

# **Estimação de Equações Diferenciais Estocásticas Usando Verossimilhança Empírica e Mínimo Contraste Generalizado**

Márcio Poletti Laurini

Luiz Koodi Hotta

**Insper Working Paper**

WPE: 173/2009



Copyright Insper. Todos os direitos reservados.

É proibida a reprodução parcial ou integral do conteúdo deste documento por qualquer meio de distribuição, digital ou impresso, sem a expressa autorização do Insper ou de seu autor.

A reprodução para fins didáticos é permitida observando-se a citação completa do documento

# ESTIMAÇÃO DE EQUAÇÕES DIFERENCIAIS ESTOCÁSTICAS USANDO VEROSSIMILHANÇA EMPÍRICA E MÍNIMO CONTRASTE GENERALIZADO

MÁRCIO POLETTI LAURINI  
INSTITUTO INSPER E IMECC-UNICAMP  
LUIZ KOODI HOTTA  
IMECC-UNICAMP

RESUMO. Neste artigo abordamos a estimação semi-paramétrica de Equações Diferenciais Estocásticas utilizando métodos de Verossimilhança Empírica Generalizada e Mínimo Contraste Generalizado. Os resultados obtidos mostram que os estimadores propostos, em especial o estimador *Exponentially Tilted Empirical Likelihood* ([Schennach, 2007]), obtêm resultados superiores aos estimadores de Métodos de Momentos Generalizados normalmente utilizados na estimação de equações diferenciais estocásticas. Estes resultados são derivados das propriedades de robustez deste método na presença de problemas de especificação incorreta, o que no contexto da estimação de equações diferenciais estocásticas ocorre pelo uso de discretizações aproximadas do processo na construção das condições de momentos. As análises são realizadas por meio de experimentos de Monte Carlo e de uma aplicação empírica estimando diversos de modelos de taxas de juros de curto prazo para uma série de *Treasury Bills* com maturidade de um mês.

Palavras Chave: Equações Diferenciais Estocásticas, Verossimilhança Empírica, Mínimo Contraste Generalizado.  
JEL Codes: C14, C22.

## 1. INTRODUÇÃO

O uso de processos estocásticos em tempo contínuo na modelagem e precificação de instrumentos financeiros é uma das bases da moderna teoria de finanças, e sua origem pode ser traçada ao trabalho seminal de [Bachelier, 1900]. O uso de processos estocásticos em tempo contínuo é justificado pela conveniência matemática em relação ao uso de processos em tempo discreto e a possibilidade do uso da teoria matemática desenvolvida para a classe geral de processos conhecidos como semi-martingales contínuos, permitindo uma aplicação de toda a teoria de precificação por não-arbitragem ([Harrison & Kreps, 1979], [Harrison & Pliska, 1981] e [Delbaen & Schachermayer, 1994]) nesse contexto.

Os objetos básicos da modelagem de processos estocásticos em tempo contínuo são as chamadas Equações Diferenciais Estocásticas (*Stochastic Differential Equations*), que são objetos representados na forma geral:

$$(1.1) \quad dX_t = \mu(t, X_t) + \sigma(t, X_t)dW_t$$

onde  $\mu(t, X_t)$  representa a parte determinística do processo (*drift* instantâneo),  $\sigma(t, X_t)$  representa o componente estocástico (volatilidade) do processo, e  $W_t$  é o chamado processo de Wiener ou Movimento Browniano. Esta representação é útil já que permite definir a evolução das trajetórias processo  $X_t$  através de uma representação dada por uma integração estocástica (e.g [Rogers & Williams, 2000], [Karatzas & Shreve, 1987], [Kloeden & Platen, 1992]):

$$(1.2) \quad X_t = X_{t_0} + \int_{t_0}^t \mu(t, X_t)dt + \int_{t_0}^t \sigma(t, X_t)dW_t.$$

Diferentes especificações dos processos de drift  $\mu(t, X_t)$  e volatilidade  $\sigma(t, X_t)$  na equação diferencial estocástica dão origem a processos com propriedades distintas. Estas propriedades permitem representar uma ampla classe de processos utilizados em finanças. Focando na modelagem de taxas de juros de curto prazo, uma série de especificações alternativas para a modelagem

Merton (1973)	$dr = \alpha dt + \sigma dW$
Vasicek (1977)	$dr = (\alpha + \beta r)dt + \sigma dW$
CIR SR (1985)	$dr = (\alpha + \beta r)dt + \sigma r^{1/2}dW$
Dothan (1978)	$dr = \sigma r dW$
GBM (1973)	$dr = \beta r dt + \sigma r dW$
Brennan-Schwartz (1980)	$dr = (\alpha + \beta r)dt + \sigma r dW$
CIR VR (1980)	$dr = \sigma r^{3/2}dW$
CEV (1975)	$dr = \beta dt + \sigma dW$
CIR Generalizado	$dr = (\alpha + \beta r)dt + \sigma r^\gamma dW$

TABELA 1. Modelos para taxas de juros de curto prazo.

de taxa de juros de curto prazo tem sido utilizadas. A Tabela 1 mostra algumas formulações empregadas na literatura, compreendendo os modelos de [Merton, 1973], [Vasicek, 1977], [Cox *et al.*, 1985], [Dothan, 1978], [Black & Scholes, 1973], [Brennan & Schwartz, 1980], [Cox *et al.*, 1980] e [Cox, 1975]. Note que na última linha está definido o modelo chamado de CIR generalizado, que contém todos os modelos anteriores como casos particulares, como mostrado em [Chan *et al.*, 1992b], que contém uma discussão geral sobre propriedades destes modelos.

A estimação de parâmetros em equações diferenciais estocásticas é um tema bastante desenvolvido na literatura econômica, e existe uma amplitude enorme de técnicas disponíveis<sup>1</sup>. Esta gama de técnicas está relacionada as dificuldades inerentes a estimação de equações diferenciais estocásticas, em especial a não existência de soluções fechadas para equações diferenciais estocásticas em casos gerais e o problema do uso de dados observados discretamente na estimação de um processo em formulado em tempo contínuo. Como exemplos de métodos de estimação neste contexto temos métodos de Máxima Verossimilhança, Métodos de Momentos Generalizados, Método de Momentos Simulados, Martingale Estimating Equations, Markov Chain Monte Carlo e Inferência Indireta e Métodos Não-Paramétricos.

Em princípio a forma mais recomendada de estimação seria a utilização da função de verossimilhança, já que na validade das condições de regularidade estimadores de máxima verossimilhança são consistentes, eficientes e assintoticamente normais. Entretanto no contexto da estimação de equações diferenciais estocásticas a não existência de soluções gerais é uma dificuldade geral para o uso de métodos baseados na verossimilhança do processo, já que a função de verossimilhança é formulada utilizando a densidade de transição resultante da solução da equação diferencial estocástica.

Na ausência de soluções analíticas é necessário o uso de aproximações na construção da função de verossimilhança, como o uso de métodos de quasi-máxima verossimilhança, que gera estimadores com mínimo erro quadrático médio ou uso de máxima verossimilhança simulada que utiliza trajetórias simuladas por discretizações de Euler ou Milstein na avaliação da verossimilhança ([Pedersen, 1995]), ou então as aproximações usando expansões de Hermite obtidas por [Ait-Sahalia, 2002]. Note que dado o uso de aproximações na avaliação da função de verossimilhança as propriedades de otimalidade deste estimador podem não se manter e assim outros estimadores menos eficientes se tornam competitivos.

Estimadores utilizando condições de momentos também são freqüentemente utilizados na estimação de equações diferenciais estocásticas. A estimação utilizando o Método de Momentos Generalizados (*Generalized Method of Moments - GMM*) de [Hansen, 1982], utilizando uma discretização simples do processo talvez seja a forma mais amplamente utilizada (e.g. [Chan *et al.*, 1992a]). Embora o método de momentos generalizados seja caracterizado por propriedades de consistência e eficiência assintótica, suas propriedades em amostras finitas e na presença de problemas de especificação podem não ser ótimas. Para tratar destes problemas discutimos o uso de duas classes de estimadores na estimação de equações diferenciais estocásticas usando dados discretos - estimadores de Verossimilhança Empírica Generalizada e estimadores de Mínimo Contraste Generalizado, comparando seu desempenho com o de estimadores baseados na estimação pelo Método de Momentos Generalizados. Estes estimadores são semi-paramétricos no sentido que a forma paramétrica da equação diferencial estocástica é utilizada através de condições de momentos, mas a densidade não-observada do processo é avaliada de forma não-paramétrica.

<sup>1</sup>Veja [Gourieroux & Monfort, 1996], [Prakasa Rao, 1999], [Bishwal, 2007] e [Zivot & Wang, 2006] para uma revisão da literatura de estimação de equações diferenciais estocásticas

Os estimadores apresentados (*Generalized Empirical Likelihood*, *Continuos Updating Empirical Likelihood*, *Exponential Tilting* e *Exponentially Tilted Empirical Likelihood*) possuem as mesmas propriedades de consistência e eficiência assintótica de primeira ordem (e.g. [Smith, 2001], [Schennach, 2007]) dos estimadores GMM comparados (GMM em dois estágios, GMM Iterativo, GMM *Continuos Updating*). Entretanto resultados teóricos mostram que estes estimadores podem ter propriedades superiores em termos de viés em amostras finitas, propriedades assintóticas de ordem superior, são assintoticamente eficientes na classe de estimadores semi-paramétricos (no sentido de [Bickel *et al.*, 1993]), e possuem propriedades ótimas em termos de testes de hipóteses (otimalidade minimax, otimalidade no sentido de *large deviations* e são testes uniformemente mais poderosos no sentido generalizado de Neyman-Person) e a classe de estimadores de Mínimo Contraste Generalizado (*Exponential Tilting* e *Exponentially Tilted Empirical Likelihood*) possuem características de robustez na presença de problemas de especificação. Estas características de robustez dos estimadores baseados em mínimo contraste generalizado são extremamente importantes para a estimação de equações diferenciais estocásticas, já que na ausência de soluções exatas para a densidade de transição do processo e a não existência de discretizações exatas todos os estimadores de processos contínuos utilizando dados observados discretamente são caracterizados por um problema de especificação incorreta.

Este artigo discute a utilização destes métodos na estimação de equações diferenciais estocásticas, e os resultados obtidos mostram que estes estimadores obtêm resultados superiores às técnicas normalmente utilizadas de métodos de momentos generalizados. Um resultado particularmente interessante é que o estimador de *Exponentially Tilted Empirical Likelihood* ([Schennach, 2007]) obtêm resultados muito superiores em termos de viés em amostras finitas, um resultado derivado das propriedades de robustez deste modelo na presença de especificação incorreta.

Este artigo está estruturado da seguinte forma - na seção 2 fazemos uma breve revisão da estimação de equações diferenciais estocásticas usando o método de momentos generalizados (GMM). Na seção 3 apresentamos os estimadores de Verossimilhança Empírica Generalizada e Mínimo Contraste Generalizado, discutindo suas propriedades, similaridades e as potenciais vantagens na estimação de equações diferenciais estocásticas. Realizamos a estimação dos modelos colocados na Tabela 1 usando estimadores de GMM, Verossimilhança Empírica Generalizada e Mínimo Contraste Generalizado na seção 4 para uma série de taxas de juros de Treasury Bills com maturidade de um mês, e a seção 5 mostra os resultados dos testes de especificação baseados em condições de sobre-identificação para os modelos estimados. Uma série de experimentos de Monte Carlo é realizada na seção 6, com o objetivo de evidenciar algumas propriedades dos estimadores discutidos no artigo. As conclusões finais estão na seção 7, e de forma resumida mostram que os estimadores propostos, até agora inéditos no contexto de estimação de equações diferenciais estocásticas, obtêm resultados superiores as técnicas de métodos de momentos generalizados normalmente utilizadas na estimação de equações diferenciais estocásticas.

## 2. ESTIMAÇÃO PELO MÉTODO DE MOMENTOS GENERALIZADOS

Como a técnica de Método de Momentos Generalizados (Generalized Method of Moments - GMM) é uma forma bastante utilizada na literatura econométrica para a estimação de equações diferenciais estocásticas e também possuem conexões profundas com os métodos de estimação de Máxima Verossimilhança Empírica e Mínimo Contraste Generalizado, iniciamos com uma revisão desta metodologia, com especial atenção para as condições de momentos utilizados na estimação.

A estimação pelo Método de Momentos Generalizados, introduzida por [Hansen, 1982], é realizada utilizando condições de momentos na forma:

$$(2.1) \quad E(g(\theta, x_t)) = 0$$

onde estas condições são avaliadas utilizando os momentos amostrais como:

$$(2.2) \quad \bar{g}(\theta) = \frac{1}{T} \sum_{t=1}^T g(\theta, x_t) = 0$$

Note que exceto no caso de um número de parâmetros igual ao número de condições de momentos (sistema exatamente identificado) o problema descrito por 2.2 admite mais que uma solução. Para obter uma solução única é necessário definir a seguinte função critério:

$$(2.3) \quad J(\theta) = \bar{g}(\theta)' W \bar{g}(\theta)$$

a e minimização desta função define a solução ótima do problema, onde  $W$  é uma matriz de ponderação positiva-definida. [Hansen, 1982] mostra que a solução assintoticamente eficiente da estimação por GMM é obtida quando esta matriz é dada por:

$$(2.4) \quad W^* = \left\{ \lim_{t \rightarrow \infty} \text{Var} \left( \sqrt{T} \bar{g}(\theta) \right) \right\}^{-1} = \Omega(\theta)^{-1}$$

e desta forma a ponderação ótima é obtida usando a inversa da matriz de variância-covariância amostral. Esta matriz é usualmente estimada usando a classe de estimadores HAC de [Newey & West, 1987], dados por:

$$(2.5) \quad \hat{\Omega} = \sum_{s=-(T-1)}^{T-1} k_h(s) \hat{\Gamma}_s(\theta^*)$$

onde  $k$  é uma função de kernel dependente da escolha de um bandwidth  $h$ , que pode ser escolhido usando os procedimentos de [Newey & West, 1987] ou [Andrews, 1991]:

$$(2.6) \quad \hat{\Gamma}_s(\theta^*) = \frac{1}{T} \sum_{t=1}^T g(\theta^*, x_t) g(\theta^*, x_{t+s})'$$

O estimador eficiente de Método de Momentos Generalizados é então obtido como a solução do problema:

$$(2.7) \quad \hat{\theta} = \arg \min_{\theta} \bar{g}(\theta)' \hat{\Omega}(\theta^*) \bar{g}(\theta)$$

Existem várias formas de se realizar a implementação do estimador GMM. A forma inicial proposta por [Hansen, 1982] é o estimador conhecido como GMM em dois estágios. Este estimador é obtido realizando-se um primeiro estágio obtendo um estimador inicial  $\hat{\theta}^* = \arg \min_{\theta} \bar{g}(\theta)' \Omega \bar{g}(\theta)$ , onde  $\Omega$  é uma matriz inicial de ponderação, normalmente uma matriz identidade. A partir deste primeiro estágio é calculada uma matriz HAC  $\hat{\Omega}(\theta^*)$  em função desta estimação inicial, e a estimativa final do estimador GMM é obtida como  $\hat{\theta} = \arg \min_{\theta} \bar{g}(\theta)' \hat{\Omega}(\theta^*) \bar{g}(\theta)$  com a matriz HAC obtida no primeiro estágio.

Note que neste caso existe uma dependência dos resultados do segundo estágio com a estimação inicial no primeiro estágio, e assim este procedimento pode criar um viés de primeira ordem prejudicando o desempenho do estimador em amostras finitas ([Hansen *et al.*, 1996]). Para resolver este problema dois procedimentos alternativos foram propostos. O primeiro procedimento é conhecido como GMM Iterativo, que é uma modificação do procedimento em dois estágios. Neste procedimento a estimação do primeiro estágio é reinicializada com o resultado da estimação do segundo estágio, e esta iteração continua até a variação no vetor de parâmetros ser menor que um épsilon escolhido.

Um outro estimador possível é conhecido como GMM com *Continuous Updating* ([Hansen *et al.*, 1996]). Nesse caso a estimação do parâmetro  $\hat{\theta}$  não é realizada em estágios, e sim simultaneamente utilizando um algoritmo de otimização numérica. Partindo de um vetor inicial  $\theta_0$  (normalmente escolhido utilizando o método de GMM em dois estágios) a estimação é realizada através de  $\hat{\theta} = \arg \min_{\theta} \bar{g}(\theta)' \hat{\Omega}(\theta^*) \bar{g}(\theta)$ , mas agora  $\theta$  e  $\hat{\Omega}(\theta^*)$  são determinados simultaneamente. Este procedimento obtém as mesmas propriedades de primeira ordem do estimador de GMM Iterativo, mas segundo [Hansen *et al.*, 1996] tem propriedades melhores em termos de viés em amostras finitas. De acordo com [Newey & Smith, 2004] e [Anatolyev, 2005] os três métodos são assintoticamente equivalentes, mas o viés de segunda ordem do estimador Continuous Updating é menor, e as iterações aumentam a eficiência do estimador. Entretanto o procedimento numérico pode estar sujeito a múltiplas modas na função objetivo, o que torna este estimador numericamente instável.

Para realizar a estimação de equações diferenciais estocásticas usando o método de momentos generalizados é necessário formular as condições de momentos em termos de alguma forma discretizada do modelo. A primeira abordagem utilizada é através da discretização simples adotada em [Chan *et al.*, 1992b] para o modelo CIR generalizado (Tabela 1) dada por:

$$(2.8) \quad r_{t+1} - r_t = \alpha_0 + \beta_0 r_t + \sqrt{r_t} \varepsilon_{t+1}$$

com as condições:  $E(\varepsilon_{t+1}) = 0$  e  $E(\varepsilon_{t+1}^2) = \sigma_0^2 r_t^2$ . Neste caso podemos formular as condições e momentos necessárias a estimação dos parâmetros  $(\alpha, \beta, \gamma, \sigma^2)$  definindo  $\varepsilon_{t+1} = r_{t+1} - r_t - \alpha_0 + \beta_0 r_t$ , e definindo quatro condições de momentos na forma:

$$(2.9) \quad g(\theta) = \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1} r_t \\ \varepsilon_{t+1}^2 - \sigma_0^2 r_t^{2\gamma} \\ (\varepsilon_{t+1}^2 - \sigma_0^2 r_t^{2\gamma}) r_t \end{bmatrix}.$$

e aplicando a estimação por GMM definida pela equação 2.7. Condições de momentos para os demais submodelos da família CIR generalizada podem ser obtidos impondo as restrições necessárias, de acordo com a Tabela I em [Chan *et al.*, 1992b]. Note que esta discretização simples não é consistente (a discretização não converge para a verdadeira solução do processo), já que ignora o intervalo de tempo entre observações. Uma forma simples de obter uma discretização consistente para este processo é utilizar uma discretização de Euler de primeira ordem, que define condições de momentos dadas por um vetor de resíduos na forma:  $\varepsilon_{t+\Delta t} = r_{t+\Delta} - r_t - (\alpha_0 + \beta_0 r_t) \Delta t$ , e assim construindo o vetor de condições de momentos como:

$$(2.10) \quad g(\theta) = \begin{bmatrix} \varepsilon_{t+\Delta t} \\ \varepsilon_{t+\Delta t} r_t \\ \varepsilon_{t+\Delta t}^2 - \sigma_0^2 r_t^{2\gamma} \Delta t \\ (\varepsilon_{t+\Delta t}^2 - \sigma_0^2 r_t^{2\gamma} \Delta t) r_t \end{bmatrix}.$$

Esta é a forma utilizada neste artigo. Note que o uso de discretizações sempre representa um problema de especificação no procedimento de inferência, já que mesmo utilizando discretizações consistentes o termo de viés causado pela discretização utilizada só tende a zero quando  $\Delta t \rightarrow 0$ . Note que o intervalo de tempo  $\Delta t$  utilizado no processo de discretização depende da frequência de observação dos dados e assim não está sob controle do pesquisador. Desta forma temos duas fontes de problemas de viés na estimação de equações diferenciais estocásticas - uma primeira forma derivada do uso de estimadores de Métodos de Momentos Generalizados e uma forma adicional gerada pela especificação incorreta dada pelo uso de uma discretização não-exata do processo. Note que no artigo original de [Chan *et al.*, 1992b] a estimação utiliza uma discretização simples do modelo e não a discretização de Euler, o que representa um aumento de viés na estimação devido a uma especificação com um erro de aproximação maior. Conseqüências deste problema podem ser vistos em [Prakasa Rao, 1999], e uma discussão suplementar deste problema está colocada na seção 6, que mostra que este problema de discretização leva a um problema de especificação incorreta na estimação de equações diferenciais estocásticas.

### 3. VEROSSIMILHANÇA EMPÍRICA GENERALIZADA E ESTIMADORES DE MÍNIMO CONTRASTE GENERALIZADOS

Enquanto que o uso do Método de Momentos Generalizados é uma forma bastante utilizada da estimação de equações diferenciais estocásticas, existe um trade-off entre a menor necessidade de pressupostos para a sua utilização e a eficiência do método em amostras finitas. Condições de regularidade para estimadores de métodos de momentos generalizados ([Hansen, 1982], [Newey & McFadden, 1994]) envolvem apenas condições para a validade assintótica das condições de momentos, e não assumem condições mais fortes como o conhecimento da distribuição do processo, e assim em amostras finitas as propriedades deste estimador podem não ser ótimas.

A situação oposta seria a estimação pelo Método de Máxima Verossimilhança, que utiliza não apenas os momentos condicionais do processo mas toda a informação existente nas densidades condicionais, e se o processo é corretamente especificado e atende as condições de regularidade ele é melhor estimador assintoticamente normal, e também alcança otimalidade em medidas como a eficiência de Badahur ([Kitamura, 2006], [DasGupta, 2008]). Entretanto o uso de máxima verossimilhança na estimação de equações diferenciais estocásticas é dificultado pela não existência em geral de formas fechadas para a solução das equações diferenciais estocásticas, e assim não é possível utilizar formas paramétricas para a estimação de máxima verossimilhança.

Uma forma alternativa e ainda não explorada na literatura de inferência em processos em tempo contínuo é o uso de uma forma de estimação *não-paramétrica* de máxima verossimilhança conhecida como Verossimilhança Empírica (*Empirical Likelihood - EL*). Seguindo [Kitamura, 2006], assumindo uma sequência de dados IID  $\{x_i\}_{i=1}^n$  de uma densidade desconhecida, e definindo  $\Delta$  como o simplex  $\{(p_1, \dots, p_n) : \sum_{i=1}^n p_i = 1, 0 \leq p_i, i = 1, \dots, n\}$ , a função de log-verossimilhança não-paramétrica é definida como:

$$(3.1) \quad \ell_{NP}(p_1, \dots, p_n) = \sum_{i=1}^n \log p_i, \quad (p_1, \dots, p_n) \in \Delta$$

que pode ser interpretada como a log-verossimilhança de uma distribuição multinomial com o suporte dado pela observações  $\{x_i\}_{i=1}^n$  da amostra, mesmo que a densidade de  $x_i$  não seja uma multinomial. Como esta formulação não envolve nenhum modelo, ela é um tanto restritiva para ser utilizada em problemas de inferência. Um avanço notável na literatura de Verossimilhança Empírica foi obtido por [Owen, 1991], que estabeleceu conexões entre a verossimilhança não-paramétrica e a estimação utilizando condições de momentos, que é também utilizada na estimação pelo Método de Momentos Generalizados, como mostrado por [Qin & Lawless, 1994]. Assumindo uma condição de momento condicional na forma:

$$(3.2) \quad E[g(\theta, x_t)] = \int g(\theta, x) d\mu = 0, \theta \in \Theta \subset \mathbb{R}^1$$

podemos transformar este problema de estimação usando condições de momentos em um problema de verossimilhança não-paramétrica utilizando probabilidades implícitas  $p_i$ , e assim a função de log-verossimilhança a ser maximizada fica sendo:

$$\ell_{NP}(p_1, \dots, p_n) = \sum_{i=1}^n \log p_i, \quad \text{tal que} \quad \sum_{i=1}^n g(\theta, x_t) p_i = 0$$

O estimador que maximiza esta expressão é o estimador de *Máxima Verossimilhança Empírica*. As probabilidades implícitas estão relacionadas a validade das condições de momentos. Estas probabilidades implícitas dão maior peso para observações aonde as condições de momentos sejam mais próximas de zero e menor peso para observações aonde as condições de momentos sejam não sejam satisfeitas. Note a similaridade com a estimação pelo método de Momentos Generalizados, que é uma forma simplificada assumindo que todos os pesos  $p_t = 1/n$ .

O uso de verossimilhança empírica é especialmente importante na estimação de equações diferenciais estocásticas, já exceto em poucos casos particulares não existem soluções exatas para as equações diferenciais estocásticas, e desta forma não é possível construir analiticamente as densidades de transição do processo, o que impossibilita a construção de uma função de verossimilhança exata. Métodos de verossimilhança empírica possibilitam avaliar de forma não paramétrica a verossimilhança do processo, e assim não dependem da existência de soluções analíticas para as equações diferenciais estocásticas, e esta avaliação não-paramétrica da função de verossimilhança é eficiente no sentido semi-paramétrico (e.g. [Bickel *et al.*, 1993]), ao mesmo tempo que utiliza a especificação paramétrica dada pela equação diferencial estocástica para construir condições de momentos.

Uma diferença com o método de Momentos Generalizados é que na metodologia de verossimilhança empírica generalizada a condição de momento pode ser um processo fracamente dependente e heterocedástico. Para tratar desta situação [Anatolyev, 2005] propõe substituir  $g(\theta, x_t)$  por uma versão suavizada (*smoothed*) definida como:

$$(3.3) \quad g^w(\theta, x_t) = \sum_{s=-m}^m w(s) p_t g(\theta, x_{t-s})$$

onde  $w(s)$  são pesos obtidos por uma função de kernel somando um, no espírito de um estimador HAC ([Andrews, 1991]). Esta modificação permite obter as mesmas condições de eficiência assintótica de primeira ordem existente nos métodos de GMM. Nesta forma as condições de momentos ficam na forma:



$$(3.4) \quad \sum_{t=1}^T p_t g^w(\theta, x_t) = 0$$

Uma interpretação da equação 3.4 em relação ao estimador GMM é que enquanto em modelos sobre-identificados estimados por GMM as condições de momentos não são exatamente iguais a zero, nos estimadores definidos por esta equação as condições de momentos são exatamente iguais a zero pela reponderação utilizando as probabilidades empíricas  $p_t$ . Note que em modelos exatamente identificados todos os estimadores propostos obtêm resultados similares, já que em todos estes estimadores as condições de momentos são sempre válidas. Em modelos sobre-identificados com condições de momentos válidas todos estes estimadores obtêm a mesma variância assintótica.

Uma interpretação alternativa do estimador de verossimilhança empírica pode ser obtida, como a de um caso particular de estimador de Mínimo Contraste Generalizado (e.g. [Bickel *et al.*, 1993]), similar a interpretação do estimador GMM como um estimador de mínimo  $\chi^2$  ou a interpretação de estimadores de Quasi-Máxima Verossimilhança como estimadores de Mínimo Contraste. Definindo uma função geral de divergência entre duas medidas de probabilidade  $P$  e  $Q$  como:

$$(3.5) \quad D(P, Q) = \int \phi \left( \frac{dP}{dQ} \right) dQ$$

aonde  $\phi$  é uma função convexa. Defina  $M$  como o conjunto de todas as medidas de probabilidade em  $\mathbb{R}^p$  e

$$(3.6) \quad \mathcal{P}(\theta) = \left\{ P \in M : \int g(\theta, x) dP = 0 \right\}$$

e  $\mathcal{P}$  o modelo estatístico de todas as medidas de probabilidade compatíveis com 3.6. O problema de otimização de mínimo contraste é dado por

$$(3.7) \quad \inf_{\theta \in \Theta} \inf_{P \in \mathcal{P}(\theta)} D(P, \mu).$$

e assim em um modelo especificado corretamente esta discrepância deve ser mínima em  $\theta = \theta_0$ . No caso de estimadores de verossimilhança empírica a estimação pontual  $\hat{\theta}$  é aquela que minimiza a discrepância entre  $\hat{p}_t$  e pesos uniformes.

Algumas medidas de divergência utilizadas na literatura são a divergência de Kullback-Leibler e a medida de entropia. Podemos formular este problema de Mínimo Contraste na forma de momentos condicionais  $E(g(\theta, x_t)) = 0$ , através do uso de uma condição modificada na forma:

$$(3.8) \quad \hat{g}(\theta) = \sum_{t=1}^T p_t g^w(\theta, x_t) = 0$$

e o estimador de Mínimo Contraste é obtido com o uso de alguma função de contraste  $h_T$ :

$$(3.9) \quad \hat{\theta}_n = \arg \min_{\theta, p_t} \sum_{t=1}^T h_T(p_t).$$

Um resultado importante é que uma escolha adequada de função de discrepância pode levar a uma representação unificada de estimadores de Máxima Verossimilhança Empírica e Mínimo Contraste. Esta representação pode ser obtida quando a função  $h_T(p_t)$  pertence a família Cressie-Read de discrepâncias dada por:

$$(3.10) \quad h_T(p_t) = \frac{[\gamma(\gamma + 1)]^{-1} (T p_t)^{\gamma+1} - 1}{T}$$

e com restrições sobre a definição da discrepância Cressie-Read temos casos particulares de várias classes de estimadores. O método de Verossimilhança Empírica é obtido com a restrição  $\gamma = 0$  na função de discrepância  $h_T(p_t)$ , o método de mínimo contraste generalizado conhecido como *Exponential Tilting* (ET) de [Kitamura & Stutzer, 1997] e [Imbens *et al.*, 1998] é obtido com  $\gamma = -1$  e o estimador de *Continuous Updating* utilizando a formulação de Verossimilhança Empírica é obtido com  $\gamma = 1$ .

[Smith, 2001] mostrou que é possível definir um outro estimador que também contém estes estimadores como casos particulares. O método de Verossimilhança Empírica Generalizado (*Generalized Empirical Likelihood - GEL*) de [Smith, 2001] é obtido como a solução do seguinte problema de ponto de sela (*saddlepoint*):

$$(3.11) \quad \hat{\theta}_n = \arg \min_{\theta} \left[ \max_{\lambda} \frac{1}{T} \sum_{t=1}^T \rho(\lambda' g^w(\theta, x_t)) \right]$$

onde  $\lambda$  define multiplicadores de Lagrange associados a restrição:

$$(3.12) \quad \sum_{t=1}^T p_t g^w(\theta, x_t) = 0$$

Estimadores são obtidos resolvendo a equação anterior com a condição de primeira ordem:

$$(3.13) \quad \sum_{t=1}^T p_t \lambda' \left( \frac{\partial g^w(\theta, x_t)}{\partial \theta} \right) = 0$$

onde:

$$(3.14) \quad p_t = \frac{1}{T} \rho'(\lambda' g^w(\theta, x_t))$$

Este estimador de verossimilhança generalizado contém o estimador de verossimilhança empírica, assumindo as mesmas condições sobre  $\gamma$  da função de divergência Cressie-Read, e modificando as funções  $h$  e  $\rho$ . O estimador de verossimilhança empírica é obtido com  $h(p) = -\ln np$  e  $\rho(\xi) = \ln(1 - \xi)$ , o estimador de *Exponential Tilting* ([Imbens *et al.*, 1998]) com  $h(p) = np \ln np$  e  $\rho(\xi) = -\exp(\xi)$ , o estimador de *Continuous Updating* como  $h(p) = (np)^2$  e  $\rho(\xi) = -(1 + \xi)^2/2^2$ . A solução para obter os estimadores dos parâmetros e dos multiplicadores de Lagrange pode ser obtida por métodos de otimização numérica ou via métodos iterativos de quasi-Newton e podem ser formuladas em um problema de dimensão menor através de uma formulação dual ([Kitamura, 2006]), e assim esta é a forma geral utilizada na estimação neste artigo.

Uma classe adicional de estimadores pode ser obtido combinando o estimador de verossimilhança empírica e o estimador de Exponencial Tilting, gerando o estimador conhecido como *Exponentially Tilted Empirical Likelihood (ETEL)* proposto por [Schennach, 2007]. Este estimador é definido como:

$$(3.15) \quad \hat{\theta} = \arg \min_{\theta} \left( n^{-1} \sum_{i=1}^n \tilde{h}(p_i(\theta)) \right)$$

onde  $\tilde{w}_i(\theta)$  é solução de :

$$(3.16) \quad \min_{\{w_i\}_{i=1}^n} n^{-1} \sum_{i=1}^n h(p_i)$$

sujeito a  $\sum_{i=1}^n p_i g(\theta, x_t) = 0$  e  $\sum_{i=1}^n p_i = 1$ , com  $\tilde{h}(\tilde{p}_t) = -\ln(np_t)$  e  $h(p_t) = np_t \ln(np_t)$ .

<sup>2</sup>Veja a tabela 1 em [Smith, 2001] para maiores detalhes.

Note que o estimador ETEL (Exponentially Tilted Empirical Likelihood) utiliza o método de Exponential Tilting para encontrar as probabilidades  $\hat{w}_i(\theta)$  e o método de verossimilhança empírica para estimar o vetor de parâmetros  $\hat{\theta}$ . Estas probabilidades estão relacionadas com os multiplicadores  $\lambda$  através da relação:

$$(3.17) \quad \hat{p}_t(\theta) = \frac{\left(\hat{\lambda}(\theta)'g(\theta, x_t)\right)}{\sum_{i=1}^n \left(\hat{\lambda}(\theta)'g(\theta, x_i)\right)}$$

Uma propriedade importante dos estimadores da classe ETEL é seu comportamento na presença de especificação incorreta. [Imbens *et al.*, 1998] apontam que o estimador de verossimilhança empírica pode ter comportamento inadequado na presença de especificação incorreta devido a presença de uma singularidade na sua função de influência, e o teorema 1 em [Smith, 2001] mostra que as propriedades assintóticas do estimador de verossimilhança empírica podem ser severamente degradadas na presença de problemas mínimos de especificação, e este efeito também afeta as estimações das probabilidades implícitas  $\hat{p}_t$ , já que na presença de problemas de especificação as probabilidades implícitas em problemas de verossimilhança tendem a se concentrar nas observações extremas, em oposição ao esperado em um estimador robusto.

O resultado obtido por [Smith, 2001] é que na classe de estimadores de mínima discrepância o único estimador com comportamento adequado na presença de problemas de especificação é o estimador de *Exponential Tilting*, já que sua função de influência não apresenta singularidades. Como o estimador ETEL é uma combinação dos estimadores de verossimilhança empírica e do estimador de *Exponential Tilting*, ele mantém as características de eficiência assintótica e mínimo viés do estimador EL e adicionalmente é robusto a presença de problemas de especificação, devido ao uso do estimador *Exponential Tilting* para estimar as probabilidades implícitas, como mostrado nos teoremas 8-10 em [Smith, 2001], indicando que este estimador é  $\sqrt{n}$  convergente mesmo na presença de problemas de especificação.

Podemos resumir algumas propriedades comuns dos estimadores discutidos neste artigo. A primeira propriedade é que todos os estimadores apresentados (GMM em dois estágios, GMM Iterativo, GMM *Continuous Updating*, *Generalized Empirical Likelihood*, *Exponential Tilting* e *Exponentially Tilted Empirical Likelihood*) possuem as mesmas propriedades de consistência e eficiência assintótica de primeira ordem (e.g. [Smith, 2001], [Schennach, 2007]), são eficientes no sentido semi-paramétrico de [Bickel *et al.*, 1993], e na validade das condições de momentos especificados todos os estimadores possuem a mesma variância assintótica, mas resultados superiores em termos de viés e propriedades assintóticas de ordem superior são válidos para os estimadores baseados em verossimilhança empírica generalizada, *Exponential Tilting* e *Exponentially Tilted Empirical Likelihood* (e.g. [Kitamura, 2006]). A classe de estimadores baseados em verossimilhança empírica também apresenta propriedades ótimas em termos de testes de hipóteses - estes testes são ótimos nos critérios minimax e de *large deviations* e uniformemente mais poderosos no sentido generalizado de Neyman-Person, como mostrado em [Kitamura, 2006].

No entanto o desempenho em amostras finitas pode ser bastante diferente. O estimador de GMM em dois estágios pode ser severamente viesado nos tamanhos de amostra utilizados em economia e finanças, e estimadores de *Continuous Updating* são numericamente instáveis devido a existência de múltiplas modas na função objetivo (e.g. [Hansen *et al.*, 1996]). [Newey & Smith, 2004] mostram que o estimador de Verossimilhança Empírica deve ter um viés em amostras finitas menor que os estimadores das classes *Exponential Tilting* e *Continuous Updating*. Em estimadores de Verossimilhança Empírica e Exponential Tilting o viés não cresce com o número de condições de momentos, o que acontece no estimador GMM. [Newey & Smith, 2004] também mostram que estimadores baseados em GEL tem boas propriedades em termos de viés de segunda ordem. . Outra propriedade interessante é que estimadores baseados em GMC e GEL são invariantes a transformações lineares no vetor de condições de momentos, o que não ocorre no estimador de GMM em dois estágios.

#### 4. ESTIMAÇÃO DE MODELOS PARA TAXAS DE JUROS DE CURTO PRAZO

Para realizar o procedimento de comparação de modelos com dados reais, seguimos a estrutura básica do artigo de [Chan *et al.*, 1992b], estimando o modelo CIR generalizados e os oito submodelos (Merton, Vasicek, CIR SR, Dothan, Brennan-Schwartz, GBM, CIR VR e CEV, na notação de [Chan *et al.*, 1992b]), com uma amostra expandida de dados de yields de *Treasury Bills* com maturidade de um mês. A amostra utiliza contém dados mensais de Julho de 1964 até Dezembro de 2003, totalizando 475 observações. Os dados utilizados são extraídos do banco de dados do Center for Research

	N	Média	Desv. Pad	$\rho_1$	$\rho_2$	$\rho_3$	$\rho_4$	$\rho_5$	$\rho_6$
$r_t$	475	0.05707511	0.0259669	0.955	0.917	0.883	0.854	0.834	0.815
$r_{t+1} - r_t$	474	-5.321519e-05	0.007433653	-0.096	-0.038	-0.072	-0.101	-0.017	0.050

TABELA 2. Estatísticas Descritivas

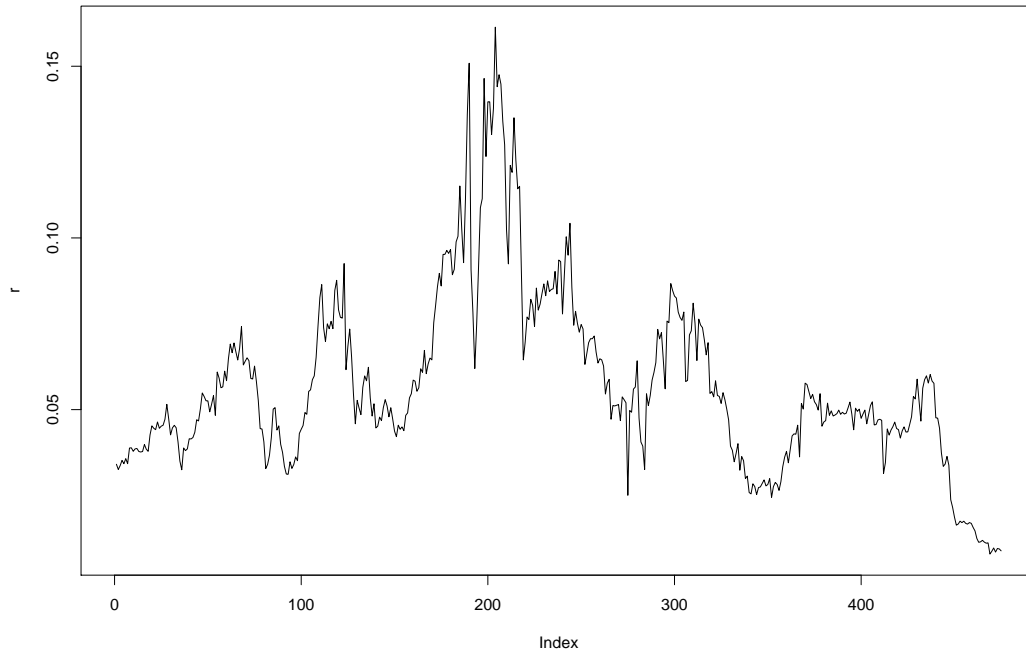


FIGURA 4.1. Taxas de Juros- Treasury Bills com maturidade de um mês

and Security Prices (CRSP Data). A Figura 4.1 mostra a série utilizada, e as estatísticas descritivas estão colocadas na Tabela 2. Podemos notar que comportamento geral é bastante similar a amostra original estudada por [Chan *et al.*, 1992b].

Para cada modelo realizamos a estimação por 11 métodos - GMM em dois estágios (GMM2S), GMM iterativo (GM-MITER), *GMM Continuous Updating* (GMMCUE), *Generalized Empirical Likelihood* (GEL), *Exponential Tilting* (ET), *Generalized Empirical Likelihood com Continuous Updating* (GELCUE) e *Exponentially Tilted Empirical Likelihood* (ETEL), bem como as versões assumindo momentos fracamente dependentes ([Anatolyev, 2005]) através de momentos suavizados dos quatro últimos estimadores (SGEL, SET, SGELCUE e SETEL), utilizando as condições de momentos descritas na seção 2, utilizando o pacote gmm de [Chausse, 2009]. Os valores iniciais utilizados na estimação são as estimativas pontuais obtidas em [Chan *et al.*, 1992b].

Os resultados obtidos pela estimação do modelo CIR Generalizado estão colocados na Tabela 3. Como o sistema é exatamente identificado, as estimações pontuais dos métodos de GMM são iguais, e em geral bastante similares as obtidas por todos os demais estimadores, exceto os estimadores ETEL que apresentam uma pequena diferença nas estimativas pontuais. Nas duas últimas linhas desta tabela temos uma variável binária que indica se houve convergência nos estimadores de parâmetros (c. par) e na estimação dos multiplicadores de Lagrange (c.  $\lambda$ ) para os estimadores fora da classe GMM. Neste problema exatamente identificado aconteceram problemas de estimação dos parâmetros no estimador GEL, SGEL e no estimador SETEL, e nenhum problema de convergência na estimação de  $\lambda$ .

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\alpha$	0.025680	0.025680	0.025680	0.025610	0.025650	0.025680	0.032970	0.025630	0.025640	0.025640	0.037580
se $\alpha$	0.013150	0.013150	0.013150	0.012690	0.012680	0.012680	0.012640	0.015990	0.016000	0.016000	0.017130
$\beta$	-0.460280	-0.460280	-0.460280	-0.457730	-0.459500	-0.460360	-0.568730	-0.459610	-0.459780	-0.459850	-0.591690
se $\beta$	0.272150	0.272150	0.272150	0.263030	0.262940	0.262940	0.260930	0.301850	0.301880	0.301900	0.305860
$\sigma^2$	1.673600	1.673600	1.673600	0.940540	0.920420	0.924040	1.332080	0.923730	0.925020	0.927560	1.313140
se $\sigma^2$	2.295810	2.295810	2.295810	1.340070	1.314290	1.318720	2.350850	2.376070	2.379200	2.385720	4.327730
$\gamma$	1.461410	1.461410	1.461410	1.343540	1.339400	1.340160	1.472900	1.340100	1.340350	1.340880	1.509510
se $\gamma$	0.262060	0.262060	0.262060	0.269900	0.270330	0.270200	0.340430	0.469460	0.469440	0.469470	0.603110
c. par	1	1	1	0	1	1	1	1	1	0	0
c. $\lambda$				1	1	1	1	1	1	1	1

TABELA 3. Estimação - Generalized CIR Model

As tabelas seguintes (4,4,5,6,7,8,5 e 10) mostram os resultados da estimação dos sistemas sobre-identificados. Na estimação do modelo de Vasicek (Tabela 4), temos três parâmetros e quatro condições de momentos. Nestas estimações podemos notar uma maior variabilidade na estimação dos parâmetros  $\alpha$  e  $\beta$ , embora pelos desvios padrões obtidos estas estimativas não são estatisticamente diferentes. Na estimação do modelo CIR SR (Tabela 5) obtemos um resultado bastante similar entre todos os métodos de estimação, com exceção da estimação de  $\beta$  para o estimador GELCUE, que pode estar relacionado a um problema de máximo local.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\alpha$	0.012770	0.009700	0.005720	0.015400	0.009320	-0.002170	0.008820	0.001950	0.005000	0.008650	0.005000
se $\alpha$	0.011490	0.011420	0.011380	0.010800	0.012040	0.010740	0.010750	0.013690	0.012510	0.010770	0.012510
$\beta$	-0.209910	-0.145250	-0.058160	-0.177900	-0.176290	0.100590	-0.168670	-0.168920	-0.180300	-0.125850	-0.180300
se $\beta$	0.242210	0.241110	0.240380	0.225640	0.249630	0.225830	0.225710	0.221900	0.229360	0.204440	0.229360
$\sigma^2$	0.000390	0.000380	0.000380	0.000400	0.003620	0.000400	0.000710	0.007560	0.005170	0.000370	0.005170
se $\sigma^2$	0.000060	0.000060	0.000060	0.000060	0.000620	0.000060	0.000070	0.003450	0.001870	0.000080	0.001870
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				0	1	1	0	0	0	1	0

TABELA 4. Estimação - Vasicek Model

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\alpha$	0.014700	0.013510	0.011000	0.012730	0.048840	0.005680	0.012190	0.012370	0.017930	0.013130	0.012060
se $\alpha$	0.011540	0.011520	0.011490	0.010850	0.011860	0.010850	0.010810	0.011110	0.012610	0.011270	0.011030
$\beta$	-0.245170	-0.219500	-0.164570	-0.220600	-0.626310	-0.058160	-0.218900	-0.224410	-0.230920	-0.214410	-0.216110
se $\beta$	0.242760	0.242370	0.241820	0.226930	0.237360	0.227960	0.226900	0.210930	0.238380	0.214710	0.208060
$\sigma^2$	0.008100	0.008050	0.008000	0.008520	0.010370	0.008160	0.004210	0.005910	0.016620	0.007830	0.002990
se $\sigma^2$	0.001110	0.001110	0.001110	0.001690	0.001230	0.001200	0.001250	0.001490	0.002410	0.001490	0.001660
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				0	1	1	0	0	1	1	0

TABELA 5. Estimação - CIR SR Model

Os resultados da estimação do modelo de Merton (Tabela 6) mostram dois padrões de comportamento, com valores de  $\alpha$  próximos de .0003 para os métodos de GMM2S, GMMITER e GELCUE e SGELCUE, e valores próximos a 0.0065 para os demais estimadores, mas notando que nestes casos não houve convergência na estimação de  $\lambda$ . Resultados similares são

obtidos na estimação do modelo de Dothan (Tabela 7). Os valores dos estimadores do parâmetro  $\alpha$  de GMM e dos demais estimadores com  $\lambda$  convergente são bastante similares.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\alpha$	0.002940	0.002960	0.003050	0.000820	0.000720	0.002610	0.000720	0.000850	0.000690	0.002800	0.000690
se $\alpha$	0.002570	0.002570	0.002570	0.002370	0.002370	0.002410	0.002370	0.002490	0.002540	0.003130	0.002540
$\sigma^2$	0.000390	0.000380	0.000380	0.001030	0.000930	0.000400	0.000930	0.001100	0.000940	0.000370	0.000940
se $\sigma^2$	0.000060	0.000060	0.000060	0.000190	0.000180	0.000060	0.000180	0.000320	0.000280	0.000080	0.000280
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				0	0	1	0	0	0	1	0

TABELA 6. Estimação - Merton Model

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\sigma^2$	0.135950	0.134590	0.134560	-0.052760	-0.052760	0.139540	-0.052760	-0.062710	-0.062710	0.129520	-0.062710
se $\sigma^2$	0.018210	0.018220	0.018220	0.027610	0.027610	0.018870	0.027610	0.039050	0.039050	0.022500	0.039050
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				0	0	1	0	0	0	1	0

TABELA 7. Estimação - Dothan Model

Na estimação dos modelos de Brennan-Schwartz (Tabela 8), CIR VR ( 9) e CEV (Tabela 10) obtivemos sucesso na convergência do vetor de parâmetros e do vetor de multiplicadores de Lagrange em todos os métodos, e como esperado obtivemos resultados bastantes similares em todos os estimadores utilizados, exceto os estimadores baseados em *Exponentially Tilted Empirical Likelihood*.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\alpha$	0.019800	0.019680	0.018880	0.020820	0.019830	0.017060	0.032160	0.025000	0.022600	0.019920	0.024570
se $\alpha$	0.011810	0.011810	0.011800	0.011320	0.011240	0.011240	0.011760	0.012850	0.012650	0.012490	0.013460
$\beta$	-0.343020	-0.340200	-0.323030	-0.322980	-0.338490	-0.287340	-0.298100	-0.446630	-0.400510	-0.347120	-0.315520
se $\beta$	0.246860	0.246820	0.246660	0.236550	0.235330	0.235330	0.240360	0.244610	0.241050	0.238700	0.250080
$\sigma^2$	0.144450	0.144100	0.143670	0.158580	0.146250	0.143280	0.127940	0.152150	0.146510	0.142520	0.134780
se $\sigma^2$	0.018530	0.018530	0.018540	0.019680	0.019650	0.019650	0.019790	0.025920	0.025780	0.025690	0.025900
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				1	1	1	1	1	1	1	1

TABELA 8. Estimação - Brennan-Schwartz Model

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\sigma^2$	1.947930	1.935290	1.973020	2.134040	2.084420	2.031730	0.930910	2.089810	1.991970	1.949750	1.894050
se $\sigma^2$	0.260370	0.260230	0.260670	0.265510	0.264360	0.263270	0.273340	0.341910	0.337720	0.336220	0.334510
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				1	1	1	1	1	1	1	1

TABELA 9. Estimação - CIR VR Model

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
$\beta$	0.061900	0.062120	0.065920	0.054140	0.052030	0.051540	0.011800	0.029150	0.028310	0.060810	-0.006150
se $\beta$	0.054340	0.054340	0.054340	0.050480	0.050480	0.050460	0.050610	0.054950	0.055490	0.056250	0.056190
$\sigma^2$	0.238460	0.149560	0.149790	0.729960	0.500060	0.213700	0.732500	0.606200	0.593460	0.143730	0.560070
se $\sigma^2$	0.437080	0.299670	0.299890	1.067110	0.773820	0.377510	1.432650	1.653030	1.748400	0.446840	2.093630
$\gamma$	1.098550	1.016000	1.016110	1.296680	1.229400	1.075610	1.365170	1.254660	1.290140	1.014000	1.355840
se $\gamma$	0.337350	0.365020	0.364740	0.278770	0.292490	0.327000	0.377480	0.503040	0.546550	0.560770	0.700410
c. par	1	1	1	1	1	1	1	1	1	1	1
c. $\lambda$				1	1	1	1	1	1	1	1

TABELA 10. Estimação - CEV Model

## 5. TESTES DE ESPECIFICAÇÃO

Uma forma de realizar testes de especificação no contexto da estimação por GMM em modelos sobre-identificados é através da distância entre as condições de momentos e zero. No caso sobre-identificado a maior proximidade das condições de momentos avaliadas em  $\hat{\theta}$  seria uma evidência da validade da especificação utilizada. Uma forma de definir uma estatística de teste é através da própria função critério, dando origem ao chamado teste J, cuja forma é dada por:

$$(5.1) \quad J = T\bar{g}(\hat{\theta})'[\hat{\Omega}(\hat{\theta})]^{-1}\bar{g}(\hat{\theta})$$

A distribuição assintótica do teste é uma distribuição  $\chi^2(m - k)$ , cujos graus de liberdade são dados pelo número de momentos em excesso ao número de parâmetros estimados. Na estimação pelos métodos de mínimos contraste generalizado/verossimilhança empírica generalizada também é possível construir dois testes alternativos de especificação, um teste de multiplicador de Lagrange (LM) e um teste de razão de verossimilhança LR, como discute [Smith, 2001], utilizando os multiplicadores de Lagrange da equação 3.11.

A intuição do teste LM é semelhante ao do teste J - se as condições de momentos são válidas os multiplicadores de Lagrange não devem ser diferentes de zero, e assim não é necessário a imposição de restrições no modelo para tornar as condições de momentos iguais a zero. A forma do teste LM neste contexto é dada por:

$$(5.2) \quad LM = T\hat{\lambda}'\hat{\Omega}(\hat{\theta})\hat{\lambda}$$

Note que também é possível realizar análises de especificação olhando para as distribuições estimadas individualmente para cada  $\hat{\lambda}$  ao invés do teste conjunto  $LM^3$ .

O teste de Razão de Verossimilhança (LR) é obtido comparando a função objetivo  $\rho$  de um modelo irrestrito com a estimação com um modelo restrito, formulado no contexto de modelos de mínimo contraste generalizado, como definido na seção 3. Esta forma é dada por:

$$(5.3) \quad LR = 2 \left[ \sum_{t=1}^T \rho \left( \hat{\lambda}' g^w(\hat{\theta}, x_t) \right) - \rho(0) \right]$$

Os testes LM e LR também são distribuídos assintoticamente como uma  $\chi^2(m - k)$ . Note que a validade dos testes LM e LR depende da convergência na estimação de  $\lambda$ , e conforme já relatado isto não acontece em todas as estimações. As Tabelas (11,12,14,15,16 e 17) mostram os resultados dos testes J, LR e LM para todos os modelos estimados, notando que a última coluna indica se houve convergência na estimação de  $\lambda$ . Embora os testes J, LM e LR sejam assintoticamente equivalentes, os resultados de otimalidade de testes em modelos de verossimilhança empírica/mínimo contraste generalizado (e.g. [Kitamura, 2006]) bem como as melhores propriedades nas estimativas pontuais em amostras finitas indicam que os testes LR e LM devem ter propriedades melhores em amostras finitas que o teste J obtido por GMM, que pode ser severamente

<sup>3</sup>Estes testes foram calculados para todos os modelos estimados, mas não são reportados por questões de espaço e são disponíveis com os autores.

viesado para baixo em amostras finitas como apontado por [Zhou, 2000], levando a uma maior probabilidade de aceitação incorreta de uma hipótese nula de especificação correta do modelo.

	J test	p-value	LR Test	LR p-value	LM test	LM p-value	$\lambda$ c.
GMM2S	6.488968	0.010855					
GMMITER	5.526426	0.018731					
GMMCUE	5.440151	0.019679					
GEL	9.287416	0.002307	0.022057	0.881933	6.435292e-05	0.993599	0
ET	386.872492	0	613.622506	1.823266e-135	1787.180837	0	1
GELCUE	6.378555	0.011551	6.378554	0.011550	6.3785546	0.011550	1
ETEL	45.481497	0	6.420884e-05	0.993606	6.4231349e-05	0.993605	0
SGEL	458.032031	0	1.375009e-05	0.997041	1.3751784e-05	0.997041	0
SET	441.455389	0	1.375500e-05	0.997040	1.3755933e-05	0.997040	0
SGELCUE	18.940141	0.000013	18.940140	1.348845e-05	18.940140	1.348845e-05	1
SETEL	441.455412	0	1.3755061e-05	0.997040	1.3755933e-05	0.997040	0

TABELA 11. Testes - Vasicek

Os resultados dos testes obtidos para as estimações no modelo de Vasicek (11) mostram que os resultados dos testes J por GMM não indicam a rejeição da especificação deste modelo a um nível de 1%, enquanto que as mesmas estatísticas J obtidas pelas estimações por GMC/GEL em geral indicam a rejeição da validade deste modelo a níveis elevados de significância, mesmo resultado obtidos pelos testes LR e LM com  $\lambda$  convergente.

	J test	p-value	LR Test	LR p-value	LM test	LM p-value	$\lambda$ c.
GMM2S	4.216564	0.040031					0
GMMITER	3.592661	0.058035					0
GMMCUE	3.551054	0.059508					0
GEL	117.453521	0	6.407120e-05	0.993613	6.409091e-05	0.993612	0
ET	37.126372	0	33.63956	6.632985e-09	53.89472	2.115245e-13	1
GELCUE	4.020585	0.044948	4.020584	0.044948	4.0205841	0.044948	1
ETEL	20.558885	0.000006	6.410327e-05	0.993611	6.412423e-05	0.993610	0
SGEL	22.747932	0.000002	1.377496e-05	0.997038	1.377572e-05	0.997038	0
SET	154.764952	0	134.979566	3.33584e-31	205.812831	1.125719e-46	1
SGELCUE	12.610579	0.000384	12.610578	0.00038	12.610578	0.000383	1
SETEL	70.456519	0	1.3764273e-05	0.99703	1.3765027e-05	0.997039	0

TABELA 12. Testes - CIR SR

Nos testes realizados para o modelo CIR SR (12) temos um resultado contrastante entre as estimações por GMM e as estimações por GMC/GEL. O resultado dos testes J por GMM são mais favoráveis a nula de validade do modelo, enquanto que estes mesmos testes pelos demais métodos (exceto GELCUE) indicam uma evidência forte para a rejeição da nula de validade da especificação. Para os modelos de Merton (13) e Dothan (14) obtemos um resultado semelhante, com os testes derivados de GMM indicando a validade da especificação e os testes obtidos por GMC/GEL indicando a rejeição da nula.

Nos testes de especificação do modelo de Brennan-Schwartz (15) as evidências são mistas. Os testes J de GMM indicam a validade da hipótese nula, e temos resultados distintos para os demais testes e estimadores. O resultado mais notável é que os testes baseados em *Exponentially Tilted Empirical Likelihood* mostram evidências fortes contrárias a validade da hipótese nula, enquanto que os demais em geral não levam a rejeição da nula a um nível de significância de 1%, notando que todas as estimações de  $\lambda$  foram convergentes. Para o modelo CIR VR (16) e CEV (17) os testes J, LR e LM de GMM, GEL, ET e GELCUE são favoráveis a nula de uma especificação correta.



	J test	p-value	LR Test	LR p-value	LM test	LM p-value	c.
GMM2S	5.782589	0.055504					
GMMITER	5.479873	0.064574					
GMMCUE	5.478272	0.064626					
GEL	403.937267	0	6.47028e-05	0.99996	6.473173e-05	0.99996	0
ET	399.352189	0	6.47194e-05	0.99996	6.473422e-05	0.99996	0
GELCUE	6.490608	0.038957	6.49060	0.03895	6.490608	0.03895	1
ETEL	399.344468	0	6.47065e-05	0.99996	6.473422e-05	0.99996	0
SGEL	457.815412	0	1.15816e-05	0.99999	1.158377e-05	0.99999	0
SET	288.648313	0	410.812	6.21210e-90	2522.883266	0	1
SGELCUE	21.766052	0.000019	21.76605	1.877421982e-05	21.76605183	1.877421e-05	1
SETEL	390.833746	0	1056.6133	3.625113710e-230	3283.8835	0	1

TABELA 13. Testes - Merton

	J test	p-value	LR Test	LR p-value	LM test	LM p-value	$\lambda$ c.
GMM2S	4.895422	0.179617					
GMMITER	4.884452	0.180456					
GMMCUE	4.884442	0.180457					
GEL	107.365956	0	6.469681e-05	0.999999	6.472510e-05	0.999999	0
ET	107.364954	0	6.470953e-05	0.999999	6.472510e-05	0.999999	0
GELCUE	3.396479	0.334439	3.396479	0.334438	3.396479	0.334438	1
ETEL	107.365956	0	6.469681e-05	0.999999	6.472510e-05	0.999999	0
SGEL	215.285851	0	1.154637e-05	0.999999	1.154808e-05	0.999999	0
SET	215.285775	0	1.154716e-05	0.999999	1.154808e-05	0.999999	0
SGELCUE	21.531162	0.000082	21.53116	8.165149e-05	21.531161	8.165149e-05	1
SETEL	215.285851	0	1.154637e-05	0.999999	1.1548084e-05	0.999999	0

TABELA 14. Testes- Dothan

	J test	p-value	LR Test	LR p-value	LM test	LM p-value	$\lambda$ c.
GMM2S	1.207473	0.271833					
GMMITER	1.110753	0.291918					
GMMCUE	1.105794	0.292998					
GEL	2.439994	0.118277	2.898200	0.088678	7.007773	0.008115	1
ET	1.272431	0.259311	1.8133323	0.178108	4.596062	0.032045	1
GELCUE	1.206081	0.272110	1.2060806	0.272109	1.206080	0.272109	1
ETEL	24.950105	0.000001	31.877056	1.642457e-08	44.896698	2.077078e-11	1
SGEL	5.346504	0.020764	6.8471325	0.008878	14.292672	0.000156	1
SET	4.295750	0.038208	5.8586040	0.015501	12.920344	0.000325	1
SGELCUE	4.006386	0.045328	4.0063857	0.045328	4.0063857	0.045328	1
SETEL	23.766350	0.000001	29.6530317	5.167175e-08	44.951707	2.019540e-11	1

TABELA 15. Testes Brennan-Schwartz

Uma forma alternativa de verificar problemas de especificação incorreta é através do uso das probabilidades implícitas estimadas (Figuras 5.1 e 5.2). Enquanto que para os estimadores com valores  $\gamma \leq 0$  na função de discrepância Cressie-Read as probabilidades implícitas são positivas por construção (GEL, ET e ETEL) nos estimadores com  $\gamma > 0$  elas podem assumir valores negativos, mas na validade da especificação essas probabilidades devem convergir para  $n^{-1}$ . Indicações de problemas

	J test	p-value	LR Test	LR p-value	LM test	LM p-value	$\lambda$ c.
GMM2S	6.207851	0.101924					
GMMITER	6.317539	0.097143					
GMMCUE	6.301818	0.097815					
GEL	5.116470	0.163465	5.474152	0.140192	6.286955	0.098453	1
ET	4.980668	0.173218	5.376636	0.146206	6.344717	0.095991	1
GELCUE	4.931533	0.176880	4.931532	0.176879	4.931532	0.176879	1
ETEL	25.751894	0.000011	74.70250	4.196299e-16	302.2736	3.203888e-65	1
SGEL	26.923453	0.000006	36.47767	5.934422e-08	58.57768	1.183193e-12	1
SET	26.157968	0.000009	33.43432	2.608053e-07	56.52854	3.239995e-12	1
SGELCUE	26.080043	0.000009	26.08004	9.176361e-06	26.08004	9.176361e-06	1
SETEL	26.216408	0.000009	38.45658	2.262388e-08	61.08300	3.450144e-13	1

TABELA 16. Testes CIR VR

	J test	p-value	LR Test	LR p-value	LM test	LM p-value	$\lambda$ c.
GMM2S	3.617753	0.057166					
GMMITER	3.442114	0.063554					
GMMCUE	3.437260	0.063741					
GEL	3.153802	0.075750	3.366812	0.066522	4.744549	0.029391	1
ET	2.847644	0.091508	3.155803	0.075657	4.453728	0.034825	1
GELCUE	2.628710	0.104947	2.628709	0.104946	2.628709	0.104946	1
ETEL	9.336512	0.002246	15.40316	8.684280e-05	25.647695	4.097845e-07	1
SGEL	22.514575	0.000002	29.84871	4.671062e-08	54.040122	1.964371e-13	1
SET	20.245597	0.000007	28.54174	9.170001e-08	69.290234	8.498936e-17	1
SGELCUE	16.127466	0.000059	16.12746	5.921894e-05	16.127466	5.921894e-05	1
SETEL	49.217782	0	122.008296	2.298746e-28	387.619783	2.729337e-86	1

TABELA 17. Testes CEV

de especificação estão relacionadas a presença de probabilidades negativas nos estimadores CUE e SCUE ( $\gamma = 1$ ), o que ocorre de forma mais evidente nos modelos de Vasicek e Dothan para este método de estimação.

Note que também é possível construir testes alternativos de especificação e quebra estrutural utilizando as probabilidades implícitas obtidas pela estimação de modelos de verossimilhança empírica generalizada, como mostrado por [Antoine *et al.*, 2007], [Ramalho & Smith, 2005] e [Guay & Lamarche, 2008], utilizando estatísticas do tipo Pearson medindo a distância quadrática entre as probabilidades implícitas de modelos restritos e irrestritos. Estas estatísticas são assintoticamente equivalente aos testes de especificação usando as condições de momentos apresentados nesta seção.

Em geral os resultados dos testes de especificação usando condições de sobre-identificação estimados por verossimilhança empírica/mínimo contraste generalizado apontam para a rejeição da validade da hipótese nula de correta especificação, enquanto que em geral os testes J por GMM são mais favoráveis a validade da nula. Os resultados dos testes de verossimilhança empírica/mínimo contraste generalizado são consistentes com a percepção que modelos de um único fator como os estimados nesse artigo podem ser excessivamente simples para modelar processos de taxas de juros ou a precificação de instrumentos de renda fixa (e.g. [Stambaugh, 1988], [Stanton, 1997], [Litterman & Scheinkman, 1991], [Longstaff & Schwartz, 1992] e [Lund & Andersen, 1997]).

## 6. EVIDÊNCIA DE MONTE CARLO

Nas estimações realizadas na seção 4 temos a comparação de diversos métodos de estimação baseados em GMM, GMC e GEL. Como todos estes métodos são primeira ordem assintoticamente equivalentes é esperado que seu comportamento seja similar em todas as estimações convergentes e assim é difícil avaliar suas propriedades na estimação com apenas uma

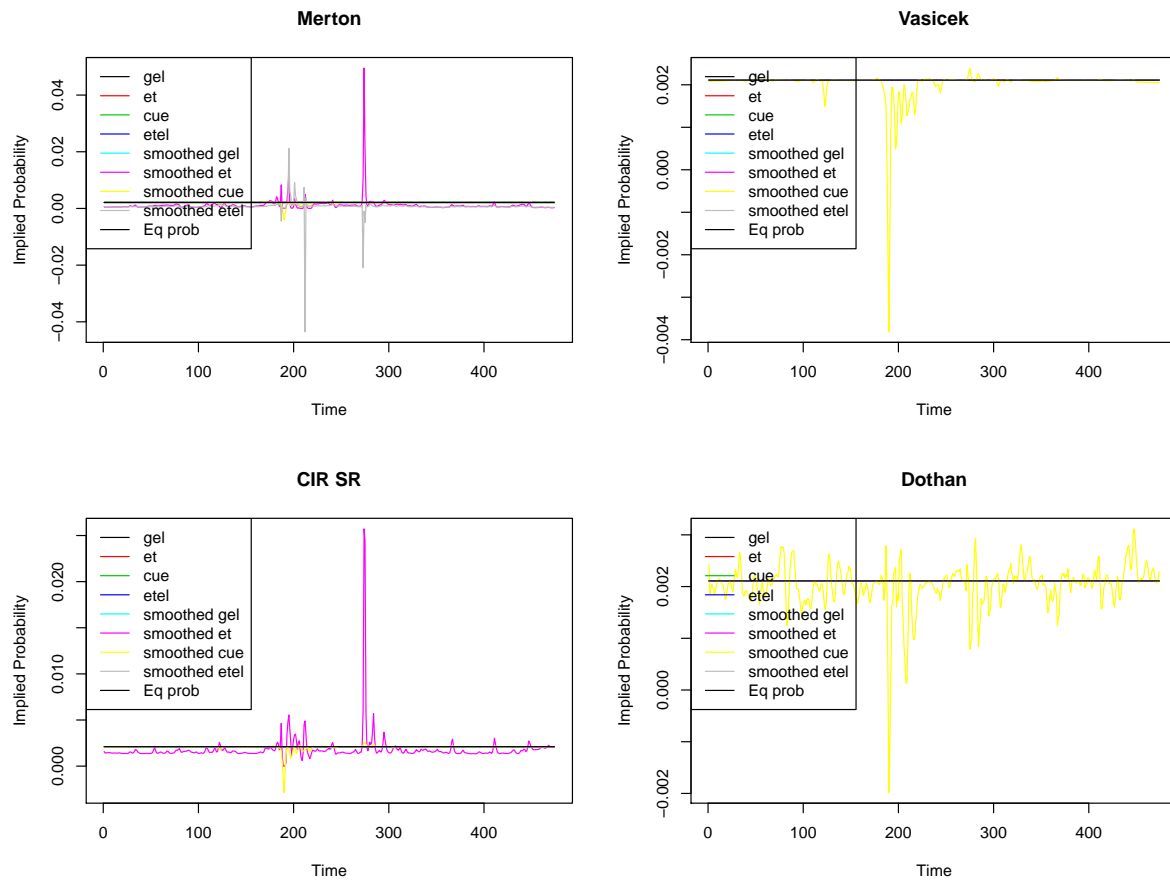


FIGURA 5.1. Probabilidades Implícitas - Merton, Vasicek, CIR SR e Dothan

amostra. Para realizar uma análise do comportamento geral destes estimadores realizamos uma análise de Monte Carlo avaliando diversas propriedades destes estimadores, em especial o viés em amostras finitas e uma discussão sobre validade na presença de especificação incorreta no contexto de estimação de equações diferenciais estocásticas.

O procedimento de Monte Carlo realizado consiste em simular os modelos CIR Generalizado, Vasicek e CIR SR<sup>4</sup> e realizar a estimação com os métodos de estimação propostos, e com base nestas estimações avaliar o viés, erro quadrático médio (mse) e o erro absoluto médio (mae) em relação a cada parâmetro estimado. O procedimento de simulação utilizado para o processo CIR Generalizado utiliza uma discretização de Milstein (e.g. [Milstein, 1974], [Kloeden & Platen, 1992]) para gerar trajetórias do processo, já que neste processo não existe solução analítica exata para a densidade de transição. Para os processos Vasicek e CIR SR utilizamos a densidade de transição exata para gerar trajetórias simuladas (e.g. [Ait-Sahalia, 2002]).

Note que este detalhe é de fundamental importância. Antes de discutir este ponto vamos introduzir a notação de convergência forte de discretizações. Suponha que estamos interessados em gerar uma trajetória da equação diferencial estocástica  $dX_t = \mu(t, X_t) + \sigma(t, X_t)dW_t$  utilizando uma discretização que gere trajetórias  $Y_t^\Delta$  deste processo, e que as trajetórias desta aproximação convirjam para a verdadeira trajetória. Uma aproximação  $Y_t^\Delta$  é dita convergente de ordem forte  $\gamma > 0$  se existem constantes positivas  $K$  e  $\gamma$  tais que para todo  $\Delta$  seja válido:

<sup>4</sup>Estes experimentos foram realizados para os demais modelos e apresentam resultados similares, mas por questões de espaço não são apresentados.

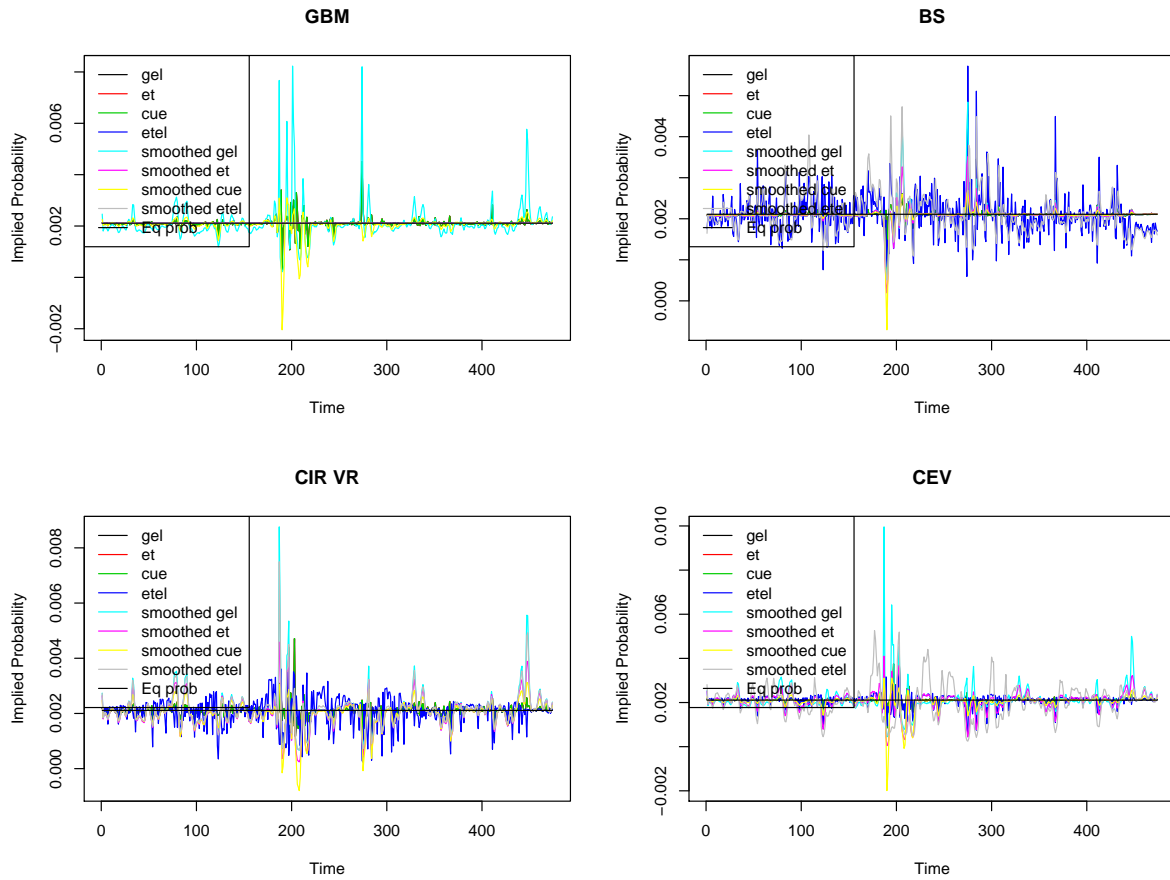


FIGURA 5.2. Probabilidades Implícitas - GBM, BS, CIR VR e CEV

$$E\left(|X_t - Y_t^\Delta|\right) \leq K\Delta^\gamma,$$

em que  $K$  não dependa do intervalo de discretização  $\Delta$ . Sobre condições usuais Lipschitz e crescimento é possível mostrar (e.g. [Kloeden & Platen, 1992], [Prakasa Rao, 1999]) que a discretização de Euler converge com ordem forte  $\gamma = 0.5$  e a discretização de Milstein ([Milstein, 1974]) é convergente de ordem forte com  $\gamma = 1$ .

Como a discretização utilizada nas condições de momentos é de ordem forte inferior a utilizada na simulação do processo temos um problema de especificação incorreta gerada pela discretização utilizada. Este problema ocorre de forma mais intensa quando podemos utilizar a solução exata da equação diferencial estocástica para gerar a trajetória do processo. O ponto fundamental é que na estimação baseada em discretizações aproximadas sempre existe um viés gerado pela discretização do processo, e um dos objetivos do estudo de Monte Carlo é verificar se algum método consegue uma redução de viés em relação a este efeito, que pode ser interpretado como um problema de especificação. Note que no artigo original de [Chan *et al.*, 1992b] a discretização utilizada é ainda mais simples que a de Euler, e assim o viés existente nos estimadores deve ser ainda maior.

O primeiro experimento de Monte Carlo correspondente a simulação de 1000 trajetórias de tamanho 474 de um processo CIR generalizado com um vetor de parâmetros dado por  $\alpha = 0.0408$ ,  $\beta = -.5921$ ,  $\sigma^2 = 1.6704$  e  $\gamma = 1.4999$ . Os resultados deste experimento estão colocados na Tabela 18. Os resultados obtidos mostram que existe um viés relevante na estimação de todos os parâmetros, em especial do parâmetro  $\sigma^2$ . Os resultados em termos do tamanho do viés e do erro quadrático médio

são bastante similares para quase todos os estimadores, exceto para o estimador ETEL e SETEL, que apresentam resultados muito superiores em termos de viés, mse e mae em relação aos demais métodos para todos os parâmetros estimados.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
média $\alpha$	0.057959	0.057959	0.057959	0.055002	0.058778	0.058774	0.042554	0.052587	0.057653	0.059102	0.041983
viés $\alpha$	0.017159	0.017159	0.017159	0.014202	0.017978	0.017974	0.001754	0.011787	0.016853	0.018302	0.001183
mse $\alpha$	0.000601	0.000601	0.000601	0.000589	0.000649	0.000652	0.000250	0.000373	0.000541	0.000647	0.000141
mae $\alpha$	0.018542	0.018542	0.018542	0.016740	0.019303	0.019505	0.005722	0.014768	0.018438	0.019485	0.006281
média $\beta$	-0.845974	-0.845974	-0.845974	-0.796893	-0.858470	-0.859247	-0.559705	-0.762769	-0.846514	-0.864585	-0.571162
viés $\beta$	-0.253874	-0.253874	-0.253874	-0.204793	-0.266370	-0.267147	0.032395	-0.170669	-0.254414	-0.272485	0.020938
mse $\beta$	0.127159	0.127159	0.127159	0.104235	0.139544	0.141990	0.006764	0.082451	0.130439	0.141021	0.007560
mae $\beta$	0.276890	0.276890	0.276890	0.250283	0.293792	0.296454	0.050187	0.219321	0.284997	0.296498	0.053798
média $\sigma^2$	1.943908	1.943908	1.943908	1.588835	1.313430	1.341354	1.703154	1.573219	1.340508	1.325817	1.705367
viés $\sigma^2$	0.273508	0.273508	0.273508	-0.081565	-0.356970	-0.329046	0.032754	-0.097181	-0.329892	-0.344583	0.034967
mse $\sigma^2$	2.164011	2.164011	2.164011	1.767383	2.036499	2.326232	0.005270	0.771392	2.145764	2.101051	0.011036
mae $\sigma^2$	0.693318	0.693318	0.693318	0.786565	0.293792	1.108167	0.045776	0.589236	0.988021	1.088776	0.064033
média $\gamma$	1.487937	1.487937	1.487937	1.440525	1.373429	1.370929	1.543712	1.456968	1.385992	1.381241	1.542676
viés $\gamma$	-0.011963	-0.011963	-0.011963	-0.059375	-0.126471	-0.128971	0.043812	-0.042932	-0.113908	-0.118659	0.042776
mse $\gamma$	0.026592	0.026592	0.026592	0.020981	0.043057	0.045600	0.004902	0.014792	0.037274	0.120674	0.005330
mae $\gamma$	0.096684	0.096684	0.096684	0.111090	0.172902	0.178556	0.047938	0.088226	0.158066	0.185138	0.051808

TABELA 18. Monte Carlo - Generalized CIR Model

No experimento de Monte Carlo para o processo Vasicek (Tabela 19), simulamos novamente mil trajetórias com um vetor de parâmetros dado por  $\alpha = 0.0154$ ,  $\beta = -0.1779$ ,  $\sigma^2 = .0004$  e  $\gamma = 0$ . Os resultados mostram que novamente temos um desempenho superior dos estimadores ETEL, e também é possível notar que neste experimento o estimador com o pior desempenho foi o estimador GMMCUE. Para o processo CIR SR (Tabela 20), simulamos mil trajetórias do processo com  $0.0189$ ,  $\beta = -0.2339$ ,  $\sigma^2 = 0.0073$ , e  $\gamma = .5$ . Observamos o mesmo comportamento de melhor desempenho dos estimadores da classe ETEL e um desempenho similar dos demais estimadores.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
média $\alpha$	0.025073	0.025095	0.037840	0.023264	0.024902	0.024964	0.018063	0.021952	0.024932	0.025063	0.017604
viés $\alpha$	0.009673	0.009695	0.022440	0.007864	0.009502	0.009564	0.002663	0.006552	0.009532	0.009663	0.002204
mse $\alpha$	0.000255	0.000256	0.095821	0.000196	0.000250	0.000253	0.000033	0.000152	0.000252	0.000257	0.000036
mae $\alpha$	0.011528	0.011555	0.024212	0.009726	0.011357	0.011461	0.003555	0.008423	0.011399	0.011560	0.003811
média $\beta$	-0.293078	-0.293337	-0.308412	-0.269012	-0.291078	-0.291721	-0.170961	-0.251174	-0.291324	-0.292897	-0.170805
viés $\beta$	-0.115178	-0.115437	-0.130512	-0.091112	-0.113178	-0.113821	0.006939	-0.073274	-0.113424	-0.114997	0.007095
mse $\beta$	0.032397	0.032552	0.157658	0.024414	0.031609	0.032012	0.000089	0.018236	0.031862	0.032543	0.000094
mae $\beta$	0.132587	0.132930	0.147679	0.108667	0.130302	0.131647	0.007034	0.091495	0.130582	0.132902	0.007454
média $\sigma^2$	0.000396	0.000396	0.009369	0.000338	0.000397	0.000396	-0.000661	0.000236	0.000397	0.000396	-0.000576
viés $\sigma^2$	-0.000004	-0.000004	0.008969	-0.000062	-0.000003	-0.000004	-0.001061	-0.000164	-0.000003	-0.000004	-0.000976
mse $\sigma^2$	0.000000	0.000000	0.060878	0.000001	0.000000	0.000000	0.000015	0.000002	0.000000	0.000000	0.000020
mae $\sigma^2$	0.000021	0.000021	0.008993	0.000136	0.130302	0.000021	0.001685	0.000234	0.000020	0.000021	0.002261

TABELA 19. Monte Carlo - Vasicek

Note que até agora o problema de especificação incorreta era causado apenas pelo uso da discretização aproximada na construção das condições de momentos do processo. Para verificar se as propriedades de melhor desempenho dos estimadores da classe ETEL são válidas em situações mais gerais de especificação incorreta, utilizamos como processo gerador de dados

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
média $\alpha$	0.028081	0.028125	0.028042	0.026700	0.027506	0.027955	0.036431	0.024417	0.027453	0.028069	0.032234
viés $\alpha$	-0.010819	-0.010775	-0.010858	-0.012200	-0.011394	-0.010945	-0.002469	-0.014483	-0.011447	-0.010831	-0.006666
mse $\alpha$	0.000270	0.000270	0.000272	0.000261	0.000269	0.000274	0.000151	0.000313	0.000269	0.000272	0.000208
mae $\alpha$	0.014132	0.014158	0.014209	0.014499	0.014344	0.014278	0.008566	0.016214	0.014363	0.014205	0.010919
média $\beta$	-0.359449	-0.360034	-0.359019	-0.254924	-0.349823	-0.357672	-0.225273	-0.243888	-0.349048	-0.359527	-0.223735
viés $\beta$	-0.125549	-0.126134	-0.125119	-0.021024	-0.115923	-0.123772	0.008627	-0.009988	-0.115148	-0.125627	0.010165
mse $\beta$	0.043979	0.044457	0.044259	0.009497	0.038638	0.043871	0.000239	0.006125	0.038912	0.044394	0.000306
mae $\beta$	0.151631	0.152107	0.151257	0.049759	0.143119	0.150894	0.010178	0.040522	0.142224	0.151973	0.012182
média $\sigma^2$	0.007231	0.007229	0.007233	0.007796	0.007260	0.007232	0.010458	0.008470	0.007256	0.007234	0.011572
viés $\sigma^2$	-0.000069	-0.000071	-0.000067	0.000496	-0.000040	-0.000068	0.003158	0.001170	-0.000044	-0.000066	0.004272
mse $\sigma^2$	0.000000	0.000000	0.000000	0.000003	0.000000	0.000000	0.000079	0.000021	0.000000	0.000000	0.000105
mae $\sigma^2$	0.000410	0.000410	0.000410	0.000791	0.143119	0.000409	0.004024	0.001569	0.000407	0.000410	0.005531

TABELA 20. Monte Carlo - CIR SR

trajetórias do processo CIR Generalizado com vetor de parâmetros  $\alpha = 0.0408$ ,  $\beta = -0.5921$ ,  $\sigma^2 = 1.6704$  e  $\gamma = 1.4999$ , mais agora utilizamos como especificação do modelo estimado um modelo CIR SR que assume que  $\gamma = .5$ .

Os resultados deste experimento (Tabela 21) mostram que neste caso geral também ocorre um melhor desempenho dos estimadores ETEL e ET, sendo que os demais estimadores tem um desempenho muito pior em relação a estimação do parâmetro  $\sigma^2$ . Note que é esperado que o problema de especificação incorreta nesta situação afete principalmente a estimação da variância do processo, já que nas classes de modelos CIR a volatilidade do processo é uma função do nível do processo com parâmetro  $\gamma$ .

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
média $\alpha$	0.057816	0.068673	0.044626	0.020599	0.030640	0.037049	0.022194	0.017033	0.028992	0.053330	0.019102
viés $\alpha$	0.017016	0.027873	0.003826	-0.020201	-0.010160	-0.003751	-0.018606	-0.023767	-0.011808	0.012530	-0.021698
mse $\alpha$	0.000893	0.001613	0.002179	0.002029	0.000732	0.003294	0.001177	0.000679	0.001008	0.002621	0.001807
mae $\alpha$	0.020967	0.033635	0.037182	0.024590	0.021927	0.048778	0.021668	0.024926	0.022467	0.043855	0.024988
média $\beta$	-0.862520	-1.049505	-0.663469	-0.246591	-0.356435	-0.550476	-0.226987	-0.228193	-0.340314	-0.812398	-0.225541
viés $\beta$	-0.270420	-0.457405	-0.071369	0.345509	0.235665	0.041624	0.365113	0.363907	0.251786	-0.220298	0.366559
mse $\beta$	0.179692	0.424533	0.568831	0.195367	0.187767	0.790453	0.157128	0.136428	0.189744	0.695255	0.146098
mae $\beta$	0.327886	0.546866	0.605406	0.378042	0.392843	0.787275	0.374096	0.368246	0.387664	0.720360	0.372610
média $\sigma^2$	0.008544	0.008212	0.008443	0.011794	0.015133	0.010043	0.016717	0.013590	0.016594	0.008669	0.019304
viés $\sigma^2$	-1.661856	-1.662188	-1.661957	-1.658606	-1.655267	-1.660357	-1.653683	-1.656810	-1.653806	-1.661731	-1.651096
mse $\sigma^2$	2.761939	2.762916	2.762168	2.756341	2.753653	2.758349	2.736745	2.745116	2.756104	2.761587	2.737121
mae $\sigma^2$	1.661856	1.662188	1.661957	1.659888	0.392843	1.660357	1.653683	1.656810	1.659657	1.661731	1.654361

TABELA 21. Monte Carlo - Especificação Incorreta

## 7. CONCLUSÕES

Neste artigo apresentamos métodos semi-paramétricos baseados em Verossimilhança Empírica /Mínimo Contraste Generalizado para a estimação de equações diferenciais estocásticas. Estes estimadores são caracterizados por propriedades de eficiência assintótica de ordem superior, propriedades de otimalidade em testes de hipóteses e robustez à problemas de especificação incorreta para os estimadores baseados em *Exponential Tilting*. Estas propriedades são especialmente importantes neste contexto de estimação de equações diferenciais estocásticas, já que em geral não é possível construir a função de verossimilhança exata do processo devido a não existência de soluções analíticas (e conseqüentemente de discretizações exatas) para equações diferenciais estocásticas. Neste contexto o uso da aproximação não-paramétrica para a densidade

do processo utilizando estes métodos é especialmente vantajosa já que permite avaliar de forma eficiente a densidade do processo, ao mesmo tempo que a especificação paramétrica dada pela equação diferencial estocástica é utilizada através de condições de momentos.

Os resultados obtidos mostram que em especial o estimador de *Exponentially Tilted Empirical Likelihood* proposto por [Schennach, 2007] obtém um desempenho superior as demais técnicas propostas, pelo suas propriedades de robustez na presença de problemas de especificação. Como é possível interpretar a estimação de equações diferenciais estocásticas utilizando dados discretos como um problema de especificação incorreta devido ao uso de uma discretização aproximada do modelo, os resultados dos experimentos de Monte Carlo mostram que o desempenho deste estimador é bastante superior aos demais métodos de estimação que utilizam condições de momentos, e em geral os estimadores baseados em verossimilhança empírica/mínimo contraste generalizado tem melhor desempenho em termos de viés e erro quadrático médio que os estimadores de Métodos de Momentos Generalizados.

Os principais resultados da aplicação empírica estão relacionados aos testes de especificação utilizando as condições de sobre-identificação. Os testes baseados nas estimações por verossimilhança empírica/mínimo contraste generalizado dão suporte a evidência de que modelos de um único fator não são adequados para descrever a dinâmica de modelos de taxas de juros de curto prazo, resultado nem sempre obtido pelo teste J obtido por GMM e relacionado ao viés para baixo existente no teste J quando utilizado em amostras pequenas.

## REFERÊNCIAS

- [Ait-Sahalia, 2002] Ait-Sahalia, Y. 2002. Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach. *Econometrica*, **70**, 223–262.
- [Anatolyev, 2005] Anatolyev, S. 2005. GMM, GEL, Serial Correlation and Asymptotic Bias. *Econometrica*, **73**, 983–1002.
- [Andrews, 1991] Andrews, Donald W K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, **59**, 817–858.
- [Antoine *et al.*, 2007] Antoine, B., Bonnal, H., & Renault, E. 2007. On the efficient use of the informational content of estimating equations: implied probabilities and euclidian empirical likelihood. *Journal Of Econometrics*, **138**, 461–487.
- [Bachelier, 1900] Bachelier, L. 1900. *Theorie de la Speculation*. English translation by A J Boness in *The Random Character of Stock Market Prices*, ed. Paul H Cootner, pg 17–78, Cambridge, Mass, MIT press 1967.
- [Bickel *et al.*, 1993] Bickel, P., Klassen, C., Ritov, Y., & Wellner, J. 1993. *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Press.
- [Bishwal, 2007] Bishwal, J. P. N. 2007. *Parameter Estimation in Stochastic Differential Equations*. Springer.
- [Black & Scholes, 1973] Black, Fisher, & Scholes, Myron S. 1973. The Pricing of options and corporate liabilities. *Journal of Political Economy*, **7**, 637–54.
- [Brennan & Schwartz, 1980] Brennan, M., & Schwartz, E. S. 1980. Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis*, **15**, 907–929.
- [Chan *et al.*, 1992a] Chan, K. G., Karolyi, G., Longstaff, F., & Sanders, A. 1992a. An Empirical Comparison of Alternative Models of Short Term Interest Rate. *Journal of Finance*, **47**, 1209–1297.
- [Chan *et al.*, 1992b] Chan, K. G., Karolyi, G.A., Longstaff, F.A., & Sanders, A. B. 1992b. An Empirical Comparison of Alternative Models of Term Structure of Interest Rates. *Journal of Finance*, **47**, 1209–1227.
- [Chausse, 2009] Chausse, P. 2009. *gmm: Generalized Method of Moments and Generalized Empirical Likelihood*. R package version 1.0-6.
- [Cox, 1975] Cox, J. 1975. *Notes on option pricing I: constant elasticity of variance diffusions*. Working Paper, Stanford University.
- [Cox *et al.*, 1980] Cox, J. C., Ingersoll, J. E., & Ross, S. A. 1980. An analysis of variable rate loan contracts. *Journal of Finance*, **35**, 389–403.
- [Cox *et al.*, 1985] Cox, J. C., Ingersoll, J. E., & Ross, S. A. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica*, **53**, 385–408.
- [DasGupta, 2008] DasGupta, A. 2008. *Asymptotic Theory of Statistics and Probability*. Springer.
- [Delbaen & Schachermayer, 1994] Delbaen, F., & Schachermayer, W. 1994. A General Version of The Fundamental Theory of Asset Pricing. *Mathematische Annalen*, **300**, 463–250.
- [Dothan, 1978] Dothan, U. L. 1978. On the term structure of interest rates. *Journal of Financial Economics*, **6**, 59–59.
- [Gourieroux & Monfort, 1996] Gourieroux, C., & Monfort, A. 1996. *Simulation-Based Econometric Models*. Oxford University Press.
- [Guay & Lamarche, 2008] Guay, A., & Lamarche, J-F. 2008. *The Information Content of Implied Probabilities to Detect Structural Change*. CIRPÉ Working Paper 08-33.
- [Hansen, 1982] Hansen, L. P. 1982. Large sample properties of Generalized Method of Moments estimators. *Econometrica*, **50**(July), 1029–1054.
- [Hansen *et al.*, 1996] Hansen, L. P., Heaton, J., & Yaron, A. 1996. Finite Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics*, **14**, 262–280.
- [Harrison & Kreps, 1979] Harrison, J M, & Kreps, D. 1979. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, **20**, 381–408.

- [Harrison & Pliska, 1981] Harrison, J. M., & Pliska, S. 1981. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, **11**, 215–260.
- [Imbens *et al.*, 1998] Imbens, G. W., Spady, R. H., & Johnson, P. 1998. Information theoretic approaches to inference in moment conditions models. *Econometrica*, **66**, 333–357.
- [Karatzas & Shreve, 1987] Karatzas, I., & Shreve, S. E. 1987. *Brownian Motion and Stochastic Calculus*. Springer-Verlag.
- [Kitamura, 2006] Kitamura, Y. 2006. *Empirical Likelihood Methods in Econometrics: Theory and Practice*. Manuscrito não publicado.
- [Kitamura & Stutzer, 1997] Kitamura, Y., & Stutzer, M. 1997. An Information-Theoretic Alternative to Generalized Method of Moments Estimation. *Econometrica*, **65**(5), 861–874.
- [Kloeden & Platen, 1992] Kloeden, P., & Platen, E. 1992. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag.
- [Litterman & Scheinkman, 1991] Litterman, R., & Scheinkman, J. 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income*, **1**, 54–61.
- [Longstaff & Schwartz, 1992] Longstaff, F., & Schwartz, E. 1992. Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. *Journal of Finance*, **47**, 1259–1282.
- [Lund & Andersen, 1997] Lund, J., & Andersen, T. 1997. Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate. *Journal Of Econometrics*, **77**, 343–377.
- [Merton, 1973] Merton, Robert C. 1973. The theory of rational option pricing. *Bell Journal*, **4**, 141–183.
- [Milstein, 1974] Milstein, G. N. 1974. Approximate integration of stochastic differential equations. *Theory of Probability and Applications*, **19**, 557–562.
- [Newey & McFadden, 1994] Newey, W., & McFadden, D. 1994. *Handbook of Econometrics*. Vol. 4. Elsevier. Chap. Large sample estimation and hypothesis testing.
- [Newey & Smith, 2004] Newey, W., & Smith, R. J. 2004. High-Order properties of GMM and generalized empirical likelihood estimators. *Econometrica*, **72**, 219–255.
- [Newey & West, 1987] Newey, Whitney K., & West, Kenneth D. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703–708.
- [Owen, 1991] Owen, A. 1991. Empirical Likelihood for Linear Models. *The Annals of Statistics*, **18**(1), 1725–1747.
- [Pedersen, 1995] Pedersen, A. R. 1995. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics*.
- [Prakasa Rao, 1999] Prakasa Rao, B. L. S. 1999. *Statistical Inference for Diffusion Type Process*. Arnold.
- [Qin & Lawless, 1994] Qin, J., & Lawless, J. 1994. Empirical Likelihood and general estimating equations. *The Annals of Statistics*, **20**(1), 300–325.
- [Ramalho & Smith, 2005] Ramalho, J. J. S., & Smith, J. 2005. *Goodness of fit tests for moment conditions models*. Documento de trabalho 2005/05.
- [Rogers & Williams, 2000] Rogers, L. C. G., & Williams, D. 2000. *Diffusions, Markov Process and Martingales: Volume 2 Itô Calculus*. Cambridge.
- [Schennach, 2007] Schennach, S. 2007. Point Estimation with Exponentially Tilted Empirical Likelihood. *Annals of Statistics*, **35**(2), 634–672.
- [Smith, 2001] Smith, R. J. 2001. *GEL Criteria for Moment Conditions Models*. Working Paper, University of Bristol.
- [Stambaugh, 1988] Stambaugh, R. 1988. The Information In Forward Rates: Implications For Models Of The Term Structure. *Journal of Financial Economics*, **21**, 41–70.
- [Stanton, 1997] Stanton, R. 1997. A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk. *Journal of Finance*, **52**, 1973–2002.
- [Vasicek, 1977] Vasicek, O. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics*, **5**, 177–88.
- [Zhou, 2000] Zhou, H. 2000. *A study of the finite sample properties of EMM, GMM, QMLE, and MLE for a square-root interest rate diffusion model*. Federal Reserve System Finance and Economics Discussion Series 2000-45.
- [Zivot & Wang, 2006] Zivot, E., & Wang, J. 2006. *Modeling Financial Time Series with S-PLUS, second edition*. Springer-Verlag.