos: explorar uma base de informação mais rica e fornecer alguma robustez contra instabilidade estrutural presente em modelos com menos parâmetros. Apresentam vários esquemas de combinação de pesos: combinação simples, erro quadrático médio invertido, previsões reduzidas, parâmetros variantes no tempo e média Bayesiana.

Clements e Hendry (2006) consideram o impacto das quebras estruturais e sua relevância nos modelos de previsão. Enfatizam que diferentes modelos e métodos apresentam reações diversas em face de uma quebra estrutural, mas provavelmente

em todos os casos a quebra conduz a um viés sistemático de previsão. Argumentam que enquanto uma quebra estrutural induz à instabilidade em alguns parâmetros de um modelo particular, as consequências para a previsão são específicas do tipo de quebra e da especificação do modelo.

Apesar de serem aplicadas em diversas áreas econômicas¹⁹, até recentemente as combinações de modelos de previsão da taxa de juros tinham sido pouco exploradas. Clark e McCracken (2006) enfatizam a incerteza presente em qualquer método de representação de instabilidade dos modelos VAR — para dados de produto, inflação e taxa de juros de curto prazo — como justificativa para o uso de combinações de previsões, com o intuito de melhorar o poder preditivo. Os métodos de combinação adotados pelos autores são média simples (MS), mediana (ME) e média aparada (MA), além de pesos baseados em: regressões Bates-Granger, fatores comuns, erro quadrático médio (EQM), regressões em quartis e médias Bayesianas.

Guidolin e Timmermann (2009) propõem uma abordagem de combinação de previsão flexível das futuras taxas de juros de curto prazo. Nesta, sempre é possível escolher um modelo econométrico entre as demais alternativas. Adicionalmente, esta abordagem permite que os pesos das combinações variem ao longo de um conjunto de estados. Estes pesos, por sua vez, podem ser identificados por um processo conjunto que comanda as taxas à vista, a termo e momentos das distribuições do processo

¹⁹Por exemplo, Stock e Watson (2004) aplicam para dados de crescimento do produto e Rapach e Strauss (2005) para dados de crescimento de emprego.

gerador dos dados. Já a combinação ótima dos pesos dependerá tanto do estado corrente das probabilidades como do horizonte de previsão. Dentre os esquemas de combinação adotados pelos autores, destacam-se os pesos variantes no tempo, os quais possuem regimes dinamicamente variáveis.

Bolder e Romanyuk (2008) defendem que mudanças estruturais ou de regimes podem dificultar qualquer modelo na captura de tendências nos dados e dominar todas as abordagens alternativas. Os autores usam os dados de taxa de juros do Canadá e examinam quatro modelos: Nelson-Siegel, *spline* exponencial, série de Fourier e uma abordagem de espaços de estado. Também utilizam vários esquemas de combinação: média simples, Mínimos Quadrados Ordinários (MQO) simples, MQO de fatores, regressões *spline* multiadaptativas, probabilidade preditiva, modelo de probabilidade marginal e probabilidade logarítmica (marginal e preditiva). A principal observação dos resultados obtidos é a evidência de que as combinações de previsões da ETTJ apresentam desempenhos superiores aos dos modelos individuais.

De Pooter *et al.* (2007) ressaltam os ganhos preditivos que podem ser obtidos quando as previsões de diferentes modelos são combinadas, argumentando que os modelos individualmente podem produzir resultados insatisfatórios. Eles dividem a amostra das taxas de juros americanas em subperíodos e aplicam os seguintes modelos: passeio aleatório, autorregressivo, vetor autorregressivo, Nelson-Siegel dinâmico e Afim. Quanto aos esquemas de combinações, examinam: pesos iguais, erro quadrático médio invertido e probabilidade preditiva Bayesiana.

Os autores argumentam que o fato de diferentes modelos apresentarem melhor desempenho em diferentes subperíodos, confirma *ex-post* que diferentes especificações de modelos exibem um papel complementar na aproximação do processo gerador dos dados. Ademais, afirmam que os resultados empíricos do estudo fornecem uma importante argumentação para o uso de técnicas de combinação.

De Pooter *et al.* (2010) atualizaram o trabalho anterior. Baseando-se na premissa da dificuldade de qualquer modelo individual produzir previsões consistentes ao longo do tempo, utilizam técnicas de combinação como alternativa para a previsão da ETTJ. Os modelos individuais são os mesmos da versão anterior. Todavia, com relação aos esquemas de combinação, há uma substituição da probabilidade preditiva Bayesiana pela combinação de modelos após exclusão, usando a abordagem do Modelo de Conjunto de Confiança (MCS, em inglês). Também não há divisão da amostra em subperíodos. Concluem, após a análise dos resultados, que a combinação de modelos individuais conduz a ganhos de previsibilidade consistentes ao longo do tempo, especialmente para maturidades mais longas.

Para os dados brasileiros, não há registro de aplicações de esquemas de combinação de modelos, o que torna este trabalho uma contribuição para a literatura.

3 Modelos da ETTJ considerados

São considerados $\tau = 1, 2, 3, 6, 9, 12, 15, 18, 21, 24, 30$ as maturidades em meses de um título sem pagamentos intermediários com valor de face \$1 e $N = 11$ o número de diferentes maturidades. Seja $y_t(\tau_i)$ a taxa do título até a maturidade do período t até o período $t + \tau_i$, para $i = 1, \dots, N$. Seja h o número de períodos à frente²⁰. A previsão da taxa de maturidade τ_i no tempo $t + h$ é dada por $\hat{y}_{t+h}(\tau_i)$.

3.1 Modelos de Previsão da ETTJ

A primeira questão relevante para implementação do trabalho reside na estimação dos seguintes modelos individuais de previsão da curva de juros. Foram considerados modelos com diferentes níveis de complexidade. Os modelos considerados variam daqueles com estrutura linear (passeio aleatório, AR e VAR) e com estrutura de fatores (DNS e FSN-ECM)²¹.

3.1.1 Passeio Aleatório

Seja um modelo de passeio aleatório puro, ou seja, sem deslocamento, das taxas de juros para cada maturidade individual τ_i :

²⁰No caso, meses.

²¹Os modelos AR pela abordagem iterada, DNS e FSN-ECM já foram implementados para os dados brasileiros em programas escritos em Ox por Daniel Cajueiro.

$$y_t(\tau_i) = y_{t-1}(\tau_i) + \varepsilon_t(\tau_i) \quad (1)$$

onde $\varepsilon_t(\tau_i) \sim N(0, \sigma^2(\tau_i))$, isto é, um processo de ruído branco.

A previsão das taxas h períodos à frente é dada por:

$$\hat{y}_{t+h}(\tau_i) = y_t(\tau_i) \quad (2)$$

ou seja, é igual ao valor $y_t(\tau_i)$ mais recentemente observado.

3.1.2 Autorregressivo Univariado

No contexto de previsão, duas abordagens do modelo autorregressivo univariado de 1ª ordem podem ser consideradas. Na abordagem iterada, o modelo a ser estimado é:

$$y_t(\tau_i) = c(\tau_i) + \phi(\tau_i)y_{t-1}(\tau_i) + \varepsilon_t(\tau_i) \quad (3)$$

onde $\varepsilon_t(\tau_i) \sim N(0, \sigma^2(\tau_i))$ e $c(\tau_i)$, $\phi(\tau_i)$ e $\sigma^2(\tau_i)$ são parâmetros escalares.

A previsão para um horizonte h meses à frente é produzida a partir de:

$$\hat{y}_{t+h}(\tau_i) = \hat{c}(\tau_i) + \hat{\phi}(\tau_i)\hat{y}_{t+h-1}(\tau_i) \quad (4)$$

Por outro lado, na abordagem direta, as previsões são obtidas ao regredir y_t diretamente sobre o seu valor y_{t-h} . Neste caso, o modelo a ser estimado é:

$$y_t(\tau_i) = c_h(\tau_i) + \phi_h(\tau_i)y_{t-h}(\tau_i) + \varepsilon_t(\tau_i) \quad (5)$$

onde $\varepsilon_t(\tau_i) \sim N(0, \sigma^2(\tau_i))$. Assim, a variável no tempo t é projetada diretamente com base no seu valor passado no período $t - h$, de tal forma que os coeficientes estimados resumam o efeito dos h meses à frente. A previsão é então obtida como:

$$\hat{y}_{t+h}(\tau_i) = \hat{c}_h(\tau_i) + \hat{\phi}_h(\tau_i)\hat{y}_t(\tau_i) \quad (6)$$

Apesar de implementadas para o caso brasileiro, a abordagem direta não foi considerada neste trabalho. Isto porque, como apontado por Carriero *et al.* (2009) para os dados americanos, os resultados de previsões obtidos foram inferiores aos da abordagem iterada.

Quanto à ordem do processo, Carriero *et al.* (2009) concluem que as previsões obtidas em modelos AR de ordem superior são próximas aos da 1ª ordem²², o que dá suporte para manter o processo de forma mais simples.

3.1.3 Vetor Autorregressivo

Os modelos VAR permitem o uso dos dados históricos de outras maturidades como informação adicional, além do dado histórico da própria maturidade. No presente trabalho, é considerado um modelo VAR com a seguinte especificação:

²²O mesmo resultado vale para os modelos VAR.

$$Y_t = c + \Phi Y_{t-1} + H \varepsilon_t(\tau) \quad (7)$$

onde $\varepsilon_t(\tau_i) \sim N(0, I)$, $Y_t = (y_t(1), \dots, y_t(30))'$ é um vetor (11×1) que contém as taxas das diferentes maturidades, c é um vetor (11×1), Φ é uma matriz (11×11) e H é a matriz da variância residual contendo $\frac{1}{2}N(N+1) = 66$ parâmetros livres.

Dado o elevado número de parâmetros a serem estimados, o vetor Y_{t_1} é substituído por um vetor de três componentes principais²³. Dessa forma, tem-se:

$$Y_t = c + \Phi F_{t-1} + \varepsilon_t(\tau) \quad (8)$$

onde $\varepsilon_t(\tau_i) \sim N(0, H)$ e Φ passa a ser uma matriz (11×3). Desse modo, a previsão pode ser obtida por:

$$\hat{Y}_{t+h} = \hat{c} + \hat{\Phi} \hat{F}_{t+h-1} \quad (9)$$

onde para se calcular \hat{F}_{t+h-1} , são primeiramente calculadas as cargas dos fatores da análise de componentes principais usando os dados até o mês t , para depois multiplicar estas cargas pelas previsões das taxas iteradas.

²³De forma similar à encontrada por Litterman e Scheinkman (1991) e De Pooter *et al*, (2010), os primeiros três componentes principais explicam quase toda a variação da ETTJ (99,91%) e têm a mesma interpretação: nível, inclinação e curvatura.

3.1.4 Nelson-Siegel Dinâmico

O modelo DNS tem como ponto de partida o modelo estático de Nelson e Siegel (1987), que definem a taxa *forward* instantânea como um polinômio multiplicado pelo termo decaimento exponencial λ ²⁴ :

$$f_t(\tau) = \beta_1 + \beta_2 e^{-\lambda\tau} + \beta_3 \lambda e^{-\lambda\tau} \quad (10)$$

Nelson e Siegel (1987) afirmam que a equação acima pode ser interpretada como uma aproximação da solução de uma equação diferencial. Além disso, argumentam que a Hipótese da Expectativa fornece motivação heurística para utilização desta abordagem. Mais especificamente, defendem que se a taxa à vista é gerada por uma equação diferencial, então as taxas *forward*, sendo previsões, serão as soluções da equação diferencial. A estrutura a termo correspondente é dada por²⁵:

$$y_t(\tau_i) = \beta_1 + (\beta_2 + \beta_3) \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} \right) + \beta_3 (-e^{-\lambda\tau_i}) \quad (11)$$

Diebold e Li (2006) adotam uma interpretação moderna da equação (11), de modo a adequar a abordagem de Nelson e Siegel (1987) a um modelo de fatores latentes com estrutura dinâmica, onde β_1 , β_2 e β_3 são, respectivamente, interpretados como nível, inclinação e curvatura, todos variantes no tempo. Assim, obtêm-se:

²⁴Pode ser vista como uma função Laguerre, que é uma classe matemática de funções de aproximações.

²⁵A relação entre a taxa *forward* e a curva de juros é obtida através de: $y(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u) du$.

$$y_t(\tau_i) = \beta_{1t} + \beta_{2t} \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} \right) + \beta_{3t} \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i} \right) \quad (12)$$

onde cada termo que multiplica os fatores é conhecido como carga do fator²⁶.

A carga de β_{1t} é uma constante (1). Assim, alterações em β_{1t} modificam as taxas das diferentes maturidades igualmente. Além disso, como a carga de β_{1t} não tende a zero no limite, o fator β_{1t} controla as taxas de maturidades mais longas, ou seja, $\lim_{\tau_t \rightarrow \infty} y_t(\tau_t) = \beta_{1t}$. A carga de β_{2t} $((1 - e^{-\lambda\tau_i})/\lambda\tau_i)$ é uma função decrescente de τ que começa em 1 e cai rapidamente para zero. Mudanças em β_{2t} impactam mais as taxas de maturidades mais curtas do que as mais longas. Dessa forma, β_{2t} comanda as taxas mais curtas. Por fim, carga de β_{3t} $((1 - e^{-\lambda\tau_i})/\lambda\tau_i) - e^{-\lambda\tau_i}$ é uma função côncava de τ , que começa em zero, atinge um máximo e tende depois para zero. Variações em β_{3t} têm pouco efeito nas taxas curtas e longas. Logo, o fator β_{3t} governa as taxas de maturidades médias. Já o parâmetro λ controla tanto o ritmo de decaimento exponencial como o local em que a carga de β_{3t} atinge seu máximo.

Neste trabalho, é utilizado λ fixo. Assim como em Cajueiro et al. (2009), a maturidade considerada é 12 meses, que corresponde ao ponto médio do trecho da curva de juros considerado médio, com extremos em 6 e 18 meses. Por conseguinte, o valor que maximiza a carga $((1 - e^{-\lambda\tau_i})/\lambda\tau_i) - e^{-\lambda\tau_i}$ em $\tau = 12$ é 0,14944.

Assim como em Diebold e Li (2006), pressupõe-se que os fatores assumem um

²⁶Definida por Litterman e Scheinkman (1991) como a sensibilidade da taxa de um título em relação ao respectivo fator.

processo AR(1):

$$\hat{\beta}_{t+h,t} = \hat{a}_0 + \hat{a}_1 \tilde{\beta}_{jt} \quad (13)$$

onde $j = 1, 2, 3$, $\tilde{\beta}_{jt}$ são estimativas obtidas por meio de MQO, \hat{a}_0 e \hat{a}_1 são os coeficientes da regressão de $\tilde{\beta}_{jt}$ sobre $\tilde{\beta}_{j,t-h}$.

Uma vez que os $\hat{\beta}_j$ são disponibilizados em $t + h$, as previsões das taxas podem ser obtidas a partir de:

$$\hat{y}_{t+h}(\tau_i) = \hat{\beta}_{1t+h} + \hat{\beta}_{2t+h} \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} \right) + \hat{\beta}_{3t+h} \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i} \right) \quad (14)$$

3.1.5 FSN-ECM

A abordagem proposta por Bowsher e Meeks (2008) baseia-se em uma função econômica²⁷, no caso uma curva de juros latente, que passa a ser observada junto com um erro de mensuração²⁸. Para modelar a curva de juros, os autores empregam uma *cubic spline* natural²⁹ (NCS) cuja evolução dinâmica é ditada por um VAR cointegrado na forma de modelo de correção de erro (ECM, em inglês), o qual

²⁷ *Functional signal*, em inglês.

²⁸ *Noise*, em inglês.

²⁹ *Cubic spline* corresponde a um conjunto de partes de uma função polinomial de grau no máximo 3, em que cada parte é uma aproximação local da função de interesse, agrupadas de maneira a formar uma função suave. Adicionalmente, se as derivadas segundas nos nós extremos foram iguais a zero, a função em questão é denominada *cubic spline* natural.

apresenta dimensão relativamente pequena. Argumentam que FSN-ECM pode ser interpretado como um tipo especial de modelo de fatores dinâmicos, onde as taxas nos nós da *spline* são os fatores e as cargas dos fatores, determinadas pela exigência da função econômica, uma *cubic spline* natural.

Seja $S_{\gamma_t}(\tau) = (S_{\gamma_t}(\tau_1), \dots, S_{\gamma_t}(\tau_N))'$ uma NCS dinâmica avaliada no vetor das maturidades observadas $\tau = (\tau_1, \dots, \tau_N)$ com dimensão N . Seja m o número de nós da NCS posicionados nas maturidades latentes $k = (1, k_2, \dots, k_m)$, que por sua vez são fixas e invariantes no tempo. Considere $\gamma_t = (\gamma_{1t}, \dots, \gamma_{mt})'$ o vetor de taxas avaliadas nos nós, ou seja, $S_{\gamma_t(k_{jt})} = \gamma_{jt}$ para $j = 1, \dots, m$, denominadas taxas-nós da *spline*. $S_{\gamma_t}(\tau)$ é interpolada a partir de γ_t , sendo portanto determinada de forma única por γ_t . Importante ressaltar que o vetor k tem dimensão menor que o vetor τ , isto é, $m < N$.

O modelo para os vetores de taxas observadas $Y_t(\tau) = (y_t(\tau_1), \dots, y_t(\tau_N))'$ é dado por:

$$\begin{aligned} Y_t(\tau) &= S_{\gamma_t}(\tau) + \varepsilon_t(\tau) \\ &= W(k, \tau)\gamma_t + \varepsilon_t(\tau) \end{aligned} \tag{15}$$

$$\Delta\gamma_{t+1} = \alpha(\beta'\gamma_t - \mu) + \Psi\Delta\gamma_t + \nu_t \tag{16}$$

onde $W(k, \tau)$ corresponde à matriz de interpolação, dependente apenas de τ e k . Já

$S_{\gamma_t}(\tau) = W(k, \tau)\gamma_t$ é dado pelo Lema 7 de Bowsher e Meeks (2008). Adicionalmente, α é uma matriz $m \times (m - 1)$ de posto completo e β é uma matriz determinada unicamente por $\beta'\gamma_t = (\gamma_{j+1,t} - \gamma_{j,t})_{j=1}^{m-1}$.

O estado inicial $(\gamma'_1, \gamma'_0)'$ possui primeiro e segundo momentos dado por γ^* e ω^* respectivamente. O vetor $u_t = (\varepsilon_t(\tau)', v_t)'$ é um processo de ruído branco.

A equação (16) descreve a equação de estado de ECM usando um VAR cointegrado. O processo γ_t é integrado de primeira ordem, I(1). A diferença entre as taxas latentes, Δ_{t+1} é cointegrada e $\beta'\gamma_t - \mu$ é um vetor das relações cointegradas estacionário com média zero. Tem-se então que $E[\gamma_t] = \mu$ é a média estacionária das relações cointegradas e $E[\Delta\gamma_{t+1}] = 0$, portanto excluindo tendência determinística.

Sobre as condições definidas acima, Bowsher e Meeks (2008) argumentam que o modelo FSN-ECM pode ser estruturado na forma de espaço de estado linear. Esta forma de representação aceita o uso do filtro de Kalman de modo a permitir implementar tanto a estimação por quase-máxima verossimilhança como a previsão um período à frente.

Dessa forma, é possível definir a especificação do modelo por meio de FSN(m)-ECM(p), onde m é o número de nós e p é a ordem máxima da defasagem de γ_{t+1} que entra na equação (16). Para fins de previsão, neste trabalho é considerado o caso $m = 5$ e $p = 2$, assim como em Cajueiro *et al.* (2009).

Para se proceder com a estimação, restrições adicionais de parâmetros são impostas em termos da matriz não singular Q , onde $\varphi = Q\gamma_t$ é o vetor de estado

transformado a seguir, contendo a taxa latente de curto prazo e os diferenciais entre as taxas latentes:

$$\varphi_t = (\gamma_{1,t}, \gamma_{2,t} - \gamma_{1,t}, \dots, \gamma_{m,t} - \gamma_{m-1,t})' = \begin{bmatrix} 1 & 0_{1 \times (m-1)} \\ \beta' \end{bmatrix} \gamma_t = Q\gamma_t \quad (17)$$

A equação de estado pode ser escrita como o VAR:

$$\Delta\varphi_{t+1} = Q\alpha(\beta'Q^{-1}\varphi_t - \mu_s) + Q\Psi Q^{-1}\Delta\varphi_t + \eta_t \quad (18)$$

onde $\eta_t = Qv_t$ e $\Omega_\eta = \text{var}[\eta_t] = Q\Omega_v Q'$. A diagonalidade da matriz de covariância Ω_η é imposta e $\Omega_\varepsilon = \sigma_\varepsilon^2 I_N$. O filtro de Kalman é inicializado ao se usar $(\gamma'_1, \gamma'_0) = (\mu^*, \Omega^*)$, onde $\Omega^* = 0$ e μ é igualado a $(y_0(k)', y_{-1}(k)')$.

Para fins computacionais³⁰, a equação (18) pode ser escrita como:

$$\Delta\varphi_{t+1} = [I + Q\alpha\beta'Q^{-1} + Q\Psi Q^{-1}] \varphi_t + Q\Psi Q^{-1}\varphi_{t-1} + Q\alpha\mu_s + Qv_t \quad (19)$$

O vetor X_t é definido como:

$$X_t = \begin{bmatrix} \varphi_t \\ \varphi_{t-1} \\ 1 \end{bmatrix} \quad (20)$$

³⁰Ver em Cajueiro *et al.* (2009).

Portanto, chega-se à seguinte representação de estado de espaços:

$$X_{t+1} = \begin{bmatrix} A & B & \bar{C} \\ I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} X_t + \begin{bmatrix} Q \\ 0 \\ 0 \end{bmatrix} \nu_t \quad (21)$$

onde $A = [I + Q\alpha\beta'Q^{-1} + Q\Psi Q^{-1}]$, $B = Q\Psi Q^{-1}$, e $\bar{C} = -Q\alpha\mu_s$. A matriz A é totalmente estimada. Contudo, assume-se que B é diagonal quando $m \leq 5$, tal qual Bowsher e Meeks (2008) e Cajueiro *et al.* (2009).

4 Métodos de Combinação de Previsões

Após a implementação dos modelos individuais, nesse capítulo são implementados dez métodos de combinação organizados em quatro classes: combinações simples (média simples, mediana e média aparada), erro quadrático médio invertido, Mínimos Quadrados Ponderados e combinações via Conjunto de Modelo de Confiança. Nesta última classe, todos os métodos anteriores são aplicados após serem excluídos os piores modelos individuais.

Para testar a capacidade preditiva dos modelos, são separados os dados dentro da amostra (primeiras R observações) e fora da amostra (as últimas P observações). Além disso, alguns esquemas³¹ requerem um período extra fora da amostra (amostra de treinamento) para o cálculo dos pesos usados na combinação das previsões dos modelos individuais. Para tanto, são usadas as primeiras P_0 observações das previsões fora da amostra.

Especificamente, os pesos são calculados usando as previsões individuais fora da amostra, $\{\hat{y}_{t+h|t}(\tau_i)\}_{t=R+P_0}^{t-h}$, e as observações, $\{y_{t+h|t}(\tau_i)\}_{t=R+P_0}^{t-h}$, ambas disponíveis no início do período extra fora da amostra até o tempo t .

Por sua vez, as observações fora da amostra posteriores P_1 são utilizadas para comparar o desempenho de previsão dos modelos individuais e dos esquemas de combinação.

³¹EQM invertido, MQP e combinações via MCS.

Seja C o número de modelos de previsão a serem combinados, com $c = 1, \dots, C$. Assim, a previsão combinada para um horizonte de h meses à frente para as taxas com maturidade τ_i é dada de forma genérica por:

$$\hat{y}_{t+h|t}(\tau_i) = f(\hat{\mathbf{y}}_{t+h|t}(\tau_i); \mathbf{w}_{t+h|t}(\tau_i)), \quad (22)$$

onde $\hat{\mathbf{y}}_{t+h|t}(\tau_i)$ denota o vetor de previsões dos modelos individuais e $\mathbf{w}_{t+h|t}(\tau_i)$ designa os parâmetros ou pesos associados à combinação. Já a função f reduz $\hat{\mathbf{y}}_{t+h|t}(\tau_i)$ na medida resumida $\hat{y}_{t+h|t}(\tau_i)$.

4.1 Combinações Simples

São considerados três métodos simples de combinação de previsões individuais: média simples, mediana e média aparada.

Média Simples: iguala os pesos para todos os modelos individuais, isto é, $w_{t+h|t,c}(\tau_i) = 1/C$. Tem-se que a previsão combinada é dada por:

$$\hat{y}_{t+h|t}(\tau_i) = \sum_{c=1}^C \frac{1}{C} \hat{y}_{t+h|t,c}(\tau_i), \quad (23)$$

Mediana: utiliza a mediana amostral das previsões individuais. Assim, a previsão combinada é:

$$\hat{y}_{t+h|t}(\tau_i) = \text{mediana de } \{\hat{y}_{t+h|t,c}\}_{c=1}^C \quad (24)$$

Média Aparada: iguala a zero ($w_{t+h|t,c}(\tau_i) = 0$) os pesos dos modelos individuais que produzem a maior e menor previsão. Para os modelos restantes, é atribuído $w_{t+h|t,c}(\tau_i) = 1/(C - 2)$, com a previsão combinada sendo:

$$\hat{y}_{t+h|t}(\tau_i) = \sum_{c=1}^{C-2} \frac{1}{C-2} \hat{y}_{t+h|t,c}(\tau_i), \quad (25)$$

Se os dados pertencem à mesma distribuição de probabilidade, a média é a medida de tendência central mais representativa. Contudo, se existem observações afastadas em relação às demais, a mediana e a média aparada são mais eficientes na representação da amostra³².

4.2 Erro Quadrático Médio Invertido

Neste esquema de combinação, os pesos na equação (22) dependem, de forma inversa, da performance relativa histórica da previsão dos modelos individuais. Assim, os pesos são baseados nos EQM invertidos de cada modelo, com relação à soma do EQM invertido de todos os modelos, calculados sobre uma janela de 24 meses anteriores³³.

³²A média aparada e a mediana são medidas resistentes, isto é, mudam pouco se for alterado um pequeno número de observações. Diferem quanto mais assimétrica for a distribuição de probabilidade.

³³A mesma janela considerada na amostra de treinamento usada no cálculo dos pesos dos modelos individuais. Este mesmo número também usado nas combinações por MQP e MCS detalhadas a seguir.

$$w_{t+h|t,c}(\tau_i) = \frac{1/EQM_{h|t,c}^{(\tau_i)}}{\sum_{c=1}^C 1/EQM_{h|t,c}^{(\tau_i)}}$$

onde $w_{t+h|t,c}(\tau_i)$ denota o peso atribuído à previsão para o período t do c –ésimo modelo ($\hat{y}_{t+h|t,c}$) e $EQM_{h|t,c}^{(\tau_i)} = \frac{1}{24} \sum_{p_0=1}^{24} (\hat{y}_{t-p_0+1|t-h-p_0+1,c}^{(\tau_i)} - y_{t-p_0+1}^{(\tau_i)})^2$. Para um modelo com menor EQM é atribuído um maior peso relativo. Portanto, a previsão combinada é obtida por:

$$\hat{y}_{t+h|t}(\tau_i) = \sum_{c=1}^C w_{t+h|t,c}(\tau_i) \hat{y}_{t+h|t,c}(\tau_i), \quad (26)$$

4.3 Mínimos Quadrados Ponderados

Diebold e Pauly (1987) argumentam que, na presença de mudanças estruturais, a combinação de previsões baseadas em pesos variantes no tempo pode promover um melhor resultado. Para implementar o esquema de combinação em questão, os autores adotam uma abordagem de Mínimos Quadrados Ponderados (MQP), detalhada abaixo.

Para a estimação inicial da previsão de combinação MQP para a observação $y_{R+P_0+h}^h(\tau_i)$, as observações passadas $\{y_{s+h}^h(\tau_i)\}_{s=R}^{R+P_0-h}$ são regredidas sobre uma constante e sobre $\{\hat{y}_{c,s+h|s}(\tau_i)\}_{s=R}^{R+P_0-h}$ para $c = 1, \dots, C$. O peso de cada modelo é igualado ao estimador de MQP:

$$\hat{\beta}_{MQP} = (X'W^{-1}X)^{-1}X'W^{-1}y$$

onde W é a matriz ($T \times T$) de média ponderada. Para a construção da segunda previsão de combinação MQP para a observação $y_{R+P_0+h}^h(\tau_i)$, os coeficientes MQP são estimados ao regredir $\{y_{s+h}^h(\tau_i)\}_{s=R+1}^{R+P_0-h+1}$ sobre uma constante e $\{\hat{y}_{c,s+h|s}(\tau_i)\}_{s=R+1}^{R+P_0-h+1}$ para $c = 1, \dots, C$. Este procedimento é repetido até o fim do período fora da amostra. Dessa maneira, os coeficientes calculados por MQP servem como peso na equação de previsão combinada:

$$\hat{y}_{t+h|t}(\tau_i) = w_{0,t} + \sum_{c=1}^C w_{t+h|t,c}(\tau_i) \hat{y}_{t+h|t,c}(\tau_i), \quad (27)$$

Diebold e Pauly (1987) recomendam a utilização de uma matriz diagonal $W = \text{diag}[\psi_{tt}] = kt^\lambda$, onde T é o número de observações usadas na regressão de MQP. Assim, os pesos podem ser declinantes para taxas crescentes, decrescentes ou constantes, dependendo do sinal de $(\lambda - 1)$. De forma similar ao EQM invertido, T é fixo e igual a 24. Já o λ e k utilizados no trabalho são iguais a 1, ou seja, os pesos das observações mais antigas decrescem linearmente. Finalmente, vale salientar que o MQP é o único método de combinação irrestrito, ou seja, os pesos atribuídos a cada modelo individual podem assumir qualquer valor. Nos demais casos, os pesos oscilam entre 0 e 1, com a soma de todos igual a 1.

4.4 Combinações via Modelo de Conjunto de Confiança

Neste esquema de combinação, os modelos com piores performances de previsão

são descartados pela abordagem do Modelo de Conjunto de Confiança³⁴(MCS, em inglês), para então as previsões dos modelos restantes serem combinadas, usando combinação simples, EQM invertido e MQP.

Dado um conjunto de modelos de previsão, C_0 , o procedimento do MCS identifica:

$$\widehat{C}_\alpha^* \subset C_0$$

onde \widehat{C}_α^* é o subconjunto que contém os melhores modelos de previsão dado um nível de confiança $(1 - \alpha)$.

Começando por $C = C_0$ e um vetor P_0 de previsões³⁵, o procedimento do MCS testa repetitivamente a hipótese nula de precisão de previsão igual a:

$$H_{0,C} : E[d_{ij,t}] = 0, \forall i, j \in C$$

onde $d_{ij,t} = L_{i,t} - L_{j,t}$ é o diferencial de perda entre os modelos i e j , com L sendo uma função de perda apropriada. Dessa forma, o procedimento do MCS elimina sequencialmente o modelo com pior desempenho, contanto que a hipótese nula seja rejeitada. Este procedimento é repetido até a hipótese nula não ser mais rejeitada, caso em que o conjunto restante é \widehat{C}_α^* . Para a realização do teste, é usada a seguinte estatística semiquadrática:

$$T_{SQ} \equiv \sum_{i,j \in C} t_{i,j}^2$$

³⁴Desenvolvido por Hansen *et al.* (2003 e 2010).

³⁵No caso $P_0=24$.

onde $t_{i,j} = \frac{\bar{d}_{i,j}}{\sqrt{\hat{\text{var}}(\bar{d}_{i,j})}}$ para $i, j \in C$ e $\bar{d}_{i,j} = \frac{1}{P_0} \sum_{t=1}^{P_0} d_{ij,t}$.

Portanto, para se obter uma previsão combinada h -meses à frente, a estatística T_{SQ} será usada para construir $\hat{C}_{0,25}^*$, utilizando um nível de confiança de $(1 - \alpha) = 75\%$ ³⁶ tal como em De Pooter *et al.* (2010). Para se obter a distribuição de probabilidade sob a hipótese nula, é utilizada a implementação por *bootstrap* através de blocos estacionários³⁷, com 1.000 reamostragens, extensão do bloco igual a 20 e o EQM como função de perda.

³⁶O conjunto de modelos é baseado no nível de significância de 75%, uma vez que em muitos casos os erros de previsão são relativamente pequenos. Assim, níveis de significância tradicionais (como 5%) não são muitos informativos, pois poucos ou nenhum modelo são excluídos do MCS.

³⁷Desenvolvido por Politis e Romano (1994).

5 Descrição dos Dados

A base de dados a ser utilizada neste trabalho para a construção da ETTJ será obtida das taxas de juros dos contratos de *swap* DI-Pré da BM&FBovespa³⁸. Neste contrato, uma parte se compromete a pagar uma taxa flutuante (DI de um dia) e em troca recebe uma taxa fixa, definida no momento da celebração do contrato. O inverso ocorre com sua contraparte. A troca de fluxo de caixa resultante é calculada sobre um mesmo principal definido pela BM&FBovespa. Por sua vez, ao atuar como contraparte central na negociação, a BM&FBovespa exige depósito de margem de garantia das partes, o que permite interpretar as taxas de juros do contrato de *swap* como taxa livre de risco de crédito.

As maturidades consideradas dos *swaps* serão de 1, 2, 3, 6, 9, 12, 15, 18, 21, 24 e 30 meses. O período analisado da amostra vai de dezembro de 1997 a julho de 2010. Cada *swap* é observado diariamente (3112 observações), com os dados mensais construídos por meio das médias diárias (152 observações). Apesar das previsões das taxas obtidas nesse trabalho representarem a média mensal, o uso de dados diários foi utilizado pois o modelo FSN-ECM necessita de mais dados para ser estimado.

Apesar de os trabalhos da literatura internacional de referência utilizarem como

³⁸Ao contrário do que ocorre nos EUA, onde os preços dos títulos é que são observados, as taxas de juros no mercado brasileiro são diretamente negociadas. Desse modo, não há necessidade de estimar a ETTJ brasileira a partir dos preços dos títulos

dados as taxas dos títulos públicos, o uso do *swap* DI-Pré possui justificativas para o presente trabalho. Primeiro, para o período considerado não há dados disponíveis suficientemente antigos para os títulos públicos. Segundo, há maior disponibilidade de maturidades de menor prazo nos *swaps*. Terceiro, diferente dos títulos públicos, o fato de os *swaps* possuírem maturidade fixa dispensa a necessidade de um método de interpolação das taxas. Quarto, em contraste com a maioria dos mercados internacionais, a formação da taxa da ETTJ se dá primariamente no mercado de derivativos e não no mercado de títulos públicos, possuindo aquele maior liquidez. Por fim, o DI diário³⁹ apresenta alto grau de associação com a Selic diária⁴⁰, o que permite adotar o DI diário como *proxy* da Selic diária.

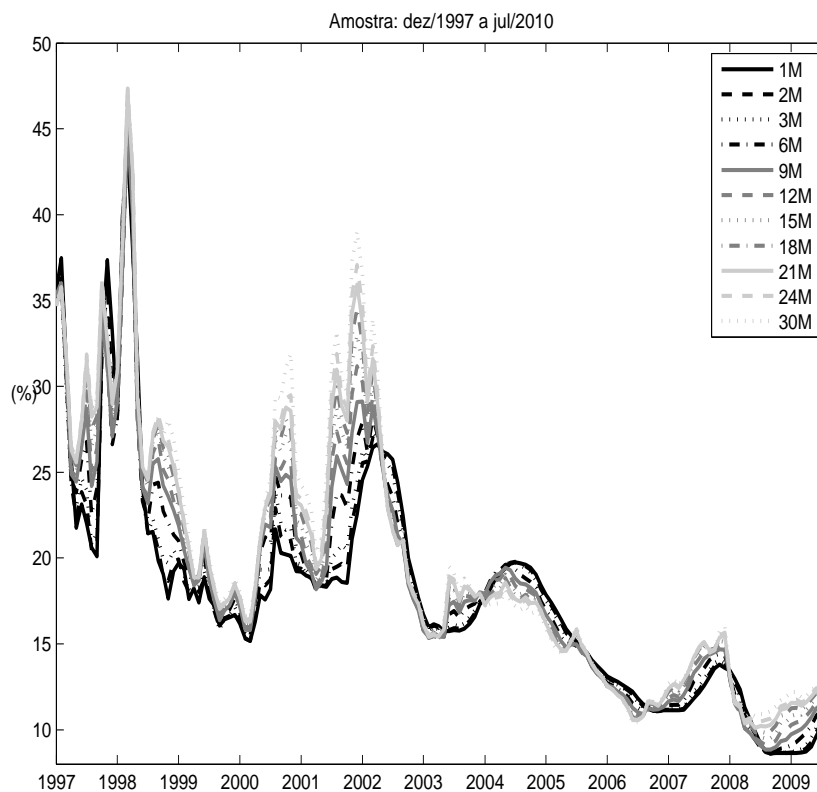
Por conseguinte, as únicas variáveis usadas para representar a ETTJ no Brasil serão as taxas de juros dos *swaps* DI-Pré da BM&FBovespa, como é feito em Vicente e Tabak (2008) e Cajueiro *et al.* (2009).

A figura 1 mostra a dinâmica das taxas de DI-Pré consideradas nesse trabalho. Nela verifica-se o período de elevada volatilidade até março de 1999, marcado pelas crises da Ásia e da Rússia e pela mudança do regime cambial (de fixo para flutuante) no Brasil. Destaca-se também o maior diferencial das taxas dos vencimentos

³⁹Taxa média de um dia de empréstimo de reservas entre instituições financeiras.

⁴⁰Taxa média de um dia das operações de empréstimos de reservas entre as instituições financeiras e entre as instituições financeiras e o Banco Central do Brasil, ambas com lastro em títulos públicos. É a taxa básica utilizada para a formação EETJ de títulos públicos no Brasil.

Figura 1: Evolução das Taxas de *swap* DI-Pré



Fonte: BM&FBovespa.

curtos e longos entre meados de 2002 e início de 2003, reflexos da incerteza eleitoral que marcou a corrida presidencial do ano de 2002. A partir de 2003, observa-se tendência de redução de nível da curva de juros, com especial contribuição da percepção pelos agentes econômicos da manutenção do “tripé” da política econômica no país estabelecido no ano de 1999: câmbio flutuante, meta de inflação e superávit primário.

Nos anos mais recentes vale destacar a elevação de nível da ETTJ, no final de

2008, em função da crise financeira internacional e a posterior redução, no ano de 2009, com o processo de afrouxamento monetário implementado pelo Banco Central do Brasil. A série de dados encerra-se em julho de 2010, mês do fim do posterior ciclo de elevação na taxa de curto prazo com o objetivo de conter as pressões inflacionárias.

Por sua vez, a tabela 1 representa estatísticas descritivas da ETTJ brasileira para toda amostra considerada. A partir desta tabela é possível estabelecer algumas observações. A curva de juros média possui inclinação positiva, em consonância com resultados encontrados na literatura. As volatilidades (desvio padrão) das taxas dos *swaps* negociadas também crescem com a maturidade. Este resultado difere do observado para a curva de juros nos EUA, cuja volatilidade decresce (para as mesmas maturidades analisadas neste trabalho).

Já a análise das autocorrelações (três últimas colunas da tabela 1) reflete a persistência da dinâmica das taxas. Por fim, a amplitude das taxas indica a elevada mudança de nível da curva de juros durante o período analisado.

Tabela 1: Estatísticas Descritivas das Taxas de *swap* DI-Pré

Maturidade	Média	Dpad	Min	Max	$\rho(1)$	$\rho(12)$	$\rho(24)$
1m	17,81	6,67	8,61	43,95	0,92	0,40	0,14
2m	17,84	6,57	8,62	43,59	0,92	0,41	0,16
3m	17,93	6,56	8,62	44,86	0,92	0,42	0,17
6m	18,27	6,66	8,69	44,88	0,93	0,46	0,20
9m	18,59	6,87	8,84	45,01	0,93	0,47	0,20
12m	18,83	7,02	9,10	45,64	0,93	0,47	0,20
15m	19,14	7,26	9,47	46,19	0,93	0,47	0,20
18m	19,39	7,45	9,82	46,92	0,94	0,47	0,19
21m	19,58	7,60	10,10	47,23	0,94	0,47	0,18
24m	19,75	7,73	10,37	47,37	0,94	0,47	0,18
30m	20,04	7,95	10,43	47,43	0,94	0,48	0,18

Nota: O período da amostra vai de dezembro/1997 a julho/2010.

6 Resultados

6.1 Procedimentos de Previsão

Assim como em Cajueiro *et al.* (2009), para realizar as previsões fora da amostra (*out of sample*) são estimados os modelos usando a primeira metade da amostra (*in sample*), ou seja, uma janela de estimação móvel de 76 meses. A primeira janela de estimação corresponde a 12/1997-03/2004. Em seguida, estes modelos são estimados novamente, no próximo mês, excluindo o primeiro mês da amostra anterior. Os parâmetros são atualizados até a última janela de estimação, 04/2001-07/2007. As previsões são obtidas para os seguintes horizontes à frente: 1, 3, 6 e 12 meses. Os modelos individuais fazem previsões diárias e depois convertem para a média mensal.

Assim como Almeida *et al.* (2008), cada mês à frente corresponde a 21 dias úteis. Dentro do período fora da amostra, os primeiros 24 meses são usados para obtenção dos pesos dos modelos individuais utilizados nas combinações de previsão, (amostra de treinamento P_0 descrita no capítulo 4). Os modelos individuais são comparados ao modelo de passeio aleatório, usado como referencial na literatura de previsão da curva de juros.

6.2 Procedimentos de Avaliação dos Resultados

Para avaliação de todo o período fora da amostra é calculada, para cada horizonte de previsão e maturidade, a Raiz do Erro Quadrático Médio.

$$REQM_c^{(\tau_i)} = \sqrt{\frac{1}{P_1} \sum_{t=1}^{P_1} (\hat{y}_{t+h|t,c}^{(\tau_i)} - y_{t+h}^{(\tau_i)})^2} \quad (28)$$

A precisão das previsões é avaliada pela Raiz do Erro Quadrático Médio Retalivo (ReREQM).

$$ReREQM_c^{(\tau_i)} = \frac{REQM_c^{(\tau_i)}}{REQM_{PA}^{(\tau_i)}} \quad (29)$$

De maneira similar a Hordal *et al.* (2006) e a De Pooter *et al.* (2010), os desempenhos de cada modelo e método de combinação são resumidos pelo uso do Traço da Raiz do Erro Quadrático Médio, proposto por Christoffersen e Diebold (1998).

$$TREQM_c = \sqrt{\frac{1}{N} \frac{1}{P_1} \sum_{i=1}^N \sum_{t=1}^{P_1} (\hat{y}_{t+h|t,c}^{(\tau_i)} - y_{t+h}^{(\tau_i)})^2} \quad (30)$$

A fim de testar se os diferenciais dos erros de previsão dos modelos são estatisticamente significantes entre si é usada a abordagem MCS descrita na seção 4.4. A mudança relevante aqui é no número de observações utilizadas, P_1 em vez de P_0 . Isto porque P_1 que são as observações fora da amostra efetivamente utilizadas para comparar o desempenho de previsão dos modelos e dos esquemas.

Com o intuito de analisar em que parte da amostra (meses) os modelos apresentaram desempenhos superiores ou inferiores ao passeio aleatório é exposto graficamente do Erro Quadrático Acumulado (EQA) para cada período acumulado T :

$$EQA_{c,T}^{(\tau_i)} = \sum_{t=1}^T [(\hat{y}_{t+h|t,PA}^{(\tau_i)} - y_{t+h}^{(\tau_i)})^2 - (y_{t+h|t,c}^{(\tau_i)} - y_{t+h}^{(\tau_i)})^2] \quad (31)$$

Esta análise foi usada em Welch e Goyal (2008) para ações e adaptada para taxas de juros por De Pooter *et al.* (2010).

6.3 Resultados dos Modelos Individuais

A discussão dos desempenhos dos modelos individuais é iniciada ao considerar os resultados em termos (do Traço) da Raiz do Erro Quadrático Médio Relativo, (T)ReREQM, disponibilizados para cada horizonte de previsão nas tabelas 2 a 5. A primeira linha de cada tabela contém o (T)REQM das previsões do passeio aleatório. As linhas restantes nos Painéis A, B e C⁴¹ reportam os (T)ReREQM dos demais modelos em relação ao passeio aleatório, de acordo com a equação (29).

Assim, um número menor que 1 indica que o modelo considerado supera o passeio aleatório, enquanto um número maior que 1 sinaliza que o modelo é superado pelo passeio aleatório. Os asteriscos à direita dos ReREQM individuais sinalizam que o modelo pertence ao conjunto do modelo $\hat{C}_{0,25}^*$ e os dados em parênteses correspondem aos percentuais de vezes que cada modelo individual é incluído em $\hat{C}_{0,25}^*$.

No que diz respeito às previsões com horizonte um mês à frente, os modelos individuais superaram o passeio aleatório para quase a totalidade das maturidades analisadas. Exceções apenas para algumas maturidades do DNS e do VAR. Desta-

⁴¹Para análise de previsão dos modelos individuais interessa somente o Painel A.

que para o melhor desempenho em geral para o modelo AR. Contudo, nos demais horizontes de previsão ($h=3$, $h=6$ e $h=12$) nenhum modelo de previsão individual produz previsão superior ao passeio aleatório de forma sistemática. Nestes horizontes, os modelos AR e DNS apresentaram desempenhos regulares, enquanto o VAR e o FSN-ECM foram superados pelo passeio aleatório.

Em relação ao teste de significância da abordagem MCS, vale salientar que embora os modelos AR e DNS apresentem $ReREQM$ acima de 1 para horizontes a partir de três meses, os resultados dos testes sugerem não ser possível afirmar a superioridade do passeio aleatório em várias maturidades, em especial nas mais longas. Todavia, apesar de não ser possível rejeitar a não superioridade do DNS, o resultado difere em parte de Cajueiro *et al.* (2009). Tal fato demonstra que os desempenhos dos modelos podem variar ao longo do tempo, resultado defendido no campo teórico por Timmermann (2006) e obtido para estudos da ETTJ americana⁴².

Analisando sob a ótica das maturidades específicas, verifica-se que os modelos FSN-ECM e VAR apresentam melhores desempenhos nas maturidades e nos horizontes mais curtos. Por outro lado, o modelo AR demonstra bom desempenho nas maturidades mais longas até o horizonte de seis meses. Já o modelo DNS possui melhor performance nos horizontes e nas maturidades mais longos.

⁴²Como em Bowsher e Meeks (2008), e De Pooter *et al.* (2010).

Tabela 2: (Traço) Raiz do Erro Quadrático Médio Relativo, horizonte de um mês

Modelos	(T)REQM	1m	2m	3m	6m	9m	12m	15m	18m	21m	24m	30m
PA	0,54	0,39	0,39	0,40	0,43	0,48	0,53	0,57	0,60	0,63	0,66	0,70
		(0,00)	(0,00)	(0,07)	(0,10)	(0,05)	(0,05)	(0,07)	(0,07)	(0,12)	(0,22)	(0,20)
Painel A: Modelos Individuais												
AR	0,75	0,78	0,71	0,69	0,74*	0,75*	0,76*	0,76*	0,76*	0,76*	0,76*	0,76*
		(0,00)	(0,02)	(0,34)	(0,34)	(0,56)	(0,66)	(0,66)	(0,68)	(0,66)	(0,66)	(0,76)
VAR	0,90	0,50	0,46*	0,51*	0,80	0,97	1,02	1,02	0,99	0,95	0,92	0,91
		(0,56)	(0,63)	(0,54)	(0,02)	(0,02)	(0,07)	(0,15)	(0,17)	(0,17)	(0,27)	(0,07)
DNS	0,85	1,00	0,91	0,87	0,84	0,86	0,88	0,88	0,86	0,82	0,79	0,79
		(0,24)	(0,44)	(0,56)	(0,61)	(0,56)	(0,51)	(0,59)	(0,59)	(0,68)	(0,71)	(0,83)
FSN-ECM	0,82	0,44*	0,48	0,54*	0,74*	0,86	0,92	0,91	0,89	0,86	0,85	0,87
		(0,73)	(0,37)	(0,27)	(0,17)	(0,12)	(0,07)	(0,07)	(0,07)	(0,12)	(0,12)	(0,12)
Painel B: Combinações de Previsões												
MS	0,79	0,67	0,63	0,64	0,75	0,81	0,84	0,85	0,84	0,81	0,80	0,80
ME	0,76	0,66	0,62	0,62	0,74	0,78	0,80	0,79	0,79	0,77	0,77	0,78
MA	0,77	0,65	0,61	0,62	0,73	0,78	0,82	0,82	0,81	0,79	0,78	0,79
EQM	0,77	0,55	0,54	0,58	0,73	0,80	0,83	0,83	0,83	0,80	0,79	0,79
MQP	0,91	0,53	0,60	0,55	0,73	0,85	0,98	0,94	0,98	0,95	0,97	1,08
Painel C: Combinações de Previsões via MCS												
MCS-MS	0,75	0,49	0,47	0,56	0,69	0,76	0,80	0,84	0,81	0,78	0,79	0,79
MCS-ME	0,75	0,48	0,47	0,57	0,70	0,76	0,78	0,82	0,81	0,77	0,78	0,79
MCS-MA	0,75	0,48	0,47	0,57	0,70	0,76	0,78	0,82	0,81	0,77	0,78	0,79
MCS-EQM	0,75	0,49	0,47	0,56	0,69	0,76	0,80	0,83	0,81	0,78	0,79	0,79
MCS-MQP	0,77	0,59	0,62	0,68	0,71	0,92	0,84	0,65	0,73	0,75	0,79	0,87

Nota: O período fora da amostra para o horizonte de um mes usado na tabela corresponde a 04/2006-08/2009. O período de treinamento usado nos esquemas de combinação compreende uma janela móvel de 24 meses, com início em 04/2004-03/2006 e término em 08/2007-07/2009. Já os parâmetros dos modelos individuais são estimados usando uma janela de móvel de 76 meses, com início 12/1997-03/2004 e término em 04/2001-07/2007.

Tabela 3: (Traço) Raiz do Erro Quadrático Médio Relativo, horizonte de três meses

Modelos	(T)REQM	1m	2m	3m	6m	9m	12m	15m	18m	21m	24m	30m
PA	1,29	1,09	1,10*	1,12*	1,18*	1,25*	1,31*	1,37*	1,39*	1,41*	1,42*	1,46*
		(0,63)	(0,44)	(0,49)	(0,76)	(0,83)	(0,85)	(0,83)	(0,83)	(0,83)	(0,83)	(0,83)
Painel A: Modelos Individuais												
AR	1,06	1,15	1,09	1,07*	1,10*	1,08*	1,07*	1,05*	1,04*	1,04*	1,03*	1,03*
		(0,00)	(0,02)	(0,34)	(0,46)	(0,61)	(0,66)	(0,66)	(0,68)	(0,66)	(0,66)	(0,85)
VAR	1,48	0,81	0,91*	1,02*	1,34	1,49	1,56	1,59	1,63	1,65	1,67	1,70
		(0,59)	(0,68)	(0,61)	(0,10)	(0,02)	(0,07)	(0,15)	(0,17)	(0,17)	(0,27)	(0,07)
DNS	1,17	1,34	1,31	1,28*	1,24*	1,20*	1,17	1,14	1,12	1,10	1,08	1,07*
		(0,78)	(0,80)	(0,83)	(0,88)	(0,80)	(0,73)	(0,78)	(0,71)	(0,73)	(0,78)	(0,83)
FNS-ECM	1,44	0,78*	0,91*	1,03*	1,30*	1,45	1,53	1,55	1,57	1,58	1,61	1,64
		(0,76)	(0,37)	(0,27)	(0,17)	(0,12)	(0,07)	(0,07)	(0,07)	(0,12)	(0,12)	(0,12)
Painel B: Combinações de Previsões												
MS	1,12	0,91	0,94	0,98	1,09	1,14	1,16	1,16	1,17	1,17	1,17	1,18
ME	1,05	0,97	1,04	1,06	1,11	1,09	1,08	1,06	1,05	1,04	1,04	1,03
MA	1,11	0,89	0,94	0,98	1,09	1,13	1,15	1,15	1,15	1,15	1,15	1,16
EQM	1,03	0,85	0,91	0,96	1,05	1,07	1,08	1,07	1,07	1,06	1,06	1,05
MQP	1,25	0,89	0,80	0,74	0,94	1,31	1,39	1,41	1,40	1,33	1,34	1,42
Painel C: Combinações de Previsões via MCS												
MCS-MS	0,97	0,83	0,83	0,92	1,04	0,99	0,99	1,00	0,99	0,99	0,99	0,99
MCS-ME	0,96	0,83	0,84	0,89	1,03	0,99	0,98	1,00	0,99	0,99	0,98	0,98
MCS-MA	0,96	0,83	0,84	0,89	1,03	0,99	0,98	1,00	0,99	0,99	0,98	0,98
MCS-EQM	0,97	0,83	0,83	0,91	1,03	0,99	0,99	1,00	0,99	0,99	0,99	0,99
MCS-MQP	1,10	0,92	0,73	0,82	1,01	1,25	1,24	1,19	1,13	1,10	1,12	1,19

Nota: O período fora da amostra para o horizonte de três meses usado na tabela corresponde a 06/2006-10/2009. O período de treinamento usado nos esquemas de combinação compreende uma janela móvel de 24 meses, com início em 06/2004-05/2006 e término em 10/2007-09/2009. Já os parâmetros dos modelos individuais são estimados usando uma janela de móvel de 76 meses, com início 12/1997-03/2004 e término em 04/2001-07/2007.

Tabela 4: (Traço) Raiz do Erro Quadrático Médio Relativo, horizonte de seis meses

Modelos	(T)REQM	1m	2m	3m	6m	9m	12m	15m	18m	21m	24m	30m
PA	2,16	1,96*	1,97*	1,99*	2,09*	2,18*	2,24 *	2,29*	2,29*	2,27*	2,24*	2,20*
		(0,76)	(0,83)	(0,88)	(0,85)	(0,83)	(0,85)	(0,83)	(0,83)	(0,83)	(0,83)	(0,83)
Painel A: Modelos Individuais												
AR	1,15	1,24	1,20*	1,18*	1,20*	1,17*	1,15*	1,12*	1,12*	1,10*	1,10*	1,08*
		(0,00)	(0,12)	(0,49)	(0,51)	(0,61)	(0,66)	(0,66)	(0,68)	(0,66)	(0,66)	(0,85)
VAR	1,77	1,15	1,25*	1,35*	1,58	1,70	1,77	1,83	1,91	1,98	2,06	2,20
		(0,59)	(0,78)	(0,61)	(0,12)	(0,12)	(0,07)	(0,15)	(0,17)	(0,17)	(0,27)	(0,07)
DNS	1,15	1,32	1,30*	1,27*	1,21*	1,16*	1,11*	1,08*	1,07*	1,06*	1,06*	1,07*
		(0,83)	(0,93)	(0,93)	(0,90)	(0,88)	(0,85)	(0,95)	(0,93)	(0,93)	(0,95)	(0,90)
FSN-ECM	1,69	1,11*	1,23*	1,33*	1,53	1,65	1,72	1,76	1,82	1,88	1,96	2,08
		(0,93)	(0,41)	(0,29)	(0,17)	(0,12)	(0,07)	(0,07)	(0,07)	(0,12)	(0,12)	(0,12)
Painel B: Combinações de Previsões												
MS	1,24	1,08	1,11	1,14	1,21	1,24	1,25	1,25	1,27	1,29	1,31	1,35
ME	1,15	1,14	1,16	1,17	1,21	1,18	1,17	1,14	1,13	1,12	1,12	1,11
MA	1,24	1,08	1,12	1,15	1,22	1,24	1,25	1,25	1,27	1,28	1,30	1,33
EQM	1,11	1,07	1,10	1,13	1,16	1,15	1,13	1,11	1,10	1,09	1,08	1,07
MQP	0,85	0,81	0,60	0,56	0,71	0,83	0,85	0,89	0,91	0,95	0,97	1,00
Painel C: Combinações de Previsões via MCS												
MCS-MS	1,05	1,12	1,15	1,13	1,15	1,11	1,03	1,00	0,98	0,97	0,97	0,97
MCS-ME	1,04	1,13	1,14	1,12	1,13	1,09	1,03	1,00	0,99	0,97	0,98	0,97
MCS-MA	1,04	1,13	1,14	1,12	1,13	1,09	1,03	1,00	0,99	0,97	0,98	0,97
MCS-EQM	1,04	1,11	1,14	1,13	1,15	1,10	1,02	1,00	0,98	0,97	0,97	0,97
MCS-MQP	0,92	0,88	0,85	0,76	0,82	0,87	0,96	0,95	0,95	1,06	0,96	0,92

Nota: O período fora da amostra para o horizonte de seis meses usado na tabela corresponde a 09/2006-01/2010. O período de treinamento usado nos esquemas de combinação compreende uma janela móvel de 24 meses, com início em 09/2004-08/2006 e término em 01/2008-12/2009. Já os parâmetros dos modelos individuais são estimados usando uma janela de móvel de 76 meses, com início 12/1997-03/2004 e término em 04/2001-07/2007.

Tabela 5: (Traço) Raiz do Erro Quadrático Médio Relativo, horizonte 12 meses

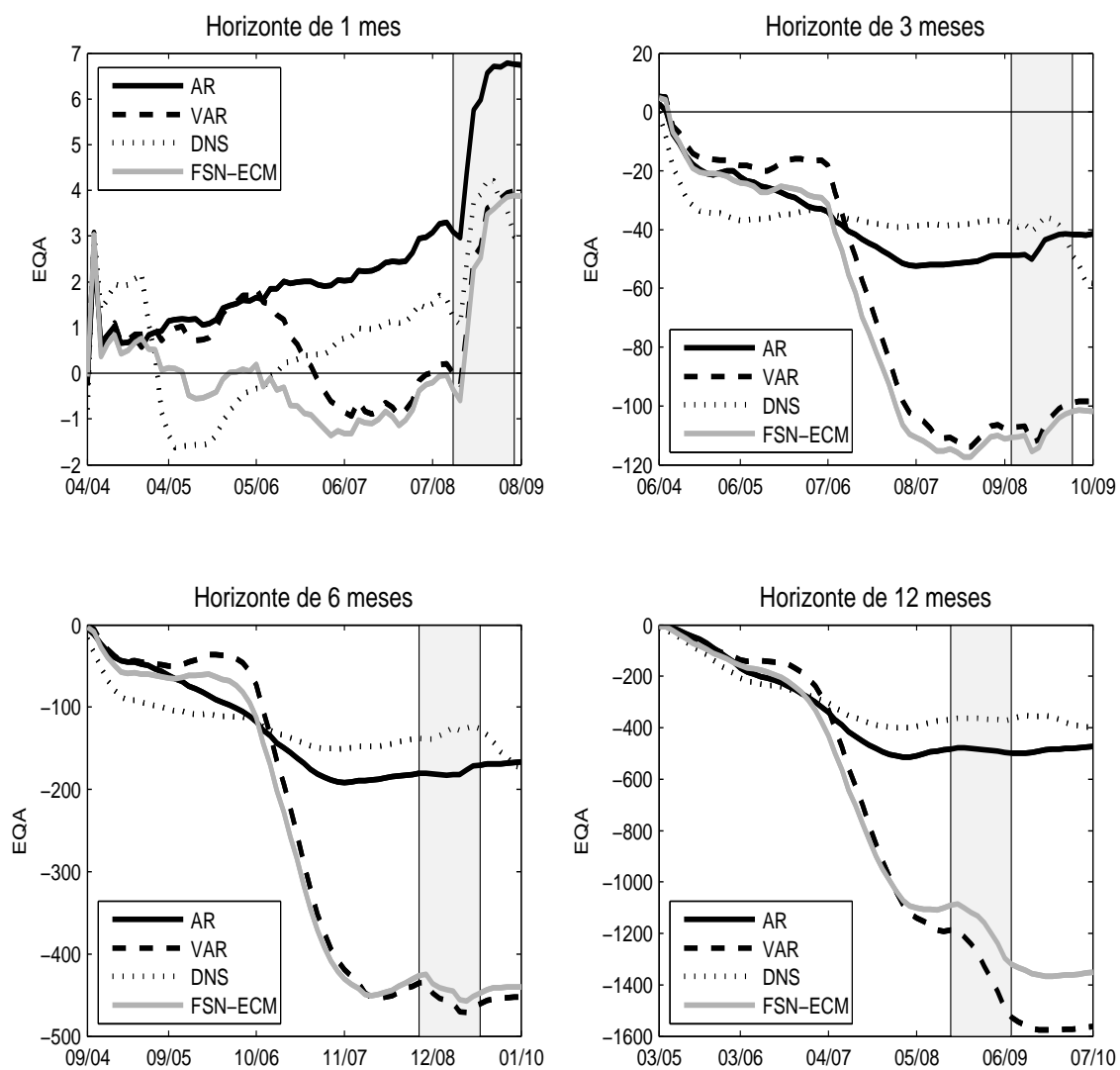
Modelos	(T)REQM	1m	2m	3m	6m	9m	12m	15m	18m	21m	24m	30m
PA	2,97	2,83*	2,84*	2,86*	2,96*	3,04*	3,09*	3,11*	3,08*	3,02*	2,96*	2,89*
		(0,98)	(0,93)	(0,98)	(0,95)	(0,93)	(0,88)	(0,83)	(0,83)	(0,83)	(0,83)	(0,83)
Painel A: Modelos Individuais												
AR	1,21	1,3	1,27	1,26	1,27	1,24	1,21	1,18	1,17	1,16	1,16	1,14*
		(0,12)	(0,17)	(0,49)	(0,51)	(0,61)	(0,66)	(0,66)	(0,68)	(0,66)	(0,66)	(0,85)
VAR	2,16	1,67	1,75	1,82	1,97	2,05	2,11	2,19	2,28	2,39	2,50	2,70
		(0,59)	(0,78)	(0,61)	(0,12)	(0,12)	(0,07)	(0,15)	(0,17)	(0,17)	(0,27)	(0,07)
DNS	1,15	1,29	1,27	1,24	1,17	1,12	1,09*	1,07*	1,07*	1,08*	1,10*	1,13*
		(0,83)	(0,93)	(0,93)	(0,95)	(0,95)	(0,88)	(0,95)	(0,93)	(0,93)	(0,95)	(0,90)
FSN-ECM	1,95	1,56	1,64	1,70	1,82	1,88	1,93	1,97	2,04	2,13	2,22	2,38
		(0,98)	(0,46)	(0,29)	(0,17)	(0,12)	(0,07)	(0,07)	(0,07)	(0,12)	(0,12)	(0,12)
Painel B: Combinações de Previsões												
MS	1,38	1,28	1,30	1,32	1,36	1,36	1,36	1,36	1,39	1,42	1,45	1,51
ME	1,23	1,30	1,30	1,30	1,30	1,26	1,23	1,19	1,18	1,17	1,17	1,16
MA	1,37	1,29	1,32	1,33	1,36	1,36	1,36	1,35	1,37	1,39	1,42	1,47
EPQM	1,17	1,24	1,25	1,25	1,23	1,19	1,16	1,14	1,12	1,11	1,11	1,10
MQP	0,57	0,59	0,55	0,52	0,53	0,55	0,57	0,59	0,59	0,59	0,58	0,56
Painel C: Combinações de Previsões via MCS												
MCS-MS	1,04	1,25	1,21	1,19	1,02	0,99	0,97	0,97	0,96	0,97	0,97	0,99
MCS-ME	1,04	1,26	1,20	1,15	1,01	0,99	0,97	0,97	0,96	0,97	0,97	0,99
MCS-MA	1,04	1,26	1,20	1,15	1,01	0,99	0,97	0,97	0,96	0,97	0,97	0,99
MCS-EQMP	1,04	1,24	1,20	1,18	1,02	0,99	0,97	0,97	0,96	0,97	0,97	0,99
MCS-MQP	0,88	0,85	0,79	0,74	1,13	1,04	0,91	0,84	0,83	0,82	0,82	0,82

Nota: O período fora da amostra para o horizonte de 12 meses usado na tabela corresponde a 03/2007-07/2010. O período de treinamento usado nos esquemas de combinação compreende uma janela móvel de 24 meses, com início em 03/2005-02/2007 e término em 07/2008-06/2010. Já os parâmetros dos modelos individuais são estimados usando uma janela de móvel de 76 meses, com início 12/1997-03/2004 e término em 04/2001-07/2007.

A análise gráfica do EQA é iniciada aproveitando o fato de que o mesmo foi gerado para todo o período fora da amostra. Como cita De Pooter *et al.* (2010), esta opção ajuda a explicitar o desempenho dos modelos individuais no período de treinamento inicial. A figura 2 mostra os gráficos de cada horizonte de previsão analisado. Cada linha no gráfico representa um modelo e ilustra o desempenho de determinado modelo em comparação ao passeio aleatório. EQA crescente indica que o modelo está superando o passeio aleatório, enquanto o contrário indica que o modelo está produzindo erros de previsão superiores ao passeio aleatório. Foi usado o Traço do Erro Quadrático Acumulado (TEQA) de modo a diminuir o número de gráficos a serem exibidos.

De forma geral, a análise do período de treinamento (P_0) sugere uma reduzida volatilidade no desempenho dos modelos, com exceção do DNS no horizonte de um mês à frente. Nos períodos posteriores (P_1), pode-se chegar a algumas considerações. Primeiro, a performance bem inferior dos modelos VAR e FSN-ECM a partir do final de 2006 para horizontes mais longos. Em segundo lugar, desempenho superior de todos os modelos em relação ao passeio aleatório no período da crise financeira internacional (final de 2008 a meados de 2009) para o horizonte de um mês. Por fim, o modelo DNS apresentou erro de previsão maior do que o passeio aleatório após o fim do auge da crise internacional (meados de 2009) para os horizontes de três e seis meses à frente.

Figura 2: Traço do Erro Quadrático Acumulado (inclui período de treinamento)



6.4 Resultados dos Métodos de Combinação

A análise dos resultados das previsões dos esquemas de combinação também tem como ponto de partida as tabelas 2 a 5, especialmente os painéis B e C. O Painel B refere-se às combinações aplicadas diretamente. O Painel C relaciona os mesmos métodos de combinações, porém, aplicados posteriormente à eliminação dos piores modelos individuais pela abordagem MCS. Estas tabelas permitem obter algumas considerações.

Primeiramente, em cada horizonte de previsão foi sempre possível encontrar pelo menos um método de combinação que tenha apresentado desempenho igual ou superior ao melhor modelo individual. Adicionalmente, quanto mais longo o horizonte, maior foi a contribuição das combinações. Nos horizontes mais curtos, as combinações simples e o EQM invertido aparadas por MCS tiveram desempenhos similares ao modelo AR em um mês e superior ao passeio aleatório três meses à frente. Nos horizontes mais longos, verifica-se que as combinações por MQP (aparada ou não por MCS) apresentam performances superiores ao passeio aleatório.

Em segundo lugar, ao se comparar os Painéis B e C verifica-se bom desempenho das combinações após o uso da abordagem MCS. Com exceção da combinação por MQP nos horizontes mais longos, para cada esquema de combinação foram encontrados resultados superiores de previsão ao aplicar os métodos posteriormente à exclusão dos modelos individuais. As percentagens inclusas nas tabelas (em

parênteses) indicam que na maior parte das vezes os esquemas pela abordagem MCS selecionam os modelos individuais com melhores desempenhos. Outra característica interessante é que nos horizontes mais curtos, onde as combinações via MCS apresentaram os melhores resultados, a diferença entre as combinações foi mínima. Tal fato permite concluir nestes casos que os procedimentos de exclusão dos piores modelos individuais por MCS foram mais benéficos para previsões do que o modo como os pesos foram calculados. De forma semelhante, Timmermann (2006) afirma que a exclusão dos piores modelos geralmente incrementa a performance das previsões.

Em terceiro lugar, a combinação por MQP (sem MCS), o único esquema de combinação irrestrito, apresentou desempenho bem distinto a depender do horizonte de previsão considerado. Enquanto nos horizontes mais curtos este esquema de combinação apresentou maior erro de previsão, nos horizontes mais longos foi o método com melhor desempenho, com destaque ainda maior no horizonte 12 meses à frente. Neste horizonte (justamente o mais longo), a superação de desempenho em relação ao passeio aleatório superou 40% em todas as maturidades. Esta característica de melhor desempenho em horizontes mais longos de um esquema de combinação irrestrito é semelhante a encontrada por Bolder e Romanyuk (2008) para os dados de taxas de juros do Canadá ao aplicar o MQO simples. Como aponta Timmermann (2006), os esquemas que utilizam regressões podem ajustar eventuais vieses dos modelos por meio do termo do intercepto. Vale ressaltar que, embora seja a combinação por MQP eficaz nos horizontes mais longos, apresenta a desvantagem

de ser mais sensível ao período de treinamento usado na determinação dos pesos.

A análise das figuras 3 a 6 também permite obter conclusões adicionais. As figuras compreendem o mesmo período das tabelas 2 a 5 de modo a confrontá-los de forma mais eficaz. Os erros dos esquemas de previsão são certamente menos voláteis que os modelos individuais. Dentre os esquemas de combinação, destaque é dado às combinações via MCS, cujos erros de previsão são ainda menos voláteis. Em relação aos eventuais benefícios do uso de parâmetros variantes no tempo no caso de quebras estruturais, foram encontradas evidências de melhoria de previsão nos horizontes curtos da combinação por EMQ invertido e nos longos da combinação por MQP (em ambos os casos, com ou sem exclusão por MCS) no período da crise financeira internacional (2008/2009). Outro ponto relevante é que o mau desempenho dos modelos VAR e FSN-ECM nos horizontes a partir do horizonte de três meses é mitigado com a introdução dos métodos de combinação⁴³. Esta constatação corrobora com argumento de Timmermann (2006) de que as combinações podem ser interpretadas como uma forma de tornar as previsões mais robustas contra os vieses de especificações dos modelos individuais.

Finalmente, as figuras 7 a 22 no anexo mostram as taxas observadas contra as taxas previstas para os períodos reportados nas tabelas 2 a 5, tanto para os modelos individuais como para os esquemas de combinação.

⁴³Em particular, nos esquemas com exclusão dos piores modelos por MCS os erros de previsão do VAR e do FSN-ECM são na maior parte das vezes praticamente dirimidos.

Figura 3: Traço do Erro Quadrático Acumulado, horizonte de um mês

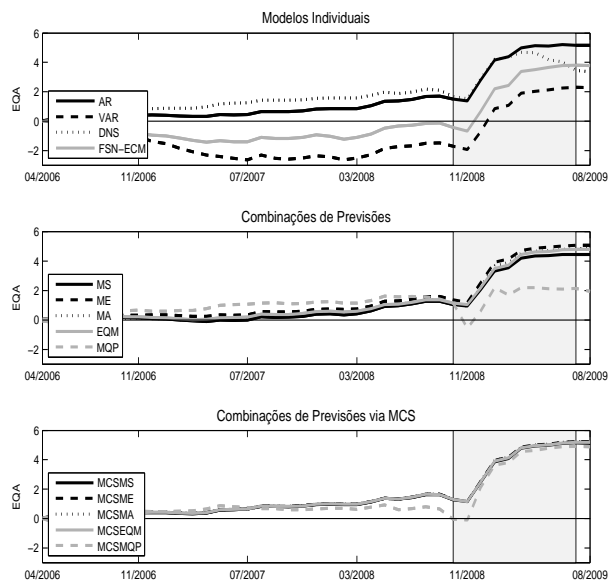


Figura 4: Traço do Erro Quadrático Acumulado, horizonte de três meses

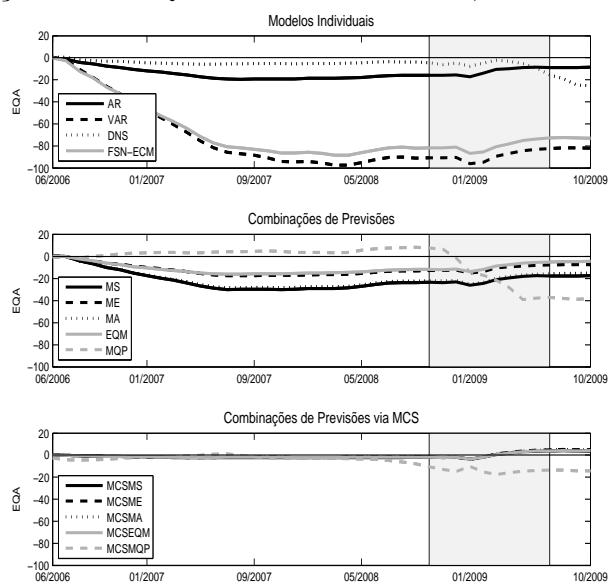


Figura 5: Traço do Erro Quadrático Acumulado, horizonte de seis meses

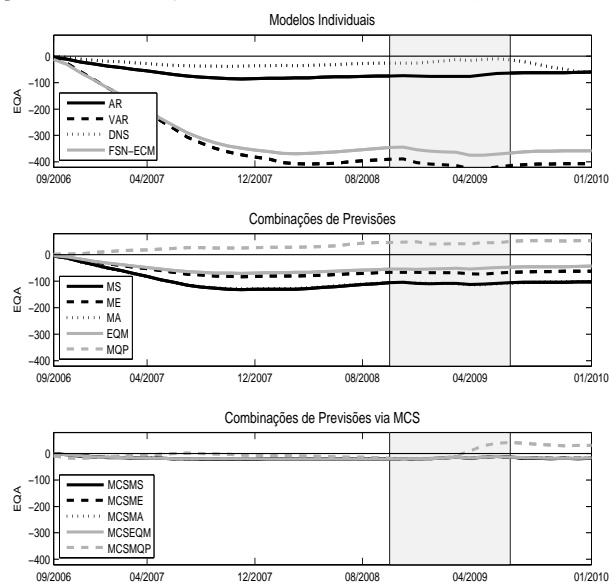
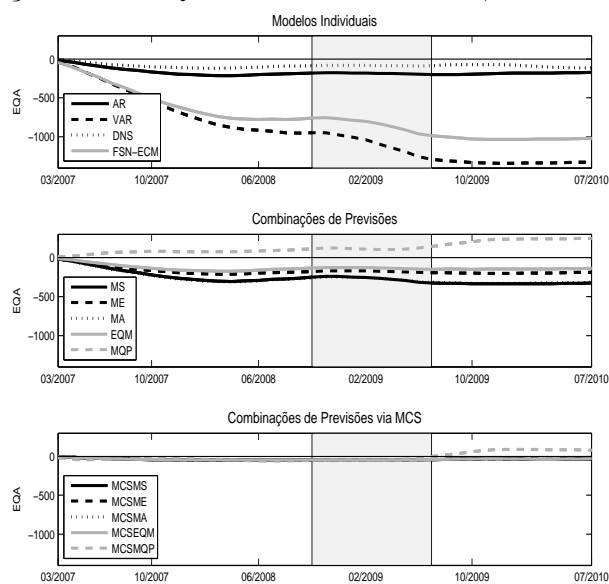


Figura 6: Traço do Erro Quadrático Acumulado, horizonte de 12 meses



7 Conclusão

Esse trabalho buscou identificar a existência de técnicas de combinação que possibilitem obter previsões da ETTJ superiores às dos modelos individuais para os dados brasileiros. Como medida da ETTJ foram utilizadas as taxas dos *swaps* DI-Pré da BM&FBovepsa.

Com base nos modelos teóricos apresentados no capítulo 3, foram obtidas as previsões dos modelos individuais. Os resultados alcançados levaram a concluir não ser possível identificar um modelo individual que consistentemente produza menor erro de previsão para o período fora da amostra, similar ao encontrado por Cajueiro *et al.* (2009) para os dados brasileiros. Porém, ao contrário de Cajueiro *et al.* (2009), o resultado do trabalho indica que modelo DNS não superou o passeio aleatório no período fora da amostra considerado neste trabalho. Tal fato demonstra que os desempenhos dos modelos podem variar ao longo do tempo, semelhante ao defendido no campo teórico e em aplicações empíricas na curva de juros americana.

As habilidades de cada modelo variam dependendo do horizonte de previsão, da maturidade e do período analisados. Tal fato confirma que modelos com diferentes especificações contribuem de forma complementar para aproximar o processo gerador de dados, forte argumento para o uso de esquemas de combinação como alternativa de confiar em um único modelo.

Os resultados dos métodos de combinação discutidos no capítulo 4 estão em con-

sonância com esta análise. Em cada horizonte de previsão considerado foi possível encontrar pelo menos um esquema de combinação que tenha apresentado desempenho igual ou superior ao melhor modelo individual. Adicionalmente, quanto mais longo o horizonte considerado, maior foi a contribuição das combinações.

Os erros de previsão dos esquemas de previsão são certamente menos voláteis que os modelos individuais, especialmente nos esquemas de combinação via MCS. Esta abordagem também apresentou resultados superiores no curto prazo (um e três meses), o que torna recomendável sua utilização nessas situações.

Nos horizontes mais curtos, as combinações simples e o EQM invertido via MCS tiveram desempenhos similares ao melhor modelo individual (AR) um mês à frente e superiores no horizonte de três meses. Outro fato relevante é que a diferença entre estas combinações foi mínima. Tal constatação permite concluir, nestes casos, que os procedimentos de exclusão de modelos individuais por MCS foram mais benéficos para previsões de combinação do que a maneira como os pesos foram calculados.

Nos horizontes mais longos, a combinação por MQP foi o método com melhor desempenho, superando o passeio aleatório. Embora seja a combinação que apresenta menor erro de previsão, tem como desvantagem demonstrar alguma instabilidade com respeito ao período da amostra de treinamento selecionado para a determinação dos pesos.

Cumprir destacar que os resultados alcançados nesse trabalho restringem-se ao período fora da amostra considerado, não sendo possível garantir que as conclusões

aqui feitas se aplicam de maneira definitiva. Assim, em função da curta série temporal disponível para ETTJ no Brasil (quando comparado a países como Estados Unidos e Canadá), faz-se necessário incorporar mais dados ao longo do tempo, de modo a ter maior certeza dos resultados obtidos nesse trabalho.

Como sugestão para trabalhos futuros, é interessante que seja analisado se a inclusão de variáveis macroeconômicas aos modelos individuais de previsão contribuem para a previsão da ETTJ brasileira. Em particular, em função da tendência a longo prazo de redução de nível da curva de juros brasileira (em um cenário de convergência da taxa de juros real no Brasil em relação a outras economias emergentes com estágio equivalente de desenvolvimento), é recomendável a incorporação de variáveis macroeconômicas que vislumbrem este cenário. Vale ressaltar que em função dos dados macros serem mensais, neste caso não seria possível utilizar o modelo FSN-ECM. Adicionalmente, cabe testar se os esquemas de combinação apresentariam resultados superiores nestes tipos de modelos individuais.

Também seria conveniente realizar o mesmo procedimento empregado neste trabalho utilizando uma base de dados que incorpore maturidades mais longas. A partir de janeiro de 2007, os vencimentos da ETTJ brasileira até dez anos passaram a ganhar maior liquidez. Apesar de a curva de juros com extensão de dez anos ser amplamente analisada para os dados americanos, no caso brasileiro é necessário obter mais observações para estimar os parâmetros de forma mais robusta.

8 Referências Bibliográficas

ALMEIDA, C.; GOMES, R.; LEITE, A.; SIMOSEN, A.; VICENTE, J. **Does Curvature Enhance Forecasting?**. International J. of Theoretical and Applied Finance, v. 12, n. 8, p. 1171-1196, 2009.

ALMEIDA, C.; GOMES, R.; LEITE, A.; SIMOSEN, A.; VICENTE, J. **Movimentos da Estrutura a Termo e Critérios de Minimização do Erro de Previsão em um Modelo Paramétrico Exponencial**, Revista Brasileira de Economia, v. 62, n. 4, p. 497-510, 2008.

ANG, A.; PIAZZESI, M. **A No-Arbitrage Vector Autoregression of Term Structure Dynamics With Macroeconomic and Latent Variables**. Journal of Monetary Economics, v. 50, p. 745-787, 2003.

BATES, J. M.; GRANGER, C. W. J. **The Combination of Forecasts**. Operational Research Quarterly v. 20, p. 451-468, 1969.

BLACK, F.; DERMAN, E.; TOY, W. **A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options**. Financial Analysts Journal, p. 33-39, Jan-Fev, 1990.

BOLDER, D. J.; ROMANYUK, Y. **Combining Canadian Interest-Rate Forecasts**. Working Paper 34, Bank of Canada, 2008.

BOWSHER, C.G.; MEEKS, R. **The Dynamics of Economic Functions: Modeling and Forecasting the Yield Curve**. Journal of American Statistical Association, v. 103, p. 1419-1437, Dez, 2008.

BRENNAN, M.; SCHWARTZ, E. **A Continuous Time Approach to the Pricing of Bonds**. Journal of Banking and Finance, v. 3, p. 133-155, 1979.

CAJUEIRO, D.O.; DIVINO, J. A.; TABAK, B.M. **Forecasting the Yield Curve for Brazil**. Working Paper 197, Banco Central do Brasil, Nov, 2009.

CAMPBELL, J.Y.; LO A.W.; MACKINLAY, A.C. **The Econometrics of Financial Markets**. Princeton: Princeton University Press, 1997.

CARRIERO, A.; KAPETANIOS, G.; MARCELLINO, M. **Forecasting Government Bond Yields**. Working Paper University of London, 2009.

CHRISTENSEN, J.H.E.; DIEBOLD, F.X.; RUDEBUSH, G.D. **The Affine Arbitrage-**

Free Class of Nelson-Siegel Term Structure Models. Manuscrito, University of Pennsylvania o Federal Reserve Bank of San Francisco, 2007 revisado em 2009a.

CHRISTENSEN, J.H.E.; DIEBOLD, F.X.; RUDEBUSH, G.D. **An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model.** The Econometrics Journal, v. 12, p. 33-64, 2009b.

CHRISTOFFERSEN, P. F.; DIEBOLD, F.X. **Cointegration and Long-Horizon Forecasting.** Journal of Business & Economic Statistics, v. 16, p. 450-458, 1998.

CLARK, T. E.; MCCRACKEN, M. W. **Averaging forecasts from VARs with uncertain instabilities.** Journal of Applied Econometrics. Aceito para publicação. Revisão de Federal Reserve Bank of Kansas City Working Paper, 06-12, 2009.

CLEMENTS, M.P.; HENDRY, D.F. **Forecasting with Breaks in Data Processes.** Handbook of Economic Forecasting, Amsterdam, North-Holland, cap. 12, p. 605-657, 2006.

COX, J.C.; INGERSOLL, J.E., ROSS, S.A **A theory of the term structure of interest rates.** Econometrica, v.53, p. 385-407, 1985.

DAI, Q.; SINGLETON, K.J., **Specification Analysis of Affine Term Structure Models**. The Journal of Finance, v. 50, Out, 2000.

DE POOTER, M. **Examining the Nelson-Siegel Class of Term Structure Models**. Discussion Paper v. 043/4, Tinbergen Institute, 2007.

DE POOTER, M.; RAVAZZOLO, F.; VAN DIJK, D. **Predicting the Term Structure of Interest Rates: Incorporating Parameter Uncertainty, Model Uncertainty and Macroeconomic Information**. Discussion Paper v. 028/4, Tinbergen Institute, 2007.

DE POOTER, M.; RAVAZZOLO, F.; VAN DIJK, D. **Term Structure Using Macro Factors and Forecast Combination**. International Finance Discussion Papers, Board of Governors of the Federal Reserve System, n. 993, Jan, 2010.

DELLA CORTE, P.; SARNO, L.; THORTON, D. L. **The Expectation Hypothesis of the Term Structure of Very Short-Term Rates: Statistical Tests and Economic Value**. Federal Reserve Bank of St. Louis Working Paper, v. 61, 2006

DIEBOLD, F. X.; LOPEZ, J. A. **Forecast Evaluation and Combination**. Research Paper 9525, Federal Reserve Bank of New York, 1995.

DIEBOLD, F. X.; PAULY, P. **Structural Change and the Combination of Forecasts.** *Journal of Forecasting*, v.6, p.21-40, 1987.

DIEBOLD, F.X; PIAZZESI, M.; RUDEBUSH, G.D. **Modeling Bond Yields in Finance and Macroeconomics.** *American Economic Review*, v. 95, p. 415-420, 2005.

DIEBOLD, F.X, RUDEBUSH, G.D.; AROUBA, B. **The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach.** *Journal of Econometrics*, v. 131, p. 309-338, 2006.

DIEBOLD, F.X.; LI, C. **Forecasting the Term Structure of Government Bond Yields.** *Journal of Econometrics*, v. 130, p. 337-604, 2006.

DIEBOLD, F.X.; LI, C.; YUE, V **Global Yield Curve Dynamics and Interactions: A Generalized Nelson-Siegel Approach.** *Journal of Econometrics*, v. 146, p. 351-363, 2008.

DUFFEE, G.R. **Forecasting with the Term Structure: The Role of No-Arbitrage.** Haas School of Business (University of California - Berkeley) Discussion

Paper, Mai, 2008.

DUFFEE, G.R. **Term Premia and Interest Rate Forecasts in Affine Models.**

The Journal of Finance, v. 57, Fev, 2002.

DUFFIE, D.; KAN, R. **A Yield-Factor Model of Interest Rate.** Mathematical

Finance, v. 6, p. 379-406, Out, 1996.

ELLIOT, G.; C.W.J. GRANGER; TIMMERMANN, A. Editores. **Handbook of Economic Forecasting.** Amsterdam: North-Holland, 2006.

EVANS, C.L.; MARSHALL, D.A. **Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory.** Carnegie-Rochester Conference Series on Public Policy v. 49, p. 53-111, 1998.

FABOZZI, F. Editor. **Interest Rate, Term Structure and Valuation Modeling.** Wiley Finance, 2002.

FISHER, I. **Appreciation and Interests.** American Economic Association, v. 11, p. 21-29, 1896.

GUIDOLIN, M.; TIMMERMAN, A. **Forecasts of US Short-Term Interest Rates: A Flexible Forecast Combination Approach.** Journal of Econometrics, v. 150, p. 297-311, Jun, 2009.

GUILLEN, O.T.; TABAK, B.M. **Characterizing the Brazilian Term Structure of Interest Rates.** International Journal of Monetary Economics and Finance, v. 2, no. 2, p. 103-114, Jan, 2009.

HANSEN, P. R.; LUNDE, A.; NASON, J.M. **Chosing the Best Volatility Model: The Model Confidence Set Approach.** Oxford Bulletin of Economics and Statistics, v.65, p. 839-861, 2003.

HANSEN, P. R.; LUNDE, A.; NASON, J.M. **The Model Confidence Set.** Aceito para publicação em Econometrica, 2010.

HEATH, D.; JARROW, R.; MORTON, A. **Bond Pricing and the Term Structure of Interest Rates: a New Methodology for Contingent Claims Valuation.** Econometrica, v.60, p. 77-105, 1992.

HENDRY, D.F.; CLEMENTS, M.P. **Pooling of Forecasts.** Econometrics Journal, v. 5, p. 1-26, 2002.

HO, S.Y.; LEE S-B. **Term Structure Movements and Pricing Interest Rate Contingent Claims.** Journal of Finance, v. 41, p. 1011-1029, 1986.

HORDAHL, P.; TRISTANI, O.; VESTIN, D. **A Joint Econometric Model of Macroeconomic and Term-Structure Dynamics.** Journal of Econometrics, v. 131, p. 405-444, 2006.

HULL, J.; WHITE, A. **Pricing Interest Rate Derivative Securities.** Review Financial Studies, v.3, p.573-92, 1990.

KOOPMAN, S.J.; MALLE M.I.P.; VAN DER WELL, M. **Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters.** Discussion Paper 095/4, Tinbergen Institute, 2007.

LEITE, A.L.; GOMES FILHO. R.; VICENTE, J. **Previsão da Curva de Juros: um Modelo Estatístico com Variáveis Macroeconômicas.** Working Paper 186, Banco Central do Brasil, Mai, 2009.

LIMA, E.J.A.; LUDUVICE, F.; TABAK, B.M. **Forecasting Interest Rates: an Application for Brazil.** Working Paper 120, Banco Central do Brasil, Out, 2006.

LITTERMAN, R.; SCHEINKMAN, J. **Common Factors Affecting Bond Returns**. Journal of Fixed Income, v. 1, p. 51-61, 1991.

LONGSTAFF, F.A.; SCHWARTZ, E.S. (1995) **Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model**. Journal of Finance, v.47, p. 1259-1282, Set, 1992.

LUDVINGSON, S. C.; NG, S., **Macro Factors in Bond Risk Premia**. Aceito para publicação em Review of Financial Studies, 2009.

MONCH, E. **Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach**. Working Paper 544, European Central Bank, 2005.

NELSON, C.R.; SIEGEL, A.F. **Parsimonious Modeling of Yield Curves**. The Journal of Business, v. 60, p. 473-489, Out, 1987.

POLITIS D. N.; ROMANO, J. P. **The Stationary Bootstrap**. Journal of the American Stastical Association, v. 89, p. 13031313.

RAPACH, D. E.; STRAUSS, J.K. **Forecasting Employment Growth in Missouri with Many Potentially Relevant Predictors: An Analysis of Forecast Combining Methods.** Federal Reserve Bank of St. Louis Regional Economic Development, v.1, n.1, p. 97-112, 2005.

SINGLETON, K.J. **A Latent Time Series Model of the Cyclical Behavior of Interest Rates.** International Economic Review, v. 21, Out, 1980.

STOCK, J.H.; WATSON, M.W. **Combination forecasts of output growth in a seven-country data set.** Journal of Forecasting v. 23, p. 405-430, 2004.

STOCK, J.H.; WATSON, M.W., **Forecasting with Many Predictors.** Handbook of Forecasting, Amsterdam: North-Holland, cap. 10, p. 515-554, 2006.

SVENSSON, L. E. O. **Estimating Forward Interest Rates with the Extended Nelson-Siegel Method.** Quarterly Review, v. 3, p. 13-26, Sveriges Riksbank, 1995.

TIMMERMANN, A. **Forecast Combinations.** Handbook of Economic Forecasting, Amsterdam: North-Holland, cap. 4, p. 135-196, 2006.

VASICEK, O. **An Equilibrium Characterization of the Term Structure.**

Journal of Financial Economics, v.5, p.177-188, 1977.

VEREDA, L.; LOPES, H., FUKUDA, R. **VAR models for the term structure of interest rates.** Insurance: Mathematics and Economics, v. 42, p. 548-559, 2008.

VICENTE, J.; TABAK, B. M. **Forecasting Bonds Yields in the Brazilian Fixed Income Market.** Aceito para publicação em International Journal of Forecasting, 2008.

WELCH, I.; GOYAL, A. **A Comprehensive Look at the Empirical Performance of Equity Premium Prediction.** Review of Financial Studies, v.21, p. 1455-1508, 2008.

9 Apêndice

9.1 Figuras

As figuras a seguir mostram as taxas observadas contra as taxas previstas para os períodos reportados nas tabelas 2 a 5, tanto para os modelos individuais como para os esquemas de combinação. Com o intuito de reduzir o número de gráficos a serem exibidos, foram selecionadas as maturidades dos *swaps* de 3, 12, 18 e 24 meses. Em cada figura são exibidos três gráficos, onde são comparadas as taxas observadas e previstas dos modelos individuais, das combinações de previsão e das combinações de previsão via MCS.

Figura 7: Taxas observadas e previstas de *swap* de 3 meses, horizonte de um mês

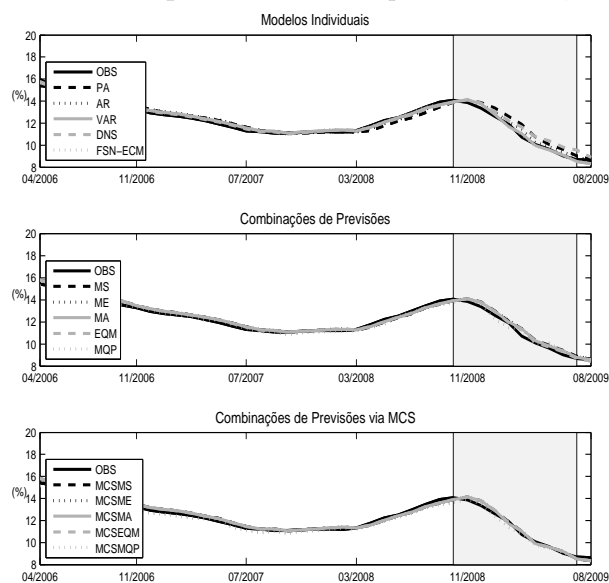


Figura 8: Taxas observadas e previstas de *swap* de 12 meses, horizonte de um mês

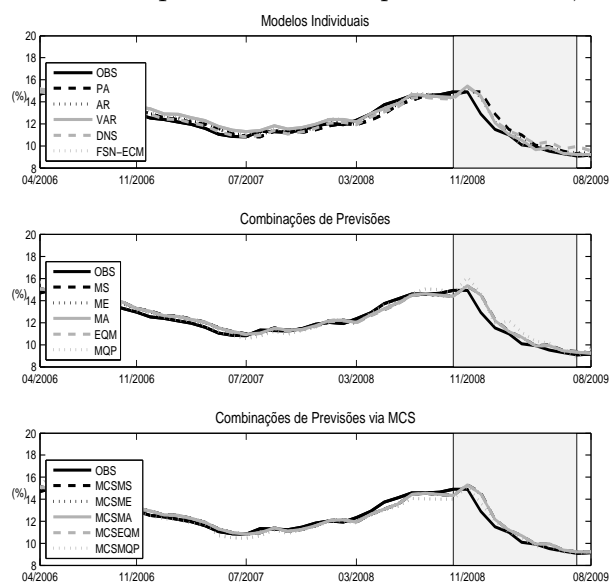


Figura 9: Taxas observadas e previstas de *swap* de 18 meses, horizonte de um mês

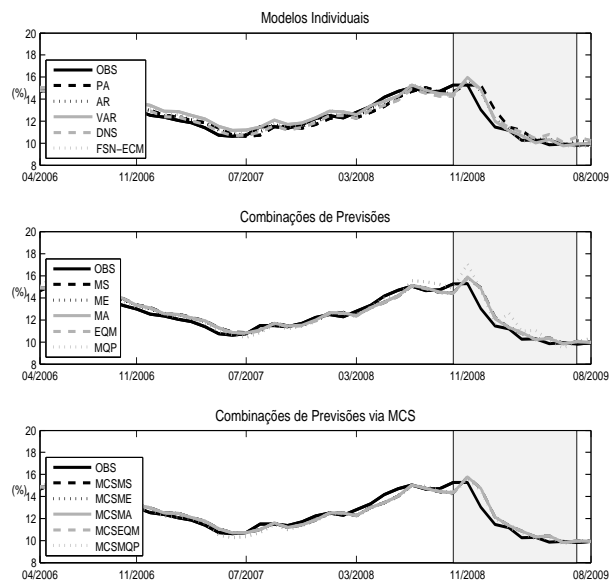


Figura 10: Taxas observadas e previstas de *swap* de 30 meses, horizonte de um mês

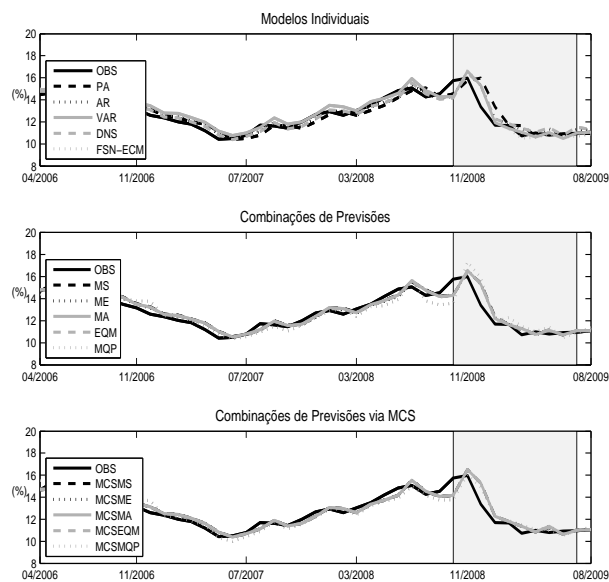


Figura 11: Taxas observadas e previstas de *swap* de 3 meses, horizonte de três meses

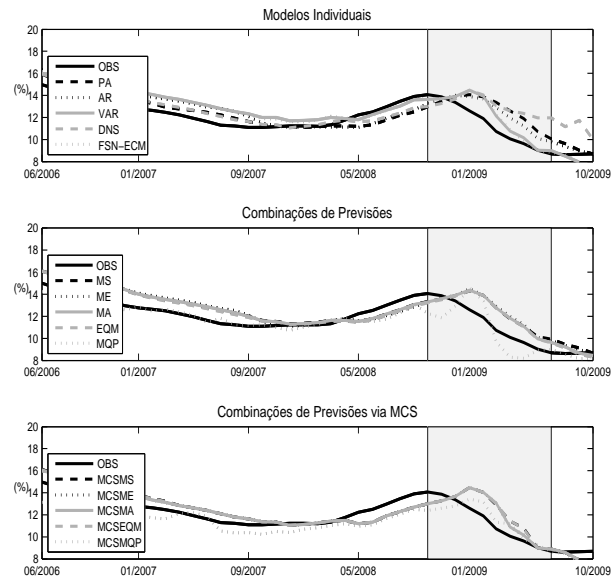


Figura 12: Taxas observadas e previstas de *swap* de 12 meses, horizonte de três meses

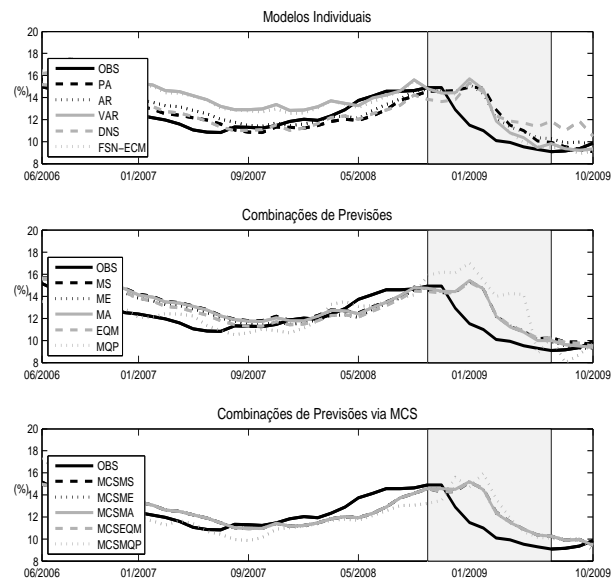


Figura 13: Taxas observadas e previstas de *swap* de 18 meses, horizonte de três meses

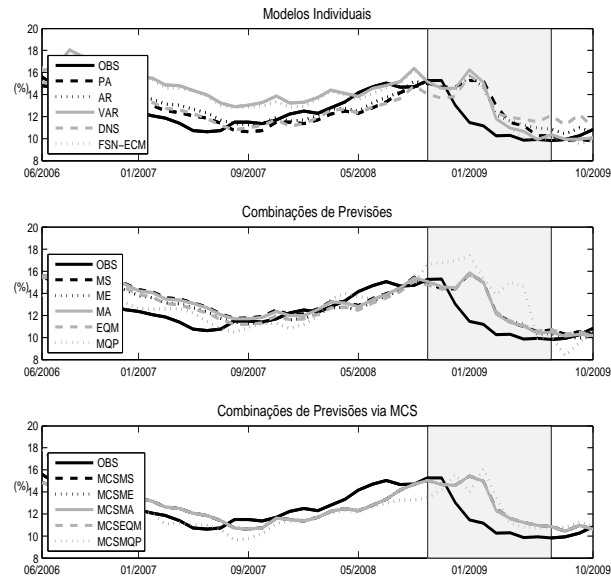


Figura 14: Taxas observadas e previstas de *swap* de 30 meses, horizonte de três meses

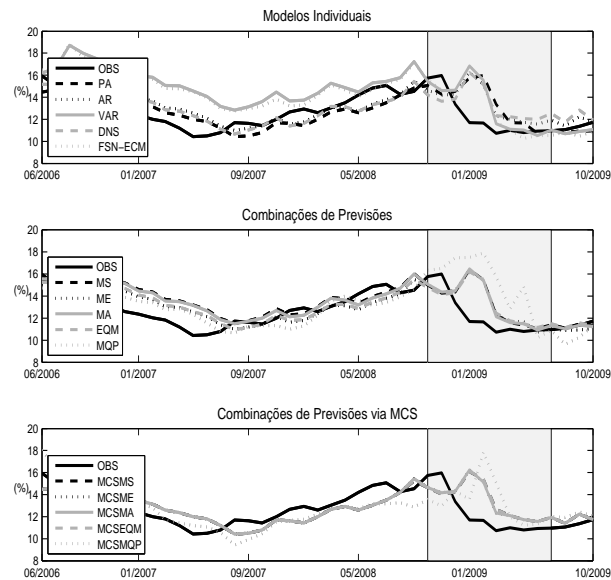


Figura 15: Taxas observadas e previstas de *swap* de 3 meses, horizonte de seis meses

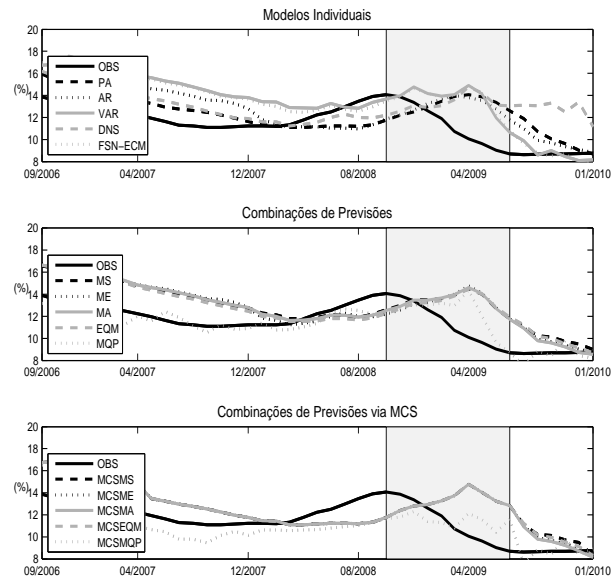


Figura 16: Taxas observadas e previstas de *swap* de 12 meses, horizonte de seis meses

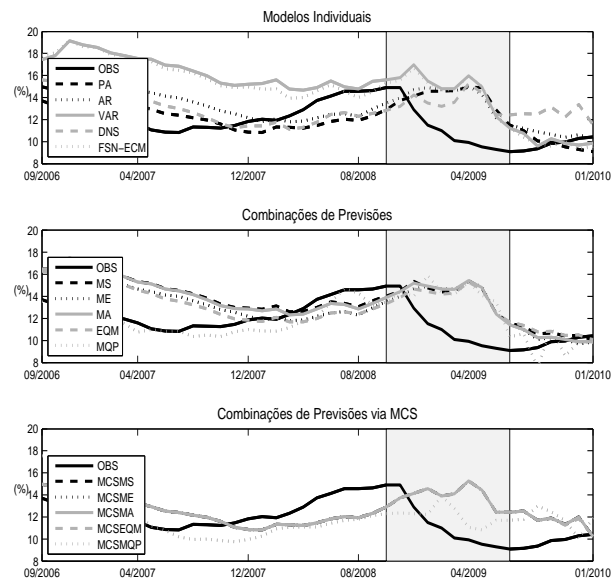


Figura 17: Taxas observadas e previstas de *swap* de 18 meses, horizonte de seis meses

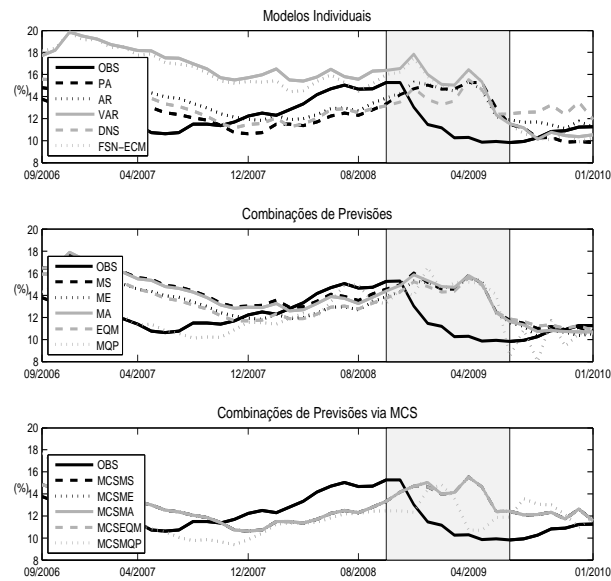


Figura 18: Taxas observadas e previstas de *swap* de 30 meses, horizonte de seis meses

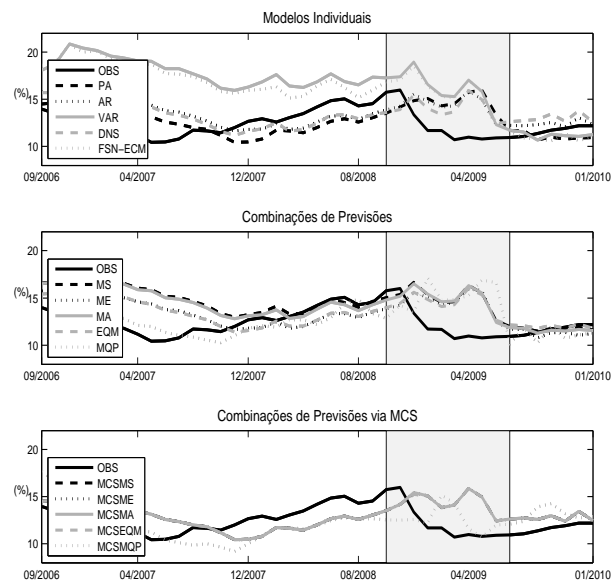


Figura 19: Taxas observadas e previstas de *swap* de 3 meses, horizonte de 12 meses

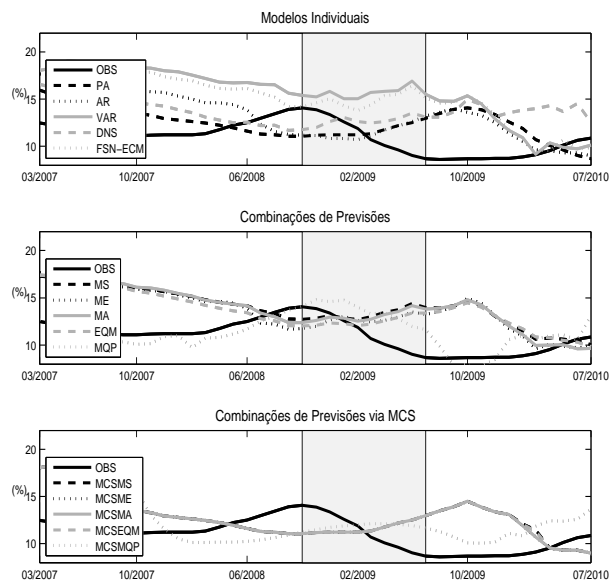


Figura 20: Taxas observadas e previstas de *swap* de 12 meses, horizonte de 12 meses

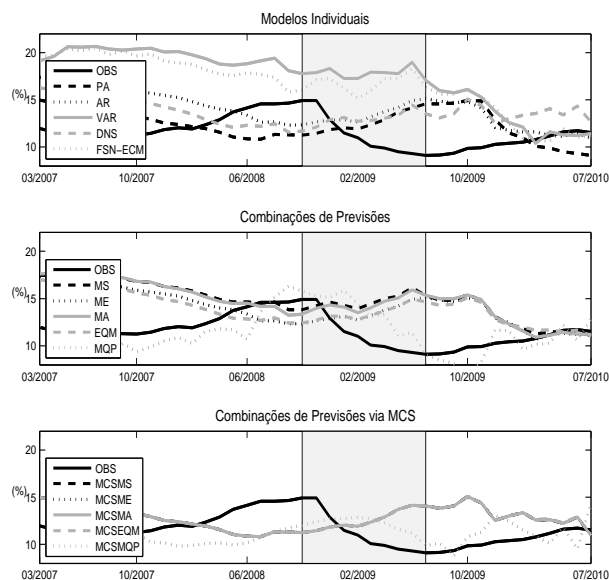


Figura 21: Taxas observadas e previstas de *swap* de 18 meses, horizonte de 12 meses

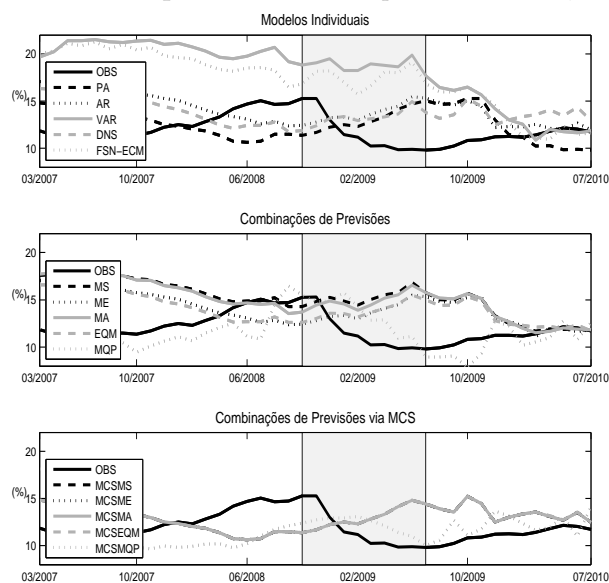


Figura 22: Taxas observadas e previstas de *swap* de 30 meses, horizonte de 12 meses

