

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE
RIBEIRÃO PRETO
DEPARTAMENTO DE ECONOMIA

RAFAEL NOGUEIRA DO PRADO

Modelagem da taxa de juros de covariância zero para a economia brasileira

Orientador: Prof. Dr. Alex Luiz Ferreira

RIBEIRÃO PRETO
2023

Prof. Dr. Carlos Gilberto Carlotti Junior
Reitor da Universidade de São Paulo

Prof. Dr. Fábio Augusto Reis Gomes
Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Prof. Dr. Milton Barossi Filho
Chefe do Departamento de Economia

RAFAEL NOGUEIRA DO PRADO

Modelagem da taxa de juros de covariância zero para a economia brasileira

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível no Serviço de Pós-Graduação da FEA-RP/USP.

Orientador: Prof. Dr. Alex Luiz Ferreira

RIBEIRÃO PRETO

2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Prado, Rafael Nogueira do

Modelagem da taxa de juros de covariância zero para a economia brasileira / Rafael Nogueira do Prado – Ribeirão Preto, 2023.

68f.: il.; 30 cm

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível no Serviço de Pós-Graduação da FEA-RP/USP. – Universidade de São Paulo

Orientador: Ferreira, Alex Luiz

1. Taxa de juros. 2. Risco. 3. Fator estocástico de desconto. 4. Precificação.

RAFAEL NOGUEIRA DO PRADO

Modelagem da taxa de juros de covariância zero para a economia brasileira

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível no Serviço de Pós-Graduação da FEA-RP/USP.

Área de Concentração: Economia Aplicada

Data de Aprovação:

____/____/____

Banca Examinadora:

Prof. Dr. Alex Luiz Ferreira
Orientador

Prof. Dr. João Victor Issler
Avaliador 1

Dr. Marco Lyrio
Avaliador 2

Dedico este trabalho à minha mãe, que sempre me apoiou na minha jornada.

AGRADECIMENTOS

Agradeço aos meus familiares, por me darem suporte emocional durante toda essa caminhada e estarem sempre presentes, nos melhores e piores momentos.

Agradeço ao meu orientador, pela sua paciência em me direcionar ao caminho certo e pela sua disponibilidade em ajudar-me a desvendar os mistérios dessa pesquisa.

Por último, agradeço aos professores e demais funcionários da FEARP por todo amparo ao longo do mestrado.

"O que é escrito sem esforço em geral é lido sem
prazer."(Samuel Johnson)

RESUMO

PRADO, Rafael Nogueira do. *Modelagem da taxa de juros de covariância zero para a economia brasileira*. 2023. 68f. Mestrado (Dissertação) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2023.

A taxa livre de risco é utilizada conceitualmente em muitas aplicações em finanças e macrofinanças. Enquanto em muitos trabalhos teóricos essa taxa é assumida como constante, e em trabalhos empíricos é comum se utilizar o retorno de títulos públicos como *proxy* a ela, há evidências contrárias a esses usos, resultado esse que ficou conhecido como *Risk-Free Rate Puzzle*. Essa dissertação preocupa-se com a taxa livre de risco usada nas aplicações empíricas da literatura de macrofinanças. O objetivo é estimar uma taxa livre de risco para o Brasil, calculada a partir da equação fundamental de apreçamento de ativos, que se utiliza de um fator estocástico de desconto, com base em um modelo CCAPM. Utilizam-se diferentes abordagens econométricas para a aplicação do modelo teórico aos dados, sendo um modelo semi-paramétrico e dois modelos não-paramétricos. Ao final, as taxas estimadas serão comparadas entre si e com a taxa DI, que é a principal referência para a taxa livre de risco utilizada na economia brasileira. Os resultados sugerem que a taxa DI e as taxas estimadas se distanciaram durante boa parte do período analisado, de 2002 a 2022, e o uso de diferentes modelos econométricos levou a diferentes conclusões acerca da existência do *Risk-Free Rate Puzzle* para o Brasil.

Palavras-chave: Taxa de juros. Risco. Fator estocástico de desconto. Precificação.

JEL: E44

ABSTRACT

PRADO, Rafael Nogueira do. Modeling the zero covariance interest rate for the Brazilian economy. 2023. 68f. Mestrado (Dissertação) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2023.

The risk-free rate is used conceptually in many applications in finance and macro finance. While in many theoretical works this rate is assumed to be constant, and in empirical works it is common to use the return on government bonds as a proxy to it, there is contrary evidence to these uses, a result that has become known as the Risk-Free Rate Puzzle. This dissertation is concerned with the risk-free rate used in empirical applications of the macro-finance literature. The objective is to estimate a risk-free rate for Brazil, calculated from the fundamental asset pricing equation, which uses a stochastic discount factor, based on a CCAPM model. Different econometric approaches are used to apply the theoretical model to the data, being one semiparametric model and two nonparametric models. At the end, the estimated rates will be compared among themselves and with the DI rate, which is the main proxy for the risk-free rate used in the Brazilian economy. The results suggest that the DI rate and the estimated rates were far apart during much of the period analyzed, from 2002 to 2022, and the use of different econometric models led to different conclusions about the existence of the Risk-Free Rate Puzzle for Brazil.

Keywords: Interest rate. Risk. Stochastic discount factor. Pricing.

JEL: E44

LISTA DE ILUSTRAÇÕES

Figura 1 – Expectativa de Inflação, IPCA e DI trimestrais (% a.a.)	33
Figura 2 – Crescimento do consumo	33
Figura 3 – Consumo em proporção do PIB (%)	34
Figura 4 – DI vs. Taxas estimadas (% a.a.)	37
Figura 5 – Correlações entre as diferentes taxas	38

LISTA DE TABELAS

Tabela 1 – Estatísticas dos ativos utilizados	32
Tabela 2 – Resultados do modelo MQO - AIF	35
Tabela 3 – Parâmetros estimados dos modelos GMM e GEL	36
Tabela 4 – Resultados do modelo MQO - GMM	36
Tabela 5 – Resultados do modelo MQO - GEL	37
Tabela 6 – Médias e desvios-padrão das taxas estimadas e da DI	38
Tabela 7 – Ranking de liquidez das ações de acordo com volume médio de negociação	46
Tabela 7 – Continuação da tabela 7	47
Tabela 7 – Continuação da tabela 7	48

SUMÁRIO

	INTRODUÇÃO	13
1	REVISÃO DA LITERATURA E METODOLOGIA	17
1.1	Revisão da literatura	17
1.2	Metodologia	19
1.2.1	CAPM e CCAPM	19
1.2.2	A equação Fundamental de Apreçamento de Ativos	20
1.2.3	O ativo livre de risco	22
1.2.4	Modelo AIF	23
1.2.5	Modelo GMM	24
1.2.6	Modelo GEL	27
1.2.7	O método de Low Frequency Econometrics	29
2	DADOS E RESULTADOS	31
2.1	Dados	31
2.2	Resultados	34
3	CONCLUSÃO	40
	REFERÊNCIAS	41
	APÊNDICE A – LISTA DE AÇÕES	46
	APÊNDICE B – CÓDIGO EM R	49

INTRODUÇÃO

Um ativo livre de risco é aquele que garante com certeza uma certa rentabilidade em um determinado período de tempo. Num modelo de equilíbrio geral Arrow-Debreu com mercados completos (ARROW, 1964; DEBREU, 1959; ARROW; DEBREU, 1954), esse ativo pode ser obtido formando-se uma carteira específica que engloba todos ativos primitivos. O ativo sintético resultante pagará uma unidade monetária da economia em questão, qualquer que seja o estado da natureza. Essa construção básica que resulta na taxa livre de risco da economia é utilizada conceitualmente em trabalhos de macro-finanças.

Essa taxa é de fundamental importância para se medir o prêmio de risco de ativos arriscados, isto é, a compensação de um investidor ao alocar recursos em ativos arriscados. Muitos trabalhos teóricos assumem que tal taxa é constante (CAMPBELL; COCHRANE, 1999). Essa hipótese é estabelecida para simplificação dos modelos, com objetivo analítico. Em aplicações empíricas, frequentemente utiliza-se do retorno de um título público como *proxy* para essa taxa (DIMSON; MARSH; STAUNTON, 2006; GOMES; COSTA; PUPO, 2013). Essa dissertação preocupa-se com a taxa livre de risco usada nas aplicações empíricas da literatura de macro-finanças. O objetivo é estimar essa taxa de acordo com a teoria econômica, utilizando-se de diferentes abordagens econométricas. O ambiente macroeconômico é descrito por um modelo de apreçamento de ativos baseado em consumo (CCAPM - *Consumption Capital Asset Pricing Model*), utilizando-se de uma função utilidade com aversão relativa ao risco constante (CRRA - *Constant Relative Risk Aversion*).

Há evidências de que: 1) essa taxa não é constante ao longo do tempo (PENNACCHI, 1991) e 2) há controvérsias sobre o uso de títulos públicos como *proxies* (KRISHNAMURTHY; VISSING-JORGENSEN, 2012; BINSBERGEN; DIAMOND; GROTTERRIA, 2022). Em primeiro lugar, pode-se citar o fato de que a taxa livre de risco é afetada pelas condições de mercado, como mudanças nas políticas de juros de bancos centrais, diretamente ligadas a crescimento econômico e expectativas de inflação. Especialmente para economias emergentes, essa taxa possui ainda maior volatilidade, dada a maior exposição dessas economias a diferentes tipos de risco, como o risco de inflação, de *default* e político, que podem estar ligados ao ciclo da economia. De tal maneira, esses riscos estabelecem dificuldades em se utilizar retornos de títulos públicos como *proxies* para a taxa livre de risco, uma vez que refletem muitas incertezas presentes nessas economias.

Mesmo para economias desenvolvidas, o uso de retornos de títulos públicos como *proxies* para a taxa livre de risco são questionados. Importante evidência desse questionamento é o trabalho de Weil (1989), que descreve o chamado *Risk-Free Rate Puzzle*. O autor encontra que a taxa livre de risco imposta pelo modelo teórico, após ser calibrado para a economia dos EUA, é de 6,5% ao ano, em comparação a uma média histórica de 0,8% ao ano. Outro exemplo é o trabalho desenvolvido por Jiang et al. (2019). Conforme argumentam os autores, o retorno esperado de um portfólio de títulos da dívida pública deve refletir o perfil de risco dos superávits primários do governo. Os autores encontram evidências de erros de precificação, tendo

em vista que os retornos desse portfólio são menores do que a taxa de juros que os investidores deveriam exigir por carregar tais títulos, indicando um quebra-cabeças que os autores intitulam como *government debt risk premium puzzle*. Embora a verificação desse quebra-cabeças não tenha sido realizada para o Brasil, surge uma questão geral importante: qual é a taxa livre de risco mais adequada para precificar títulos públicos de um determinado país? A partir dessa questão geral, a questão específica dessa dissertação relaciona-se com a verificação do *risk-free rate puzzle* para o caso particular da economia brasileira, em que é comum se utilizar a taxa DI como *proxy* para a taxa livre de risco.

A utilização da DI como taxa livre de risco da economia brasileira apresenta alguns problemas, possivelmente devidos à sua correlação com a taxa Selic. Como principal dificuldade, pode-se citar a existência de um risco soberano, também conhecido como risco de *default* (BONOMO; DOMINGUES, 2002). Nesse sentido, evidências apontam que podem existir diferenças entre os valores das taxas referentes a títulos públicos e os da taxa livre de risco para o Brasil. Conforme argumentado por Simonassi (2006), que realiza uma estimação da taxa livre de risco para o Brasil e a compara com a taxa Selic, a diferença entre essas duas taxas fornece um bom referencial para mensurar o *spread* pago pelo país em virtude dos riscos que a economia brasileira oferece. Na mesma linha, o presente trabalho tem como objetivo estimar uma taxa livre de risco para o Brasil e compará-la com a taxa DI, a partir de diferentes métricas.

A fim de se aplicar o modelo CCAPM para dados da economia brasileira, utilizam-se dois modelos não-paramétricos e um semi-paramétrico. A opção da não utilização de modelos paramétricos se dá pelo fato de estes serem muito sujeitos a erros de especificação quanto ao processo gerador de dados e sobre os parâmetros dos modelos. Conforme argumenta Horowitz (2009), no caso de má especificação das funções de distribuição para os dados em modelos paramétricos, pode-se chegar a estimações inconsistentes. Sendo assim, modelos não-paramétricos são capazes de lidar com esse problema devido à ausência da necessidade da imposição de hipóteses sobre o processo gerador de dados. No caso de modelos semi-paramétricos, as vantagens se dão pela possibilidade de que, mesmo não sendo necessária a imposição de hipóteses sobre o processo gerador de dados, pode-se fazer hipóteses acerca de determinados parâmetros de dimensão finita da economia. Desse modo, tais modelos possuem a capacidade de testar hipóteses acerca de parâmetros advindos da teoria econômica.

Inicialmente utiliza-se como ponto de referência o trabalho realizado por Simonassi (2006), que faz uso de um fator estocástico de desconto (SDF) estimado a partir do modelo não-paramétrico desenvolvido por Araújo, Issler e Fernandes (2005). Em seguida, realiza-se uma estimação semi-paramétrica do SDF, utilizando o modelo GMM (*Generalized Method of Moments*), que tem como base o trabalho de Hansen (1982). Por último, utiliza-se uma estimação via GEL (*Generalized Empirical Likelihood*) (NEWHEY; SMITH, 2004), que é um método não-paramétrico mais robusto a erros de especificação e que possui melhores propriedades em amostras finitas quando comparado ao GMM.

Com as estimações do SDF feitas pelas diferentes abordagens, calcula-se as taxas livres de risco para o Brasil com base na relação entre a taxa de juros livre de risco e o SDF descrita

pelo CCAPM. Dada a possibilidade de erros de especificação para a forma funcional do SDF, é justificado o uso de modelos não-paramétricos e semi-paramétricos, como tentativa inicial de se livrar de vieses gerados por modelos paramétricos (FAN; YAO, 2003). Ademais, são realizados testes de sobreidentificação, para garantir que os modelos estão corretamente especificados. O uso de diferentes abordagens é justificado também para testar a robustez dos resultados encontrados. Após calculadas, as taxas são comparadas entre si, e também com os valores da taxa DI ao longo do tempo.

Embora o CCAPM ainda seja amplamente utilizado na teoria de apreçamento de ativos, seu uso traz desafios empíricos, uma vez que os diferentes riscos observados em diferentes economias, não capturados totalmente pelo modelo empírico, tornam difíceis a recuperação das condições de equilíbrio que o modelo implica. Evidências que apontam falhas empíricas do CCAPM se acumulam desde os trabalhos de Hansen e Singleton (1983) e do *Equity Premium Puzzle*, de Mehra e Prescott (1985). Apesar das diferenças metodológicas, os autores encontram que o ambiente econômico descrito pelo CCAPM é incapaz de explicar os prêmios de risco observados para dados referentes à economia americana. Mesmo assim, o CCAPM apresenta uma boa reputação por trás do seu paradigma teórico. Nesse sentido, o modelo teórico a ser aqui testado empiricamente será uma versão condicional do CCAPM, conforme apresentado por Lettau e Ludvigson (2001), a partir do uso de variáveis instrumentais. Segundo os autores, o uso de um CCAPM condicional traz vantagens sobre o CCAPM incondicional, uma vez que o desempenho de retornos de ativos estão diretamente ligados aos diferentes estados da natureza futuros. Isso se dá pois, em estados da natureza adversos, a aversão ao risco aumenta e a demanda por ativos arriscados diminui, fazendo aumentar a demanda por consumo, que possui alta utilidade marginal (MEHRA et al., 2007).

Adicionalmente, para buscar resultados mais próximos à relação de equilíbrio de longo prazo imposta pelo CCAPM, utiliza-se um método de suavização da série de consumo utilizada, com o objetivo de extrair apenas informações relacionadas à sua tendência de longo prazo, excluindo seu componente cíclico. Este método foi desenvolvido por Müller e Watson (2015), e se baseia em projeções trigonométricas da série de consumo. A opção por não se utilizar o Filtro de Hodrick-Prescott, ou Filtro HP (HODRICK; PRESCOTT, 1997), se dá pelo argumento exposto por Hamilton (2018). De acordo com o autor, os componentes de tendência e cíclicos resultantes da aplicação do Filtro HP não são característicos da verdadeira dinâmica do processo gerador de dados em si, e sim de uma caracterização particular do uso do próprio filtro. Tal fato gera uma artificialidade nos resultados, uma vez que os componentes resultantes são, por construção, funções de realizações futuras das variáveis, que não são conhecidas em tempo real.

A aplicação empírica realizada nessa dissertação se vale de dados trimestrais de retornos de ações listadas na B3 e de retornos do próprio Ibovespa, de consumo em nível e de consumo em proporção do PIB, desde o primeiro trimestre de 2002 até o segundo trimestre de 2022. Além desses dados, foram também utilizados dados referentes à taxa DI para o mesmo período, com a finalidade de comparação entre as taxas estimadas e a DI. Os resultados apontaram um dis-

tanciamento entre as estimações e a taxa DI ao longo de todo o período da amostra. Enquanto o modelo GMM resultou em uma taxa livre de risco média acima da taxa DI, os modelos AIF e GEL resultaram em taxas livres de risco médias abaixo da taxa DI. Os resultados para os três modelos evidenciam a existência de um *puzzle* referente à taxa livre de risco mais apropriada para o Brasil.

Além dessa introdução, este trabalho possui três capítulos. No capítulo seguinte, será realizada uma revisão da literatura e apresentada a metodologia utilizada na pesquisa. No capítulo 2, serão descritos a base de dados utilizada e os resultados encontrados pelas diferentes abordagens. Por último, o capítulo 3 concluirá.

1 REVISÃO DA LITERATURA E METODOLOGIA

1.1 Revisão da literatura

Por muitos anos, desde que foi concebido, o modelo conhecido hoje como CAPM (*Capital Asset Pricing Model*) foi um dos pilares da literatura de finanças, e foi desenvolvido nos trabalhos de Sharpe (1964) e Budgets e Lintner (1965). Nesse modelo, o prêmio de risco de um investidor avesso ao risco é dado pela diferença entre o retorno esperado da carteira de mercado e a taxa livre de risco. Apesar de ter sido referência por décadas na literatura, evidências apontam para falhas empíricas na utilização do CAPM para explicar o prêmio de risco (BOLLERSLEV; ENGLE; WOOLDRIDGE, 1988; ROLL, 1977), uma vez que este pode estar correlacionado com outras variáveis que não a covariância entre o retorno do ativo e o da carteira de mercado.

Em resposta às falhas do CAPM para explicar retornos de ativos, Fama e French (1993) desenvolveram um modelo em que os retornos são funções de fatores relacionados ao tamanho da empresa, à razão *book-to-market* e ao risco agregado do mercado de ações. Esse modelo, conhecido como Modelo de Três Fatores de Fama-French, conseguiu superar os resultados do CAPM em capturar o prêmio de risco. Os autores explicam que essa superação se dá pelo fato de o modelo conseguir expressar uma aproximação para um risco comum não-observável nos retornos de portfólios. Contudo, de acordo com Lettau e Ludvigson (2001), essa interpretação é controversa, uma vez que o modelo não esclarece como esses fatores se relacionam com eventos macroeconômicos subjacentes. Ainda de acordo com esses autores, uma das razões para a falha do CAPM se dá pela sua abordagem estática. Nesse sentido, foram desenvolvidos modelos intertemporais para apreçamento de ativos, sendo o mais notável deles o modelo CCAPM (*Consumption Capital Asset Pricing Model*), desenvolvido por Breeden (1979), que toma como base o ambiente econômico descrito por Lucas (1978).

No entanto, conforme apontam Mehra e Prescott (1985) e Hansen e Singleton (1983), o modelo CCAPM também apresenta falhas para explicar o prêmio de risco. Os autores aplicam o modelo para o caso de dados da economia americana, e suas contribuições ficaram conhecidas como o *Equity Premium Puzzle*. Inspirados nesse *puzzle*, Weil (1989) discute ainda a incapacidade desse modelo em explicar os retornos do ativo livre de risco gerados por essa abordagem, o que ficou conhecido como o *Risk-Free Rate Puzzle*. Importante resolução desses *puzzles* se dá pelo trabalho de Bansal e Yaron (2004). Os autores modelam preferências de acordo com a representação para a função utilidade desenvolvida por Epstein e Zin (1989), que permite separabilidade entre a aversão ao risco e a elasticidade de substituição intertemporal, e modelam a incerteza por um componente cíclico e outro de tendência (estocástico), em contraposição ao uso de uma função CRRA, como em Mehra e Prescott (1985).

Ao buscar a relação entre o mercado de ativos financeiros e flutuações macroeconômicas, origina-se o campo da literatura conhecido macro-finanças (COCHRANE, 2017). Essa literatura tem como base a Equação Fundamental de Apreçamento de Ativos, em que o preço

de qualquer ativo é determinado pelo valor esperado do fluxo de caixa gerado por esse ativo (*payoff*) descontado por um fator estocástico de desconto (SDF) (HANSEN; JAGANNATHAN, 1991; HARRISON; KREPS, 1979; HANSEN; RICHARD, 1987). Como argumenta Campbell (2014), a teoria moderna do SDF se origina nos trabalhos de Ross (1978) e Harrison e Kreps (1979).

Do ponto de vista teórico, o SDF pode ser derivado a partir de uma função utilidade. Nesse sentido, pode-se utilizar a função potência, que expressa o SDF como uma relação entre a elasticidade de substituição intertemporal do indivíduo com a sua aversão relativa ao risco (RUBINSTEIN, 1976; GOURIEROUX; MONFORT, 2007). Há, contudo, críticas em relação ao uso do SDF em modelos empíricos semi-paramétricos (HOROWITZ, 2009), muito por conta das restrições impostas por esses modelos, que podem levar à especificação incorreta do seu formato e a estimações de parâmetros pouco plausíveis (MEHRA; PRESCOTT, 1985; BROWN; GIBBONS, 1985; CAMPBELL, 1992). Essa abordagem semi-paramétrica tem como principal referência econométrica o modelo GMM (Generalized Method of Moments (HANSEN, 1982; HANSEN; SINGLETON, 1982; HANSEN; SINGLETON, 1983).

Visando contornar estes problemas, a literatura empírica buscou outras formas de se estimar o fator estocástico de desconto. Rosenberg e Engle (2002) e Araújo, Issler e Fernandes (2005), por exemplo, desenvolvem modelos semelhantes para o cálculo do SDF, sem o uso de uma função utilidade, em que ele é uma função dos retornos dos ativos existentes na economia e consistem em estimações não-paramétricas. Além desses trabalhos, na literatura de macro-financeiras empírica mais recente, é crescente a quantidade de trabalhos que se utiliza de abordagens não-paramétricas para apreçamento de ativos, sendo uma delas o modelo GEL (*Generalized Empirical Likelihood*) (NEWHEY; SMITH, 2004; NODA et al., 2011; ITO; NODA, 2012; ALMEIDA; GARCIA, 2012; LAURINI; PEREIRA, 2016).

Utilizando-se da abordagem descrita por Araújo, Issler e Fernandes (2005), Simonassi (2006) realiza uma estimação não-paramétrica do SDF para estimar a taxa livre de risco para o Brasil, a fim de compará-la com a Selic, e encontra diferentes valores para ambas. Outros trabalhos já questionaram o uso de retornos de títulos públicos como proxies para a taxa livre de risco. Krishnamurthy e Vissing-Jorgensen (2012), por exemplo, argumentam que os baixos retornos de títulos do Tesouro americano não são um *benchmark* apropriado para a taxa livre de risco em um modelo CAPM, dado que empresas com um beta igual a zero não são capazes de se financiarem a essas taxas. Segundo os autores, a taxa livre de risco utilizada pelo modelo, que representa os custos de capital para essas empresas, deve ser maior que as taxas do Tesouro americano. Binsbergen, Diamond e Grotteria (2022) desenvolvem um modelo em que o *spread* entre a taxa livre de risco e o retorno de títulos públicos é explicado pelo *convenience yields* de tais ativos. Esse tipo de *yield*, como explicam, surge da característica desses ativos de possuírem alta liquidez e funcionarem como seguros para perdas nominais.

Na literatura brasileira, destacam-se os trabalhos de Issler e Piqueira (2000) e Bonomo e Domingues (2002) na verificação de *puzzles* para dados brasileiros. Em ambos os trabalhos, não se encontram evidências de um *equity premium puzzle* para o Brasil. Apesar disso, Bonomo

e Domingues (2002) concluem que há um *risk-free rate puzzle* invertido. Esse último fato é resultado de uma taxa livre de risco estimada menor que a taxa de títulos públicos brasileiros. Cysne (2006) encontra evidências contrárias a essas. O autor conclui que existem evidências de um *equity premium puzzle*, mas que não há evidências de um *risk-free rate puzzle* invertido para dados da economia brasileira.

O objetivo do presente trabalho é contribuir para essa literatura, ao se estimar uma taxa livre de risco para o Brasil a partir das abordagens descritas na Introdução. Até onde se sabe, não há outro trabalho na literatura que tenha aplicado o modelo GEL com o intuito de estimar os parâmetros de uma CRRA e utilizá-los para uma estimação da taxa livre de risco para o Brasil.

1.2 Metodologia

1.2.1 CAPM e CCAPM

O principal resultado do modelo CAPM, que tem como principal referência o trabalho de Sharpe (1964), é dado por

$$\mathbb{E}\{R_i - R_f\} = \beta_i \mathbb{E}\{R_p - R_f\}, \quad (1)$$

em que R_i é o retorno de um ativo arriscado i , R_f é a taxa livre de risco da economia, R_p é o retorno da carteira de mercado e β_i é o beta do ativo i , uma medida do risco desse ativo em relação ao risco de mercado, dada por $\beta_i = \frac{Cov(R_i, R_p)}{Var(R_i)}$. Essa equação mostra que o excesso de retorno de um ativo arriscado qualquer depende tanto do risco desse ativo em relação ao risco da carteira de mercado quanto do excesso de retorno dessa mesma carteira. A ausência de incerteza quanto aos valores para a taxa livre de risco e ausência de risco de *default* implicam que essa taxa é determinística, de tal forma que $Cov(R_f, R_i) = 0$, $Cov(R_p, R_f) = 0$ e $\mathbb{E}\{R_f\} = R_f$.

Buscando estabelecer uma relação entre a economia real, fundamentos macroeconômicos e os mercados financeiros, o modelo CCAPM vai além da hipótese de que os riscos de um ativo se relacionam apenas com oscilações no mercado financeiro (LETTAU; LUDVIGSON, 2001). Nesse sentido, a alocação de recursos em ativos financeiros afeta também as decisões de consumo do investidor. Tendo como principal referência o trabalho desenvolvido por Breeden (1979), que se utiliza do ambiente econômico descrito por Lucas (1978), esse modelo pode ser derivado do seguinte problema de otimização:

$$\begin{aligned} \max_{C_t} \quad & \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t u(C_t) \right\} \\ \text{s.t.} \quad & C_t + p_t q_t \leq x_t q_{t-1} + W_t \\ & \forall t, \end{aligned} \quad (2)$$

em que $u_t(\cdot)$ (com $u'_t(\cdot) > 0$ e $u''_t(\cdot) < 0$) é a função de utilidade no período t , $0 < \beta < 1$ é a taxa de desconto intertemporal (subjativa), C_t é o nível de consumo do indivíduo, p_t é o preço de um ativo genérico no período t , q_t é a quantidade comprada desse ativo em t , x_t é o *payoff* do ativo em t obtido pela compra desse ativo em $t - 1$ e W_t é uma renda exógena recebida

pelo indivíduo na data t . Conforme argumentado por Cochrane (2009), para o caso de um ativo arriscado com preço p_t , tem-se que o payoff desse ativo é $x_{t+1} = p_{t+1} + d_{t+1}$, em que d_{t+1} são os dividendos pagos no período $t + 1$, e para o caso do ativo livre de risco, tem-se que $p_t = 1$ e $x_{t+1} = R_f$.

Como demonstrado por Mehra et al. (2007), o problema de otimização em (2) resulta na equação de Euler

$$p_t = \mathbb{E}_t \left\{ \beta \frac{u'(C_{t+1})}{u'(C_t)} x_{t+1} \right\}. \quad (3)$$

A equação acima descreve a solução para a condição de primeira ordem do problema do indivíduo, e define a dinâmica das variáveis ao longo de um caminho ótimo. No caso específico da sua aplicação na teoria de apreçamento de ativos, ela determina que o preço de um ativo é igual ao que se espera receber de *payoffs* desse ativo, descontado por um fator subjetivo β e pela taxa à qual o indivíduo troca consumo presente por consumo futuro, a taxa marginal de substituição intertemporal, $\frac{u'(C_{t+1})}{u'(C_t)}$.

1.2.2 A equação Fundamental de Apreçamento de Ativos

Ao definir $\beta \frac{u'(C_{t+1})}{u'(C_t)} \equiv M_{t+1}$ como o fator estocástico de desconto (SDF) em $t + 1$, obtém-se a chamada Equação Fundamental de Apreçamento de Ativos, descrita por

$$\mathbb{E}_t \{ M_{t+1} x_{t+1} \} = p_t, \text{ ou} \quad (4)$$

$$\mathbb{E}_t \{ M_{t+1} R_{t+1} \} = 1, \quad (5)$$

em que $R_{t+1} \equiv \frac{x_{t+1}}{p_t}$ é o retorno bruto do ativo genérico. As equações 4 e 5 são generalizações da equação 3 e servem para precificar qualquer ativo da economia. Desenvolvida por Hansen e Jagannathan (1991)¹, a construção dessa equação dispensa a necessidade de se assumir mercados completos ou uma especificação para a função utilidade, nem mesmo uma distribuição para os *payoffs* dos ativos. De acordo com os autores, essa equação indica que o preço de qualquer ativo pode ser representado como a expectativa do produto entre o *payoff* desse ativo e a taxa marginal de substituição intertemporal de qualquer consumidor, nesse caso representada pelo SDF.

Conforme argumenta Cochrane (2009), mesmo que o SDF possa ser definido como a taxa marginal de substituição intertemporal descontada pelo fator subjetivo β , ele pode ser entendido como qualquer variável aleatória que satisfaça $p = \mathbb{E} \{ Mx \}$. Nesse sentido, a existência de um SDF M é garantida pela Lei do Preço Único, que diz que portfólios com o mesmo *payoff* esperado devem possuir o mesmo preço.

¹ Outros trabalhos importantes no desenvolvimento da Equação Fundamental de Apreçamento de Ativos se devem a Harrison e Kreps (1979) e Hansen e Richard (1987).

Considere uma economia com S estados da natureza, $s = 1, 2, \dots, S$, todos com probabilidade (objetiva) positiva $\pi(s)$. Para todo e qualquer ativo da economia, o *payoff* desse ativo pode ser escrito como $x(s) \in \underline{X}$, em que \underline{X} é o conjunto de possíveis *payoffs*. A Lei do Preço Único implica que, para um portfólio X com preço P , formado pelos ativos cujos *payoffs* são $x(s)$, $s = 1, 2, \dots, S$, deve-se ter que $P = \sum_{s=1}^S q(s)x(s)$, em que $q(s)$ é o preço do ativo no estado s (CAMPBELL, 2014). Assim, existe um SDF $M(s)$, que pode ser escrito como $M(s) = \frac{q(s)}{\pi(s)}$, de tal forma que $P = \sum_{s=1}^S \pi(s) \frac{q(s)}{\pi(s)} x(s) = \sum_{s=1}^S \pi(s) M(s) x(s) = \mathbb{E} \{MX\}$. De acordo com Cochrane (2009), a não ser que os mercados sejam completos, existe um número infinito de variáveis aleatórias M que satisfazem $P = \mathbb{E} \{MX\}$.

Adicionalmente, pela hipótese de ausência de arbitragem, tem-se que existe um SDF estritamente positivo, $M > 0$. A condição de não-arbitragem diz que, para um dado *payoff* x , em que $x \geq 0$ (quase certamente), e $x > 0$ com alguma probabilidade positiva, tem-se que $p(x) > 0$ (COCHRANE, 2009). Em outras palavras, essa condição diz que o investidor não consegue comprar um portfólio com preço zero, que paga um *payoff* positivo com probabilidade positiva, mas que certamente não lhe custará nada².

A equação 4 pode ainda ser interpretada sob a ótica da hipótese de mercados eficientes (FAMA, 1970), isto é, que os preços dos ativos refletem completamente toda informação disponível a cada período. Para isso, considere um conjunto formado por J investidores. Seja I_t^j o conjunto informacional do j -ésimo investidor, para $j = 1, 2, \dots, J$, na data t . O conjunto informacional de todos os agentes que transacionam no mercado de ativos na data t pode ser escrito como $I_t = \cap I_t^j$, $j = 1, 2, \dots, J$. Assim, os preços dos ativos em t devem fazer parte do conjunto informacional de cada investidor, I_t^j , e, conseqüentemente, devem fazer parte de I_t . Como é dispensada a hipótese de mercados completos, os investidores podem transacionar em um mercado incompleto de ativos. Dessa forma, o SDF pode ser calculado a partir da utilidade marginal de qualquer investidor que transaciona ativos livremente, e cada investidor pode possuir uma variação idiossincrática na sua utilidade marginal. Nesse caso, pode existir um SDF para cada investidor, $M_{t+1}^j = \beta \frac{u^j(C_{t+1})}{u^j(C_t)}$, $j = 1, 2, \dots, J$. Assim, uma versão mais geral dessa equação pode ser escrita como $p_t = \mathbb{E}_t \{M_{t+1}^j x_{t+1} | I_t^j\}$, $j = 1, 2, \dots, J$ (HANSEN; JAGANNATHAN, 1991).

Em uma economia com conjuntos informacionais comuns a todos os agentes e mercados completos, tem-se que $\frac{u^j(C_{t+1})}{u^j(C_t)} = \frac{u'(C_{t+1})}{u'(C_t)}$ para todo $j \in J$ (CAMPBELL, 2014; HANSEN; JAGANNATHAN, 1991). Assim sendo, $M_{t+1}^j = M_{t+1}$ para todo $j \in J$, e $p_t = \mathbb{E}_t \{M_{t+1}^j x_{t+1} | I_t^j\} = \mathbb{E}_t \{M_{t+1} x_{t+1} | I_t^j\}$ para todo $j \in J$. Por último, pela Lei das Expectativas Iteradas, como os preços são observados por todos os agentes, $p_t = \mathbb{E}_t \{M_{t+1} x_{t+1} | I_t^j\} = \mathbb{E}_t \{M_{t+1} x_{t+1} | I_t\}$. Desse modo, em mercados completos, existe um único SDF estritamente positivo que precifica todos os ativos da economia³.

Embora não seja necessária a especificação de uma função utilidade, é frequente o uso da

² Para uma demonstração formal da existência de um SDF $M > 0$, consultar Cochrane (2009), seções 4.1 e 4.2.

³ Para uma demonstração formal da unicidade de SDF $M > 0$ em mercados completos, consultar Cochrane (2009), seção 4.2.

função CRRA (função com coeficiente de aversão relativa ao risco constante) para se determinar o SDF. Em uma economia com mercados completos e um agente representativo, essa função toma a forma $u(C) = \frac{C^{1-\sigma}-1}{1-\sigma}$, com $u'(C) = C^{-\sigma}$, e resulta em um SDF expresso por

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma}. \quad (6)$$

O parâmetro σ indica o grau de aversão ao risco do consumidor, o seu coeficiente de aversão relativa ao risco. Quando $\sigma = 0$, o consumidor é neutro ao risco, quando $\sigma < 0$, ele é amante do risco, e quando $\sigma > 0$, avesso ao risco (ANDERSEN et al., 2008), de modo que, quanto maior for σ , mais avesso ao risco é o consumidor. Particularmente, para o caso em que $\sigma = 1$, $u(C) = \ln(C)$ (MEHRA; PRESCOTT, 1985; COCHRANE, 2009).

1.2.3 O ativo livre de risco

Como a equação 5 é válida para qualquer tipo de ativo, em particular ela é válida também para o ativo livre de risco. Denotando-se então R_{t+1}^f como o retorno bruto do ativo livre de risco e assumindo que $Cov(M_{t+1}, R_{t+1}^f) = 0$ e que $\mathbb{E}_t \{ R_{t+1}^f \} = R_{t+1}^f$, obtém-se $\mathbb{E}_t \{ M_{t+1} R_{t+1}^f \} = R_{t+1}^f \mathbb{E}_t \{ M_{t+1} \} = 1$ e, portanto,

$$R_{t+1}^f = \frac{1}{\mathbb{E}_t \{ M_{t+1} \}}. \quad (7)$$

Conforme argumentam Jiang et al. (2019), o retorno esperado de uma carteira de títulos da dívida pública deve refletir o perfil de risco do resultado primário do governo, sendo esta uma variável pró-cíclica. Isso quer dizer, por exemplo, que em recessões, quando a utilidade marginal dos consumidores é alta, os resultados primários tendem a ser negativos, e a emissão de dívida pública aumenta. Desse modo, tem-se que os fluxos de caixa de um portfólio de títulos de dívida pública estão sujeitos a um risco relacionado ao ciclo da economia. Nesse sentido, os retornos de títulos públicos contém um prêmio de risco que refletem a exposição a esse tipo de risco. Assim, ao supor que $Cov(M_{t+1}, R_{t+1}^f) = 0$, abstrai-se desse risco relacionado ao ciclo econômico para R_{t+1}^f .⁴

A hipótese discutida acima é amplamente aceita na literatura e normalmente é dada em modelos de apreçamento de ativos (FARNSWORTH et al., 2002; CAMPBELL, 2014; ARAÚJO; ISSLER; FERNANDES, 2005). Conforme argumenta Cochrane (2009), essa hipótese implica que, mesmo que o investidor compre uma unidade a mais desse ativo, isso não surtirá efeito algum sobre o fluxo de consumo desse investidor.

Com o propósito final de estimar a taxa implícita na equação 7, inicialmente é necessário estimar-se o SDF. Para isso, utiliza-se em primeiro lugar o modelo desenvolvido por Araújo, Issler e Fernandes (2005), doravante denominado AIF. Em seguida, utiliza-se o modelo GMM,

⁴ De um outro ponto de vista, Ferreira e Matos (2020) argumentam que, em um contexto de risco cambial, essa hipótese se relaciona à ausência de covariância entre o SDF e mudanças na inflação entre o período t e $t + 1$, esta que é a variável desconhecida para o retorno real esperado em $t + 1$ no modelo proposto pelos autores.

de Hansen (1982) e, por último, o modelo GEL, de Newey e Smith (2004), que serão descritos adiante. A escolha por esses modelos se dá pelo fato de eles dispensarem uma abordagem completamente paramétrica para a estimação do SDF. Conforme argumentado no capítulo introdutório, o uso de modelos paramétricos está sujeito a erros de especificação, justificando a escolha de modelos não-paramétricos e semi-paramétricos para contornar tais erros.

Com as estimações do SDF feitas pelas três abordagens, os valores de $\mathbb{E}_t \{M_{t+1}\}$ serão calculados conforme o modelo a seguir:

$$M_{t+1} = \alpha + \gamma X_t + \lambda \ln \left(\frac{C_t}{Y_t} \right) + \varepsilon_{t+1}, \quad (8)$$

em que X_t é uma variável representativa do mercado financeiro, $\ln \left(\frac{C_t}{Y_t} \right)$ é o logaritmo da razão consumo/PIB e ε_{t+1} é o componente de erro do modelo. Conforme argumenta Simonassi (2006), as escolhas das variáveis para este modelo devem-se ao fato de que o SDF pode ser analisado como uma variável representativa dos fundamentos da economia, de modo que para ser estimado empiricamente, é comum assumir que $SDF = f(\text{dados}; \text{fundamentos})$ (COCHRANE, 2009).

Como $R_{t+1} = \frac{1+i_{t+1}}{1+\pi_{t+1}}$, da equação 5, tem-se que

$$i_{t+1} = \frac{1}{\mathbb{E}_t \left\{ \frac{M_{t+1}}{1+\pi_{t+1}} \right\}} - 1, \quad (9)$$

em que π_{t+1} é a inflação no período $t + 1$ e i_{t+1} é a taxa (nominal) livre de risco a ser estimada.

1.2.4 Modelo AIF

Este modelo propõe uma abordagem não-paramétrica para o cálculo do SDF sem o uso de uma função utilidade, em que o estimador é uma função dos retornos de ativos. O principal benefício disso é a ausência de risco de especificação incorreta para a forma funcional adequada para estimar-se o SDF. Ademais, o estimador proposto não depende de dados de consumo para ser construído. Este benefício surge uma vez que estimações baseadas em consumo geraram importantes *puzzles* na literatura de macro-finanças, como o já citado *Equity Premium Puzzle*. Por último, dispensa-se a hipótese de homocedasticidade para a construção do estimador. Ou seja, pode ser feito o uso de dados de alta e altíssima frequência, que normalmente possuem heterocedasticidade, e a um custo computacional baixo.

Para desenvolver este estimador, Araújo, Issler e Fernandes (2005) argumentam que o SDF é a característica comum⁵ da correlação serial de todos os ativos. Na argumentação dos autores, tem-se que os retornos de qualquer ativo podem ser escritos como $r_{i,t} = -m_t - \gamma_{i,t}^2 + \varepsilon_{i,t}$, $i = 1, \dots, N$, em que $r_{i,t} = \ln(R_{i,t})$, $m_t = \ln(M_t)$, $\gamma_{i,t}^2 = \mathbb{E}_{t-1} \{ \ln(M_t R_{i,t}) \}$ e $\varepsilon_{i,t}$ é um componente de inovação do processo. Como $\gamma_{i,t}^2$ e $\varepsilon_{i,t}$ são idiossincráticos, a única fonte de

⁵ Para duas séries econômicas quaisquer, uma característica comum existe se ela está presente em ambas as séries e pode ser removida por uma combinação linear.

correlação serial é m_t .

A construção do estimador é feita a partir de quatro hipóteses:

1. A equação 5 é válida;
2. $M_t > 0$;
3. Existe um ativo livre de risco;
4. O processo $\ln(M_t R_t)$ é estacionário em covariância, com primeiro e segundo momentos finitos. Seja $\varepsilon_{i,t} = \ln(M_t R_{i,t}) - \mathbb{E}_{t-1} \{\ln(M_t R_{i,t})\}$ o processo de inovação para prever $\ln(M_t R_{i,t})$. Assume-se que $\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N |\mathbb{E} \{\varepsilon_{i,t} \varepsilon_{j,t}\}| = 0$.

O estimador proposto é consistente⁶ à medida que $N, T \rightarrow \infty$, e é dado por⁷

$$M_t^{AIF} = \frac{\bar{R}_t^G}{\frac{1}{T} \sum_{i=1}^T (\bar{R}_t^G \bar{R}_t^A)}, \quad (10)$$

em que $\bar{R}_t^G = \prod_{i=1}^N R_{i,t}^{-\frac{1}{N}}$ e $\bar{R}_t^A = \frac{1}{N} \sum_{i=1}^N R_{i,t}$ representam a média geométrica e a média aritmética dos N ativos, respectivamente.

Ainda de acordo com Araújo, Issler e Fernandes (2005), é importante que seja escolhida uma quantidade suficientemente grande de ativos, com o objetivo de diversificar o portfólio analisado. Isso pois, ao passo que cresce o número de ativos utilizado, o grau de correlação dos retornos diminui. Tal medida é importante devido ao fato de que muita correlação pode levar à não convergência em probabilidade de \hat{M}_t para M_t .

1.2.5 Modelo GMM

A segunda técnica de estimação utilizada neste trabalho é o modelo GMM. Este modelo consiste na estimação de parâmetros finitos, com base em condições de ortogonalidade advindas da teoria econômica que o econometrista deseja testar. No caso de modelos de série temporal, a estimação dos parâmetros de interesse consiste na elaboração de condições de momentos amostrais, análogas aos momentos populacionais, que se aproximam de zero à medida que $T \rightarrow \infty$, em que T é o tamanho da série temporal. Conforme argumenta Hansen (1982), a identificação do modelo requer que a quantidade de condições a serem testadas seja pelo menos igual à quantidade de coordenadas do vetor de parâmetros a ser estimado.

Como não é necessária a especificação completa da natureza da correlação serial ou da heterocedasticidade, e dada a existência de um vetor de parâmetros, essa abordagem consiste em uma estimação semi-paramétrica (HOROWITZ, 2009). Uma vantagem do GMM é que ele possui boas propriedades assintóticas, ao passo que permite uma estimação consistente dos parâmetros na presença de heterocedasticidade condicional e autocorrelação serial nas condições

⁶ Embora não haja a necessidade de impor homocedasticidade para os erros, este estimador se torna eficiente apenas sob essa hipótese.

⁷ Para uma prova completa desta equação, consulte Araújo, Issler e Fernandes (2005).

de ortogonalidade, e tendo em vista que o estimador proposto possui distribuição assintoticamente Normal (HANSEN, 1982). Como não são impostas restrições de linearidade sobre essas condições, o GMM consegue lidar de maneira direta com modelos não lineares (CAMPBELL, 2014), o que se torna um benefício principalmente no caso de um SDF como especificado na equação 6.

Nos moldes de Hansen (1982), para obter as condições de momento, suponha um modelo econométrico dado por $u_t = F(x_t, \beta_0)$, $z_t = G(x_t, \beta_0)$ e $\mathbb{E}[u_t \otimes z_t] = 0$. Nesse modelo, x_t é uma série de tempo, β_0 é um vetor de parâmetros, u_t é um vetor de choques não observável, z_t é um vetor de variáveis instrumentais, \otimes é o produto de Kronecker e F e G são funções especificadas *a priori*⁸. Seja ainda $f(x_t, \beta_0) = F(x_t, \beta_0) \otimes G(x_t, \beta_0)$. Tem-se então que a condição de ortogonalidade é dada por $\mathbb{E}[f(x_t, \beta_0)] = 0$. O vetor de condições de ortogonalidade é utilizado em conjunto com uma matriz de ponderação aleatória⁹ para obter uma função objetiva amostral, cujo argumento que a minimiza é justamente o estimador de β_0 . Defina então as condições de momentos amostrais como $g_T(x_t, \beta_0) = \frac{1}{T} \sum_{t=1}^T f(x_t, \beta_0)$. Para uma matriz de ponderação W , o problema a ser resolvido pelo GMM é dado por (COCHRANE, 2009)

$$\min_{\beta_0} g_T(x_t, \beta_0)' W g_T(x_t, \beta_0). \quad (11)$$

Neste modelo, existe um teste de sobreidentificação para se testar o quão próximas as condições de ortogonalidade amostrais estão de zero. Seja R o número de condições de ortogonalidade e Q o número de parâmetros a serem estimados. A estimação do modelo é feita ao se igualar Q combinações lineares das R condições de ortogonalidade a zero. Assim, para o caso em que $R > Q$, existem $R - Q$ combinações lineares independentes das condições de ortogonalidade que deveriam ser zero, mas não o são. O teste de sobreidentificação proposto por Hansen (1982) consiste em uma extensão do teste J de Sargan (SARGAN, 1958; SARGAN, 1959) e é dado por

$$J = g_T(x_t, \beta_0)' W g_T(x_t, \beta_0) \sim \chi_{R-Q}^2, \quad (12)$$

em que a hipótese nula é $H_0 : \mathbb{E}[f(x_t, \beta_0)] = 0$. Neste teste, rejeitar a hipótese nula indica que o modelo está mal especificado.

Para que as condições de ortogonalidade utilizadas na estimação em (11) sejam satisfeitas apenas no vetor verdadeiro de parâmetros, é necessário que seja escolhida uma matriz de ponderação simétrica, positiva definida e de dimensão $R \times R$, e a escolha dessa matriz define a métrica de distância utilizada para aproximar $g_T(\cdot)$ de zero (HANSEN, 1982; HANSEN; SINGLETON, 1982). Como exposto em Hansen e Singleton (1982), a estimação de uma matriz de covariância ótima pode ser feita em dois estágios. Essa estratégia de estimação é conhecida como *Two-Step GMM* e seu primeiro estágio consiste em uma escolha sub-ótima da matriz W

⁸ A equação $\mathbb{E}[u_t \otimes z_t] = 0$ surge da hipótese de que o vetor de choques da economia é ortogonal a um conjunto de variáveis observadas pelo economista.

⁹ O fato de permitir que essa matriz seja aleatória se dá pela sua possível dependência das informações contidas na amostra utilizada.

para se encontrar um estimador $\hat{\beta}_1$ para β_0 em (11), sendo comum o uso da matriz identidade I_R (COCHRANE, 2009).

Como argumentado por Newey e West (1986), uma estimação consistente para a matriz de covariância utilizada no segundo estágio é essencial para a construção de intervalos de confiança assintóticos e testes de hipótese. Nesse sentido, os autores propõem um método de estimação para essa matriz, conhecido como estimador HAC (*Heteroskedasticity and Autocorrelation Consistent*). A matriz em questão é simétrica e positiva definida por construção, possui dimensão $R \times R$ e é dada por

$$\hat{S} = \hat{\Omega}_0 + \sum_{j=1}^m w(j, m)[\hat{\Omega}_j + \hat{\Omega}'_j], \text{ em que} \quad (13)$$

$$w(j, m) = 1 - [j/(m + 1)] \text{ e}$$

$$\hat{\Omega}_j = \frac{1}{T} \sum_{t=j+1}^T f(x_t, \hat{\beta}_1) f(x_{t-j}, \hat{\beta}_1)',$$

em que m é o número de autocovariâncias amostrais $\hat{\Omega}_j$, determinado pela ordem do termos de médias móveis de u_t (HANSEN; SINGLETON, 1982; NEWEY; WEST, 1986). No segundo estágio, utiliza-se então $W = \hat{S}^{-1}$ como a matriz de covariância ótima para se encontrar

$$\hat{\beta}_2 = \arg \min_{\beta_0} g_T(x_t, \beta_0)' \hat{S}^{-1} g_T(x_t, \beta_0). \quad (14)$$

Conforme argumentam Hansen e Singleton (1982), pode-se entender as condições de ortogonalidade $\mathbb{E}[f(x_t, \beta_0)] = 0$ como equações de Euler estocásticas resultantes de um problema de maximização de utilidade de um consumidor representativo em um ambiente com incerteza. Assim, para o caso específico deste trabalho, considere $u_t = \beta \frac{u'(C_{t+1}; \beta_0)}{u'(C_t; \beta_0)} R_{t+1} - 1$ e z_t um conjunto de variáveis observadas em t . As condições de ortogonalidade populacionais podem então ser escritas como

$$\mathbb{E} \left\{ \left(\beta \frac{u'(C_{t+1}; \beta_0)}{u'(C_t; \beta_0)} R_{t+1} - 1 \right) \otimes z_t \right\} = 0. \quad (15)$$

Seja $\beta_0 = (\beta, \sigma)$. Ao assumir uma função utilidade do tipo CRRA, o SDF toma a forma descrita pela equação 6, de modo então que

$$g_T(x_t, \beta_0) = g_T(\beta, \sigma) = \frac{1}{T} \sum_{t=0}^T \left\{ \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} R_{t+1} - 1 \right] \otimes z_t \right\}, \quad (16)$$

em que o vetor x_t foi omitido de $g_T(\beta, \sigma)$ para fins de simplificação de notação. Após a estimação dos parâmetros, os valores encontrados para $\hat{\beta}^{GMM}$ e $\hat{\sigma}^{GMM}$ serão utilizados na equação 6 de modo a se obter

$$M_{t+1}^{GMM} = \hat{\beta}^{GMM} \left(\frac{C_{t+1}}{C_t} \right)^{-\hat{\sigma}^{GMM}}. \quad (17)$$

1.2.6 Modelo GEL

Embora o estimador de GMM possua boas propriedades assintóticas, evidências apontam para um alto viés nas estimações em amostras pequenas (SMITH, 1997). De acordo com Newey e Smith (2004), uma desvantagem do GMM é que o viés¹⁰ do estimador costuma aumentar conforme aumentam as condições de momento utilizadas na estimação. Essa é uma propriedade importante corrigida pelos modelos da classe *Generalized Empirical Likelihood Estimators* (GEL), uma vez que é comum o uso de muitos instrumentos em modelos econométricos, e conforme o número de instrumentos aumenta, cresce a quantidade de condições de ortogonalidade a serem testadas. É o caso, por exemplo, do trabalho de Hansen e Singleton (1982), que testam o modelo utilizando diferentes instrumentos, sendo eles 1, 2, 4 ou 6 defasagens das variáveis observadas.

Desenvolvida no trabalho de Newey e Smith (2004), a classe de estimadores GEL sugere alternativas para o GMM em um esforço de aprimorar as propriedades em amostras pequenas desses estimadores, mas que possuem as mesmas propriedades assintóticas. Para construir o modelo GEL, considere uma função $\rho(v)$ estritamente côncava que satisfaça as seguintes condições: $\rho(0) = 0$, $\frac{\partial \rho(0)}{\partial v} = \frac{\partial^2 \rho(0)}{\partial v^2} = -1$. O estimador GEL $\hat{\beta}_0$ é obtido ao se resolver

$$\min_{\beta_0} \sup_{\lambda \in \Lambda_T} \sum_{t=1}^T \rho(\lambda' g_t(\beta_0)), \quad (18)$$

em que $\Lambda_t = \{\lambda : \lambda' g_t(b) \in \Upsilon\}$, Υ é algum conjunto aberto que contém zero e $g_t(\beta_0) \equiv f(x_t, \beta_0)$, em que $f(x_t, \beta_0)$ é construída da mesma maneira que na Subseção 1.2.5 (NEWHEY; SMITH, 2004).

De acordo com os autores, essa é uma abordagem genérica, uma vez que diferentes escolhas para $\rho(v)$ levam a casos especiais de estimadores alternativos ao GMM. Para o caso em que $\rho(v) = -e^v$, tem-se o chamado estimador de *Exponential Tiltening* (KITAMURA; STUTZER, 1997). Para $\rho(v) = -\frac{1}{2}v^2 - v$, tem-se um estimador equivalente ao *Continuous Updating Estimator* GMM, desenvolvido por Hansen, Heaton e Yaron (1996). Por último, para $\rho(v) = \ln(1 - v)$ e $\Upsilon = (-\infty, 1)$, tem-se o estimador de *Empirical Likelihood* (OWEN, 1988; IMBENS, 1997), que será o utilizado neste trabalho.

Para construir esse estimador, considere o argumento exposto por Anatolyev e Gospodinov

¹⁰ Os autores argumentam que uma das principais fontes desse viés se dá pela correlação das funções de momentos com seus Jacobianos na presença de endogeneidade. Para modelos com dados em painel, outra importante fonte de viés surge da estimação da matriz de ponderação ótima.

(2011), a seguir. Seja uma amostra aleatória $\{x_t\}_{t=1}^T$ de uma população cuja distribuição é dada por $F(x)$ e densidade $f(x)$. Para x_t i.i.d., a função de verossimilhança conjunta é dada por $\prod_{t=1}^T f(x_t)$. A abordagem de verossimilhança usual, paramétrica, assume uma forma particular para $f(x_t|\eta)$, em que η é o vetor de parâmetros que caracteriza essa distribuição. A função de verossimilhança é então maximizada com respeito a η e recupera estimações de $f(x)$ e $F(x)$ da amostra observada.

Ao contrário da abordagem acima, a estimação via verossimilhança empírica (*empirical likelihood*) não parametriza a densidade $f(x)$ e, como alternativa a η , ajusta um vetor T -dimensional de probabilidades a cada uma das observações: $\mathbf{p} = (p_1, p_2, \dots, p_T)$, em que $p_t = f(x_t)$, $t = 1, \dots, T$ e $\sum_{t=1}^T p_t = 1$, consistindo então de uma abordagem não-paramétrica. Adicionalmente, o modelo incorpora uma restrição às condições de momento $g_t(\beta_0)$ com respeito a \mathbf{p} . Conforme argumenta Smith (1997), essa restrição tem como objetivo suavizar o modo como o modelo lida com as condições de momentos impostas, ao associar pesos a cada uma delas. O problema de verossimilhança empírica se torna então:

$$\begin{aligned} \max_{\mathbf{p}, \beta_0} \quad & \frac{1}{T} \sum_{t=1}^T \log(p_t) \\ \text{s.t.} \quad & \sum_{t=1}^T p_t g_t(\beta_0) = \mathbf{0} \\ & \sum_{t=1}^T p_t = 1, \end{aligned} \quad (19)$$

Para encontrar a solução para este problema, seja λ o vetor $Q \times 1$ de multiplicadores de Lagrange para a equação 19. O Lagrangiano pode ser escrito como

$$\mathcal{L}(p_1, p_2, \dots, p_t, \lambda, \mu) = \frac{1}{T} \sum_{t=1}^T \log(p_t) - \lambda' \sum_{t=1}^T p_t g_t(\beta_0) - \mu \left(\sum_{t=1}^T p_t - 1 \right). \quad (20)$$

Ao resolver as condições de primeira ordem para este problema, encontra-se as probabilidades de verossimilhança empírica, dadas por

$$\hat{p}_t = \frac{1}{T} \frac{1}{1 + \lambda' g_t(\hat{\beta}_0)}, t = 1, \dots, T. \quad (21)$$

Por último, após certa manipulação¹¹, tem-se que

$$\hat{\beta}_0 = \arg \max_{\beta_0} \min_{\lambda} \frac{1}{T} \sum_{t=1}^T -\log(1 + \lambda' g_t(\beta_0)). \quad (22)$$

As condições de momentos utilizadas são as mesmas da equação 16, e após a estimação dos parâmetros, os valores $(\beta, \sigma) = (\hat{\beta}^{GEL}, \hat{\sigma}^{GEL})$ serão inseridos na equação 6 de modo a se obter

$$M_{t+1}^{GEL} = \hat{\beta}^{GEL} \left(\frac{C_{t+1}}{C_t} \right)^{-\hat{\sigma}^{GEL}}. \quad (23)$$

¹¹ O leitor encontrará uma demonstração completa dos passos que se sucedem a partir da equação 19 na seção 2.2.2 de Anatolyev e Gospodinov (2011).

Uma outra vantagem do modelo GEL é que ele consta com o mesmo teste de especificação do modelo GMM, especificado na equação 12, dadas as equivalências das propriedades assintóticas desses estimadores (ANATOLYEV; GOSPODINOV, 2011; SMITH, 2011).

1.2.7 O método de Low Frequency Econometrics

Como a construção do CCAPM (BREEDEN, 1979) se baseia no ambiente econômico descrito por Lucas (1978), este modelo implica em uma relação de equilíbrio de longo prazo. Dessa forma, é importante que a modelagem empírica desse modelo seja feita utilizando-se dados relativos à tendência de longo prazo da série de consumo, para que o modelo empírico seja capaz de recuperar a relação de equilíbrio imposta pelo modelo teórico.

De maneira a extrair apenas este componente de longo prazo da série de consumo, utilizou-se o método de suavização desenvolvido por Müller e Watson (2015). Essa necessidade se dá uma vez que que a série de consumo costuma apresentar um padrão sazonal, muito ligado a componentes cíclicos (de curto prazo) da economia. Um exemplo disso é a representação da utilidade marginal do consumo como um passeio aleatório com tendência (HALL, 1978). Tais oscilações de curto prazo podem se dar, por exemplo, quando há um choque temporário no preço de ações de uma economia. Conforme argumenta Hamilton (2018), preços de ações possuem alguma habilidade em antecipar mudanças no consumo agregado. Dessa forma, um choque temporário de preços de ações leva a mudanças de curto prazo no consumo agregado.

O método proposto por Müller e Watson (2015) conta com projeções da série temporal em um número pequeno de funções trigonométricas para extrair o componente de tendência da série analisada, que os autores denominam de variabilidade de baixa frequência. De acordo com os autores, o uso de métodos de projeção é justificado pelo baixo custo computacional da aplicação de tais métodos, e pela simplicidade da inferência estatística resultante dessa aplicação.

Seja x_t uma série temporal observada em $t = 1, \dots, T$, e $\Psi_j(s) = \sqrt{2}\cos(js\pi)$. $\Psi_j(t/T)$ possui então período $2T/j$. Seja $\Psi(s) = [\Psi_1(s), \dots, \Psi_q(s)]' \in \mathbb{R}^q$. Além disso, $\Psi_T = \left[\Psi_1\left(\frac{1-\frac{1}{2}}{T}\right), \Psi\left(\frac{2-\frac{1}{2}}{T}\right), \dots, \Psi\left(\frac{T-\frac{1}{2}}{T}\right) \right]'$, uma matriz $T \times q$ obtida ao se avaliar $\Psi(\cdot)$ em $s = \frac{t-\frac{1}{2}}{T}$. As projeções de baixa frequência (a tendência de longo prazo) são os valores ajustados da seguinte regressão em MQO (Mínimos Quadrados Ordinários):

$$\hat{x}_t = \bar{x} + \Psi\left(\frac{t-\frac{1}{2}}{T}\right)' X_T, \quad (24)$$

em que $\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t$ e $X_T = \frac{1}{T} \sum_{t=1}^T \Psi((t-1/2)/T)x_t$ é o vetor de coeficientes da regressão. O j -ésimo coeficiente, X_{jT} , é a chamada j -ésima transformada de cosseno de $[x_1, x_2, \dots, x_T]'$.

Conforme argumentado no capítulo de Introdução, a opção por não usar o Filtro HP se dá por este produzir resultados artificiais, uma vez que o componente de tendência resultante da sua aplicação não é característico da verdadeira dinâmica do processo gerador de dados em si, e sim de uma caracterização particular do uso do próprio filtro (HAMILTON, 2018).

Considere uma série temporal $y_t = g_t + c_t$, em que g_t é o componente de tendência da série

e c_t é o componente cíclico. O problema de otimização resolvido pelo Filtro HP, conforme apresentado em Hodrick e Prescott (1997) é dado por

$$\min_{g_{t=1}^T} \left\{ \sum_{t=1}^T c_t^2 + \lambda \sum_{t=1}^T [(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})]^2 \right\}, \quad (25)$$

em que $c_t = y_t - g_t$, λ é um parâmetro (positivo) de penalização para a variabilidade da tendência da série e $t = 1, 2, \dots, T$. Para encontrar a solução para esse problema, considere $\tilde{T} = T + 2$, $y = (y_T, y_{T-1}, \dots, y_1)'$, um vetor $T \times 1$, e $g = (g_T, g_{T-1}, \dots, g_{-1})'$, um vetor $\tilde{T} \times 1$. Considere ainda as matrizes

$$H = \begin{bmatrix} I_T & 0 \end{bmatrix} \text{ e}$$

$$Q = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 0 \end{bmatrix},$$

em que H é uma matriz $T \times \tilde{T}$, I_T é $T \times T$, 0 é $T \times 2$ e Q é $T \times \tilde{T}$. A solução para a equação 25 é dada então por $g^* = (H'H + \lambda Q'Q)^{-1} H'y = A^*y$, em que $A^* = (H'H + \lambda Q'Q)^{-1} H'$. Assim, a tendência g_t^* para qualquer data t é uma função linear do conjunto completo das observações de y para todo $t = 1, \dots, T$ (HAMILTON, 2018)¹². Logo, o componente de tendência resultante do Filtro HP possui uma habilidade artificial em recuperar a tendência verdadeira do processo gerador de dados, uma vez que ele é, por construção, uma função de realizações futuras de y , podendo essas observações não serem conhecidas em tempo real.

¹² Em apêndice, o autor fornece uma derivação completa da solução da equação 25 no formato matricial.

2 DADOS E RESULTADOS

2.1 Dados

O período de coleta foi do primeiro trimestre de 2002 ao segundo trimestre de 2022. Utilizou-se dados de consumo final das famílias e PIB a preços de mercado, ambos com frequência trimestral, encadeados a preços de 1995, e retirados do Sistema IBGE de Recuperação Automática (SIDRA) do Instituto Brasileiro de Geografia e Estatística (IBGE). Os dados da razão consumo/PIB foram utilizados na equação 8. Para o cálculo do crescimento do consumo (C_{t+1}/C_t) foram utilizados dados dessazonalizados do consumo final das famílias encadeados a preços de 1995, também retirados do SIDRA. Estes dados foram utilizados no cálculo do SDF extraído da utilidade CRRA, de acordo com a equação 6. Os dados de expectativa de inflação (expectativa média de Inflação - IPCA - taxa acumulada para os próximos doze meses) foram retirados do Banco Central do Brasil, Boletim, Seção Atividade Econômica (Bacen / Boletim / Ativ. Ec.) e foram utilizados na equação 9. Os dados da taxa DI foram retirados do Sistema Gerador de Séries Temporais (SGS) do Banco Central do Brasil.

Para representar os retornos utilizados na equação 5 e 8, coletou-se dados trimestrais dos retornos das 150 ações mais líquidas na B3. As ações foram classificadas em ordem decrescente de acordo com o volume de negociação médio durante o período de análise. Em seguida, foram retirados os papéis cujos retornos não estavam disponíveis para algum sub-período da amostra. Ao final, restaram 76 papéis¹. Os retornos do Ibov (Índice Bovespa) foram retirados da Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (Anbima)². Para o cálculo dos retornos reais, utilizou-se dados de inflação, medidos pelo IPCA acumulado em 12 meses, que foram retirados do Instituto Brasileiro de Geografia e Estatística, Sistema Nacional de Índices de Preços ao Consumidor (IBGE/SNIPC). Todos os dados de inflação, retornos e os da taxa DI foram anualizados.

Para a estimação do modelo MQO (Mínimos Quadrados Ordinários) na equação 8, foram utilizados dados do logaritmo natural da razão consumo/PIB e, no vetor X_t , foram testadas quatro combinações de retornos. Essas combinações consistiram em utilizar-se: os retornos do Ibovespa, a média das cinco, das dez e das quinze ações de melhor *performance* no período, de acordo com o volume médio de negociação. Foram elas, em ordem decrescente no *ranking*: PETR4, VALE3, ITUB4, BBDC4, BBAS3, PETR3, ABEV3, ITSA4, GGBR4, USIM5, CSNA3, CMIG4, ELET3, ELET6, EMBR3. Estes portfólios foram utilizados também nas condições de momentos do GMM e do GEL e nos seus respectivos vetores de instrumentos.

A tabela 1 apresenta algumas estatísticas descritivas das séries de retornos reais utilizadas, em ordem decrescente de classificação no *ranking*, com exceção dos retornos do Ibovespa, que não faz parte do *ranking*. Para ilustrar as dinâmicas das demais variáveis ao longo do tempo, apresentam-se as figuras 1, 2 e 3.

¹ Para uma lista completa das ações restantes, ver Apêndice A.

² Foram utilizadas médias trimestrais dos retornos do Ibovespa.

Ativo	Retorno real médio (% a.a.)	Desvio-padrão (p.p.)
PETR4	-0,88	19,65
VALE3	0,10	18,01
ITUB4	-1,38	14,11
BBDC4	-1,14	15,87
BBAS3	0,12	19,52
PETR3	-0,90	20,23
ABEV3	-1,57	13,60
ITSA4	-0,92	14,26
GGBR4	0,32	22,17
USIM5	1,07	28,16
CSNA3	2,13	28,51
CMIG4	-0,66	19,22
ELET3	-0,61	25,24
ELET6	-0,77	21,12
EMBR3	-3,51	18,88
Média TOP5	-0,64	14,81
Média TOP10	-0,52	14,89
Média TOP15	-0,57	14,50
IBOV	18,54	60,29

Fontes: Economatica e Ambima

Tabela 1 – Estatísticas dos ativos utilizados

A figura abaixo ilustra o comportamento da taxa DI, da expectativa de inflação e do IPCA ao longo do tempo. Os dados foram transformados da frequência mensal para a trimestral utilizando a média do período referente a um trimestre.

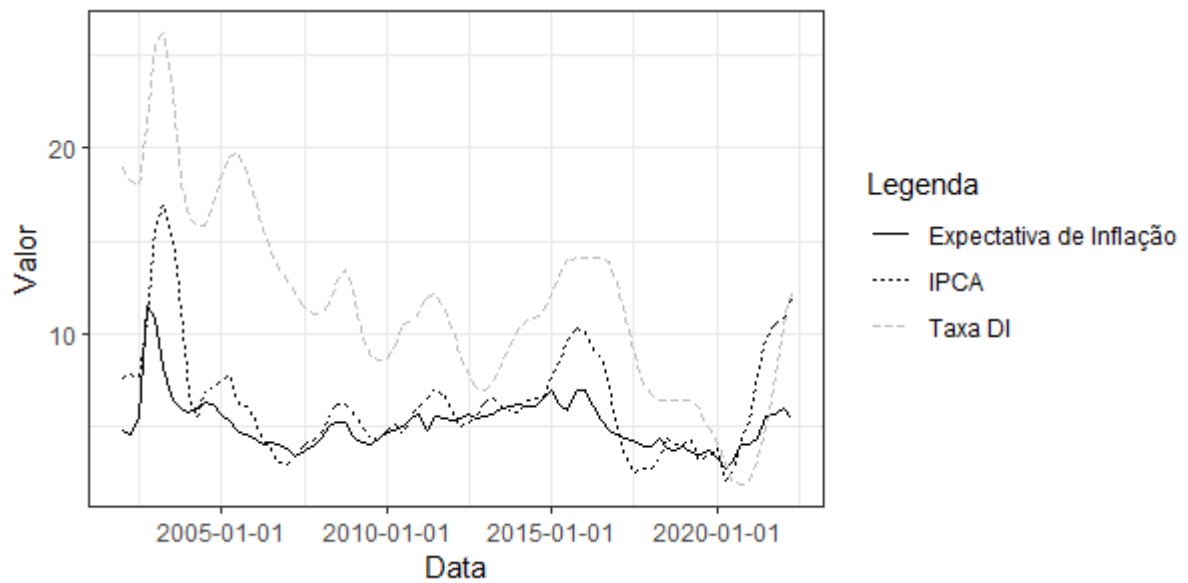


Figura 1 – Expectativa de Inflação, IPCA e DI trimestrais (% a.a.)
Fontes: Banco Central do Brasil e IBGE

O crescimento (bruto) do consumo foi utilizado na equação 16 para representar o termo C_{t+1}/C_t e também no vetor de instrumentos z_t dessa mesma equação.

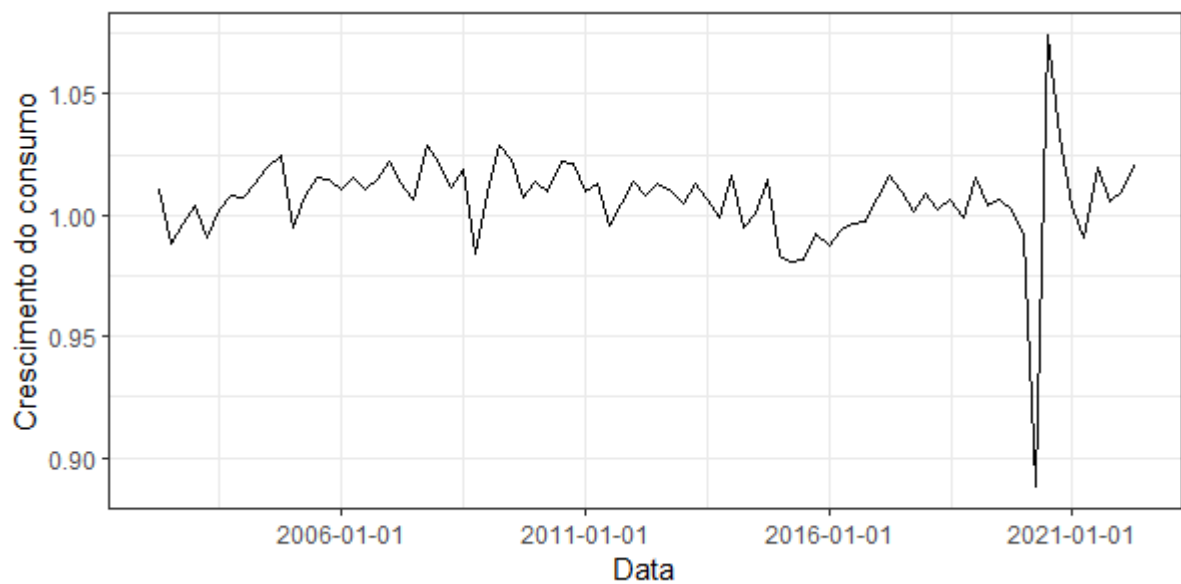


Figura 2 – Crescimento do consumo
Fonte: SIDRA IBGE

O consumo em proporção do PIB, por sua vez, foi utilizado como uma das variáveis explicativas na equação 8. A figura a seguir apresenta os dados razão consumo/PIB e do consumo/PIB suavizado, após ser aplicado o método de suavização descrito na subseção 1.2.7.

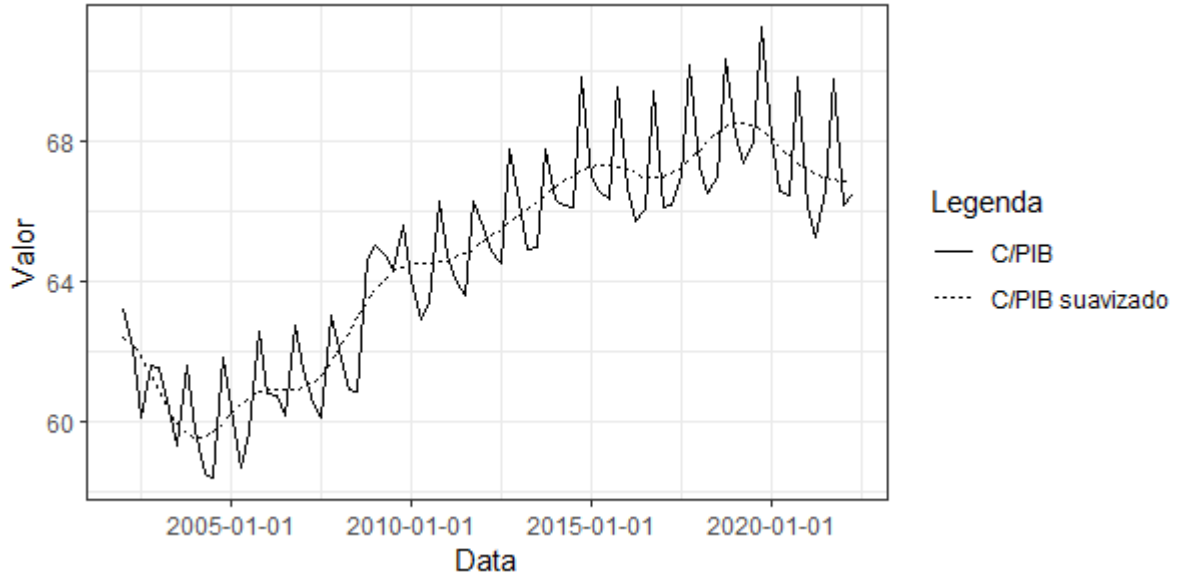


Figura 3 – Consumo em proporção do PIB (%)

Fonte: SIDRA IBGE e Elaboração Própria a partir de dados do SIDRA

2.2 Resultados

Esta seção dedica-se à apresentação dos resultados dos modelos AIF, GMM e GEL. Serão apresentados os resultados do modelo MQO da equação 8 para os três modelos em questão, bem como as estimações dos parâmetros β e σ nos modelos GMM e GEL, e as taxas livres de risco estimadas por cada um dos três modelos.

Na equação 8, o vetor X_t foi construído de forma a conter algum dos seguintes retornos: $\overline{RET5}_t$, $\overline{RET10}_t$, $\overline{RET15}_t$, que são, respectivamente, a média dos cinco, dez e quinze retornos das ações mais bem posicionadas no *ranking* descrito na seção 2.1, ou $Ibov_t$, que são os retornos do Ibovespa. A escolha do modelo de MQO que foi levado adiante se deu com base no seu R^2 ajustado e/ou na significância dos coeficientes, ou seja, no p-valor de cada um deles.

Na estimação do modelo AIF, o termo \bar{R}_t incluiu os 76 papéis que fizeram parte do *ranking* descrito na seção 2.1 para a construção de \bar{R}_t^A e \bar{R}_t^G . Após a construção de \bar{R}_t , os valores encontrados para M_t^{AIF} de acordo com a equação 10 foram utilizados na regressão em (8) para estimar-se $\mathbb{E}_t \{M_{t+1}\}$. A tabela 2 representa os resultados dessa regressão em MQO para o modelo AIF.

	MQO 1		MQO 2		MQO 3		MQO 4	
	Estim.	p-valor	Estim.	p-valor	Estim.	p-valor	Estim.	p-valor
$\overline{RET5}_t$	-0,0004007	0,997						
$\overline{RET10}_t$			-0,02167	0,838				
$\overline{RET15}_t$					-0,02870	0,792		
$Ibov_t$							-0,019418	0,46
$\ln(C_t/PIB_t)$	0,2465076	5,79e-15***	0,25160	2,25e-15***	0,25327	4,69e-15***	0,251974	<2e-16***
R^2 ajustado	0,9814		0,9814		0,9814		0,9815	

Nota: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Fonte: Elaboração própria

Tabela 2 – Resultados do modelo MQO - AIF

Pela tabela acima, nota-se que em nenhum modelo os retornos foram significantes, mas em todos eles a variável $\ln(C_t/PIB_t)$ apresentou p-valor < 0.05 . Os modelos também apresentaram bons valores para o R^2 ajustado, todos acima de 0,98. Com base no p-valor de $\ln(C_t/PIB_t)$ e no R^2 das quatro regressões, escolheu-se então o modelo MQO 4 para prosseguir com as estimações da taxa livre de risco no modelo AIF.

Para a estimação dos parâmetros $\hat{\beta}$ e $\hat{\sigma}$ nos modelos GMM e GEL, foram testadas diferentes combinações de portfólios para representar o termo R_{t+1} nas condições de momentos (equação 16). Estes portfólios foram os mesmos testados nos modelos de MQO da equação 8, isto é, $\overline{RET5}_{t+1}$, $\overline{RET10}_{t+1}$, $\overline{RET15}_{t+1}$ ou $Ibov_{t+1}$. Para o vetor de instrumentos, testou-se os modelos com $z = (1, R_t, C_t/C_{t-1})$, $z = (1, C_t/C_{t-1}, C_{t-1}/C_{t-2})$ e $z = (1, C_t/C_{t-1})$. No vetor z , o termo R_t representa $\overline{RET5}_t$, $\overline{RET10}_t$, $\overline{RET15}_t$ ou $Ibov_t$. As combinações dos retornos foram feitas utilizando-se tanto os quatro portfólios em questão quanto apenas um deles, ou uma combinação entre dois e três dos portfólios construídos. Para fins de ilustração, considere o caso em que $R_{t+1} = (\overline{RET5}_{t+1}, \overline{RET10}_{t+1}, \overline{RET15}_{t+1}, Ibov_{t+1})$. Neste caso, $R_t = (\overline{RET5}_t, \overline{RET10}_t, \overline{RET15}_t, Ibov_t)$. Considere agora o caso em que $R_{t+1} = (\overline{RET5}_{t+1}, \overline{RET10}_{t+1})$. Tem-se então que $R_t = (\overline{RET5}_t, \overline{RET10}_t)$. A mesma lógica aplica-se ao caso em que utiliza-se três portfólios ou apenas um deles.

As escolhas para R_{t+1} e z se deram com base no critério de significância econômica dos parâmetros estimados. Isto é, procurou-se combinações de retornos e instrumentos que gerassem estimações para β tais que $0 < \hat{\beta} < 1$, e estimações para σ tais que $\hat{\sigma} > 0$. Para o GMM, a combinação que gerou os resultados mais aderentes foi $R_{t+1} = Ibov_{t+1}$ e $z = (1, C_t/C_{t-1}, C_{t-1}/C_{t-2})$. Para o GEL, as escolhas para os retornos foram $R_{t+1} = \overline{RET15}_{t+1}$, e para o vetor de instrumentos, utilizou-se $z = (1, \overline{RET15}_t, C_t/C_{t-1})$. A tabela 3 abaixo ilustra as estimações realizadas para β e σ para os modelos GMM e GEL, bem como os resultados para o Teste J de sobreidentificação.

	GMM		GEL	
	Estim.	p-valor	Estim.	p-valor
$\hat{\beta}$	0,85927	<2e-16***	0,986941	<2e-16***
$\hat{\sigma}$	1,35257	0,908	-2,187177	0,171
Estatística J	2,5121		0,33384	
DF	1		1	
p-valor (J)	0,11297		0,56341	

Nota: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. DF: graus de liberdade para a construção da estatística J.

Fonte: Elaboração própria

Tabela 3 – Parâmetros estimados dos modelos GMM e GEL

Nota-se pela tabela acima que não há evidências de especificação incorreta para os modelos GMM e GEL, uma vez que não se rejeita a hipótese nula para o teste J ao nível 0,05 para nenhum deles. Entretanto, apenas o estimador $\hat{\beta}$ foi ao mesmo tempo estatisticamente e economicamente significativo em ambos os modelos. O estimador $\hat{\sigma}$, por outro lado, não foi estatisticamente significativo em nenhum deles, e foi economicamente significativo apenas no GMM. Para fins de comparação com alguns resultados já existentes na literatura, cita-se Cysne (2006), que encontra valores para σ entre 14,15 e 19,12. Para o fator de desconto, o autor encontra $\beta = 0,995$ em simulações feitas com $\sigma > 7$, e $\beta = 0,9740$ em simulações feitas com $\sigma > 9$. Issler e Piqueira (2000) encontram o valor mediano de 0,96 para β para dados anualizados em frequência trimestral, e de 1,70 para σ .

As comparações com estes trabalhos da literatura brasileira indicam que os modelos não apresentaram resultados robustos. Quer dizer, para o GMM, embora o coeficiente de aversão ao risco estimado esteja próximo ao encontrado na literatura, ele não é significativo, e o valor para o fator de desconto subjetivo se mostrou muito baixo. Para o GEL, embora o fator de desconto tenha se mostrado significativo e em linha com a literatura, o coeficiente de aversão ao risco não foi significativo e possui sinal negativo, indicando um comportamento de afinidade ao risco do investidor brasileiro.

Com as estimativas acima, os valores calculados para M_{t+1}^{GMM} e M_{t+1}^{GEL} de acordo com as equações 17 e 23, respectivamente, foram utilizados na regressão em (8) para se estimar $\mathbb{E}_t \{M_{t+1}\}$. As tabelas 4 e 5 apresentam os resultados dos modelos MQO para o GMM e para o GEL, respectivamente.

	MQO 1		MQO 2		MQO 3		MQO 4	
	Estim.	p-valor	Estim.	p-valor	Estim.	p-valor	Estim.	p-valor
\overline{RET}_t	-0,056015	0,00104**						
\overline{RET}_{10_t}			-0,059790	0,000418***				
\overline{RET}_{15_t}					-0,060923	0,000447***		
I_{bovt}							-0,010567	0,0141*
$\ln(C_t/PIB_t)$	0,217743	< 2e-16***	0,218666	< 2e-16***	0,218920	< 2e-16***	0,207408	< 2e-16***
R^2 ajustado	0,9993		0,9994		0,9994		0,9993	

Nota: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Fonte: Elaboração própria

Tabela 4 – Resultados do modelo MQO - GMM

	MQO 1		MQO 2		MQO 3		MQO 4	
	Estim.	p-valor	Estim.	p-valor	Estim.	p-valor	Estim.	p-valor
$\overline{RET5}_t$	0,130359	4,19e-05***						
$\overline{RET10}_t$			0,142284	5,68e-06***				
$\overline{RET15}_t$					0,138025	1,97e-05***		
I_{bov}_t							0,028235	0,000421***
$\ln(C_t/PIB_t)$	0,209190	< 2e-16***	0,206291	< 2e-16***	0,207349	< 2e-16***	0,232186	< 2e-16***
R^2 ajustado	0,9984		0,9985		0,9985		0,9983	

Nota: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Fonte: Elaboração própria

Tabela 5 – Resultados do modelo MQO - GEL

Os coeficientes de todos os retornos e de $\ln(C_t/PIB_t)$ foram estatisticamente significantes para todos os modelos de MQO, tanto para o GMM quanto para o GEL, e apresentaram bons valores para o R^2 ajustado, todos acima de 0,99. Com base no R^2 e no p-valor, para o GMM o modelo que foi levado adiante foi o MQO 3, e para o GEL a escolha foi o MQO 2.

Finalmente, com os resultados de $\mathbb{E}_t \{M_{t+1}\}$ encontrados pelas três abordagens (AIF, GMM e GEL), estimou-se a taxa livre de risco nominal da equação 9. A figura 4 ilustra o comportamento das taxas estimadas pelos três modelos e o comportamento da taxa DI ao longo do tempo, para fins de comparação entre os resultados.

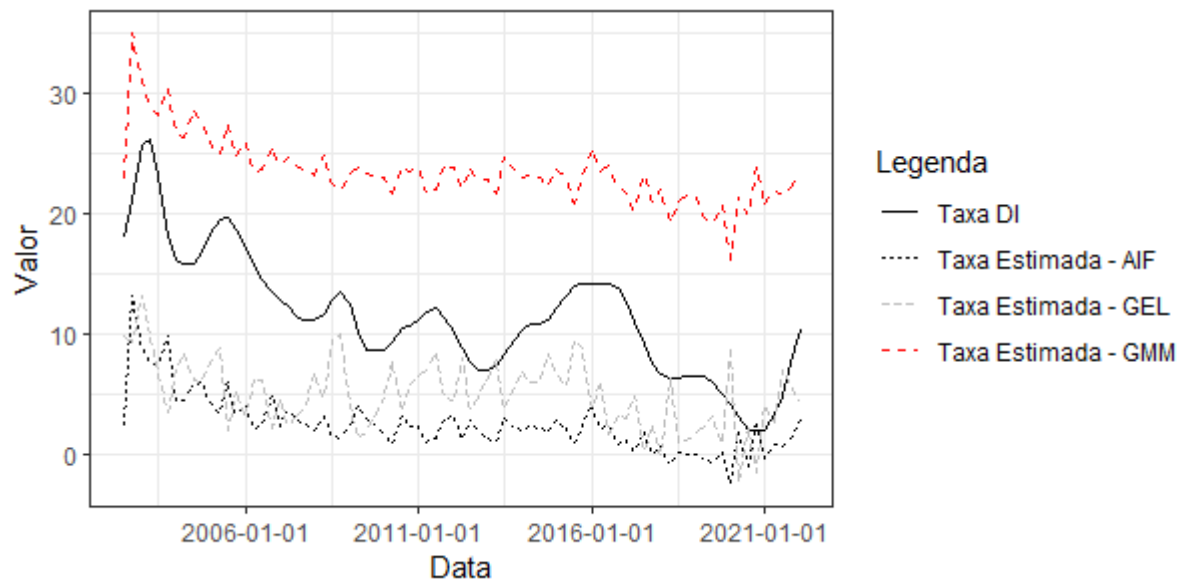


Figura 4 – DI vs. Taxas estimadas (% a.a.)

Fonte: Elaboração própria

Os resultados encontrados pelos modelos AIF e GEL vão ao encontro dos resultados de Bonomo e Domingues (2002), que encontram evidências de um *risk-free rate puzzle* invertido para o Brasil, em que as taxas estimadas pelos modelos são menores que aquelas observadas na economia. No modelo dos autores, essa discrepância entre a taxa resultante do modelo e a taxa observada surge ao se utilizar um fator de desconto subjetivo (β) compatível com dados trimestrais anualizados, isto é, no intervalo entre 0,9 e 1.

Por outro lado, o modelo GMM vai ao encontro do *risk-free rate puzzle* clássico de Weil

(1989), em que as taxas estimadas pelo modelo são maiores que aquelas observadas na economia. No modelo do autor, para encontrar taxas livre de risco estimadas próximas às observadas, seria necessário ajustar o fator de desconto subjetivo para algo maior que um, o que contraria as hipóteses do modelo teórico.

A tabela abaixo ilustra o comportamento médio dessas taxas, bem como o desvio-padrão de cada uma delas.

	Média (% a.a.)	Desvio-padrão (p.p.)
AIF	2,47	2,45
GMM	23,52	2,80
GEL	5,03	2,94
DI	11,44	5,15

Fonte: Elaboração própria

Tabela 6 – Médias e desvios-padrão das taxas estimadas e da DI

Apesar das divergências com a literatura, nota-se pela figura 5 abaixo, que as taxas estimadas foram positivamente correlacionadas entre si e com a DI, o que traz robustez às metodologias utilizadas nas estimações. Mesmo assim, pela figura 4 e pela tabela 6, nota-se que essas taxas diferem muito em nível entre si.

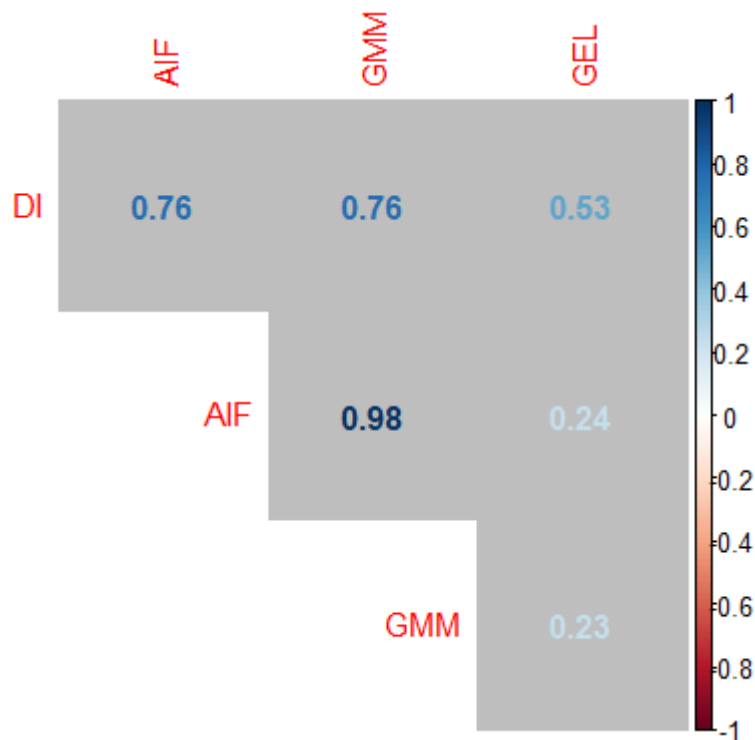


Figura 5 – Correlações entre as diferentes taxas

Fonte: Elaboração própria

Os resultados encontrados neste trabalho podem servir como referência para pesquisas futuras acerca da estimação da taxa livre de risco com base em um ambiente com especificações

diferentes sobre a economia brasileira. Tais modelos podem incluir, por exemplo, a presença de heterogeneidade na função utilidade, com a presença de mercados incompletos, e até mesmo diferentes funções utilidades que não a CRRA.

3 CONCLUSÃO

A partir de dados trimestrais de retornos de ações listadas na B3 e de retornos do próprio Ibovespa, de consumo em nível e de consumo em proporção do PIB, desde o primeiro trimestre de 2002 até o segundo trimestre de 2022, foram realizadas estimações da taxa livre de risco para o Brasil a partir do modelo CCAPM com função utilidade CRRA, utilizando-se de diferentes abordagens econométricas. As taxas livres de risco estimadas pelas três abordagens descritas, AIF, GMM e GEL, apresentaram divergências entre si e em relação à taxa DI, utilizada hoje como *proxy* para a taxa livre de risco brasileira, apesar de serem todas elas positivamente correlacionadas. O modelo GMM resultou em uma taxa livre de risco estimada maior que a taxa DI, na média, indicando a existência de um *risk-free rate puzzle* para a economia brasileira, conforme apontado por Weil (1989) para dados da economia americana. Por outro lado, os modelos AIF e GEL resultaram em taxas livres de risco estimadas menores que a taxa DI, também na média, apontando para a existência de um *risk-free rate puzzle* invertido, conforme Bonomo e Domingues (2002) apontam para dados brasileiros. As estimativas para os parâmetros β e σ nos modelos GMM e GEL indicaram pouca aderência dos dados brasileiros ao modelo teórico descrito pelo CCAPM, ao se comparar os resultados com outros importantes trabalhos na literatura brasileira de macro-finanças. A fim de aprimorar a metodologia desta dissertação, futuras pesquisas podem se utilizar de diferentes especificações para a função utilidade escolhida, e até mesmo incorporar heterogeneidade e mercados incompletos na análise. Pode-se fazer uso também de diferentes dados de consumo, como a separação entre consumo de bens duráveis, não-duráveis e consumo de serviços, ou ainda utilizar-se de diferentes dados de retornos, diversificando o portfólio e aumentando a abrangência dos dados.

REFERÊNCIAS

- ALMEIDA, Caio; GARCIA, René. Assessing misspecified asset pricing models with empirical likelihood estimators. *Journal of Econometrics*, Elsevier, v. 170, n. 2, p. 519–537, 2012.
- ANATOLYEV, Stanislav; GOSPODINOV, Nikolay. *Methods for estimation and inference in modern econometrics*. [S.l.]: CRC Press, 2011.
- ANDERSEN, Steffen et al. Eliciting risk and time preferences. *Econometrica*, Wiley Online Library, v. 76, n. 3, p. 583–618, 2008.
- ARAÚJO, Fabio; ISSLER, Joao Victor; FERNANDES, Marcelo. Estimating the stochastic discount factor without a utility function. Society for Computational Economics, 2005.
- ARROW, Kenneth J. The role of securities in the optimal allocation of risk-bearing. *The review of economic studies*, JSTOR, v. 31, n. 2, p. 91–96, 1964.
- ARROW, Kenneth J; DEBREU, Gerard. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 265–290, 1954.
- BANSAL, Ravi; YARON, Amir. Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, Wiley Online Library, v. 59, n. 4, p. 1481–1509, 2004.
- BINSBERGEN, Jules H Van; DIAMOND, William F; GROTTERIA, Marco. Risk-free interest rates. *Journal of Financial Economics*, Elsevier, v. 143, n. 1, p. 1–29, 2022.
- BOLLERSLEV, Tim; ENGLE, Robert F; WOOLDRIDGE, Jeffrey M. A capital asset pricing model with time-varying covariances. *Journal of political Economy*, The University of Chicago Press, v. 96, n. 1, p. 116–131, 1988.
- BONOMO, Marco; DOMINGUES, Gabriela Bertol. Os puzzles invertidos no mercado brasileiro de ativos. *Finanças aplicadas ao Brasil*, FGV Editora Rio de Janeiro, p. 105–120, 2002.
- BREEDEN, Douglas T. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of financial Economics*, Elsevier, v. 7, n. 3, p. 265–296, 1979.
- BROWN, David P; GIBBONS, Michael R. A simple econometric approach for utility-based asset pricing models. *The Journal of Finance*, Wiley Online Library, v. 40, n. 2, p. 359–381, 1985.
- BUDGETS, Capital; LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios. *The Review of Economics and Statistics*, v. 47, n. 1, p. 13–37, 1965.
- CAMPBELL, John Y. *Intertemporal asset pricing without consumption data*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1992.
- _____. Empirical asset pricing: Eugene fama, lars peter hansen, and robert shiller. *The Scandinavian Journal of Economics*, Wiley Online Library, v. 116, n. 3, p. 593–634, 2014.
- CAMPBELL, John Y; COCHRANE, John H. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, The University of Chicago Press, v. 107, n. 2, p. 205–251, 1999.
- COCHRANE, John. *Asset pricing: Revised edition*. [S.l.]: Princeton university press, 2009.

- COCHRANE, John H. Macro-finance. *Review of Finance*, Oxford University Press, v. 21, n. 3, p. 945–985, 2017.
- CYSNE, Rubens Penha. Equity-premium puzzle: evidence from brazilian data. *Economia Aplicada*, SciELO Brasil, v. 10, p. 161–180, 2006.
- DEBREU, Gerard. *Theory of value: An axiomatic analysis of economic equilibrium*. [S.l.]: Yale University Press, 1959. v. 17.
- DIMSON, Elroy; MARSH, Paul; STAUNTON, Mike. The worldwide equity premium: A smaller puzzle. *Chapter*, v. 11, p. 467–514, 2006.
- EPSTEIN, Larry G; ZIN, Stanley E. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, v. 57, n. 4, p. 937–969, 1989.
- FAMA, Eugene F. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, JSTOR, v. 25, n. 2, p. 383–417, 1970.
- FAMA, Eugene F; FRENCH, Kenneth R. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, Elsevier, v. 33, n. 1, p. 3–56, 1993.
- FAN, Jianqing; YAO, Qiwei. *Nonlinear time series: nonparametric and parametric methods*. [S.l.]: Springer, 2003. v. 20.
- FARNSWORTH, Heber et al. *Performance evaluation with stochastic discount factors*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 2002.
- FERREIRA, Alex; MATOS, Paulo. Precautionary risks for an open economy. *International Review of Economics & Finance*, Elsevier, v. 70, p. 154–167, 2020.
- GOMES, Fábio Augusto Reis; COSTA, Luciana de Andrade; PUPO, Ruth Carolina Rocha. The equity premium puzzle: analysis in brazil after the real plan. *BAR-Brazilian Administration Review*, SciELO Brasil, v. 10, p. 135–157, 2013.
- GOURIEROUX, Christian; MONFORT, Alain. Econometric specification of stochastic discount factor models. *Journal of Econometrics*, Elsevier, v. 136, n. 2, p. 509–530, 2007.
- HALL, Robert E. Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of political economy*, The University of Chicago Press, v. 86, n. 6, p. 971–987, 1978.
- HAMILTON, James D. Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 100, n. 5, p. 831–843, 2018.
- HANSEN, Lars Peter. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the econometric society*, JSTOR, p. 1029–1054, 1982.
- HANSEN, Lars Peter; HEATON, John; YARON, Amir. Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 14, n. 3, p. 262–280, 1996.

HANSEN, Lars Peter; JAGANNATHAN, Ravi. Implications of security market data for models of dynamic economies. *Journal of political economy*, The University of Chicago Press, v. 99, n. 2, p. 225–262, 1991.

HANSEN, Lars Peter; RICHARD, Scott F. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 587–613, 1987.

HANSEN, Lars Peter; SINGLETON, Kenneth J. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 1269–1286, 1982.

_____. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of political economy*, The University of Chicago Press, v. 91, n. 2, p. 249–265, 1983.

HARRISON, J Michael; KREPS, David M. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic theory*, Citeseer, v. 20, n. 3, p. 381–408, 1979.

HODRICK, Robert J; PRESCOTT, Edward C. Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, JSTOR, p. 1–16, 1997.

HOROWITZ, Joel L. *Semiparametric and nonparametric methods in econometrics*. [S.l.]: Springer, 2009. v. 12.

IMBENS, Guido W. One-step estimators for over-identified generalized method of moments models. *The Review of Economic Studies*, Wiley-Blackwell, v. 64, n. 3, p. 359–383, 1997.

ISSLER, João Victor; PIQUEIRA, Natália Scotto. Estimando a aversão ao risco, a taxa de desconto intertemporal, e a substitubilidade intertemporal do consumo no brasil usando três tipos de função utilidade (versão preliminar). Escola de Pós-Graduação em Economia da FGV, 2000.

ITO, Mikio; NODA, Akihiko. The gel estimates resolve the risk-free rate puzzle in japan. *Applied financial economics*, Taylor & Francis, v. 22, n. 5, p. 365–374, 2012.

JIANG, Zhengyang et al. *The us public debt valuation puzzle*. [S.l.], 2019.

KITAMURA, Yuichi; STUTZER, Michael. An information-theoretic alternative to generalized method of moments estimation. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 861–874, 1997.

KRISHNAMURTHY, Arvind; VISSING-JORGENSEN, Annette. The aggregate demand for treasury debt. *Journal of Political Economy*, University of Chicago Press Chicago, IL, v. 120, n. 2, p. 233–267, 2012.

LAURINI, Márcio Poletti; PEREIRA, Caio A Vigo. Portfolio efficiency tests with conditioning information using empirical likelihood estimation. 2016.

LETTAU, Martin; LUDVIGSON, Sydney. Resurrecting the (c) capm: A cross-sectional test when risk premia are time-varying. *Journal of political economy*, The University of Chicago Press, v. 109, n. 6, p. 1238–1287, 2001.

LUCAS, Robert E. Asset prices in an exchange economy. *Econometrica: journal of the Econometric Society*, JSTOR, p. 1429–1445, 1978.

- MEHRA, Rajnish et al. The equity premium puzzle: A review. *Foundations and Trends® in Finance*, Now Publishers, Inc., v. 2, n. 1, p. 1–81, 2007.
- MEHRA, Rajnish; PRESCOTT, Edward C. The equity premium: a puzzle. *Journal of monetary Economics*, Elsevier, v. 15, n. 2, p. 145–161, 1985.
- MÜLLER, Ulrich K; WATSON, Mark W. *Low-frequency econometrics*. [S.l.], 2015.
- NEWAY, Whitney K; SMITH, Richard J. Higher order properties of gmm and generalized empirical likelihood estimators. *Econometrica*, Wiley Online Library, v. 72, n. 1, p. 219–255, 2004.
- NEWAY, Whitney K; WEST, Kenneth D. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. National Bureau of Economic Research Cambridge, Mass., USA, 1986.
- NODA, Akihiko et al. Testing the "catching up with the joneses" model with consumption externality in japan". *Economics Bulletin*, Economics Bulletin, v. 31, n. 2, p. 1648–1658, 2011.
- OWEN, Art B. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, Oxford University Press, v. 75, n. 2, p. 237–249, 1988.
- PENNACCHI, George G. Identifying the dynamics of real interest rates and inflation: Evidence using survey data. *The Review of Financial Studies*, Oxford University Press, v. 4, n. 1, p. 53–86, 1991.
- ROLL, Richard. A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of financial economics*, Elsevier, v. 4, n. 2, p. 129–176, 1977.
- ROSENBERG, Joshua V; ENGLE, Robert F. Empirical pricing kernels. *Journal of Financial Economics*, Elsevier, v. 64, n. 3, p. 341–372, 2002.
- ROSS, Stephen A. A simple approach to the valuation of risky streams. *Journal of business*, JSTOR, p. 453–475, 1978.
- RUBINSTEIN, Mark. The valuation of uncertain income streams and the pricing of options. *The Bell Journal of Economics*, JSTOR, p. 407–425, 1976.
- SARGAN, John D. The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the econometric society*, JSTOR, p. 393–415, 1958.
- SARGAN, J Denis. The estimation of relationships with autocorrelated residuals by the use of instrumental variables. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 21, n. 1, p. 91–105, 1959.
- SHARPE, William F. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, Wiley Online Library, v. 19, n. 3, p. 425–442, 1964.
- SIMONASSI, Andrei Gomes. Estimando a taxa de retorno livre de risco no brasil. Encontro Nacional de Economia, 2006.
- SMITH, Richard J. Alternative semi-parametric likelihood approaches to generalised method of moments estimation. *The Economic Journal*, Oxford University Press Oxford, UK, v. 107, n. 441, p. 503–519, 1997.

_____. Gel criteria for moment condition models. *Econometric Theory*, Cambridge University Press, v. 27, n. 6, p. 1192–1235, 2011.

WEIL, Philippe. The equity premium puzzle and the risk-free rate puzzle. *Journal of monetary economics*, Elsevier, v. 24, n. 3, p. 401–421, 1989.

APÊNDICE A – LISTA DE AÇÕES

Tabela 7 – Ranking de liquidez das ações de acordo com volume médio de negociação

Ticker na B3	Empresa	Volume médio de negociação (em reais)
PETR4	Petrobras	33.366.758,02
VALE3	Vale	20.680.830,02
ITUB4	Itau Unibanco	15.986.332,30
BBDC4	Bradesco	13.370.310,66
BBAS3	Banco do Brasil	8.806.497,87
PETR3	Petrobras	8.568.331,60
ABEV3	Ambev	5.933.319,35
ITSA4	Itausa	5.663.351,66
GGBR4	Gerdau	5.539.758,96
USIM5	Usiminas	4.828.720,22
CSNA3	Sid Nacional	4.586.704,75
CMIG4	Cemig	3.034.713,17
ELET3	Eletrobras	2.858.765,58
ELET6	Eletrobras	2.044.273,17
EMBR4	Embraer	2.038.276,59
GOAU4	Gerdau Met	1.992.197,06
BRKM5	Braskem	1.938.969,31
BBDC3	Bradesco	1.807.842,55
OIBR3	Oi	1.309.793,24
OIBR4	Oi	865.402,92
LIGT3	Light	812.248,21
ALPA4	Alpargatas	655.666,87
RAPT4	Randon Part	494.534,63
VIVT3	Telef Brasil	491.492,40
MYPK3	Iochoy-Maxion	413.144,23
POMO4	Marcopolo	403.224,87
KLBN4	Klabin	354.955,13
USIM3	Usiminas	342.848,05
ITUB3	Itau Unibanco	321.151,57
CMIG3	Cemig	239.073,35
ETER3	Eternit	218.172,68
TASA4	Taurus Armas	216.759,30
UNIP6	Unipar	216.302,17
TUPY3	Tupy	212.798,21

Tabela 7 – Continuação da tabela 7

GUAR3	Guararapes	211.766,26
FESA4	Ferbasa	142.268,92
CLSC4	Celesc	119.933,73
GGBR3	Gerdau	115.664,33
ITSA3	Itausa	114.453,36
ROMI3	Romi	111.332,92
PNVL3	Dimed	71.525,71
INEP4	Inepar	43.509,80
CTNM4	Coteminas	36.766,66
RCSL4	Recrusul	35.412,38
BOBR4	Bombril	19.576,14
UNIP3	Unipar	17.981,65
TASA3	Taurus Armas	15.893,91
BLUT4	B Tech Eqi	10.410,32
BRKM3	Braskem	9.800,94
BAZA3	Amazonia	8.843,90
CGRA4	Grazziotin	8.793,98
WHRL4	Whrilpool	8.270,96
BDLL4	Bardella	7.967,97
CBEE3	Ampla Energ	7.722,46
ALPA3	Alpargatas	7.137,64
MGEL4	Mangels Indl	5.057,35
TEKA4	Teka	4.781,16
BMEB4	Merc Brasil	4.211,94
MOAR3	Mont Aranha	4.191,31
BRIV4	Alfa Invest	3.980,14
CTNM3	Coteminas	3.743,00
MTSA4	Metisa	3.204,15
CRIV4	Alfa Financ	3.039,26
PTNT4	Pettenati	2.920,03
CTKA4	Karsten	2.545,91
MWET4	Wetzel S/A	2.347,47
WLMM4	Wlm Ind Com	1.730,96
BRIV3	Alfa Invest	1.437,35
BGIP4	Banese	1.087,63
CRIV3	Alfa Financ	897,41
UNIP5	Unipar	863,28
RPAD5	Alfa Holding	510,20

Tabela 7 – Continuação da tabela 7

RPAD6	Alfa Holding	475,20
BNBR3	Nord Brasil	358,67
RPAD3	Alfa Holding	344,82
BRGE3	Alfa Consorc	292,90

APÊNDICE B – CÓDIGO EM R

```

# Rodando pacotes necessarios ----

library(readxl)
library(dplyr)
library(stringr)
library(tidyr)
library(ggplot2)
library(lubridate)
library(parsnip)
library(timetk)
library(momentfit)
library(tseries)
library(plotly)
library(corrplot)

# Importando dados ----

retornos_acoes <- read.csv("C:\\Users\\rafae\\Desktop\\Dados\\R
  ↪ \\VF_script\\retorno_acoes.csv", na.strings = "-", check.
  ↪ names = F) # , na.strings = "-"
rank_liq <- read_excel("C:\\Users\\rafae\\Desktop\\Dados\\R\\VF
  ↪ _script\\ranking_liquidez_b3.xlsx")
dados <- read_excel("C:\\Users\\rafae\\Desktop\\Dados\\R\\VF_
  ↪ script\\dados_vf.xlsx")
dados$date <- ymd(dados$date)

#*Ranking de liquidez ----

rank_liq <- rank_liq %>% dplyr::select(-1)

colnames(rank_liq) <- c('Empresa', 'Ticker_na_B3', 'Volume_
  ↪ medio_de_negociacao_(em_R$)')

rank_liq <- rank_liq %>% dplyr::slice(1:150)

#*Retornos ----

```

```

colnames(retornos_acoef) <- c('ativo', 'data', 'retornos')

strings_1_2 <- as.data.frame(str_sub(retornos_acoef$data, start
  ↪ = 1, end = 2))

strings_3_6 <- as.data.frame(str_sub(retornos_acoef$data, start
  ↪ = 3, end = 6))

colnames(strings_1_2) <- 'tri'

colnames(strings_3_6) <- 'ano'

date_col <- as.data.frame(paste(strings_3_6$ano, strings_1_2$
  ↪ tri))

retornos_acoef <- cbind(date_col, retornos_acoef)

colnames(retornos_acoef)[1] <- 'date'

retornos_acoef <- retornos_acoef %>% dplyr::select(-3)

retornos_acoef <- retornos_acoef %>%
  spread('ativo', 'retornos')

retornos_acoef <- retornos_acoef %>% dplyr::slice(66:147)

colnames(retornos_acoef) <- gsub(".XBSP.", "", colnames(retornos_
  ↪ acoes))

retornos_acoef <- retornos_acoef %>%
  select(matches(rank_liq$`Ticker na B3`))

retornos_acoef <- data.frame(dados[1], retornos_acoef)

retornos_acoef <- retornos_acoef[, colSums(is.na(retornos_acoef)
  ↪ ) == 0]

col_names <- as.data.frame(c(colnames(retornos_acoef)))

```

```

col_names <- col_names %>% dplyr::slice(-1)
colnames(col_names) <- 'Ticker_na_B3'

#*Ranking de liquidez filtrado ----

rank_liq_filtrado <- left_join(col_names, rank_liq, by = "
  ↪ Ticker_na_B3")
write.csv(rank_liq_filtrado, "C:\\Users\\rafae\\Desktop\\Dados
  ↪ \\R\\VF_script\\rank_liq_filtrado_new.csv", row.names =
  ↪ FALSE)

# O metodo de Low-Frequency Econometrics para suavizacao da
  ↪ serie de consumo ----

mean_conspib <- mean(dados$consumo_pib)
T <- length(dados$consumo_pib)
q <- 12

s <- seq((1-(1/2))/T, (T-(1/2))/T, length.out=T)
Psi_T <- sapply(1:q, function(j) sqrt(2)*cos(j*s*pi))

X_T <- cbind(rep(1, T), Psi_T)
X_T <- cbind(X_T, rep(mean_conspib, T))

fit1 <- lm(dados$consumo_pib ~ X_T - 1)

summary(fit1)

dados$consumo_pib_smooth <- fit1$fitted.values

dados %>%
  dplyr::select(1,5,8) %>%
  pivot_longer(cols = c(2,3), names_to = "Legenda", values_to =
    ↪ "Valor") %>%
  ggplot(aes(date, Valor)) +
  geom_line(aes(color = Legenda, linetype = Legenda)) +

```

```

scale_color_manual(values = c("black", "black")) + ggplot2::
  ↪ theme_bw() +
scale_x_date(date_breaks = "5_years")

# Modelo AIF ----

#*Calculo das medias geometricas e aritmeticas dos retornos
  ↪ segundo AIF ----

ipca_t <- 1 + 0.01 * dados$ipca

retorno_bruto_acoes <- 1 + 0.01 * retornos_acoes[-1]
retorno_bruto_acoes <- retorno_bruto_acoes/ipca_t
retorno_bruto_acoes <- cbind(retornos_acoes[1], retorno_bruto_
  ↪ acoes)

exp_geom_mean = -1/76
geom_mean_base <- retorno_bruto_acoes[-1]^exp_geom_mean
geom_mean <- matrix(geom_mean_base$PETR4, nrow(geom_mean_base),
  ↪ 1)

for (j in 2:ncol(geom_mean_base)) {
  geom_mean = geom_mean*geom_mean_base[j]
}

colnames(geom_mean) <- 'geom_mean'

mult_arit_mean = 1/76
arit_mean_base <- mult_arit_mean*retorno_bruto_acoes[-1]
arit_mean <- matrix(arit_mean_base$PETR4, nrow(arit_mean_base),
  ↪ 1)

for (j in 2:ncol(arit_mean_base)) {
  arit_mean = arit_mean+arit_mean_base[j]
}

colnames(arit_mean) <- 'arit_mean'

```

```

aif_base <- cbind(retorno_bruto_acoes[1], arit_mean, geom_mean)

##Estimador AIF para o SDF ----

mult_rg_ra <- aif_base[2] * aif_base[3]

colnames(mult_rg_ra) <- 'mult_rg_ra'

sum_mult_rg_ra <- sum(mult_rg_ra$mult_rg_ra)

aif_base$sdf <- aif_base$geom_mean / (sum_mult_rg_ra / nrow(aif
  ↪ _base))

summary(aif_base$sdf)

##Modelo OLS para SDF de AIF ----

# passando para retornos reais
top5_retornos_t <- as.data.frame(rowMeans(retornos_acoes[2:6]))
top5_retornos_t <- 1 + 0.01*top5_retornos_t
top5_retornos_t <- top5_retornos_t/ipca_t

top10_retornos_t <- as.data.frame(rowMeans(retornos_acoes
  ↪ [2:11]))
top10_retornos_t <- 1 + 0.01*top10_retornos_t
top10_retornos_t <- top10_retornos_t/ipca_t

top15_retornos_t <- as.data.frame(rowMeans(retornos_acoes
  ↪ [2:16]))
top15_retornos_t <- 1 + 0.01*top15_retornos_t
top15_retornos_t <- top15_retornos_t/ipca_t

ibov_t <- 1 + 0.01*dados[7]
ibov_t <- ibov_t/ipca_t

sdf_aif_t_mais_1 <- lead(aif_base$sdf)
# ao fazer isso, perde-se a ultima observacao - SDF comea no 1
  ↪ T2002,
# mas termina agora no 1T2022

```

```

df_aif <- cbind(dados[1], sdf_aif_t_mais_1,
               top5_retornos_t, top10_retornos_t, top15_retornos_t,
               ibov_t, dados[8])

colnames(df_aif) <- c('date', 'sdf_t_mais_1',
                    'ret_5_t', 'ret_10_t', 'ret_15_t',
                    'ibov_t', 'cons_pib_smooth_t')

df_aif <- df_aif %>% na.omit()


$$\text{***OLS 1: } X_t = \text{ret}_{5_t} \text{ ----}$$


lm1_aif <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_5_t+log(cons_pib_smooth_t)-1, data = df_aif)
  ↪ aif)

summary(lm1_aif$fit)


$$\text{***OLS 2: } X_t = \text{ret}_{10_t} \text{ ----}$$


lm2_aif <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_10_t+log(cons_pib_smooth_t)-1, data = df_aif)
  ↪ _aif)

summary(lm2_aif$fit)


$$\text{***OLS 3: } X_t = \text{ret}_{15_t} \text{ ----}$$


lm3_aif <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_15_t+log(cons_pib_smooth_t)-1, data = df_aif)
  ↪ _aif)

summary(lm3_aif$fit)


$$\text{***OLS 4: } X_t = \text{ibov}_t \text{ ----}$$


lm4_aif <- linear_reg() %>% set_engine('lm') %>%

```

```

fit(sdf_t_mais_1~ibov_t+log(cons_pib_smooth_t)-1, data = df_
  ↪ aif)

summary(lm4_aif$fit)

# Calculo de E(m_t+1) nominal

e_ipca_t_mais_1 <- dados[4] %>% dplyr::slice(-82)

e_ipca_t_mais_1 <- 1 + 0.01*e_ipca_t_mais_1

sdf_fitted_aif <- as.data.frame(lm4_aif$fit$fitted.values)

rf_bruta_aif <- 1/(sdf_fitted_aif/e_ipca_t_mais_1)

rf_liquida_aif <- as.data.frame(100*(rf_bruta_aif - 1))

di_t_mais_1 <- as.data.frame(dados[2] %>% dplyr::slice(-82))

df_aif$rf_aif_t_mais_1 <- rf_liquida_aif$`lm4_aif$fit$fitted.
  ↪ values`
df_aif$di_t_mais_1 <- di_t_mais_1$taxa_di

df_aif %>% dplyr::select(1,8,9) %>%
  pivot_longer(cols = 2:3, names_to = "Legenda") %>%
  plot_time_series(.date_var = date, .value = value, .color_var
  ↪ = Legenda, .smooth = F)

cor(df_aif$di_t_mais_1,df_aif$rf_aif_t_mais_1)

# Modelo GMM ----

ibov_t_mais_1 <- lead(ibov_t)

top5_retornos_t_mais_1 <- lead(top5_retornos_t)

```



```

top10_retornos_t_mais_1 <- lead(top10_retornos_t)

top15_retornos_t_mais_1 <- lead(top15_retornos_t)

cons_growth_t <- dados[6]/lag(dados[6])
cons_growth_t_mais_1 <- lead(cons_growth_t)

ibov_t_menos_1 <- lag(ibov_t)

top5_retornos_t_menos_1 <- lag(top5_retornos_t)

top10_retornos_t_menos_1 <- lag(top10_retornos_t)

top15_retornos_t_menos_1 <- lag(top15_retornos_t)

cons_growth_t_menos_1 <- lag(cons_growth_t)

# df com NLAG==2
df_gmm <- cbind(ibov_t_mais_1,top5_retornos_t_mais_1,top10_
  ↪ retornos_t_mais_1,top15_retornos_t_mais_1,
               cons_growth_t_mais_1,
               ibov_t,top5_retornos_t,top5_retornos_t,top5_
                 ↪ retornos_t,
               cons_growth_t,
               ibov_t_menos_1,top5_retornos_t_menos_1,top10_
                 ↪ retornos_t_menos_1,top15_retornos_t_menos_1,
               cons_growth_t_menos_1)

colnames(df_gmm) <- c('ibov_t_mais_1','top5_t_mais_1','top10_t_
  ↪ mais_1','top15_t_mais_1',
                    'cons_growth_t_mais_1',
                    'ibov_t','top5_t','top10_t','top15_t',
                    'c_growth_t',
                    'ibov_t_menos_1','top5_t_menos_1','top10_t_
                      ↪ menos_1','top15_t_menos_1',
                    'c_growth_t_menos_1')

```

```

df_gmm <- df_gmm %>% na.omit()
# df_gmm começa em 3T2002 e acaba em 1T2022

adf.test(df_gmm$ibov_t_mais_1)
adf.test(df_gmm$top5_t_mais_1)
adf.test(df_gmm$top10_t_mais_1)
adf.test(df_gmm$top15_t_mais_1)
adf.test(df_gmm$cons_growth_t_mais_1)
adf.test(df_gmm$ibov_t)
adf.test(df_gmm$top5_t)
adf.test(df_gmm$top10_t)
adf.test(df_gmm$top15_t)
adf.test(df_gmm$c_growth_t)
adf.test(df_gmm$ibov_t_menos_1)
adf.test(df_gmm$top5_t_menos_1)
adf.test(df_gmm$top10_t_menos_1)
adf.test(df_gmm$top15_t_menos_1)
adf.test(df_gmm$c_growth_t_menos_1)

#*Funcao de momentos ----

crra_moments <- function(parm,x=NULL) {

  sdf_gmm = parm[1] * x[,5]^(-parm[2])

  euler1 = sdf_gmm*x[,1] - 1
  euler2 = sdf_gmm*x[,2] - 1
  euler3 = sdf_gmm*x[,3] - 1
  euler4 = sdf_gmm*x[,4] - 1

  euler5 = euler1*x[,6]
  euler6 = euler2*x[,6]
  euler7 = euler3*x[,6]
  euler8 = euler4*x[,6]

  euler9 = euler1*x[,7]
  euler10 = euler2*x[,7]
  euler11 = euler3*x[,7]

```

euler12 = euler4*x[, 7]

euler13 = euler1*x[, 8]

euler14 = euler2*x[, 8]

euler15 = euler3*x[, 8]

euler16 = euler4*x[, 8]

euler17 = euler1*x[, 9]

euler18 = euler2*x[, 9]

euler19 = euler3*x[, 9]

euler20 = euler4*x[, 9]

euler21 = euler1*x[, 10]

euler22 = euler2*x[, 10]

euler23 = euler3*x[, 10]

euler24 = euler4*x[, 10]

euler25 = euler1*x[, 11]

euler26 = euler2*x[, 11]

euler27 = euler3*x[, 11]

euler28 = euler4*x[, 11]

euler29 = euler1*x[, 12]

euler30 = euler2*x[, 12]

euler31 = euler3*x[, 12]

euler32 = euler4*x[, 12]

euler33 = euler1*x[, 13]

euler34 = euler2*x[, 13]

euler35 = euler3*x[, 13]

euler36 = euler4*x[, 13]

euler37 = euler1*x[, 14]

euler38 = euler2*x[, 14]

euler39 = euler3*x[, 14]

euler40 = euler4*x[, 14]

euler41 = euler1*x[, 15]

euler42 = euler2*x[, 15]

```

euler43 = euler3*x[,15]
euler44 = euler4*x[,15]

# z_t = (1, c_t, c_t_menos_1)

moments = cbind(euler1,
                euler21,
                euler41) # NPORTFOLIOS = 1 - melhor resultado

return(moments)
}

theta <- c(0.8,3)
names(theta) <- c('beta','gamma')

x <- as.matrix(df_gmm)

# vcov = c("iid", "HAC", "MDS", "CL")
moment_model <- momentModel(g = crra_moments,
                            x = x,
                            theta0 = theta,
                            vcov = "HAC")

# GMM fit
# type=c("twostep", "iter", "cue", "onestep")
crra_gmm_fit <- gmmFit(moment_model, type = "twostep")
summary(crra_gmm_fit)

crra_gmm_fit@theta

sdf_gmm <- 0.8592718 * df_gmm[5]^(-1.3525680)
colnames(sdf_gmm) <- 'sdf'
summary(sdf_gmm)

df_gmm2 <- df_aif %>% dplyr::select(1,3:7,9) %>% dplyr::slice
  ↪ (-1:-2)

df_gmm2$sdf_t_mais_1 <- sdf_gmm$sdf

```

```

##Modelo OLS para SDF de GMM ----

***OLS 1: X_t = ret_5_t ----

lm1_gmm <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_5_t+log(cons_pib_smooth_t)-1, data = df_
    ↪ gmm2)

summary(lm1_gmm$fit)

***OLS 2: X_t = ret_10_t ----

lm2_gmm <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_10_t+log(cons_pib_smooth_t)-1, data = df
    ↪ _gmm2)

summary(lm2_gmm$fit)

***OLS 3: X_t = ret_15_t ----

lm3_gmm <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_15_t+log(cons_pib_smooth_t)-1, data = df
    ↪ _gmm2)

summary(lm3_gmm$fit)

***OLS 4: X_t = ibov_t ----

lm4_gmm <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ibov_t+log(cons_pib_smooth_t)-1, data = df_
    ↪ gmm2)

summary(lm4_gmm$fit)

# Calculo de E(m_t+1) nominal

e_ipca_t_mais_1 <- dados[4] %>% dplyr::slice(-1,-2,-82)

```

```

e_ipca_t_mais_1 <- 1 + 0.01*e_ipca_t_mais_1

sdf_fitted_gmm <- as.data.frame(lm3_gmm$fit$fitted.values)

rf_bruta_gmm <- 1/(sdf_fitted_gmm/e_ipca_t_mais_1)

rf_liquida_gmm <- as.data.frame(100*(rf_bruta_gmm - 1))

df_gmm2$rf_gmm_t_mais_1 <- rf_liquida_gmm$`lm3_gmm$fit$fitted.
  ↳ values`

df_gmm2 %>% dplyr::select(1,7,9) %>%
  pivot_longer(cols = 2:3, names_to = "Legenda") %>%
  plot_time_series(.date_var = date, .value = value, .color_var
  ↳ = Legenda, .smooth = F)

cor(df_gmm2$di_t_mais_1,df_gmm2$rf_gmm_t_mais_1)

# Modelo GEL ----

df_gel <- df_gmm

crra_moments_gel <- function(parm,x=NULL) {

  sdf_gel = parm[1] * x[,5]^(-parm[2])

  euler1 = sdf_gel*x[,1] - 1
  euler2 = sdf_gel*x[,2] - 1
  euler3 = sdf_gel*x[,3] - 1
  euler4 = sdf_gel*x[,4] - 1

  euler5 = euler1*x[,6]
  euler6 = euler2*x[,6]
  euler7 = euler3*x[,6]
  euler8 = euler4*x[,6]

  euler9 = euler1*x[,7]
  euler10 = euler2*x[,7]

```

```
euler11 = euler3*x[, 7]
euler12 = euler4*x[, 7]
```

```
euler13 = euler1*x[, 8]
euler14 = euler2*x[, 8]
euler15 = euler3*x[, 8]
euler16 = euler4*x[, 8]
```

```
euler17 = euler1*x[, 9]
euler18 = euler2*x[, 9]
euler19 = euler3*x[, 9]
euler20 = euler4*x[, 9]
```

```
euler21 = euler1*x[, 10]
euler22 = euler2*x[, 10]
euler23 = euler3*x[, 10]
euler24 = euler4*x[, 10]
```

```
euler25 = euler1*x[, 11]
euler26 = euler2*x[, 11]
euler27 = euler3*x[, 11]
euler28 = euler4*x[, 11]
```

```
euler29 = euler1*x[, 12]
euler30 = euler2*x[, 12]
euler31 = euler3*x[, 12]
euler32 = euler4*x[, 12]
```

```
euler33 = euler1*x[, 13]
euler34 = euler2*x[, 13]
euler35 = euler3*x[, 13]
euler36 = euler4*x[, 13]
```

```
euler37 = euler1*x[, 14]
euler38 = euler2*x[, 14]
euler39 = euler3*x[, 14]
euler40 = euler4*x[, 14]
```

```
euler41 = euler1*x[, 15]
```

```

euler42 = euler2*x[,15]
euler43 = euler3*x[,15]
euler44 = euler4*x[,15]

# z_t = (1,r_t,c_t)

moments = cbind(euler4,
                euler20,
                euler24) # NPORTFOLIOS = 1 - melhor resultado

return(moments)
}

xx <- as.matrix(df_gel)

theta1 <- c(0.8,3)
names(theta1) <- c('beta','gamma')
# vcov = c("iid", "HAC", "MDS", "CL")
moment_model_gel <- momentModel(g = crra_moments_gel,
                                x = xx,
                                theta0 = theta1,
                                vcov = "HAC")

# gelType <- c("EL", "ET", "EEL", "HD", "ETEL" or "ETHD")
crra_gel_fit <- gelFit(moment_model_gel,
                       gelType = "EL")

crra_gel_fit@theta

summary(crra_gel_fit)

sdf_gel <- 0.9869407 * df_gel[5]^(2.1871772)
colnames(sdf_gel) <- 'sdf'
summary(sdf_gel)

#*Modelo OLS para SDF de GMM ----

```



```

df_gel2 <- df_aif %>% dplyr::select(1,3:7,9) %>% dplyr::slice
  ↪ (-1:-2)

df_gel2$sdf_t_mais_1 <- sdf_gel$sdf

##Modelo OLS para SDF de GMM ----

###OLS 1: X_t = ret_5_t ----

lm1_gel <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_5_t+log(cons_pib_smooth_t)-1, data = df_
  ↪ gel2)

summary(lm1_gel$fit)

###OLS 2: X_t = ret_10_t ----

lm2_gel <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_10_t+log(cons_pib_smooth_t)-1, data = df
  ↪ _gel2)

summary(lm2_gel$fit)

###OLS 3: X_t = ret_15_t ----

lm3_gel <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ret_15_t+log(cons_pib_smooth_t)-1, data = df
  ↪ _gel2)

summary(lm3_gel$fit)

###OLS 4: X_t = ibov_t ----

lm4_gel <- linear_reg() %>% set_engine('lm') %>%
  fit(sdf_t_mais_1~ibov_t+log(cons_pib_smooth_t)-1, data = df_
  ↪ gel2)

summary(lm4_gel$fit)

```

```

# Calculo de E(m_t+1) nominal

e_ipca_t_mais_1 <- dados[4] %>% dplyr::slice(-1,-2,-82)

e_ipca_t_mais_1 <- 1 + 0.01*e_ipca_t_mais_1

sdf_fitted_gel <- as.data.frame(lm2_gel$fit$fitted.values)

rf_bruta_gel <- 1/(sdf_fitted_gel/e_ipca_t_mais_1)

rf_liquida_gel <- as.data.frame(100*(rf_bruta_gel - 1))

df_gel2$rf_gel_t_mais_1 <- rf_liquida_gel$`lm2_gel$fit$fitted.
  ↪ values`

df_gel2 %>% dplyr::select(1,7,9) %>%
  pivot_longer(cols = 2:3, names_to = "Legenda") %>%
  plot_time_series(.date_var = date, .value = value, .color_var
  ↪ = Legenda, .smooth = F)

cor(df_gel2$di_t_mais_1,df_gel2$rf_gel_t_mais_1)

# Comparacao resultados ----

df_agregado <- cbind(df_gmm2[1],df_gmm2[7],
  df_aif[8] %>% dplyr::slice(-1,-2),
  df_gmm2[9],df_gel2[9])

mean(df_agregado$di_t_mais_1)
sd(df_agregado$di_t_mais_1)

mean(df_agregado$rf_aif_t_mais_1)
sd(df_agregado$rf_aif_t_mais_1)

mean(df_agregado$rf_gmm_t_mais_1)
sd(df_agregado$rf_gmm_t_mais_1)

```

```

mean(df_agregado$rf_gel_t_mais_1)
sd(df_agregado$rf_gel_t_mais_1)

colnames(df_agregado) <- c('Data', 'Taxa_DI', 'Taxa_Estimada_',
  ↪ AIF',
                        'Taxa_Estimada_-_GMM', 'Taxa_Estimada_-_GEL'
  ↪ ')

g_comparativo <- df_agregado %>%
  pivot_longer(cols = c(2:5), names_to = "Legenda", values_to =
  ↪ "Valor") %>%
  ggplot(aes(Data, Valor)) +
  geom_line(aes(color = Legenda, linetype = Legenda)) +
  scale_color_manual(values = c("black", "black", "grey", "red"))
  ↪ + ggplot2::theme_bw() +
  scale_x_date(date_breaks = "5_years")

ggplotly(g_comparativo)

colnames(df_agregado) <- c('Data', 'DI', 'AIF',
  ↪ 'GMM', 'GEL')
df_agregado %>% select(-Data) %>%
  cor() %>%
  corrplot(method = c('number'),
           type = c('upper'),
           diag = F,
           bg = "grey")

# Estatsticas descritivas ----

#* Grafico do consumo/PIB ----

colnames(dados)[c(1, 5, 8)] <- c('Data', 'C/PIB', 'C/PIB_suavizado'
  ↪ )
dados %>% dplyr::select(1, 5, 8) %>%

```

```

pivot_longer(cols = c(2,3), names_to = "Legenda", values_to =
  ↪ "Valor") %>%
ggplot(aes(Data,Valor)) +
geom_line(aes(color = Legenda, linetype = Legenda)) +
scale_color_manual(values = c("black", "black")) + ggplot2::
  ↪ theme_bw() +
scale_x_date(date_breaks = "5_years")

#*Grafico do crescimento do consumo ----

colnames(dados)[1] <- 'Data'
dados %>% dplyr::select(1,6) %>%
  dplyr::mutate(`Crescimento do consumo` = consumo_nivel/lag(
    ↪ consumo_nivel)) %>% na.omit() %>%
ggplot(aes(Data,`Crescimento do consumo`)) +
geom_line() +
scale_color_manual(values = c("black", "black")) + ggplot2::
  ↪ theme_bw() +
scale_x_date(date_breaks = "5_years")

#*Grafico da DI, IPCA e E(IPCA) ----

colnames(dados)[1:4] <- c('Data','Taxa_DI','IPCA','Expectativa_
  ↪ de_Inflacao')
dados %>% dplyr::select(1:4) %>%
  pivot_longer(cols = c(2,3,4), names_to = "Legenda", values_to
    ↪ = "Valor") %>%
ggplot(aes(Data,Valor)) +
geom_line(aes(color = Legenda, linetype = Legenda)) +
scale_color_manual(values = c("black", "black", "grey")) +
  ↪ ggplot2::theme_bw() +
scale_x_date(date_breaks = "5_years")

retornos_acoas[2:ncol(retornos_acoas)] <- 1 + 0.01*retornos_
  ↪ acoes[2:ncol(retornos_acoas)]
retornos_acoas[2:ncol(retornos_acoas)] <- retornos_acoas[2:ncol
  ↪ (retornos_acoas)]/ipca_t
retornos_acoas[2:ncol(retornos_acoas)] <- 100*(retornos_acoas
  ↪ [2:ncol(retornos_acoas)]-1)

```

```
##Summary das quinze acoes mais bem posicionadas ----
summary(retornos_acoes[2:16])
sd(retornos_acoes$PETR4)
sd(retornos_acoes$VALE3)
sd(retornos_acoes$ITUB4)
sd(retornos_acoes$BBDC4)
sd(retornos_acoes$BBAS3)
sd(retornos_acoes$PETR3)
sd(retornos_acoes$ABEV3)
sd(retornos_acoes$ITSA4)
sd(retornos_acoes$GGBR4)
sd(retornos_acoes$USIM5)
sd(retornos_acoes$CSNA3)
sd(retornos_acoes$CMIG4)
sd(retornos_acoes$ELET3)
sd(retornos_acoes$ELET6)
sd(retornos_acoes$EMBR3)

##Summary da media das TOP5 acoes ----
summary(rowMeans(retornos_acoes[2:6]))
sd(rowMeans(retornos_acoes[2:6]))

##Summary da media das TOP10 acoes ----
summary(rowMeans(retornos_acoes[2:11]))
sd(rowMeans(retornos_acoes[2:11]))

##Summary da media das TOP15 acoes ----
summary(rowMeans(retornos_acoes[2:16]))
sd(rowMeans(retornos_acoes[2:16]))

##Summary do IBOV ----

summary(100*(ibov_t$ibov - 1))
sd(100*(ibov_t$ibov - 1))
```