



This is a repository copy of *The time-varying GARCH-in-mean model*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/144634/>

Version: Accepted Version

---

**Article:**

Dias, G.F. (2017) The time-varying GARCH-in-mean model. *Economics Letters*, 157. pp. 129-132. ISSN 0165-1765

<https://doi.org/10.1016/j.econlet.2017.06.005>

---

© 2017 Elsevier. This is an author-produced version of a paper subsequently published in *Economics Letters*. Uploaded in accordance with the publisher's self-archiving policy. Article available under the terms of the CC-BY-NC-ND licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# The time-varying GARCH-in-mean model

Gustavo Fruet Dias<sup>a,1</sup>

<sup>a</sup>*CREATES and Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark.  
E-mail address: gdias@econ.au.dk*

---

## Abstract

1 I propose an estimation strategy for the stochastic time-varying risk premium  
2 parameter in the context of a time-varying GARCH-in-mean (TVGARCH-in-  
3 mean) model. A Monte Carlo study shows that the proposed algorithm has good  
4 finite sample properties. Using monthly excess returns on the CRSP index, I  
5 document that the risk premium parameter is indeed time-varying and shows  
6 high degree of persistence.

JEL classification numbers: C13, C15, C22, G12

*Keywords:* risk-return tradeoff, time-varying coefficients, iterative estimators,  
GARCH-type models

---

---

<sup>1</sup>I thank George Kapetanios, Cristina Scherrer and an anonymous referee for constructive suggestions. I acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

7 **1. Introduction**

8 Asset pricing theories suggest that riskier assets should demand higher ex-  
 9 pected returns. Using Merton’s (1973) theoretical framework, the conditional  
 10 expectation of the market excess returns reads

$$\mathbb{E}(r_{t+1}^m | \mathcal{F}_t) - r_t^f = \lambda_t \text{Var}(r_{t+1}^m | \mathcal{F}_t), \quad (1)$$

11 where  $r_{t+1}^m$  and  $r_t^f$  are the returns on the market portfolio and risk-free asset,  
 12  $\mathcal{F}_t$  is the market-wide information available at time  $t$ , and  $\lambda_t$  is the coefficient  
 13 of relative risk aversion defined as the elasticity of marginal value with respect  
 14 to wealth. Most studies assume the risk-return trade-off is constant over time  
 15 and linear in the variance, which is usually associated with the reasons behind  
 16 mixed empirical evidences when estimating the risk-return trade-off (Linton  
 17 and Perron (2003), Brandt and Wang (2010), Christensen, Dahl, and Iglesias  
 18 (2012), among others). To address this issue, I adopt the time-varying GARCH-  
 19 in-mean (TVGARCH-in-mean) model in the spirit of Anyfantaki and Demos  
 20 (2016) which allows  $\lambda_t$  to be a time-varying stochastic process and put forward  
 21 a feasible estimation strategy for  $\lambda_t$  (see references in Anyfantaki and Demos  
 22 (2016) for variants of the TVGARCH-in-mean models). Specifically, I com-  
 23 bine Giraitis, Kapetanios, and Yates’s (2013) time-varying kernel least squares  
 24 estimator with Linton and Perron’s (2003) semiparametric iterative approach  
 25 to estimate the time-varying risk premium coefficient. A Monte Carlo study  
 26 shows that the proposed algorithm has good finite sample properties. Using the  
 27 excess returns of the Center for Research on Security Prices (CRSP) index, I  
 28 document that the risk premium parameter is indeed time-varying, alternating  
 29 positive (statistically significant) and nonsignificant values over time.

30 **2. The time-varying GARCH-in-mean**

31 The generic TVGARCH-in-mean(p,q) is defined as:

$$r_t = \lambda_t \sigma_t + \epsilon_t, \quad (2)$$

$$\epsilon_t = \sigma_t \eta_t, \quad (3)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2, \quad (4)$$

$$\epsilon_t^2 = \psi_0 + u_t + \sum_{i=1}^{\infty} \psi_i u_{t-i}, \quad (5)$$

32 where  $\eta_t$  is an independent and identically distributed (*iid*) zero mean process  
 33 with unit variance;  $\sigma_t$  is a latent conditional standard deviation; (5) is the  
 34  $MA(\infty)$  representation of the conditional variance equation;  $u_t = \epsilon_t^2 - \sigma_t^2$  is a  
 35 martingale difference sequence process;  $\phi = (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)'$  collects the  
 36 free parameters in (4); and  $\psi_i := \varrho_i(\phi)$   $i = 1, 2, \dots$  are deterministic functions

37 of the elements in  $\phi$ . Similarly as in Giraitis, Kapetanios, and Yates (2013),  
 38 the time-varying risk premium parameters are assumed to evolve smoothly over  
 39 time, so that it satisfies a local stability condition in the form of  $\sup_{s: \|s-t\| \leq h} \|\lambda_t + \lambda_s\|_2 = O_p(h/t)$ .  
 40

41 Estimating the free parameters in (2) and (4) by maximum-likelihood is not  
 42 a feasible alternative, as the class of TVGARCH-in-mean(p,q) models involves  
 43 two unobserved processes:  $\lambda_t$  and  $\epsilon_t$ . Anyfantaki and Demos (2016) address  
 44 this issue in the context of the time-varying EGARCH(1,1)-in-mean model.  
 45 Specifically, their work differs from mine in two ways. First, they parameterize  
 46 the conditional variance as an EGARCH(1,1) model and, most importantly,  $\lambda_t$   
 47 as a stationary AR(1) process. By contrast,  $\lambda_t$  in (2) is assumed to satisfy  
 48  $\sup_{s: \|s-t\| \leq h} \|\lambda_t + \lambda_s\|_2 = O_p(h/t)$ , which encompasses the case of the driftless  
 49 random walk process considered in Chou, Engle, and Kane (1992). Second,  
 50 while I propose a kernel-based nonparametric method to estimate the time-  
 51 varying risk premium parameter, Anyfantaki and Demos's (2016) estimation  
 52 strategy is based on Bayesian methods (Markov chain Monte Carlo (MCMC)  
 53 likelihood based estimation procedure).

54 I combine Linton and Perron's (2003) iterative semiparametric estimator  
 55 with Giraitis, Kapetanios, and Yates's (2013) kernel-based least squares frame-  
 56 work to estimate the free parameters  $\theta = (\underline{\lambda}, \phi)'$ , where  $\underline{\lambda} = (\lambda_1, \dots, \lambda_T)'$ . This  
 57 method consists of recursively updating estimates of  $\sigma_t$  and  $u_t$  on each iteration,  
 58 and then computing estimates of  $\underline{\lambda}$  and  $\phi$ . To this end, consider moment  
 59 conditions based on (2) and (5),

$$\mathbb{E}[\sigma_t (r_t - \lambda_t \sigma_t)] = 0, \quad \text{for each } t = 1, 2, \dots, T, \quad (6)$$

$$\mathbb{E}[z_t u_t] = 0, \quad \text{with } z_t := \frac{\partial \left( \psi_0 + \sum_{i=1}^{\bar{q}} \psi_i u_{t-i} \right)}{\partial \phi}, \quad (7)$$

60 where (7) is truncated at some lag-order  $\bar{q}$  with  $\bar{q} > p + q + 1$ . Notably, (7) holds  
 61 because  $u_t$  is a martingale difference sequence and  $z_t$  is a function of lagged  
 62 values of  $u_t$ . It follows that estimating  $\theta$  by the standard generalized method of  
 63 moments (GMM) using the moments defined in (6) and (7) is not operational,  
 64 as  $z_t$  and  $\sigma_t$  are latent variables. Using Linton and Perron's (2003) approach,  
 65 rewrite (6) and (7) using estimates of  $\sigma_t$  and  $u_t$  obtained at some  $j$  iteration,

$$\mathbb{E}[\sigma_{j,t} (r_t - \lambda_{j+1,t} \sigma_{j,t})] = 0, \quad \text{for each } t = 1, 2, \dots, T, \quad (8)$$

$$\mathbb{E}[z_{j,t} u_{j+1,t}] = 0, \quad (9)$$

66 where  $\sigma_{j,t}$  and  $z_{j,t}$  denote the filtered estimates of  $\sigma_t$  and  $z_t$  based on  $\widehat{\theta}_j$ , and  
 67  $u_{j+1,t} = \epsilon_{j,t}^2 - \psi_{j+1,0} - \sum_{i=1}^{\bar{q}} \psi_{j+1,i} u_{j,t-i}$  with  $\epsilon_{j,t}^2 = (r_t - \lambda_{j+1,t} \sigma_{j,t})^2$ . While the  
 68 finite sample counterpart of (9) is given by the usual sample mean, computing  
 69 the sample counterpart of (8) is less obvious. The work of Giraitis, Kapetanios,  
 70 and Yates (2013) suggests the use of local kernels to construct operational

71 sample counterparts of (8). In turn, a feasible moment condition based on (8)  
 72 reads

$$K_t^{-1} \sum_{\tau=1}^T k_{t,\tau} \sigma_{j,\tau} \left( r_\tau - \widehat{\lambda}_{j+1,t} \sigma_{j,\tau} \right) = 0, \quad \text{for each } t = 1, 2, \dots, T, \quad (10)$$

73 where  $k_{t,\tau} = K((t-\tau)/H)$  denotes a kernel function such that  $K(x) \geq 0$  for  
 74 any  $x \in \mathbb{R}$  is a continuous bounded function with a bounded first derivative and  
 75  $\int K(x) dx = 1$ ;  $H$  is the bandwidth parameter satisfying  $H = o(T/\ln(T))$  as  
 76  $H \rightarrow \infty$ ; and  $K_t = \sum_{\tau=1}^T k_{t,\tau}$ . Notably, writing the moment conditions as in  
 77 (10) is consistent with previous studies in the time-varying parameter literature  
 78 which maximizes kernel weighted log-likelihood functions (see Robinson (1989),  
 79 Giraitis, Kapetanios, Wetherilt, and Žikeš (2016), among others).

80 I use the fact that (10) is exactly identified for each  $t$ , and hence estimates  
 81 of  $\lambda_t$  can be obtained independently of  $\phi$ . In turn, estimates of  $\theta$  are computed  
 82 iteratively by a two-step procedure. The first step consists of solving (10) for  
 83 each  $t$ , while the second step mimics the work of Linton and Perron (2003) and  
 84 consists of estimating  $\phi$  using the sample counterpart of (9). In practise, the  
 85 kernel-based iterative estimator is as follows:

86 **Step 1:** Choose starting values  $\widehat{\lambda}_0$  and  $\widehat{\phi}_0$ , such that  $\widehat{\phi}_0$  satisfies the second-  
 87 order stationarity conditions of the GARCH(1,1) model. Using  $\widehat{\theta}_{0,t} =$   
 88  $(\widehat{\lambda}_0, \widehat{\phi}_0)'$ , compute recursively  $\{\sigma_{0,t}^2\}_{t=1}^T$ , and  $\{u_{0,t}\}_{t=1}^T$  from (2)-(5).

89 **Step 2:** Given  $\{\sigma_{0,t}^2\}_{t=1}^T$ , calculate

$$\widehat{\lambda}_{1,t} = \left( \sum_{\tau=1}^T k_{t,\tau} \sigma_{0,\tau}^2 \right)^{-1} \sum_{\tau=1}^T k_{t,\tau} \sigma_{0,\tau} r_\tau, \quad \text{for each } t = 1, 2, \dots, T. \quad (11)$$

90 **Step 3:** Solving the sample counterpart of (9) is equivalent to estimate  $\widehat{\phi}_1$  by  
 91 nonlinear least squares. Calculate

$$\widehat{\phi}_1 = \arg \min_{\widehat{\phi}_1} \sum_{t=1}^T \left\{ \left( r_t - \widehat{\lambda}_{1,t} \sigma_{0,t} \right)^2 - \widehat{\psi}_{1,0} - \sum_{i=0}^{\bar{q}} \widehat{\psi}_{1,i} u_{0,t-1-i} \right\}^2. \quad (12)$$

92 **Step 4:** Update recursively  $\{\sigma_{1,t}^2\}_{t=1}^T$  and  $\{u_{1,t}\}_{t=1}^T$  based on  $\widehat{\theta}_1$ .

93 Repeat steps 2-4  $j$  times until  $\widehat{\theta}_j$  converges. Convergence occurs when  
 94  $\|\widehat{\lambda}_j - \widehat{\lambda}_{j-1}\|_2 \leq \varepsilon$  and  $\|\widehat{\phi}_j - \widehat{\phi}_{j-1}\|_2 \leq \varepsilon$ , with  $\varepsilon$  set to  $10^{-5}$ . Parameters on

95 the  $j^{\text{th}}$  iteration are given by:

$$\widehat{\lambda}_{j,t} = \left[ \sum_{\tau=1}^T k_{t,\tau} \sigma_{j-1,\tau}^2 \right]^{-1} \sum_{\tau=1}^T k_{t,\tau} \sigma_{j-1,\tau} r_{\tau}, \quad \text{for each } t = 1, 2, \dots, T, \quad (13)$$

$$\widehat{\phi}_j = \arg \min_{\widehat{\phi}_j} \sum_{t=1}^T \left[ r_t - \widehat{\lambda}_{j,t} \sigma_{j-1,t} \right]^2 - \widehat{\psi}_{j,0} - \sum_{i=0}^{\bar{q}} \widehat{\psi}_{j,i} u_{j-1,t-1-i} \Big]^2. \quad (14)$$

96 Finally, three inputs are still necessary to implement the above algorithm: the  
 97 kernel function, the bandwidth parameter  $H$ , and the truncation lag  $\bar{q}$ . As in  
 98 Giraitis, Kapetanios, and Yates (2013), three kernel functions are used: the  
 99 Epanechnikov, Gaussian, and flat kernels. The choice of  $H$  follows from the  
 100 Monte Carlo study conducted in Section 3.1, while  $\bar{q}$  is chosen to be proportional  
 101 to  $\ln(T)$ , (Dufour and Jouini (2005)).

102 Asymptotic theory for the Quasi-Maximum Likelihood (QMLE) estimator in  
 103 the GARCH-in-mean models is yet to be fully established. Conrad and Mammen  
 104 (2016) give an important step forward and prove the asymptotic distribution  
 105 of the QMLE estimator for the simple GARCH(1,1)-in-mean. As discussed in  
 106 Linton and Perron (2003), the semiparametric GARCH-in-mean models offer  
 107 additional complications compared to the standard GARCH-in-mean models,  
 108 and hence rigorous inference is still not available. Similar difficulties arise in  
 109 the TVGARCH-in-mean specification. In turn, this note follows Linton and  
 110 Perron's (2003) approach as it briefly discusses the general conditions required  
 111 for consistency and asymptotic normality; uses the wild bootstrap to conduct  
 112 inference; and adopts a Monte Carlo study to assess the finite sample properties  
 113 of the proposed iterative estimator.

114 The concept of asymptotic contraction mapping (ACM) developed in Do-  
 115 minitz and Sherman (2005) is useful to guide the discussion on the asymptotic  
 116 properties of the kernel iterative estimator. If a collection is an ACM, then it  
 117 will have a unique fixed point that depends on the sample characteristics and  
 118 hence the iterative procedure converges.<sup>2</sup> While the two-step procedure given  
 119 in (13) and (14) is seen as the sample mapping, (8) and (9) are their popula-  
 120 tion counterpart. Consistency and asymptotic normality require the population  
 121 mapping to be an ACM, which implies, under some uniform convergence condi-  
 122 tions, that the sample mapping is also an ACM and hence has an unique fixed  
 123 point (regardless of the initial values). Combining Theorem 4 in Dominitz and  
 124 Sherman (2005) with Giraitis, Kapetanios, and Yates's (2013) results, estimates  
 125 of  $\lambda_t$  are expected to be  $\sqrt{H}$  consistent and asymptotically normally distributed,  
 126 and estimates of  $\phi$  are expected to be consistent and asymptotically normally  
 127 distributed at the usual  $\sqrt{T}$  rate.

---

<sup>2</sup>See Dominitz and Sherman, 2005, p. 841 for a formal definition of the ACM.

128 **3. Numerical illustrations**

129 *3.1. Monte Carlo*

130 I simulate data from (2)-(4) where  $p = q = 1$ ,  $\eta_t$  is normally distributed with  
131 zero mean and variance equal to one, and  $\lambda_t$  follows a bounded random walk  
132 process (see detailed discussion in the online Supplement). The sample size and  
133 the number of replications are set to 2,000 and 1,000, respectively.

134 Table 1 displays the results for the kernel-based iterative estimator computed  
135 with alternative bandwidth choices and the Epanechnikov, Gaussian, and flat  
136 kernels. Results are reported in terms of the root mean squared error (RMSE)  
137 and pointwise correlation between  $\lambda_t$  and the kernel-based estimates. The best  
138 choices of bandwidth parameters, in terms of minimizing the RMSE, are  $H =$   
139  $T^{1/2}$  and  $H = T^{6/10}$ . These are also the bandwidths that deliver the highest  
140 pointwise correlation (about 0.85) between the kernel-based estimates and the  
141 true latent time-varying risk premium parameter. All combinations of kernel  
142 methods and bandwidth parameters deliver unbiased estimates of  $\phi = (\omega, \alpha, \beta)'$   
143 (apart from  $H = T^{2/10}$ ). Finally, convergence rates are greater than 98% for all  
144 specifications, suggesting that (8) and (9) are ACMs.

145 *3.2. Empirical results*

146 I use excess returns of the CRSP value-weighted index aggregated on a  
147 monthly basis. Figure 1 plots monthly estimates of  $\lambda_t$  and their 90% wild  
148 bootstrap confidence bands from a TVGARCH(1,1)-in-mean model with band-  
149 width  $H = T^{6/10}$  (see the Supplement material for a Monte Carlo study showing  
150 that the wild bootstrap produces valid inference). Not surprisingly, likewise the  
151 semiparametric GARCH-in-mean models, the empirical confidence bands are  
152 relatively wide, which reflects the difficulties associated with estimating the risk-  
153 return trade-off (Linton and Perron (2003) and Christensen, Dahl, and Iglesias  
154 (2012)). I find that the risk premium parameter is indeed time-varying, with  $\hat{\lambda}_t$   
155 assuming both positive (generally significant) and negative (insignificant) val-  
156 ues. This finding sheds light on the mixed evidence on the risk-return literature  
157 regarding the sign and significance of the risk premium parameter. Addition-  
158 ally, in periods where  $\hat{\lambda}_t$  is statistically significant, market volatility is low. On  
159 contrary, when  $\hat{\lambda}_t$  is not statistically significant, market volatility is high. This  
160 indicates that identification of the risk premium parameter is problematic in  
161 periods of high volatility (Rossi and Timmermann (2010)).

162 **4. Conclusion**

163 I introduce a kernel-based iterative estimator that combines the estimators in  
164 Giraitis, Kapetanios, and Yates (2013) and Linton and Perron (2003) to estimate  
165 the stochastic time-varying risk premium parameter in the TVGARCH(1,1)-in-  
166 mean model. The Monte Carlo study shows that the kernel-based estimator  
167 presents a good finite sample performance. I investigate the time-varying risk  
168 premium for the CRSP index and find strong evidence that  $\lambda_t$  is indeed time-  
169 varying.

- 170 ANYFANTAKI, S., AND A. DEMOS (2016): “Estimation and Properties of a  
171 Time-Varying EGARCH(1,1) in Mean Model,” *Econometric Reviews*, 35(2),  
172 293–310.
- 173 BRANDT, M. W., AND L. WANG (2010): “Measuring the Time-Varying Risk-  
174 Return Relation from the Cross-Section of Equity Returns,” Discussion paper,  
175 Duke University.
- 176 CHOU, R., R. F. ENGLE, AND A. KANE (1992): “Measuring risk aversion from  
177 excess returns on a stock index,” *Journal of Econometrics*, 52(1), 201 – 224.
- 178 CHRISTENSEN, B. J., C. M. DAHL, AND E. M. IGLESIAS (2012): “Semipara-  
179 metric inference in a GARCH-in-mean model,” *Journal of Econometrics*, 167,  
180 458–472.
- 181 CONRAD, C., AND E. MAMMEN (2016): “Asymptotics for parametric GARCH-  
182 in-Mean models,” *Journal of Econometrics*, 194(2), 319 – 329, Financial  
183 Statistics and Risk Management.
- 184 DOMINITZ, J., AND R. SHERMAN (2005): “Some Convergence Theory for Iter-  
185 ative Estimation Procedures with an Application to Semiparametric Estima-  
186 tion,” *Econometric Theory*, 21, 838–863.
- 187 DUFOUR, J.-M., AND T. JOUINI (2005): “Asymptotic Distribution of a Simple  
188 Linear Estimator for Varma Models in Echelon Form,” in *Statistical Mod-  
189 eling and Analysis for Complex Data Problems*, ed. by P. Duchesne, and  
190 B. Rémillard, pp. 209–240. Springer US.
- 191 GIRAITIS, L., G. KAPETANIOS, A. WETHERILT, AND F. ŽIKEŠ (2016): “Es-  
192 timating the Dynamics and Persistence of Financial Networks, with an Ap-  
193 plication to the Sterling Money Market,” *Journal of Applied Econometrics*,  
194 31(1), 58–84, jae.2457.
- 195 GIRAITIS, L., G. KAPETANIOS, AND T. YATES (2013): “Inference on stochastic  
196 time-varying coefficient models,” *Journal of Econometrics*.
- 197 LINTON, O., AND B. PERRON (2003): “The Shape of the Risk Premium: Ev-  
198 idence From a Semiparametric Generalised Autoregressive Conditional Het-  
199 eroscedasticity Model,” *Journal of Business & Economic Statistics*, 21, 354–  
200 367.
- 201 MERTON, R. C. (1973): “An Intertemporal Capital Asset Pricing Model,”  
202 *Econometrica*, 41(5), 867–887.
- 203 ROBINSON, P. (1989): “Nonparametric Estimation of Time-Varying Parame-  
204 ters,” in *Statistics Analysis and Forecasting of Economic Structural Change*,  
205 ed. by P. Hackl, chap. 15, pp. 253–264. Springer, Berlin.
- 206 ROSSI, A., AND A. TIMMERMANN (2010): “What is the Shape of the Risk-  
207 Return Relation?,” Discussion paper, The Rady School of Management, Uni-  
208 versity of California, San Diego.

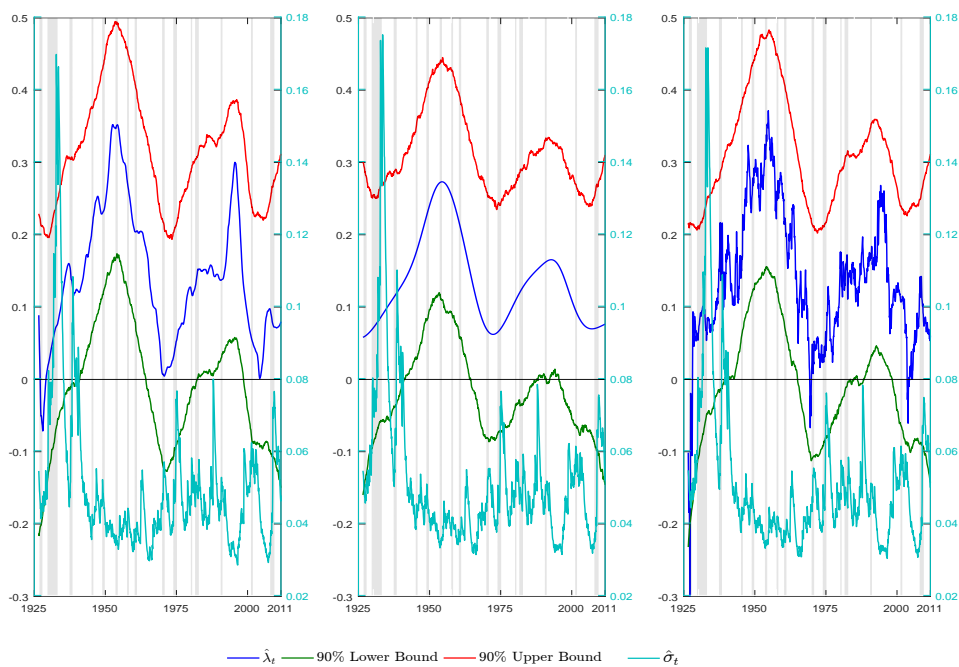


Table 1: Finite sample performance of the kernel-based iterative estimator

Bandwidth - H	Kernel	$\lambda_t$		RMSE			Mean		
		RMSE	Corr	$\omega$	$\alpha$	$\beta$	$\omega$	$\alpha$	$\beta$
$T^{2/10}$	Epanechnikov	0.39	0.60	0.01	0.03	0.10	0.02	0.07	0.84
$T^{2/10}$	Gaussian	0.25	0.71	0.01	0.02	0.06	0.01	0.05	0.88
$T^{2/10}$	Flat	0.36	0.61	0.01	0.02	0.07	0.01	0.06	0.86
$T^{3/10}$	Epanechnikov	0.26	0.72	0.01	0.02	0.05	0.01	0.05	0.88
$T^{3/10}$	Gaussian	0.18	0.81	0.01	0.02	0.05	0.01	0.05	0.89
$T^{3/10}$	Flat	0.24	0.73	0.01	0.02	0.05	0.01	0.05	0.89
$T^{4/10}$	Epanechnikov	0.18	0.80	0.01	0.02	0.05	0.01	0.05	0.90
$T^{4/10}$	Gaussian	0.13	0.86	0.01	0.02	0.06	0.01	0.05	0.90
$T^{4/10}$	Flat	0.17	0.81	0.01	0.02	0.05	0.01	0.05	0.90
$T^{5/10}$	Epanechnikov	0.13	0.86	0.01	0.02	0.05	0.01	0.05	0.90
$T^{5/10}$	Gaussian	0.11	0.88	0.01	0.02	0.05	0.01	0.05	0.90
$T^{5/10}$	Flat	0.13	0.86	0.01	0.02	0.05	0.01	0.05	0.90
$T^{6/10}$	Epanechnikov	0.12	0.88	0.01	0.02	0.06	0.01	0.05	0.89
$T^{6/10}$	Gaussian	0.13	0.86	0.01	0.02	0.06	0.01	0.05	0.89
$T^{6/10}$	Flat	0.12	0.86	0.01	0.02	0.06	0.01	0.05	0.89
$T^{7/10}$	Epanechnikov	0.13	0.85	0.01	0.02	0.05	0.01	0.05	0.90
$T^{7/10}$	Gaussian	0.16	0.79	0.01	0.02	0.05	0.01	0.05	0.90
$T^{7/10}$	Flat	0.15	0.81	0.01	0.02	0.05	0.01	0.05	0.90
$T^{8/10}$	Epanechnikov	0.17	0.77	0.01	0.02	0.05	0.01	0.05	0.89
$T^{8/10}$	Gaussian	0.21	0.68	0.01	0.02	0.05	0.01	0.05	0.89
$T^{8/10}$	Flat	0.19	0.67	0.01	0.02	0.05	0.01	0.05	0.89

RMSE accounts for root mean squared error; Corr is the pointwise correlation between  $\lambda_t$  and  $\widehat{\lambda}_t$ ;  $\lambda_t$  is defined as a bounded random walk process with upper and lower bounds given by 0.90 and -0.90, respectively; and  $\phi = (0.01, 0.05, 0.9)'$ .

Figure 1: Time-varying risk premium estimation and conditional standard deviation



The left-hand-side, center and right-hand-side graphs display estimates of  $\lambda_t$  with the Epanechnikov, Gaussian, and the flat kernel functions, respectively. Estimates of  $\lambda_t$  and the 90% confidence intervals are on the left axis. The bandwidth parameter is equal to  $H = T^{6/10}$ . The conditional standard deviation is in light blue on the right axis. I perform 1000 replications in the bootstrap algorithm. The shaded areas account for the recession periods from the National Bureau of Economic Research (NBER).