How risky is consumption in the long-run? Benchmark estimates from a robust estimator*

Ian Dew-Becker

Northwestern University

July 16, 2015

Abstract

The long-run standard deviation of consumption growth is a key moment in determining risk premia when agents have Epstein–Zin preferences. This paper studies a new estimator of the long-run standard deviation shown to provide a superior bias/variance trade-off and better confidence interval coverage than previous methods. In the postwar period the long-run standard deviation of consumption growth is estimated to be 2.5 percent per year with an upper bound to the 95-percent confidence interval of 4.9 percent. The analogous values in the longest available sample are 4 and 5.6 percent. These values can be taken as benchmarks for future calibrations.

1 Introduction

The goal of endowment- and production-economy asset pricing is to find a process for the stochastic discount factor that is consistent with both asset prices and the observed consumption process (or the dynamics of other variables determining state prices, e.g. leisure or

*Email: ian.dew-becker@kellogg.northwestern.edu. This paper was previously circulated under the title, "Estimates of the volatility of the permanent component of consumption and their implications for asset pricing." I appreciate helpful comments and discussions from John Campbell, Stefano Giglio, Cam Harvey, and Lars Lochstoer, and participants at the SFS Cavalcade.

a reference point of some sort). In representative-agent endowment-economy models, there are two degrees of freedom for determining the pricing kernel: preferences and the endowment process. This paper measures the volatility of the endowment process to help guide future calibrations.

Epstein–Zin (1991) preferences have recently become dominant in preference-based asset pricing models. Under Epstein–Zin preferences, risk premia are driven by the behavior of consumption growth at the very lowest frequencies (see Dew-Becker and Giglio, 2013, for an extensive discussion). But while it is widely understood that it is important to match the persistence in consumption growth when calibrating models, there are many ways to measure persistence. Should we ask that models match quarterly or annual autocorrelations? Should they match just the first autocorrelation or a number of longer-term autocorrelations?

Hansen and Jagannathan (1991) show that the maximal Sharpe ratio available in the economy is equal to the standard deviation of the pricing kernel divided by the gross riskless rate (which is nearly equal to 1). For a consumer with Epstein–Zin preferences and log-normal, homoskedastic consumption growth, the standard deviation of the stochastic discount factor is approximately

$$std\left(M_{t+1}\right) \approx std\left(\rho \Delta E_{t+1} \Delta c_{t+1} + (\alpha - \rho) \Delta E_{t+1} \sum_{j=0}^{\infty} \theta^{j} \Delta c_{t+1+j}\right) \tag{1}$$

where c_t is log consumption, ρ is the inverse elasticity of intertemporal substitution (EIS), α is the coefficient of relative risk aversion, and θ is a parameter slightly less than 1 that comes from a log-linearization.[1] E_t is the expectation operator conditional on information available on date t, Δ is the first-difference operator, and ΔE_{t+1} denotes the change in expectations.[2] In many recent calibrations of macro and finance models, the inverse EIS, ρ, is small relative to risk aversion, α, and so the long-run component drives the variance of the SDF, thus

[1] θ is approximately equal to the time discount factor in the agent's utility function. The approximation is derived by Restoy and Weil (1998). Hansen, Heaton, and Li (2008) derive a similar result.

[2] Equation (1) is exact when $\rho = 1$.

determining the Hansen–Jagannathan (1991) bound. In the limiting case where $\rho \to 0$ and $\theta \to 1$,

$$std\,(M_{t+1}) \to \alpha \times std\left(\Delta E_{t+1}\sum_{j=0}^{\infty}\Delta c_{t+1+j}\right) \tag{2}$$

The standard deviation term in (2) is often referred to as the long-run standard deviation (LRSD) of consumption growth. We can think of (2) as the case where households are completely indifferent to when consumption occurs, since the EIS is infinite and the pure rate of time preference approaches zero. The standard deviation of the SDF is equal to risk aversion – the price of risk – multiplied by the long-run standard deviation of consumption growth – the quantity of risk.

While a number of papers have discussed the behavior of the first few autocorrelations of consumption growth (e.g. Bansal, Kiku, and Yaron, 2012; Beeler and Campbell, 2012), the long-run standard deviation in (2) depends on the sum of every autocorrelation. And in fact, studying only the first handful of autocorrelations can lead to highly misleading results. For example, an MA(4) process with the same four quarterly autocovariances as the ARMA(1,1) process used by Bansal and Yaron (2004) would yield a pricing kernel with a standard deviation 33 percent smaller than what is obtained under their actual ARMA(1,1) process.

The consumption-based asset pricing literature uses an extraordinarily broad range of calibrations for the LRSD. Table 1 summarizes some representative publications. The calibrated values of the LRSD range from 1.5 to 9 percent per year. To obtain a Sharpe ratio of 30 percent per year (consistent with the post-war equity premium), the lowest LRSD's would require risk aversion of 20, while the highest require risk aversion of only 3.3. This paper asks what a good benchmark value and confidence interval is to help understand how reasonable these calibrations are and to guide future work.

There is a large literature on the estimation of the LRSD in both parametric and non-parametric settings. In this paper I restrict my attention to non-parametric methods both because they are generally robust to heteroskedasticity and non-normality in ways that parametric methods often are not, and also because they take seriously the fact that *we do not know the true dynamic process driving consumption growth.*

The LRSD is the square root of the spectral density at frequency zero, so non-parametric methods use low frequency information from the sample spectrum to estimate the LRSD, while leaving the higher-frequency behavior of consumption growth unconstrained. A good estimator of the LRSD should be accurate for a wide range of specifications of consumption growth. The long-run risks literature following Bansal and Yaron (2004) argues that there may be a small and highly persistent component in consumption growth. That type of process leads to poor small-sample properties for many spectral density estimators because it induces a large and narrow peak in the spectrum. The goal of this paper is not to estimate the long-run risk model, but it will be frequently used as a benchmark test of the accuracy of the estimators. I show that standard methods (e.g. Newey and West, 1987) deliver estimates in this setting that have large mean squared errors and misleading confidence intervals with poor coverage of the true value of the LRSD.

While there are recent estimators in the literature (Kiefer and Vogelsang, 2005; Müller, 2007) that are designed to be robust to strongly peaked spectra as in the long-run risks model, they tend to have high variances in their most robust form. I therefore use an alternative estimator that I show has superior performance to other recent proposals across a wide range of specifications for consumption growth, including when there is a highly persistent component.

It is well known that the quadratic spectral (QS) kernel gives the smallest mean squared error among all non-negative spectral density estimators. However, in the presence of a large peak in the spectrum, non-negative estimators tend to be biased downward and their confidence intervals have poor coverage. The estimator I propose in this paper is similar to the QS kernel, but it dispenses with the non-negativity constraint, which helps reduce bias. I therefore refer to it as the reduced-bias quadratic spectral (RQS) kernel.

Both asymptotically and in small samples, the RQS estimator improves the bias/variance trade-off compared to the QS kernel and other recent estimators. In small-sample simulations models with long-run risks, long memory, and jumps in consumption growth, I show that the RQS estimator has better confidence interval coverage than the other robust estimators that have been proposed recently.

The next section of the paper estimates the LRSD of consumption growth in a range of data samples. In the post-war sample, the LRSD is estimated to be 2.5 percent per year, which is on the lower end of the calibrations reported in Table 1. When we move to annual data going back to 1929, the estimate of the LRSD rises to 3.5 percent per year. Barro and Ursua's (2010) annual series on consumption growth since 1834 delivers an LRSD of 4.1 percent. While an annual LRSD of 4.1 percent is substantially higher than what we observe in the post-war data, it is still smaller than every calibration in the long-run risks literature reviewed in Table 1, and less than half the magnitude of the most volatile calibrations. The upper end of the confidence interval when we use the Barro–Ursua data is 5.6 percent, while the upper limit with post-war data is 4.9 percent per year. An upper bound of 5.6 percent for the LRSD captures more recent calibrations (e.g. Bansal, Kiku, and Yaron, 2012), but rejects the original long-run risks calibration.

Given the large changes in the estimated LRSD in the sample, it is impossible to make a single conclusive statement about how models should be calibrated. Production models are often designed to match the post-war economy. For example, many assume that the economy is not on a gold standard, and when price stickiness is included it is usually calibrated based on data from the past few decades. The behavior of interest rates and inflation, in particular, differ substantially between pre- and post-war data. For a model that is calibrated to match recent features of the data, it is natural to use the post-war estimates of the LRSD as a benchmark. On the other hand, some studies are meant to capture longer-run shifts in the structure of the economy. For example, endogenous growth models featuring large structural shifts, or models yielding events like the Industrial Revolution, would sensibly want to match the estimates of the LRSD obtained from the Barro–Ursua data.

Volatility clearly varies substantially over time. But the approximation for the Hansen–Jagannathan bound in (2) holds with homoskedastic consumption growth, and the majority of the models in the literature (e.g. summarized in table 1) are homoskedastic. So how does heteroskedasticity affect the analysis? First, I show that the approximation (2) has similar accuracy in Bansal and Yaron's (2004) calibration regardless of whether we include their heteroskedasticity or not.[3] In Bansal, Kiku, and Yaron's (2012) calibration, stochastic volatility has much larger effects, but the LRSD is still important for determining how risk prices vary across calibrations – variation in the LRSD induces similarly sized movements in the Hansen–Jagannathan bound regardless of whether volatility is constant. Moreover, the non-parametric estimators I study are fully robust to heteroskedasticity and non-normality.

Nevertheless, it is important to note here that the LRSD is certainly not the *only* feature of the consumption process that determines the standard deviation of the pricing kernel. Substantial heteroskedasticity, as in Bansal, Kiku, and Yaron (2012), or time-variation in disaster risk can certainly also affect the volatility of the pricing kernel, in which case equation (2) does not hold exactly. In any case, though, the LRSD is a key moment, even if it is not always the only feature of the world that matters.

The remainder of the paper is organized as follows. Section 2 reviews recent calibrations of the LRSD. Sections 3 and 4 introduce estimators of the LRSD and study their performance in the long-run risks calibration and section 5 estimates the LRSD empirically. Finally, section 6 discusses implications of the estimates of the LRSD for risk premia and the effects of stochastic volatility, and section 7 concludes.

[3] Bansal and Yaron (2004) do not report the Hansen–Jagannathan bound, which is the focus of this paper. Rather, they report an equity premium in their model. When they include stochastic volatility, the equity premium rises substantially. They obtain that result because the addition of stochastic volatility increases the correlation between market returns and the pricing kernel. Specifically, there are three shocks in the model – short-term consumption growth, persistent consumption growth, and volatility. Dividends are only driven by the persistent component and volatility. When the volatility shock is non-zero, then, the volatility of equity returns rises and their correlation with the SDF rises. But the overall variance of the SDF, and hence the maximal Sharpe ratio, is only changed by roughly 10 percent.

2 Recent calibrations

The calibrations of consumption processes in the literature can be divided into three categories. At the low end are the calibrations based on the annual standard deviation of post-war consumption growth. These papers generally assume that consumption or technology follows a random walk and that the annual standard deviation of the permanent innovations is roughly 2 percent.[4] In a middle range is a set of papers that also treat consumption as following a random walk but that are calibrated based on longer time series. Mehra and Prescott (1985), for example, calibrate a two-state Markov chain for consumption to match the empirical annual variance and autocorrelation of US consumption growth since the 19th century. At 3.16, their long-run standard deviation is higher by half than the values used in papers calibrated to match post-war data.

Finally, Table 1 lists a number of papers studying long-run risks, both in production and endowment economies. With a standard deviation of 5.54, Bansal, Kiku, and Yaron (BKY; 2012) are on the lower end, while Bansal and Yaron (BY; 2004) choose a value of 6.28, and Kaltenbrunner and Lochstoer (KL; 2010) are at the upper end with a standard deviation of 8.22.[5] The highest calibration in Table 1, at 9.02, is from Croce and Colacito (2011), who calibrate a long-run risks model to match the behavior of interest rates. Their calibration is larger than the smallest in Table 1 (Campbell and Cochrane, 1999), by a factor of 6. Under the approximation for Epstein–Zin preferences with an infinite EIS discussed above, this difference would also induce a difference in the Hansen–Jagannathan bound of a factor of 6, holding risk aversion constant. So these calibrations have substantial differences in their implications for the size of risk premia.

There is a large literature that estimates consumption dynamics using both parametric and non-parametric methods.[6] The benchmark results I obtain using the robust estimators

[4] See, recently, Tallarini (2000), Barro (2006), and Gourio (2012)

[5] Kaltenbrunner and Lochstoer (2010) study a range of different calibrations, but the one with an LRSD of 8.22 seems to fit the data best.

[6] See, e.g., Harvey (1985); Clark (1987); Poterba and Summers (1988); Lo and MacKinlay (1988); Cochrane (1988); Cochrane and Sbordone (1988); Cochrane (1994); and Morley (2007).

are on the upper end of the range of past results. That finding is consistent with the simulation evidence below: if there is a persistent component to consumption growth, standard estimators will be biased downwards.

3 Spectral density estimators

Because there is no agreement on the true specification for the consumption growth process, this paper looks for a non-parametric estimator that is valid for a wide variety of possible driving processes for consumption growth. Non-parametric methods estimate the LRSD while placing minimal restrictions on the high-frequency characteristics of the data. The estimator I propose is valid as long as the consumption process has a full set of autocovariances and a finite spectral density with two derivatives (finding the optimal bandwidth requires four derivatives). The method is, in theory, robust to time aggregation (Working, 1960; Campbell and Mankiw, 1989), stochastic volatility (as in the long-run risks literature), and non-normality (e.g. disasters; Barro, 2006).

If consumption growth is uncorrelated over time, then estimating its long-run variance is simple, since the long-run variance is the same as the unconditional variance. However, if consumption growth has a persistent component, then estimation of the long-run variance is more difficult. A persistent trend in consumption growth induces a peak in its spectral density around frequency zero. In the presence of peaks, standard non-parametric estimators tend to be biased downwards. Not only does that bias make point estimates unreliable, but it can also cause confidence intervals to fail to include the true value of the long-run variance. Given that a key hypothesis in the consumption-based asset pricing literature is that there might be a persistent component in consumption growth, it is important that the estimator I use here be robust to the possibility of a persistent component. I therefore focus especially on the issue of bias.

3.1 Estimation

Suppose for the moment that consumption growth follows a moving average process with potentially infinite order

$$\Delta c_t = \mu + b(L)\,\varepsilon_t \tag{3}$$

where Δc_t is log consumption growth, μ is its mean, $b(L) = \sum_{j=0}^{\infty} b_j L^j$ is a power series in the lag operator L, and ε_t is a martingale difference sequence with an unconditional standard deviation of 1. The change in the expectation of the long-run level of consumption on date t is then $b(1)\,\varepsilon_t$, with standard deviation $b(1)$. We can define the spectral density as

$$f(\kappa) = \sum_{j=-\infty}^{\infty} \gamma_j \cos(\kappa j) \tag{4}$$

$$\textit{where } \gamma_j \equiv cov(\Delta c_t, \Delta c_{t-j}) \tag{5}$$

The spectral density at frequency zero is $\omega^2 \equiv f(0) = \sum_{j=-\infty}^{\infty} \gamma_j$. Furthermore, note that

$$\sum_{j=-\infty}^{\infty} \gamma_j = \sum_{j=-\infty}^{\infty} \sum_{k=0}^{\infty} b_k b_{k-j} = b(1)^2 \tag{6}$$

$$\text{and } std\left(\Delta E_{t+1} \sum_{j=0}^{\infty} \Delta c_{t+1+j}\right) = b(1) = f(0)^{1/2} \tag{7}$$

which shows that the LRSD that appears in Epstein–Zin preferences, $std\left(\Delta E_{t+1} \sum_{j=0}^{\infty} \Delta c_{t+1+j}\right)$, is equal to $f(0)^{1/2}$. This result is easily extended to a setting where consumption growth is driven by multiple shocks.[7]

I consider non-parametric estimators of $f(0)$ based on kernel smoothers in the frequency domain. Given a sample of consumption growth of length T, $\{\Delta c_0, \Delta c_1, ..., \Delta c_{T-1}\}$, the

[7] Suppose $\Delta c_t = \mu + B(L)\,\boldsymbol{\varepsilon}_t$, where $\boldsymbol{\varepsilon}_t$ is a $j \times 1$ vector and $B(L) = \sum_{k=0}^{\infty} B_k L^k$, where B_k is a $1 \times j$ vector. The calculation of the covariances and LRSD then yield the same result that the LRSD is equal to $f(0)^{1/2}$.

periodogram is the square of the finite Fourier transform of the data,

$$I_T(\lambda) \equiv T^{-1} \left| \sum_{t=0}^{T-1} \exp(-i\lambda t) \Delta c_t \right|^2 \tag{8}$$

for $-\infty < \lambda < \infty$. The periodogram can be thought of as the sample spectrum. Nonparametric estimates of $f(0)$ smooth observations of the periodogram for λ near zero.

The derivation of the estimator follows Brillinger (1981) closely. I consider estimates of the form,

$$f_T(0) \equiv \frac{2\pi}{T} \sum_{s=1}^{T-1} W_T\left(\frac{2\pi s}{T}\right) I_T\left(\frac{2\pi s}{T}\right) \tag{9}$$

where $-\infty < W_T(\lambda) < \infty$ is a weighting function. The standard method to obtain asymptotic results for kernel smoothers is to assume that as T grows, the mass of W_T becomes concentrated in a smaller region around zero so that $f_T(0)$ eventually only depends on the behavior of $f(\lambda)$ local to 0. Specifically,

$$W_T(\lambda) \equiv \sum_{j=-\infty}^{\infty} B_T^{-1} W\left(B_T^{-1}(\lambda + 2\pi j)\right) \tag{10}$$

where B_T determines the bandwidth for a sample size T and will shrink asymptotically. The summation over j ensures that $W_T(\lambda)$ has period 2π. Following Brillinger (1981), I make the following assumptions about the fixed kernel function $W(\lambda)$,

Assumption 1 $W(\lambda)$ *is real-valued, even, of bounded variation,* $-\infty < W(\lambda) < \infty$,

$$\int_{-\infty}^{\infty} W(\lambda)\, d\lambda = 1 \tag{11}$$

$$\int_{-\infty}^{\infty} |W(\lambda)|\, d\lambda < \infty \tag{12}$$

$$W(\lambda) = 0 \text{ for } |\lambda| > 2\pi \tag{13}$$

$$\int_{-\infty}^{\infty} |\lambda|^P |W(\lambda)|\, d\lambda < \infty \text{ for } P \leq 4 \tag{14}$$

$f_T(0)$ is a weighted average of the values of the periodogram between frequencies $-2\pi B_T$

and $2\pi B_T$, so B_T determines the bandwidth of the estimator. As B_T shrinks, the estimate of $f_T(0)$ is driven by values of the periodogram closer to frequency zero.

Since $f_T(0)$ is a linear combination of values of the periodogram, and since the periodogram is non-negative, a sufficient condition for the estimate of $f(0)$ to be non-negative is that the kernel, $W(\lambda)$, is non-negative. That condition has been imposed in most of the leading long-run variance estimators in the past (e.g. the Newey–West (1987) estimator and the quadratic spectral estimator analyzed by Priestley (1981) and Andrews (1991)).

To derive the estimator, I use the following from Brillinger (1981)

Assumption 2 *Δc_t is real-valued with mean μ, covariances $\gamma_j \equiv cov(\Delta c_t, \Delta c_{t-j})$ for $j = 0, \pm 1, ...$ (i.e. consumption growth is second-order stationary), and $\sum_{j=-\infty}^{\infty} |u| \left|\gamma_j\right| \leq \infty$*

Assumption 3 *Denote the cumulants of Δc_t as $\kappa_j(t_1, t_2, ...t_j) = cum\left(\Delta c_{t_1}, \Delta c_{t_2}, ..., \Delta c_{t_j}\right)$. For all $k = 2, 3, ...$*

$$\sum_{v_1, v_2, ..., v_{k-1}=-\infty}^{\infty} (1 + |v_j|) \left|\kappa_k(v_1, ..., v_{k-1}, v_k)\right| < \infty \tag{15}$$

for $j = 1, ..., k-1$.

Second, I require that the spectral density have at least four derivatives

Assumption 4 *$f(0)$ has continuous bounded derivatives of order ≤ 4*

The assumption about the cumulants is needed when finding the variance and asymptotic distribution of the estimator. Andrews (1991) discusses ways that this assumption may be weakened. In Gaussian models, it will be satisfied if the other assumptions about differentiability are satisfied. However, assumption 3 does potentially rule out models in which the higher moments of consumption growth vary over time and are sufficiently autocorrelated. Assumption 4 is most likely to be restrictive in the finance literature. The assumption that the spectrum has multiple derivatives at frequency zero rules out long-memory in consumption growth. Simulations with long memory are examined below.

Taking a Taylor expansion of the expression in theorem 5.6.1 from Brillinger (1981) (see theorem 7.4.2 in Brillinger (1981); for details see, for example, Wand and Jones (1994), Sections 2.4.2 and 2.5), the following result is obtained:

Proposition 5 *Under assumptions 1, 2, and 4, and if* $B_T \to 0$ *as* $T \to \infty$

$$Ef_T(\lambda) = f(\lambda) + \frac{B_T^2}{2}\left(\int_{-\infty}^{\infty} \omega^2 W(\omega)\, d\omega\right)\frac{d^2 f(\lambda)}{d\lambda^2} + o\left(B_T^2\right) + O\left(B_T^{-1}T^{-1}\right) \tag{16}$$

where $O(\beta_T)$ *and* $o(\beta_T)$ *are Landau notation such that* $\alpha_T = O(\beta_T)$ *indicates that* $|\alpha_T/\beta_T|$ *is bounded for sufficiently large* T*, and* $\alpha_T = o(\beta_T)$ *indicates that* $\alpha_T/\beta_T \to 0$*. Furthermore, if* $B_T^3 T \to \infty$ *as* $T \to \infty$*,*

$$\lim_{T\to\infty} B_T^{-2}\left(Ef_T(\lambda) - f(\lambda)\right) = \frac{1}{2}\left(\int_{-\infty}^{\infty} \omega^2 W(\omega)\, d\omega\right)\frac{d^2 f(\lambda)}{d\lambda^2} \tag{17}$$

Proposition 6 *(Brillinger (1981) theorem 5.6.2 and corollary 5.6.2) Under assumptions 1, 2, 3, and if* $B_T T \to \infty$ *as* $T \to \infty$*,*

$$\lim_{T\to\infty} B_T T var\left(f_T(0)\right) = 4\pi f(0)^2 \int_{-\infty}^{\infty} W(\beta)^2\, d\beta \tag{18}$$

Propositions 5 and 6 are standard limiting results for kernel estimators and are used extensively in the literature. Intuitively, the bias depends on the curvature of the spectrum, while the variance depends on how concentrated the weight of the kernel is. When the curvature of the spectrum is higher, kernel estimators will tend to be less accurate, while a kernel that puts more weight on a small number of periodogram ordinates will tend to be more variable.

Using results of the same form as the two propositions here, Priestley (1962), Epanechnikov (1969), and Andrews (1991) show that the quadratic spectral kernel has the minimum mean-squared error among non-negative estimators. When the long-run variance is being estimated for the purpose of using in the denominator of a test statistic, it may be natural to use a non-negative estimator. However, we are concerned here with simply estimating its value for its own interest. It is thus worthwhile to try to reduce the bias. As is well known in the literature on nonparametric kernel density estimation, for smoother spectral densities, picking a higher order kernel can reduce the bias while not changing the order of the variance, thus reducing the mean squared error.

I therefore look for a kernel that minimizes the variance of the spectral density estimator conditional on setting the leading term of the bias to zero. This can be achieved only if the restriction on $W(\cdot)$ being nonnegative is removed. The primary concern in the consumption-based asset pricing literature is that consumption growth might have a highly persistent component, which would induce a peak in the spectral density at frequency zero, and hence $f''(0) < 0$. The bias term is thus a serious potential source of estimation error, so a good estimator in this setting should try to reduce its order.

The asymptotic variance of the estimator is proportional to $\int_{-\infty}^{\infty} W(\omega)^2 \, d\omega$. It is straightforward, through direct minimization of $\int_{-\infty}^{\infty} W(\omega)^2 \, d\omega$, to obtain the following result,

Proposition 7 *Given the assumptions in propositions 5 and 6, the bandlimited kernel W satisfying assumption 1 that minimizes the asymptotic variance,* $\lim_{T\to\infty} B_T T var(f_T(0))$, *holding $\int \omega^2 W(\omega)\, d\omega = 0$, is*

$$W_{RQS}(\omega) = \begin{cases} \frac{1}{2\pi}\left(\frac{9}{8} - \frac{15}{8}(\omega/2\pi)^2\right) \ \textit{if} \ |\omega| < 2\pi \\ 0 \ \textit{otherwise} \end{cases} \tag{19}$$

W_{RQS} is the minimum-variance fourth-order kernel obtained by Gasser, Muller, and Mammitzsch (1985). Note that the three propositions thus far rely only on a second-order Taylor approximation to the spectrum. So the derivation of the RQS kernel here requires only the assumption that the spectrum has two derivatives.

Priestley (1981) and Andrews (1991) find that the quadratic spectral kernel, which takes the form

$$W_{QS}(\omega) = \begin{cases} \frac{1}{2\pi}\frac{3}{4}\left(1-(\omega/2\pi)^2\right) \text{ if } |\omega| < 2\pi \\ 0 \text{ otherwise} \end{cases} \tag{20}$$

has the lowest asymptotic mean-squared error among non-negative kernels. W_{RQS}, which I refer to as the unbiased quadratic spectral (RQS) kernel, is also an inverted parabola, but it is no longer non-negative. The fact that it is negative in some regions is what allows it to eliminate the leading term of the asymptotic bias.

The assumptions required to derive the RQS kernel are relatively weak in the context of macroeconomics and finance. In particular, consumption growth can have very general forms of heteroskedasticity or time variation in higher moments and there is no assumption of normality. The part of the derivation most likely to come into conflict with the data is the idea that the behavior of the spectrum around frequency zero is well approximated by a Taylor approximation.

W_{RQS} is optimal in the sense that it minimizes the asymptotic variance in equation (18) while setting the leading term of the bias to zero. The idea to use a kernel that is potentially negative in order to reduce bias is certainly not novel to this paper – as noted, the exact kernel was derived by Gasser, Muller, and Mammitzsch (1985). Andrews (1991) examines kernels with negative weight at certain frequencies, and Politis and Romano (1995) give an extensive analysis. However, in neither of the latter cases is a minimum-variance unbiased kernel derived, while Gasser, Muller, and Mammitzsch (1985) do not examine the performance of such kernels in simulations or empirical analyses.

Figure 1 plots the weights for $W_{RQS,T}$ for a bandwidth that isolates only cycles that last longer than 8 years ($B_T = 2\pi/32$ in quarterly data; a standard definition of the lower limit of business cycle frequencies that is suggested as a limit in the context of long-run variance estimation in Müller (2007); see also Christiano and Fitzgerald (2003)) along with the weights for the Newey–West and QS kernels. $W_{RQS,T}$ looks similar to the kernel for the Newey–West estimator at low frequencies, but it becomes negative at higher frequencies.

Intuitively, if the spectrum is upward sloping, then $W_{RQS,T}$ is able to extrapolate so that its estimate $f_T(0)$ can be above any of the values of $I_T(\lambda)$ that we measure. Compared to the QS kernel with the same bandwidth, we can see that the RQS kernel is relatively more peaked, and also is negative at higher frequencies. The fact that it is more peaked means that we should expect it to have a higher variance for a given bandwidth.

3.2 Distribution theory

Finally, we need a distribution of the estimate using W_{RQS} so that we can construct confidence intervals. In any particular sample, $f_{T,RQS}(0)$ is a weighted average of the periodogram. The number of periodogram ordinates used for $f_T(0)$ depends on $B_T T$ (since the periodogram ordinates are spaced in proportion to $1/T$ and the frequency range used is limited by B_T). In the derivation of the RQS kernel (and for asymptotic expressions for the bias more generally), we must assume that $B_T T \to \infty$, so that the number of periodogram ordinates used is infinite. In realistic samples, though, that number will often be small. For example, in the main results below, the estimator uses only eight points on the periodogram. We would thus expect that any approximation assuming that $B_T T \to \infty$ will do a poor job of describing the asymptotic distribution.

To derive the distribution theory, rather than approximating $f_T(0)$ as though it uses an infinite number of periodogram ordinates, I approximate it as though it uses a fixed number, which is equivalent to assuming that $B_T T$ converges to a constant – that is, the bandwidth shrinks in proportion to the sample size.[8] In the end, then, the sense in which the RQS kernel is optimal requires relatively more powerful asymptotics under which the number of periodogram ordinates used in the estimator is infinite, while the distribution theory uses a weaker assumption that is likely more realistic in small samples. Both of these asymptotic assumptions are standard in the literature and neither is novel to this paper.

Under the assumption that $B_T = bT^{-1}$ for an integer b, plugging the formula for W_T as

[8] The assumption that $B_T T$ converges to a constant is similar to the fixed-b asymptotics used recently by Kiefer and Vogelsang (2005) and Phillips, Sun, and Jin (2008), among others.

a function of W into the formula for $f_T(0)$ and using the periodicity and symmetry of the periodogram, we have

$$f_T(0) = \frac{2\pi}{T}\sum_{s=1}^{T-1} W_T\left(\frac{2\pi s}{T}\right) I_T\left(\frac{2\pi s}{T}\right) \tag{21}$$

$$= \left(2b^{-1}\right)(2\pi)\sum_{s=1}^{b} W\left(b^{-1}2\pi s\right) I_T\left(\frac{2\pi s}{T}\right) \tag{22}$$

The distribution of such an estimate then satisfies the following limit

Proposition 8 *(Brillinger (1981) theorem 5.5.3) Suppose assumptions 2 and 3 hold. For a fixed set of weights* $W(b^{-1}2\pi s)$, $s \in \{0, 1, ..., b\}$, $f_T(0)$ *converges in distribution to*

$$f_T(0) \Rightarrow f(0)\frac{\sum_{s=1}^{b} W\left(b^{-1}2\pi s\right)\chi_2^2(s)/2}{\sum_{s=1}^{b} W\left(b^{-1}2\pi s\right)} \tag{23}$$

where $\Rightarrow$ *denotes convergence in distribution and* $\chi_2^2(s)$ *is a set of independent* χ_2^2 *variates.*[9]

The distribution (23) follows from the standard result that the periodogram ordinates are asymptotically independent $\chi_2^2/2$ random variables. This result does not require strong assumptions that the number of periodogram ordinates used for the kernel grows asymptotically, which is how a Gaussian distribution for the estimate of the spectrum is usually derived. Because this distribution takes seriously the fact that the number of periodogram ordinates used in the estimation is small we may expect it to be more accurate than a normal distribution.

3.3 Bandwidth selection

There are two common approaches to determining the bandwidth for an estimate of the LRSD. The first is to use economic intuition to restrict the set of frequencies used in the

[9] The theorem in Brillinger (1981) has a small typo – the denominator should not be squared. Referring to Brillinger's appendix, the result is derived using the continuous mapping theorem applied to the result that the periodogram ordinates are asymptotically independent $\chi_2^2/2$ random variables.

estimation. For example, Müller (2007) proposes in setting with macroeconomics to use information on the spectrum at frequencies lower than the business cycle.[10] The second method is to find an optimal bandwidth that minimizes the asymptotic mean squared error (AMSE)

Following Andrews (1991), we can obtain the AMSE of the estimator under specific assumptions about the shrinkage rate of the bandwidth. Since the leading term of the bias has been set to zero, by taking a Taylor expansion of a higher order than was needed for Proposition 5, the following result is obtained.

Proposition 9 *Under assumptions 1, 2, and 4, and if $B_T \to 0$ as $T \to \infty$*

$$\begin{aligned} Ef_T(\lambda) &= f(\lambda) + \frac{B_T^2}{2}\left(\int_{-\infty}^{\infty} \omega^2 W(\omega)\, d\omega\right)\frac{d^2 f(\lambda)}{d\lambda^2} \\ &\quad + \frac{B_T^4}{24}\left(\int_{-\infty}^{\infty} \omega^4 W(\omega)\, d\omega\right)\frac{d^4 f(\lambda)}{d\lambda^4} + o\left(B_T^4\right) + O\left(B_T^{-1}T^{-1}\right) \end{aligned} \tag{24}$$

Furthermore, if $\int_{\infty}^{\infty} \omega^2 W(\omega)\, d\omega = 0$ and $B_T^5 T \to \infty$,

$$\lim_{T\to\infty} B_T^{-4}\left(Ef_T(\lambda) - f(\lambda)\right) = \frac{1}{24}\left(\int_{-\infty}^{\infty} \omega^4 W(\omega)\, d\omega\right)\frac{d^4 f(\lambda)}{d\lambda^4} \tag{25}$$

Using proposition (6) for the variance, we obtain

Proposition 10 *Assume $B_T = \eta^{-1/9}T^{-1/9}$ for some positive constant η. Then*

$$\begin{aligned} \lim_{T\to\infty} B_T T E\left[\left(f_{T,RQS}(0) - f(0)\right)^2\right] &= \lim_{T\to\infty} B_T T\left(var\left(f_{T,RQS}(0)\right) + E\left[f_{T,RQS}(0) - f(0)\right]^2\right) && (26) \\ &= \lim_{T\to\infty}\left(B_T T var\left(f_{T,RQS}(0)\right) + \eta^{-1}B_T^{-8}E\left[f_{T,RQS}(0) - f(0)\right]^2\right) && (27) \\ &= 4\pi f(0)^2 \int_{-\infty}^{\infty} W(\beta)^2 d\beta + \eta^{-1}\left(\frac{1}{24}\left(\int_{-\infty}^{\infty} \omega^4 W(\omega)\, d\omega\right)\frac{d^4 f(\lambda)}{d\lambda^4}\right)^2 && (28) \end{aligned}$$

[10] Specifically, "Knowledge about the form of the spectrum... then suggests appropriate values for [the bandwidth]; for macroeconomic time series, for instance, one might want to pick [a bandwidth] small enough not to dip into business cycle frequencies."

which implies

$$\lim_{T\to\infty} T^{8/9} E\left[\left(f_{T,RQS}(0)-f(0)\right)^2\right] = \eta^{1/9} 4\pi f(0)^2 \int_{-\infty}^{\infty} W(\beta)^2 d\beta + \eta^{-8/9}\left(\frac{1}{24}\int_{-\infty}^{\infty} \omega^4 W(\omega)\, d\omega\right)^2 \left(\frac{d^4 f(\lambda)}{d\lambda^4}\right)^2 \quad (29)$$

The bandwidth that minimizes the asymptotic mean squared error of $f_{T,RQS}$ is then obtained by minimizing (29) over η. We obtain

$$\eta^*_{RQS} = 2\pi^{-1} \frac{\left(\frac{1}{24}\int_{-\infty}^{\infty} \omega^4 W(\omega)\, d\omega\right)^2}{\int_{-\infty}^{\infty} W(\beta)^2 d\beta} \left(\frac{\frac{d^4 f(\lambda)}{d\lambda^4}}{f(0)}\right)^2 \quad (30)$$

The unknown factor in this formula is a normalized measure of curvature, $\frac{d^4 f(\lambda)}{d\lambda^4}/f(0)$. Newey and West (1994) propose to estimate the normalized curvature using a kernel smoother. Following them, I use a flat kernel in the time domain and estimate the normalized curvature as

$$\frac{d^4 f\widehat{(\lambda)}}{d\lambda^4}/f(0) = \frac{2\sum_{j=1}^{n_\gamma} j^4 \hat{\gamma}_j}{\hat{\gamma}_0 + 2\sum_{j=1}^{n_\gamma} \hat{\gamma}_j} \quad (31)$$

where $\hat{\gamma}_j$ is the jth sample autocovariance of consumption growth and n_γ is a first-stage bandwidth that must be chosen.

The results in the remainder of the paper examine both fixing the bandwidth B_T based on economic knowledge and also using the estimated minimum-AMSE bandwidth, $\left(\eta^*_{RQS} T\right)^{-1/9}$. I refer to the estimates using the fixed bandwidth – set to use cycles only lasting 32 quarters or more – as the RQS estimates and those using the minimum-AMSE bandwidth as $\mathrm{RQS}_{\mathrm{mAMSE}}$. I set $n_\gamma = 20$ for the benchmark results but also explore other choices for the first-stage bandwidth.[11] Newey and West (1994) discuss the optimal asymptotic growth rate of n_γ, but asymptotic results provide no specific formula for n_γ in small samples, so I explore the

[11] It is not possible to escape having to choose a bandwidth or a fixed model at some point. Newey and West (1994) suggest, though, that the selection of the bandwidth in the estimation of the normalized curvature is less consequential for the final estimates of the LRSD than the selection of the bandwidth for estimating the LRSD itself, so using an arbitrary choice in the first stage may be more justified.

effects of different choices in simulations.

3.4 Cointegration

Previous work has taken advantage of the implication of balanced growth theory that consumption is cointegrated with other aggregate series like investment and output to help obtain more powerful estimates of the LRSD (e.g. Cochrane and Sbordone, 1988). For non-parametric estimates such as those used here, though, adding a series that is cointegrated with consumption growth does not improve the asymptotic behavior of the estimators.

Specifically, suppose consumption and output are cointegrated, so that their joint spectral density matrix at frequency zero has rank 1. Intuitively, since the non-parametric estimators use only the very low frequency features of the data, and consumption and output are assumed to be perfectly correlated at the lowest frequencies, adding output to the estimation adds no new information.[12] When Cochrane and Sbordone (1988) find that cointegration improves estimates, it is because they use a parametric method that also requires estimating the high-frequency features of the data. Because consumption and output growth may differ at higher frequencies, using both can improve power. But since here we focus on non-parametric methods that ignore high frequencies, output growth is asymptotically redundant to consumption growth.

4 Performance of estimators in simulations

To examine the bias, variance, and accuracy of the confidence intervals for various estimators, I now simulate a broad range of potential data generating processes for consumption growth. I consider five estimators: the RQS and $\text{RQS}_{\text{mAMSE}}$ estimators, the standard QS

[12] Brillinger (1969) derives the asymptotic distribution of the multivariate periodogram and shows that it follows a Wishart distribution that is independent across frequencies with scale matrix equal to $f_{XX}(\lambda)$ (where f_{XX} is the multivariate spectral density matrix). Since $f_{XX}(0)$ has rank 1 when output and consumption growth are cointegrated, the periodogram ordinates local to zero are perfectly correlated. A similar result holds under Mueller's (2007) asymptotics.

kernel, the Newey–West estimator, and Müller's (2007) estimator based on a Karhunen–Loeve transformation of the data (with a modification for the fact that in this paper we do not assume that the true mean of the time series is known ex ante).[13,14] For the Newey–West and QS kernels, I used the fixed-b asymptotics from Kiefer and Vogelsang (2005).

I simulate four different specifications for consumption growth – Bansal and Yaron's (2004) long-run risks model, an AR(1), a model with long memory, a model that combines both persistence and anti-persistence in consumption growth, and a model with jumps in consumption growth. Overall, I find that the RQS estimator outperforms the Newey–West, QS, and Müller estimators in terms of both bias/variance trade-offs and confidence interval coverage. Furthermore, the performance of the RQS and $\mathrm{RQS}_{\mathrm{mAMSE}}$ estimators is highly similar.

4.1 Bansal and Yaron's (2004) long-run risk model

4.1.1 Simulation strategy

Bansal and Yaron's (2004) long-run risk model for consumption growth is

$$\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \tag{32}$$

$$x_{t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1} \tag{33}$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu_1 \left(\sigma_t^2 - \sigma^2\right) + \frac{\sigma_w}{\sigma^2} w_{t+1} \tag{34}$$

$$e_{t+1}, \eta_{t+1}, w_{t+1} \sim i.i.d.N\,(0,1) \tag{35}$$

[13] $\omega_{Mueller,j}$ is defined as $\omega_{Mueller,j} \equiv j^{-1} \sum_{k=1}^{j} \chi_k^2$, where $\chi_j^2 \equiv (\pi j)^2 \left(T^{-1} \sum_{t=1}^{T} u_t \sqrt{2} \sin\left(j\pi t/T\right)\right)^2$ and $u_t = T^{-1/2} \sum_{j=1}^{t} \Delta c_j$. The formula for χ_j^2 here differs slightly from Müller (2007) because that paper assumes the data are not demeaned. When the data are demeaned, the trigonometric transforms must be the eigenfunctions for a Brownian bridge (Phillips (1998)).

[14] The estimator in Müller (2007) is positive-semidefinite, so it can be expected to inherit the bias that the other positive-semidefinite estimators have. And in fact, his reported simulation results show that for persistent time series, his estimator is biased downward. On the other hand, Müller (2007) shows that the KLT-based estimator is robust to contaminations in models that the usual Newey–West estimator is not. His estimator is designed to be robust and optimize a bias/variance trade-off in the presence of those contaminations in the class of positive-semidefinite estimators.

In addition to generating long-run risk, the model is also heteroskedastic (making consumption growth unconditionally leptokurtic), which adds another layer of difficulty for the estimators.

I consider a range of calibrations with different ratios of the LRSD to the unconditional standard deviation of consumption growth, with the ratio $LRSD/std\,(\Delta c_t)$ ranging from 1.05 to 2.3.[15] In the original calibration the ratio is 2.26, while a ratio of 1 corresponds to white-noise consumption growth, in which case we would expect all estimators to perform well. I use a sample of 67 years of data, as in the post-war empirical sample below, because that is the shortest and thus most difficult sample size for the estimators.[16] The simulated monthly series for consumption growth is aggregated into quarterly observations for the estimation. I simulate 20,000 samples for each value of the LRSD, and calculate each estimator along with its confidence interval in each sample.

4.1.2 Results

Figures 2 and 3 plot the results from the simulations. Figure 2 plots the bias of the estimators against the variance as the bandwidth of each estimator is varied (there is no bias/variance trade-off for $\text{RQS}_{\text{mAMSE}}$ since its bandwidth is estimated rather than chosen). The four panels of the figure correspond to different ratios of the LRSD to the unconditional standard deviation (i.e. to different degrees of peakedness in the spectrum). As we would expect, the estimators are all biased downwards, and the bias is relatively larger when the LRSD is higher. However, in all four cases, the bias of the RQS estimator is smaller for a given level of

[15] Given values for $LRSD$, $std\,(\Delta c_t)$, and $corr\,(\Delta c_t, \Delta c_{t-1})$, the parameters ρ, φ_e, and σ^2 are identified. To study the effects of a change in the LRSD, I hold $std\,(\Delta c_t)$ and $corr\,(\Delta c_t, \Delta c_{t-1})$ fixed at their original values from Bansal and Yaron (2004) and solve for the implied values of ρ, φ_e, and σ^2 given a value of the $LRSD$. In Bansal and Yaron's calibration, $std\,(\Delta c_t) = 0.0080$ and $corr\,(\Delta c_t, \Delta c_{t-1}) = 0.043$. The calibration is in monthly terms, so $std\,(\Delta c_t)$ corresponds to an annualized standard deviation of 2.76 percent. Note that in the simulations we retain the original specification for stochastic volatility, and its persistence and volatility (v_1 and σ_w) are held fixed at their original values. The original calibration is: $\rho = 0.979$; $\varphi_e = 0.044$; $\sigma^2 = 0.0078^2$; $\nu_1 = 0.987$; $\sigma_w = 0.0000023^2$.

[16] I also ran similar simulations using sample sizes corresponding to the longer periods available in annual data. The results are not reported here, but they yield qualitatively similar conclusions. Quantitatively, with the longer samples every estimator has better confidence interval coverage and less bias.

variance. Moreover, the reductions are economically large. At a given bias, the RQS kernel reduces the variance by approximately 20 percent on average across the four simulations.

The RQS kernel does not simply represent a movement along the same bias-variance frontier that is obtained by the QS kernel. Rather, because it dispenses with the non-negativity constraint it is a shift of the frontier. This result is true both in the simulations and also asymptotically, due to the construction of the RQS kernel to have minimum variance for a given bias.

I next study in more detail the specific performance of the estimator that will be used in the empirical analysis below. Following Müller (2007) and Christiano and Fitzgerald (2003), I limit the bandwidth of the RQS estimator to only cycles lasting longer than 32 quarters, which is a standard definition of the end of business-cycle fluctuations. The other estimators are implemented with bandwidths that generate the same variance when the ratio of the LRSD to the unconditional standard deviation is 2.3, so as to make the estimators comparable. This corresponds to using 32 lags for the QS kernel and seven terms in Müller's (2007) estimator.

Figure 3 plots two measures of the accuracy of the four estimators: the coverage of their confidence intervals and their bias. In the two panels of the figure, the LRSD ranges from 1.05 to 2.3 times the unconditional standard deviation. The top panel plots the coverage of each estimator's 95-percent confidence interval. Coverage is the key measure of confidence interval accuracy since it tells us whether the true value of the LRSD is actually contained in the confidence interval as often as we expect. Here and below I use a one-sided confidence interval (on the upper side) since we are primarily concerned with finding an upper bound for the LRSD rather than a lower bound, but the results in the simulations are not sensitive to the choice of one-sided versus two-sided tests.

Across the range of LRSD's, from nearly white noise (where the LRSD is equal to the unconditional standard deviation) to the original long-run risks calibration, we see that the coverage of the confidence interval for the RQS and $\mathrm{RQS}_{\mathrm{mAMSE}}$ estimators is always nearly 95 percent, whereas it is well below 95 percent for the other estimators. When the LRSD is 2.3 times the unconditional standard deviation, Müller's estimator has coverage of 84 percent, while the QS kernel has coverage of 80 percent. While all the estimators perform well in the white noise case, when there is a substantial trend component in consumption growth, the confidence intervals quickly become inaccurate for all the estimators except the RQS estimator. The success of the RQS estimator in producing an accurate confidence interval is the strongest reason to use it in estimation – whatever the merits of the point estimate, we ultimately care about what range of estimates is plausible, and it is the confidence interval that tells us that.

The bottom panel of figure 3 plots the bias of the estimators. As would be expected given that the bandwidths for the estimators are set so that they all have the same variance, the bias of the RQS estimator is the smallest of all. The RQS estimator has a bias that is 25 percent smaller (in relative terms) on average than the other three estimators, so the reduction is economically significant. $\mathrm{RQS}_{\mathrm{mAMSE}}$ has bias similar to the QS and Müller estimators.

The minimum-AMSE RQS estimator used in figure 3 uses 20 lags of consumption growth to estimate the optimal bandwidth. To examine the robustness of that choice, figure A1 replicates figure 3 using different lag choices. For $n_\gamma \geq 20$, the bias of the RQS estimator is less than or equal to that of Müller's estimator. The confidence interval coverage across all the values of n_γ is almost always closer to 95 percent than what is obtained by Müller's estimator. So while the best results are obtained with $n_\gamma = 20$, the findings are not particularly sensitive to that choice.

4.2 AR(1) model

I next consider behavior in a simple AR(1) model, which is often taken as a basic benchmark. These simulations also let me directly compare my results to those of Müller (2007), who also simulates AR(1) models. In this model, consumption growth follows

$$\Delta c_t = \rho \Delta c_{t-1} + \varepsilon_t \tag{36}$$

$$\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right) \tag{37}$$

where the variance σ_ε^2 is irrelevant for the simulation results since all the variances and autocovariances simply scale with σ_ε^2.

Figure A2 plots the spectrum of consumption growth in an AR(1) with an autocorrelation of 0.9 compared to the baseline long-run risk model (Bansal and Yaron's (2004) calibration). We see that the AR(1) spectrum is less peaked around frequency zero, suggesting that it should be easier to estimate the LRSD in the AR(1) model.

Figure A3 reports bias/variance trade-offs as in figure 2 for four different degrees of autocorrelation. Again, the RQS estimator gives the smallest bias for a given level of variance across all the estimators and degrees of autocorrelation, though the benefits of the RQS estimator now seem smaller than we observed for the long-run risk model. Figure A4 shows that the RQS and $\text{RQS}_{\text{mAMSE}}$ estimators deliver a confidence intervals with coverage far closer to the nominal 95 percent level than the other estimators and with less bias. Table A1 also replicates simulation results for AR(1) and MA(1) models from Müller (2007).

4.3 Long memory

The two classes of models examined so far both satisfy the assumptions justifying the RQS estimator. I now examine a situation where consumption growth has a highly persistent component and the assumptions are violated. In particular, I assume that the autocovariances of consumption growth take the form

$$cov\left(\Delta c_t, \Delta c_{t-j}\right) = \sigma_{\Delta c}^2 \left(1 + |j|\right)^{-(1+\phi)} \tag{38}$$

for a parameter $0 < \phi < 1$. Consumption growth in this model displays long memory in the sense that the autocovariances decline more slowly than geometrically (Bailie (1996)).

The spectral density at frequency zero for this model is finite as long as $\phi > 0$. The second derivative of the spectrum is proportional to

$$f''(0) \propto \sum_{j=1}^{\infty} j^2 (1+j)^{-1-\phi} \tag{39}$$

which diverges when $\phi < 2$. The model of consumption growth (38) thus has a well defined long-run standard deviation, but it fails to satisfy assumption 4 because the spectrum does not have a second or third derivative. Figure A2 plots the spectrum for (38) when $\phi = 0.4$. The spectrum for the long memory model is sharply peaked at frequency zero, and it is clear that the second derivative does not exist. While the model (38) violates the assumptions justifying the RQS estimator, though, it has broadly similar properties to the long-run risk and AR(1) models, in the sense that the spectrum rises as the frequency approaches zero. Since the RQS kernel is meant to accommodate a spectrum that increases as the frequency falls, we may expect it to still perform well in this setting.

As with the long-run risk model, I vary the calibrations of the long-memory model based on the ratio of the LRSD to the unconditional standard deviation, which is governed by the parameter ϕ. I set $\sigma_{\Delta c} = 0.0138$.

Figures A5 and A6 report results for simulations of various parameterizations of the long memory model. In all four cases examined in figure A5, the RQS estimator again displays a superior bias/variance trade off to the other estimators. Figure A6 shows that the 95 percent confidence interval for the RQS and $\mathrm{RQS}_{\mathrm{mAMSE}}$ estimators have superior coverage to the others. So even when the technical assumptions used to derive the RQS estimator are violated, we still find that it works well in a case where consumption growth has a highly persistent component.

4.4 A notched spectrum

The results of the three sets of simulations so far suggest that the RQS estimator performs well when the spectrum is peaked at frequency zero, even if the differentiability requirements are violated. This section shows a scenario in which the RQS estimator underperforms the others.

Suppose consumption growth follows the process

$$\Delta c_t = \sum_{j=0}^{120} b_j \varepsilon_{t-j} \tag{40}$$

$$\varepsilon_t \sim N(0,1) \tag{41}$$

In the simulations so far, the coefficients b_j have all been positive, inducing a persistent component in consumption growth. Here I consider a scenario where they are positive initially but then turn negative. Specifically,

$$b_j = \begin{cases} \sigma_{\Delta c} \text{ if } j = 0 \\ \sigma_{\Delta c}\frac{R-1}{20} \text{ if } 1 \leq j \leq 20 \\ \sigma_{\Delta c}\frac{R-1}{100} \text{ if } 21 \leq j \leq 120 \end{cases} \tag{42}$$

The LRSD is thus $\sigma_{\Delta c}$, but the first 20 lag coefficients are all significantly positive. Figure A2 plots the spectrum for this model, and we see that while it rises at relatively low frequencies, it then falls again as the frequency approaches zero. A spectral estimator that averages the periodogram ordinates any significant distance from frequency zero should thus be expected to encounter problems. Specifically, in the simulated samples of 65 years of data, the first periodogram ordinate corresponds to a cycle length of 65 years, and the second to 33 years. At those two points, the notched spectrum in figure A2 is above 0.8, whereas its value at frequency zero is only 0.31, suggesting that kernel estimates will be biased upward enormously.

Figures A7 and A8 confirm that intuition. The estimators are all substantially positively biased – by anywhere from 20 to 130 percent – and their confidence intervals have very poor coverage – often well below 50 percent. Looking at the bias/variance trade-offs, unlike the previous simulations, there is no single estimator that outperforms all the others. For larger bandwidths, the RQS estimator has a smaller bias for a given level of variance than the other estimators. But when the bandwidth shrinks, raising the variance, the bias of the RQS estimator rises higher than those of the other estimators. In the end, then, none of the four estimators performs particularly well in these simulations, all of them being biased upward substantially. The RQS estimator performs worst in terms of confidence interval coverage, and in terms of its bias/variance trade-off delivers mixed results.[17]

So the situation when we should expect the RQS estimator to perform worst seems to be when consumption displays initial persistence over a period of years, but then anti-persistence in the longer term, especially if the spectrum decreases at frequencies below the smallest available periodogram ordinate.

4.5 Excess kurtosis

The simulations above all focus on the low-frequency behavior of consumption growth, but they leave the distribution of consumption growth conditionally normal (though in the long-run risk model consumption growth is unconditionally leptokurtic). The last set of simulations that I consider examines a scenario where the persistent component of consumption growth can display large jumps.

I assume consumption growth follows

$$\Delta c_t = x_t + \varepsilon_t \quad (43)$$

$$\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right) \quad (44)$$

and where x_t follows a Markov switching process, taking on values $\pm\bar{x}$. I assume that the

[17] Since all the estimators are biased up in this case, I examine two-sided instead of one-sided confidence intervals (with one-sided intervals, all estimators have 100 percent coverage).

process is symmetrical with a probability $1-\rho_x$ of staying in the same state as in the previous period and probability ρ_x of transitioning to the other state. When $\rho_x = 0.03$, the conditional kurtosis of x_t is 32, so this specification allows us to test the performance of the estimators in the presence of extreme excess kurtosis in the persistent component of consumption growth. I simulate the model at a monthly frequency, letting ρ_x vary between 0.5 and 0.0125. σ_ε is set to 0.0078 and $\bar{x} = 0.0017$.

Figures A9 and A10 report results for simulations of the Markov switching model. As in the other specifications, the RQS estimator again displays less bias for a given level of variance. Figure A10 shows that its confidence intervals are now somewhat conservative, having coverage that is generally above the nominal level. The RQS estimator will thus tend to under-reject excessively high calibrations of the LRSD here. Otherwise, though, the basic conclusions about the performance of the RQS estimator compared to previous proposals are unchanged from the four other simulations considered.

Overall, then, this section shows that the RQS estimator displays a superior bias/variance trade-off and confidence intervals that have coverage rates closer to their nominal levels compared to the QS kernel, the Newey–West estimator, and Müller's (2007) estimator in all examined cases except when the impulse response function of consumption growth displays initial persistence which is then cancelled out by antipersistence over the very long run. The RQS kernel can be expected to work better than the other available estimators in the long-run risk model, an AR(1), with long memory, and with excess kurtosis in the consumption trend. Moreover, the results using the RQS kernel are essentially unchanged regardless of whether the bandwidth is chosen based on economic knowledge or on a minimum-AMSE criterion.

5 Empirical estimates

This section takes the RQS kernel estimator and applies it to observed consumption growth in the United States to obtain estimates and confidence intervals for the LRSD.

5.1 Data

I study three data samples. The first is the standard post-war quarterly non-durable goods and services per capita measure of consumption growth. The quarterly data covers 1947q1–2015q1. The second sample is the longest annual sample available from the BEA, which runs from 1929 to 2014. Finally, to form the longest possible sample, I take data on aggregate per capita consumption (durables, non-durables, and services) from Barro and Ursua (2010). Their sample is 1834–2009, and I add updated data from the BEA to extend it to 1834–2014. In results not reported here, I find similar LRSD estimates from the three different series in the post-war period where they overlap.

5.2 Results

Figure 4 reports the estimates. The plot is divided into three sections for the three data samples. Each of the squares is a point estimate, while the small vertical lines give one-sided 90- and 95-percent confidence intervals for each estimator. The horizontal lines are placed for reference at the calibrations used by Bansal and Yaron (BY; 2004), Bansal, Kiku, and Yaron (BKY; 2012), and Kaltenbrunner and Lochstoer (KL; 2010). As in the simulations, I set the bandwidth of the RQS kernel estimator so that it only takes into account fluctuations lasting 8 years or more (i.e. $B = 2\pi/32$ in quarterly data and $2\pi/8$ in annual data). In each sample I also estimate the optimal bandwidth for $\text{RQS}_{\text{mAMSE}}$.

In the post-war data, the point estimate for the LRSD with RQS is 2.45 percent per year, with a 90-percent CI at 3.99 and 95-percent CI at 4.91. These estimates reject both BY and KL at the 5-percent level and BKY at the 10-percent level. The two longer samples yield substantially higher point estimates – 3.50 and 4.05 percent per year for the post-1929 and the post-1834 data, respectively. In all three samples we reject the KL and BY calibrations with both the RQS and RQS_{mAMSE} estimates, while BKY is rejected at the 10-percent level in the post-war sample with RQS and the 5-percent level with RQS_{mAMSE}. The BKY calibration of the LRSD thus seems to be weakly supported by the data (especially the long samples), while the other two calibrations are soundly rejected.

In the previous section, we saw that the bias in the estimators rose as the ratio of the LRSD to the unconditional standard deviation rose. In the post-war sample, that ratio is 2.3; in the 1929–2014 sample it is 1.59; and in the 1834–2014 sample it is 1.07. We should thus expect the bias in the point estimates to be minimal except possibly in the post-war sample. If we use the 18-percent correction implied by the simulations for the ratio of 2.3 in the post-war sample, then the point estimate should be revised from 2.45 to 2.89 percent per year. That said, recall that the confidence interval coverage for the RQS estimator is essentially correct for all values of the LRSD in Figure 3, so regardless of any bias in the point estimates, the confidence intervals can be trusted.

In addition to the RQS kernel estimator, I also consider estimates from a number of standard parametric estimators. These estimators impose much tighter assumptions on the dynamics of consumption growth, but they are useful for making sure that the RQS estimator is not an outlier compared to more well known estimators. I therefore estimate ARMA models on each sample, using the Bayesian information criterion to choose the top two preferred ARMA models in each data sample. I also estimate an ARMA(1,1) model for each sample since that is a homoskedastic version of what BY and BKY use.[18] The point estimates and 90- and 95-percent confidence intervals for the various parametric estimators are all summarized in Figure 4. As with the RQS estimator, the estimates are higher in the annual data. Not surprisingly, the confidence intervals for the parametric estimators are

[18]Note that the parametric estimators do not account for time aggregation, which makes them potentially misspecified. The presence of heteroskedasticity makes them inefficient, but still consistent as quasi-maximum likelihood estimators.

much tighter than those for the RQS kernel. Intuitively, the parametric estimators are able to take advantage of high-frequency variation in consumption growth to identify the low-frequency dynamics, which helps improve power. That is why the non-parametric estimators are preferred – they estimate the LRSD without constraining the high frequency dynamics, which could easily change over time, e.g. due to shifts in monetary and fiscal policy or the structure of labor markets. That said, other than the tightness of the confidence intervals, the conclusions from the parametric estimators are not qualitatively different from those obtained with the RQS estimator.

Finally, figure 4 lets us compare the RQS and $\text{RQS}_{\text{mAMSE}}$ estimators. The two estimates, especially for the post-war and Barro–Ursua samples, are rather similar. In the post-war sample, the estimated optimal bandwidth is wider than the business cycle and includes higher frequencies, which both reduces the point estimates and narrows the confidence bands. We obtain similar results in the 1843–2014 sample. In the shorter 1929–2014 sample, we obtain the opposite result – the optimal bandwidth is narrower than the business cycle. This causes the confidence intervals to become larger and less uninformative. We thus see that in practice, when possible, selecting the bandwidth based on economic intuition (as is often done with the Newey–West estimator, for example) seems to lead to more stable results across samples.

5.3 International data

To help understand whether the experience in the United States has been anomalous, table 2 reports estimates of LRSDs from Barro and Ursua's (2010) panel of international data on real consumption growth for 42 countries. For each country, I use data for the longest continuously available sample. Because the data is often anomalously volatile when countries initially enter the panel, I eliminate the first ten years of each country's sample. Both of these filters are likely to bias estimates of risk downward.

Table 2 reports percentiles of the distribution of estimates of the LRSD across countries and also the rank of the US in that distribution. In both the full and the post-war sample, the US estimate of the LRSD is well below the median, ranking 33rd and 36th out of 42 countries, respectively. The median estimate of the LRSD in the full sample is 5.14 percent. Consistent with the US experience, though, the distribution of the LRSD shifts substantially downward in the post-war sample, with the 25th, 50th, and 75th percentiles falling by 0.8 to 1.5 percentage points. So while the US has had an anomalously low LRSD compared to most of the rest of the world, its post-war decline is similar in direction and magnitude.

5.4 Testing the benchmark calibrations

The analysis so far focuses purely on estimating the LRSD of consumption. The results therefore depend on the validity of the assumptions underlying the various estimators. If the only thing we want to do is test the calibrations used in particular models, though, whether the estimators and their confidence intervals are accurate is largely irrelevant. We can simply ask how likely we would be to see the estimates we obtained with the various estimators in the data if we simulated the original calibrations in the long-run risks models.

I consider the performance of simulations of the BY and BKY models in comparison to two empirical estimates: the RQS kernel and an ARMA(1,1). I choose the ARMA(1,1) because it is the exact model (though without the stochastic volatility) that BY and BKY use. I simulate the BY and BKY models (in the forms with stochastic volatility) for samples equivalent to the three empirical samples – 272 quarters, 85 years, and 177 years. The initial states of the model – expected consumption growth and the conditional variance – are drawn from their unconditional distributions. I then form each of the estimators in 10,000 samples.

Table 3 summarizes the results of the simulations of the benchmark calibrations. Both calibrations struggle to match the post-war sample. For the RQS estimator, the BY calibration generates a value as small as we observe in the empirical sample in only 2.6 percent of the simulations, while in the BKY calibration the simulated RQS LRSD estimate is as small as the empirical value 13.0 percent of the time. If these calibrations are true, then, the post-war period was one of extraordinarily low volatility – one that should have occurred only once every 2600 years for BY or once every 500 years for BKY.

In the post-1834 sample, the BY calibration matches the data 9.2 percent of the time, while the BKY calibration does 34.5 percent of the time. We obtain similar results with the ARMA(1,1) measure of the LRSD and in the post-1929 sample. Table 3 thus suggests that while the BY calibration is on the edge of being reasonable given the data, the BKY calibration is much more plausible, especially for the longer samples. While the post-war period is challenging for these models, Table 3 and Figure 4 show that the LRSD in the pre-1930 period is well matched by long-run risks models.

6 Implications of stochastic volatility

Some long-run risks models place a special emphasis on stochastic volatility. In section 2 we showed that the RQS kernel estimator is theoretically robust to stochastic volatility. Moreover, in simulations of Bansal and Yaron's (2004) calibration with stochastic volatility, figures 2 and 3 show that the point estimates and confidence intervals for the RQS estimator are accurate. So for the purposes of estimating the LRSD, standard calibrations of stochastic volatility have little effect.

An important question, though, is how accurate the approximation in equation (2) remains when consumption growth has stochastic volatility. Shocks to volatility affect the pricing kernel and hence the Hansen–Jagannathan (HJ) bound and price of risk. So when volatility varies, can we still use risk aversion multiplied by the LRSD as a good approximation for the HJ bound?

To answer that question, figure 5 plots the HJ bound in the BY and BKY models, both with and without stochastic volatility, for varying amounts of long-run risk. I allow the LRSD to vary between 2.5 and 6.5 percent per year. As in the simulations summarized in figures 2 and 3, I hold the unconditional standard deviation and autocorrelation (USD and $AC1$) identical to their original values in BY and BKY (though note that in Figure 5 the x-axis gives the LRSD instead of the ratio of the LRSD to the USD as in figure 3). When we retain stochastic volatility, σ_w in equation (34) is set to its original value in BY and BKY so that the volatility of volatility is held fixed in proportional terms (without stochastic volatility, $\sigma_w = 0$).

Figure 5 shows that the approximation in equation (2) is accurate across a range of values of the LRSD in the BY model, and that the inclusion of stochastic volatility does little to affect the results.

For the BKY model without stochastic volatility, the approximation $\alpha \times LRSD$ also works well. For the BKY model *with* stochastic volatility, the approximation is less accurate – the HJ bound is substantially higher than predicted, but the LRSD still describes very accurately how the HJ bound varies across calibrations.

In BKY, then, knowing only the LRSD is not sufficient for knowing the price of risk in the economy. But that statement is true more generally: if there were disaster risk, or time-varying risk aversion, or any number of other sources of variation in the economy, the LRSD would also not be a sufficient statistic for the HJ bound. What figure 5 shows, though, is that even in the BKY model, where volatility is extremely volatile and persistent – its half-life is 57 years and its standard deviation is 1.2 times its mean – the LRSD is useful for understanding how the price of risk varies across calibrations. The extent to which variation in the level of volatility affects the HJ bound obviously depends on the average level of volatility. So for a large range of important models in the literature, the LRSD is a key statistic to match, even if it is not the only important number. Moreover, the majority of the literature considers homoskedastic Gaussian models, where the LRSD is in fact a sufficient statistic.

7 Conclusion

This paper argues that the long-run standard deviation (LRSD) of consumption growth is a critical moment in determining the behavior of asset-pricing models. In a limiting approximation where a household with Epstein–Zin preferences is indifferent to the date on which consumption occurs (has an infinite EIS and a rate of time preference approaching zero), the standard deviation of the pricing kernel, which determines the price of risk and the maximal Sharpe ratio, is simply the product of the coefficient of relative risk aversion and the LRSD of consumption growth.

The LRSD is difficult to estimate when consumption growth has a persistent component, and I show that the confidence intervals from many estimators have poor coverage. However, the new estimator proposed here displays nearly exactly correct confidence interval coverage in simulations in addition to a superior bias/variance trade-off compared to other estimators.

I present a range of estimates of the LRSD using the new estimator. In the post-war samples, a reasonable upper limit for admissible values is 4–5 percent per year and the point estimate is 2.5 percent. In longer samples, the point estimates rise close to 4 percent per year with a 95-percent confidence interval rising as high as 6.0 percent in the longest sample. Figure 4 summarizes the results and gives a guide to future calibrations of models with Epstein–Zin preferences.

References

Abel, Andrew B., "Asset Prices under Habit Formation and Catching up with the Joneses," *The American Economic Review, Papers and Proceedings*, 1990, *80(2)*, 38–42.

Bailie, Richard T., "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 1996, *73(1)*, 5–59.

Bansal, Ravi and Amir Yaron, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 2004, *59* (4), 1481–1509.

___ **, Dana Kiku, and Amir Yaron**, "Long-Run Risks, the Macroeconomy, and Asset Prices," *American Economic Review*, 2010, *100(2)*, 542–546.

___ **,** ___ **, and** ___ , "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 2012, *1(1)*, 183–221.

Barberis, Nicholas, Ming Huang, and Tano Santos, "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 2001, *116(1)*, 1–53.

Barro, Robert J., "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 2006, *121(3)*, 823–866.

___ **and Jose F. Ursua**, "Barro–Ursua Macroeconomic Dataset," 2010. http://scholar.harvard.edu/barro/publications/barro-ursua-macroeconomic-data.

Beeler, Jason and John Y. Campbell, "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 2012, *1(1)*, 141–182.

Boldrin, Michele, Lawrence J. Christiano, and Jonas D. M. Fisher, "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 2001, *91(1)*, 149–166.

Brillinger, David R, "Asymptotic properties of spectral estimates of second order," *Biometrika*, 1969, *56(2)*, 375–390.

Campanale, Claudio, Rui Castro, and Gian Luca Clementi, "Asset Pricing in a Production Economy with Chew-Dekel Preferences," *Review of Economic Dynamics*, 2010, *13(2)*, 379–402.

Campbell, John Y., "Intertemporal Asset Pricing Without Consumption Data," *American Economic Review*, 1993, *83(3)*, 487–512.

—— **and John H. Cochrane**, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 1999, *107* (2), 205–251.

—— **and N. Gregory Mankiw**, "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," in Olivier Jean Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual*, 1989.

Christiano, Lawrence J. and Terry J. Fitzgerald, "The Band Pass Filter," *International Economic Review*, 2003, *44(2)*, 435–465.

Clark, Peter. K., "The Cyclical Component of U. S. Economic Activity," *Quarterly Journal of Economics*, 1987, *102(4)*, 797–814.

Cochrane, John H., "How Big Is the Random Walk in GNP?," *The Journal of Political Economy*, 1988, *96(5)*, 893–920.

—— , "Permanent and Transitory Components of GNP and Stock Prices," *The Quarterly Journal of Economics*, 1994, *109(1)*, 241–265.

—— **and Argia Sbordone**, "Multivariate Estimates of the Permanent Components of GNP and Stock Prices," *Journal of Economic Dynamics and Control*, 1988, *12(2–3)*, 255–296.

Colacito, Riccardo and Mariano M. Croce, "Risks for the Long Run and the Real Exchange Rate," *Journal of Political Economy*, 2011, *119(1)*, 153–181.

Croce, Massimiliano, "Long-Run Productivity Risk: A New Hope For Production-Based Asset Pricing?," 2010. Working paper.

—— **, Martin Lettau, and Sydney C. Ludvigson**, "Investor Information, Long-Run Risk, and the Duration of Risky Cash Flows," 2010.

Dew-Becker, Ian and Stefano Giglio, "Asset Pricing in the Frequency Domain: Theory and Empirics," 2013. Working paper.

Epanechnikov, V. A., "Non-Parametric Estimation of a Multivariate Probability Density," *Theory of Probability & Its Applications*, 1969, *14(1)*, 153–158.

Epstein, Larry G. and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *The Journal of Political Economy*, 1991, *99(2)*, 263–286.

Gourio, Francois, "Disaster Risk and Business Cycles," *American Economic Review*, 2012, *102(6)*, 2734–2766.

Hansen, Lars P., John C. Heaton, and Nan Li, "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 2008, *116(2)*, 260–302.

Hansen, Lars Peter and Ravi Jagannathan, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 1991, *99(2)*, 225–262.

Harvey, A. C., "Trends and Cycles in Macroeconomic Time Series," *Journal of Business & Economic Statistics*, 1985, *3(3)*, 216–227.

Kaltenbrunner, Georg and Lars A. Lochstoer, "Long-Run Risk Through Consumption Smoothing," *The Review of Financial Studies*, 2010, *23(8)*, 3141–3189.

Kiefer, Nicholas M. and Timothy J. Vogelsang, "A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests," *Econometric Theory*, 2005, *21(6)*, 1130–1164.

Lo, Andrew W. and A. Craig MacKinlay, "Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test," *Review of Financial Studies*, 1988, *1(1)*, 41–66.

Mehra, Rajnish and Edward C. Prescott, "The Equity Premium: A puzzle," *Journal of Monetary Economics*, 1985, *15(2)*, 145–161.

Morley, James C., "The Slow Adjustment of Aggregate Consumption to Permanent Income," *Journal of Money, Credit, and Banking*, 2007, *39(2–3)*, 615–638.

Müller, Ulrich K., "A theory of robust long-run variance estimation," *Journal of Econometrics*, 2007, *141(2)*, 1331–1352.

Newey, Whitney K. and Kenneth D. West, "Automatic Lag Selection in Covariance Matrix Estimation," *The Review of Economic Studies*, 1994, *61* (4), 631–653.

Phillips, Peter C.B., "New Tools for Understanding Spurious Regressions," *Econometrica*, 1998, *66(6)*, 1299–1325.

Politis, Dimitris N. and Joseph P. Romano, "Bias-corrected Nonparametric Spectral Estimation," *Journal of Time Series Analysis*, 1995, *16(1)*, 67–103.

Porterba, James M. and Lawrence H. Summers, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 1988, *22(1)*, 27–59.

Priestley, M. B., "Bias Considerations in Estimation of Spectra," *Technometrics*, 1962, *4(4)*, 551–564.

Restoy, Fernando and Philippe Weil, "Approximate Equilibrium Asset Prices," 1998. NBER Working Paper.

Sun, Yixiao and Peter C. B. Phillips andSainan Jin, "Optimal Bandwidth Selection in HeteroskedasticityŰAutocorrelation Robust Testing," *Econometrica*, 2008, *76* (1), 175–194.

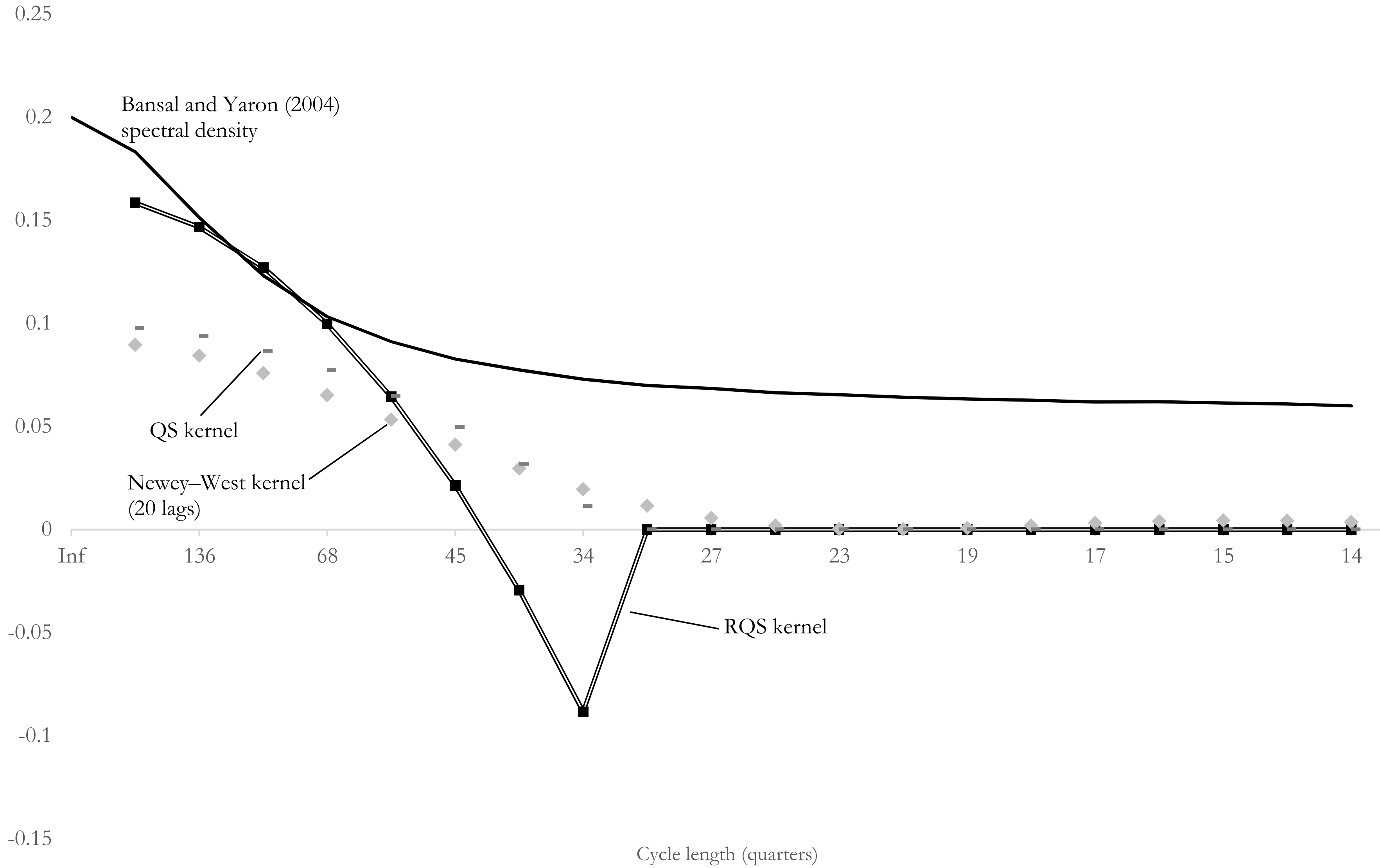
Tallarini, Thomas D., "Risk-Sensitive Real Business Cycles," *Journal of Monetary Economics*, 2000, *45(3)*, 507–532.

Wachter, Jessica A., "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?," 2010. Working Paper.

Wand, M.P. and M.C. Jones, *Kernel Smoothing*, CRC Press, 1994.

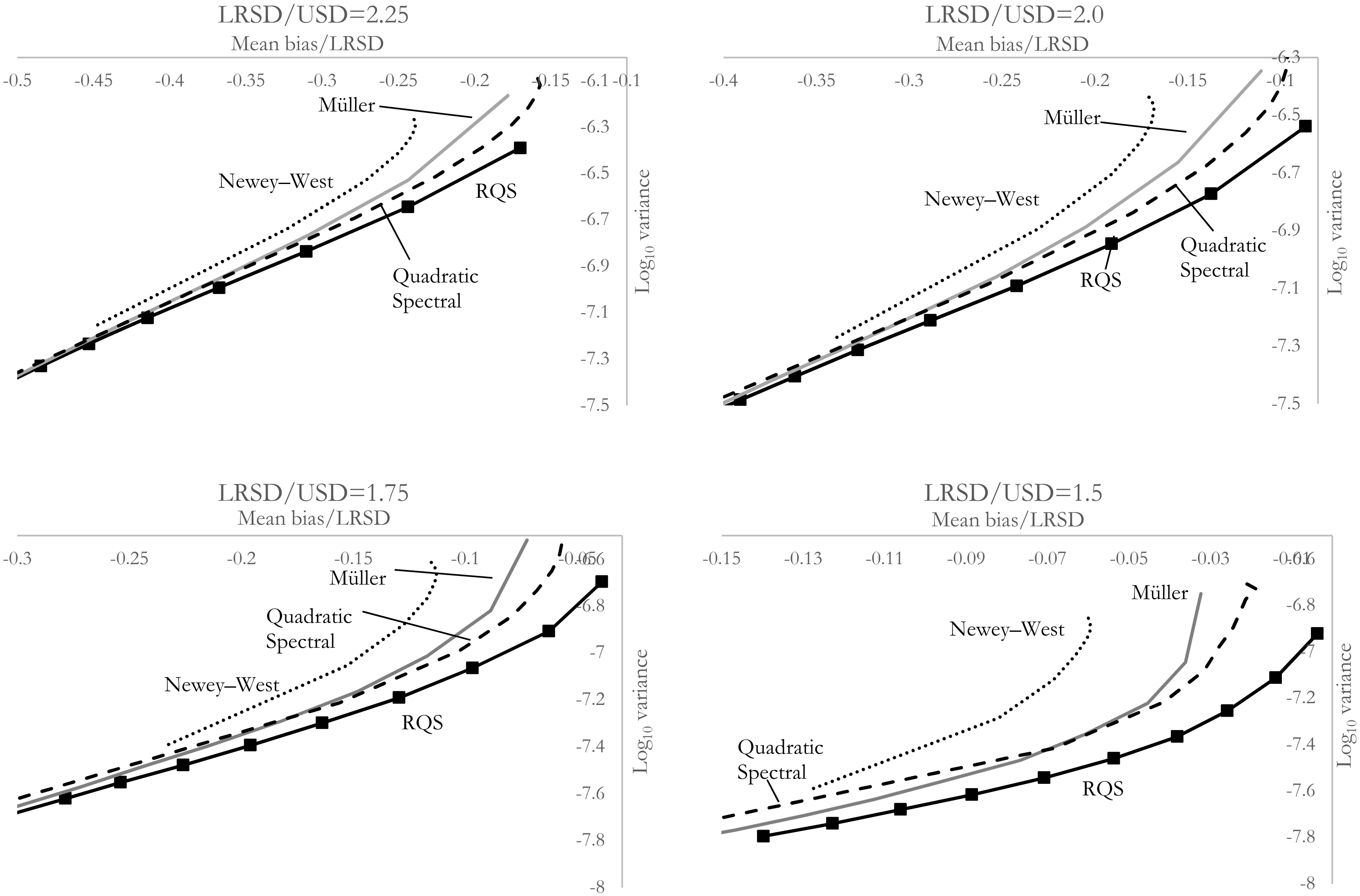
Working, Holbrook, "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica*, 1960, *28(4)*, 916–918.

Figure 1. Spectral kernels



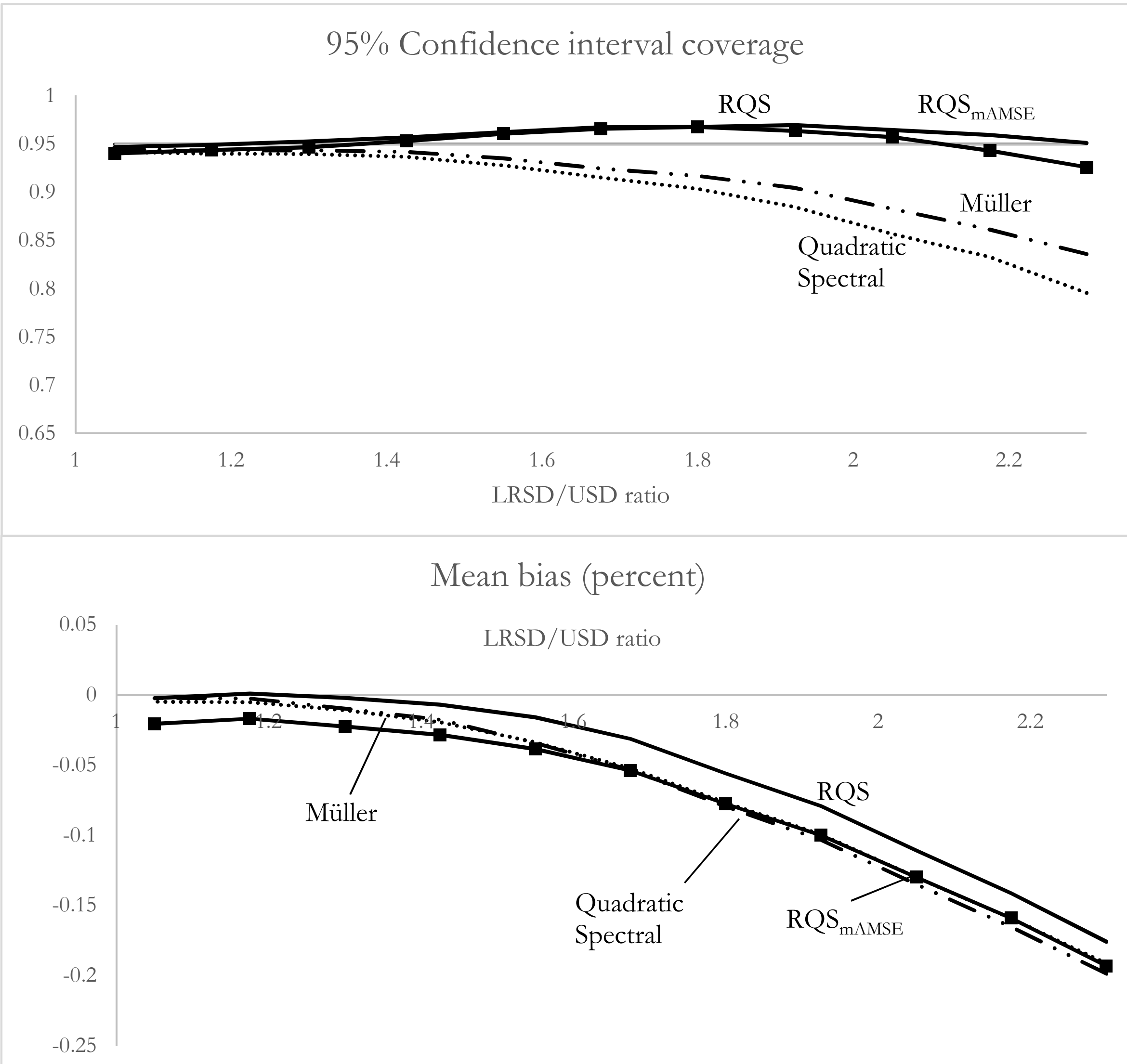
Notes: The x-axis gives cycle lengths (2π/frequency) for a 272-quarter sample. The solid line is the spectral density of consumption growth in Bansal and Yaron's (2004) model of consumption growth (rescaled to fit the plot). The diamonds are the frequency-domain kernel weights for the Newey–West estimator with 20 lags. The dashes are for the QS kernel. The double line is the UQS kernel. The points plotted for the kernels are the Fourier frequencies.

Figure 2. Bias/variance tradeoffs for different LRSDs



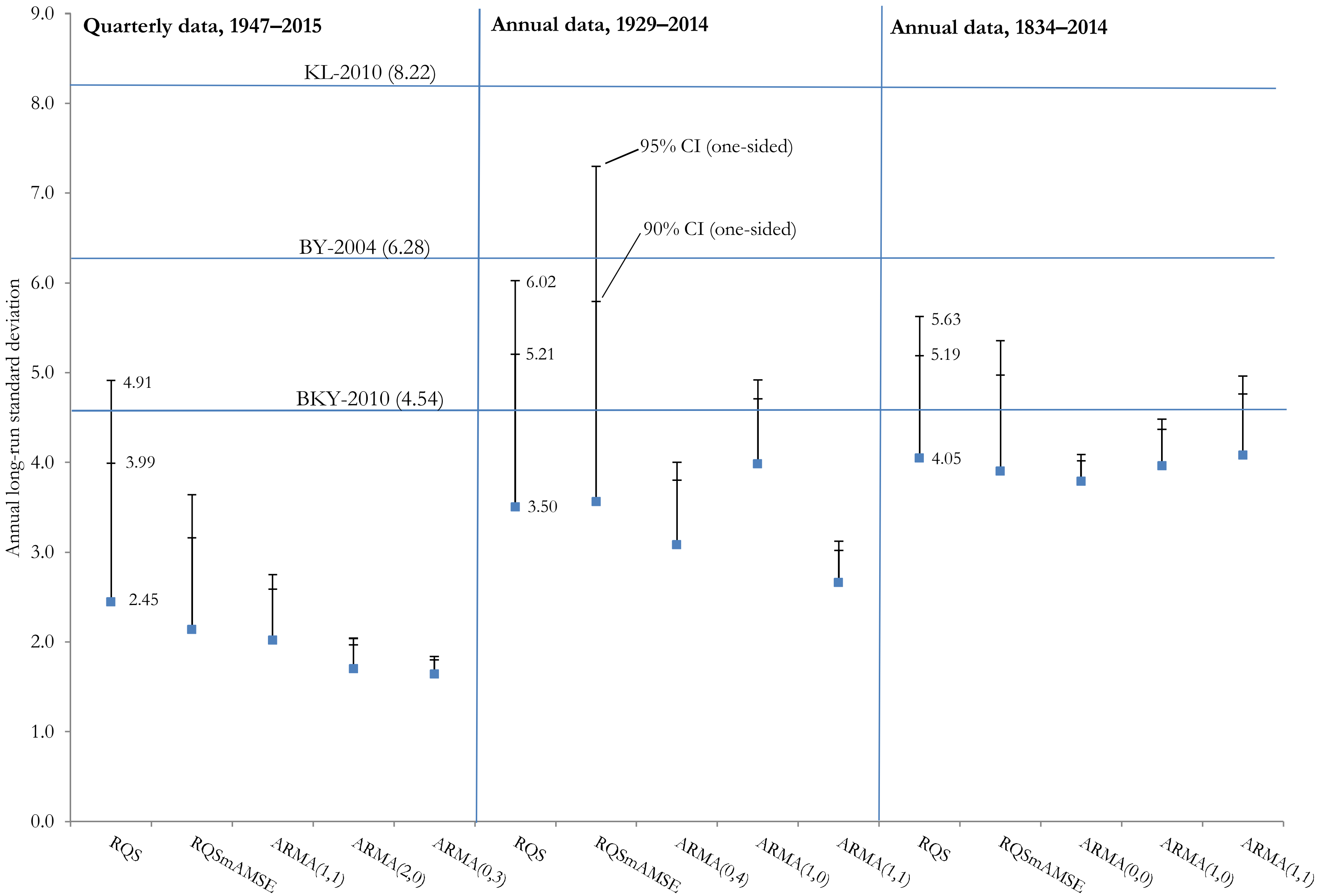
Notes: Each panel plots the bias/variance tradeoff for the four estimators of the LRSD. The lines result from varying the bandwidth of each estimator. The results are from simulations of Bansal and Yaron's (2004) model allowing the long-run/unconditional standard deviation ratio to vary.

Figure 3. Simulated estimator performance



Notes: Confidence interval coverage and mean bias for four LRSD estimators. The bandwidths are chosen so that each estimator has the same variance as the UQS estimator. The confidence interval coverage is for a one-sided (on the high side) confidence interval. UQS-mMSE uses the estimated minimum-MSE bandwidth in each sample.

Figure 4. Long-run standard deviation estimates and confidence intervals

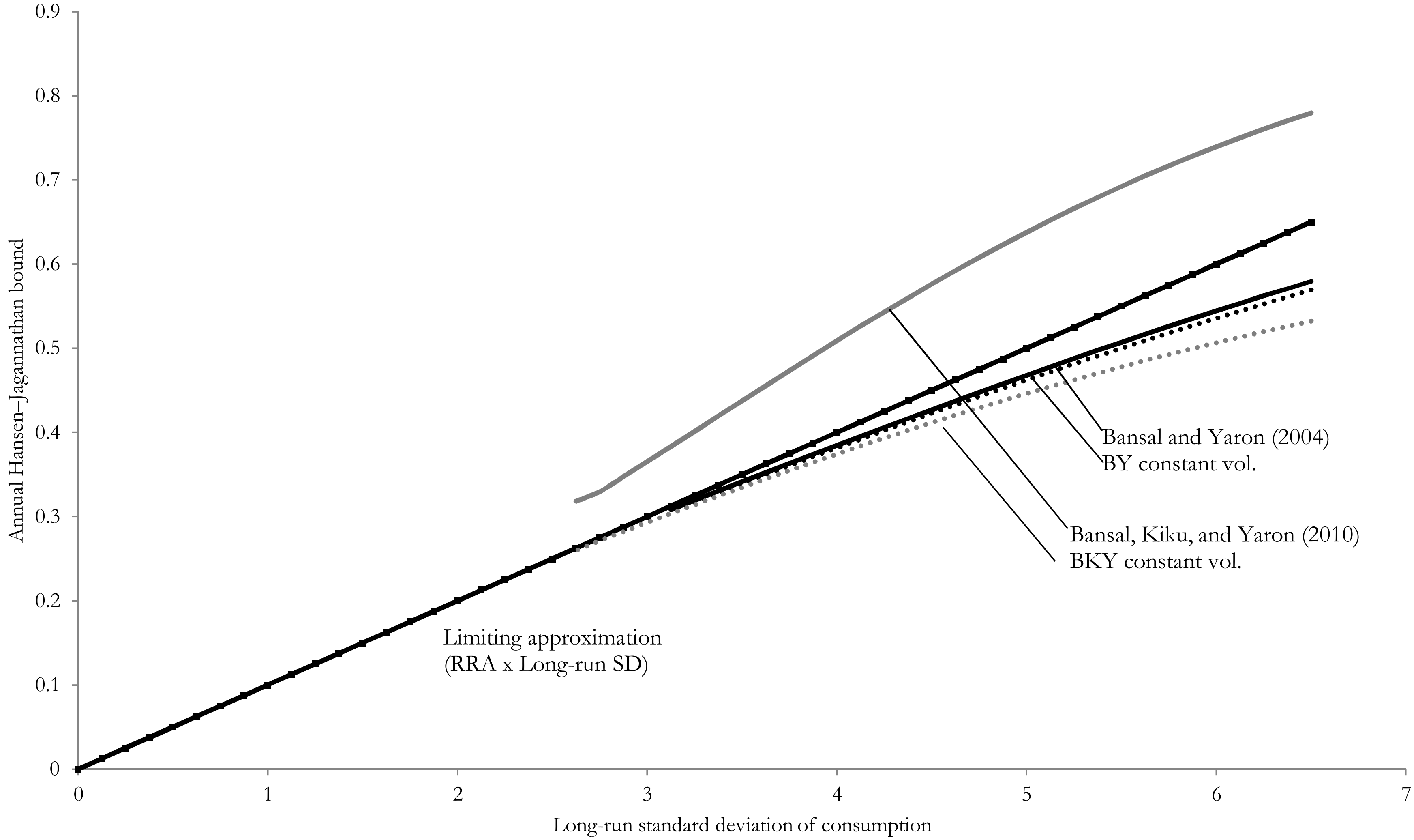


See notes on the following page.

Notes to Figure 4.

This figure summarizes point estimates and confidence intervals for the long-run standard deviation of consumption growth from a variety of estimators. The vertical lines separate results from the three data samples: post-war quarterly, post-1929 annual, and post-1934 annual (Barro and Ursua, 2010). The first two samples are non-durables and services consumption per capita; the third is total consumption per capita. The three horizontal lines give LRSD's from benchmark calibrations in the long-run risks literature, Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2010), and Kaltenbrunner and Lochstoer (2011). The squares give point estimates, and the vertical lines give the one-sided 90- and 95-percent confidence intervals. The RQSmAMSE estimator estimates the optimal bandwidth in each sample using 20 autocovariances. The three ARMA models used for each sample are chosen as those that minimize the Bayesian information criterion.

Figure 5. The long-run standard deviation of consumption and the Hansen–Jagannathan bound



Notes: The quarterly Hansen–Jagannathan bound in the models in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2010). The x-axis gives the LRSD for consumption. The parameters driving the consumption process in the model are identified by three moments: the LRSD, the unconditional standard deviation, and the first autocorrelation of consumption growth. The second two moments are held fixed as the LRSD varies. The limiting approximation gives the Hansen–Jagannathan bound for an agent with Epstein–Zin preferences and a high EIS.

Table 1. Recent calibrations of the long-run standard deviation of consumption growth (annualized)

	Long-run SD	Moments matched
Campbell and Cochrane (1999)	1.50	SD(dc) 1947-1995
Gourio (2010)	2.00	SD(TFP), 1947-2010
Barro(2006), Wachter (2010)	2.00	SD(dy) 1954-2004, international
Tallarini (2000)	2.30	SD(dc), 1948-1993
Mehra and Prescott (1985)	3.16	SD(dc) 1889-1978
Boldrin, Christiano, and Fisher (2001)	3.60	Various unconditional SDs, 1964-1988
Abel (1990)	3.60	SD(dc) 1889-1978
Barberis, Huang, and Santos (2001)	3.80	SD(dc), 1889-1985
Bansal, Kiku, and Yaron (2008)	4.54	Annual SD(dc), autocorrelations, 1929-2008
Drechsler and Yaron (2011)	4.83	Annual SD(dc), autocorrelations, 1929-2006
Campanale, Castro, and Clementi (2010)	5.20	SD(dy), 1947-2005
Bansal and Yaron (2004); Croce, Lettau, and Ludvigson (2010)	6.28	Annual SD(dc), autocorrelations, 1929-1998
Croce (2010)	8.05	Annual SD(dTFP), 1947-2010
Kaltenbrunner and Lochstoer (2010)	8.22	SD(dc), SD(dc)/SD(dy)
Colacito and Croce (2011)	9.02	SD(dc), currency movements

Table 2. Distribution of international LRSD estimates

	25%	50%	75%	Rank of US estimate (out of 42)
Full sample	4.30	5.14	6.72	33
Post-war	2.95	3.61	5.93	36

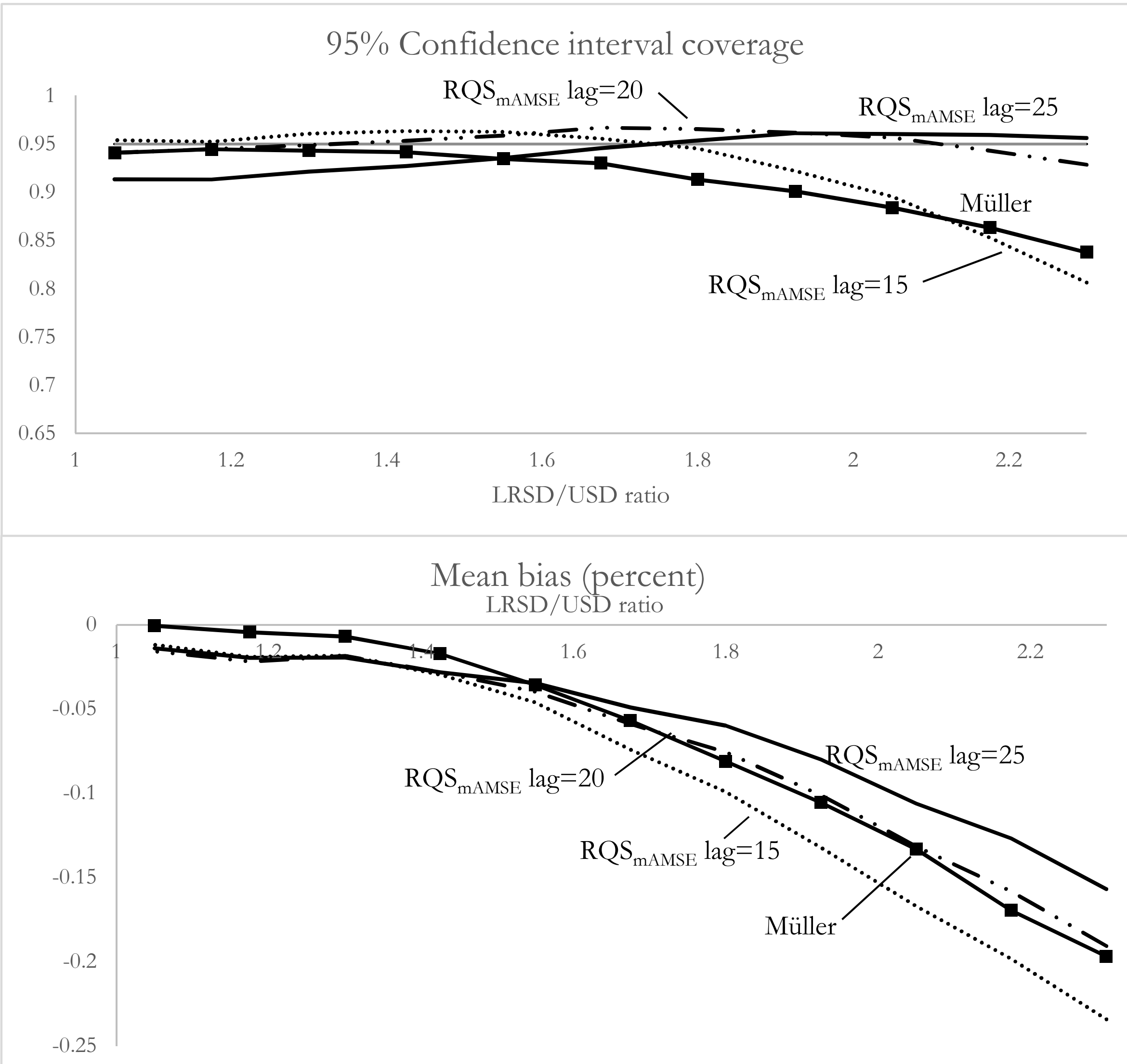
Notes: Distribution of estimates of the LRSD from Barro and Ursua's (2010) international panel using the RQS estimator.

Table 3. Fraction of simulated samples with LRSDs as low as observed empirically

	post-war		post-1929		post-1834	
	RQS	ARMA(1,1)	RQS	ARMA(1,1)	RQS	ARMA(1,1)
Bansal and Yaron (2004)	2.3	0.5	9.9	12.3	7.8	12.4
Bansal, Kiku, and Yaron (2012)	12.2	18.9	30.2	35.4	33.9	40.0

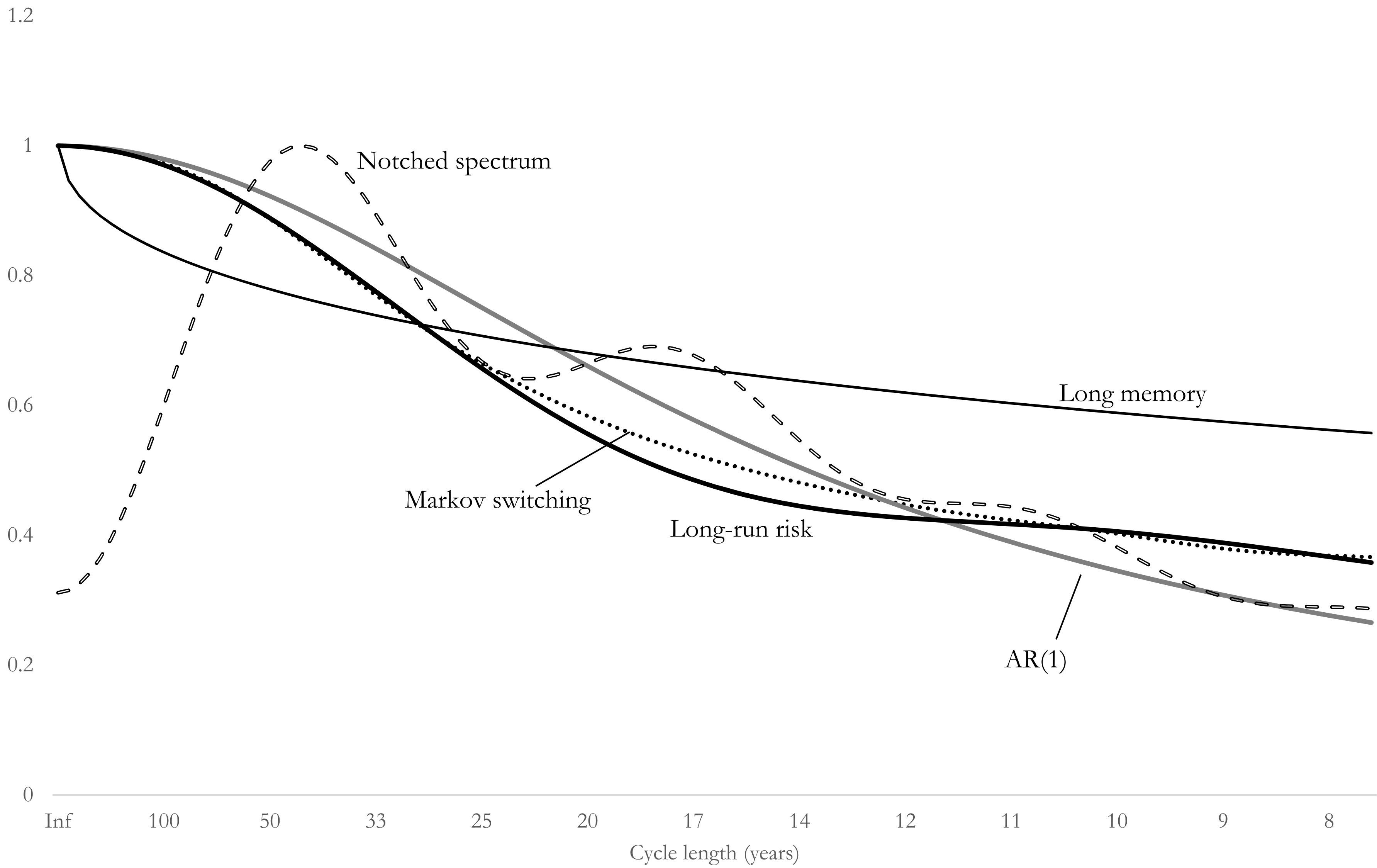
Notes: The numbers are the percentages of simulated samples that yield estimates of the LRSD as low as observed empirically. For each sample, the simulations are of identical length to the data.

Figure A1. Estimating normalized curvature



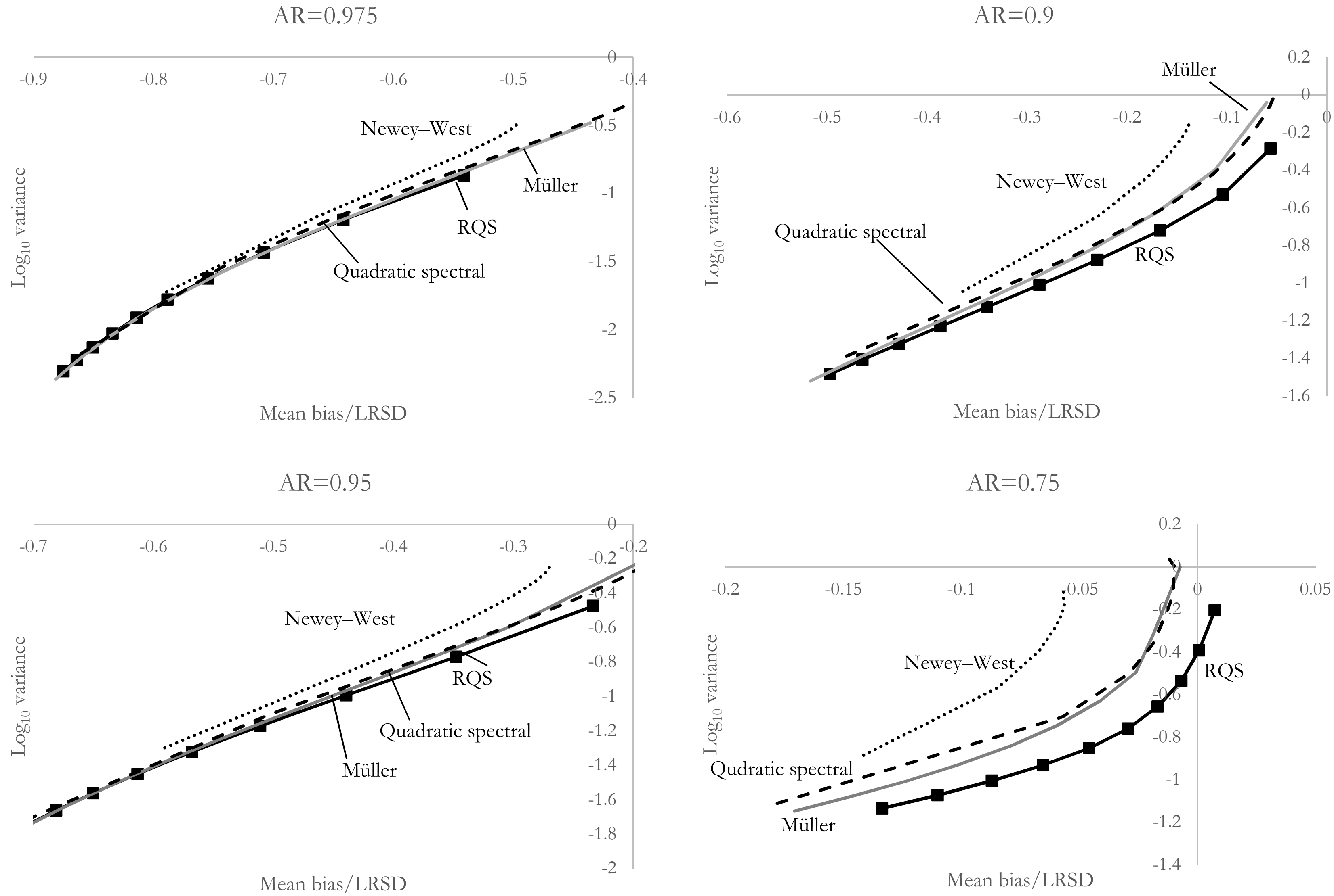
Notes: Confidence interval coverage and mean bias for four LRSD estimators. The three versions of the RQS estimator use different lags in estimating the normalized curvature. The Müller estimator uses seven Karhunen–Loeve points as in figure 3.

Figure A2. Spectra for simulations



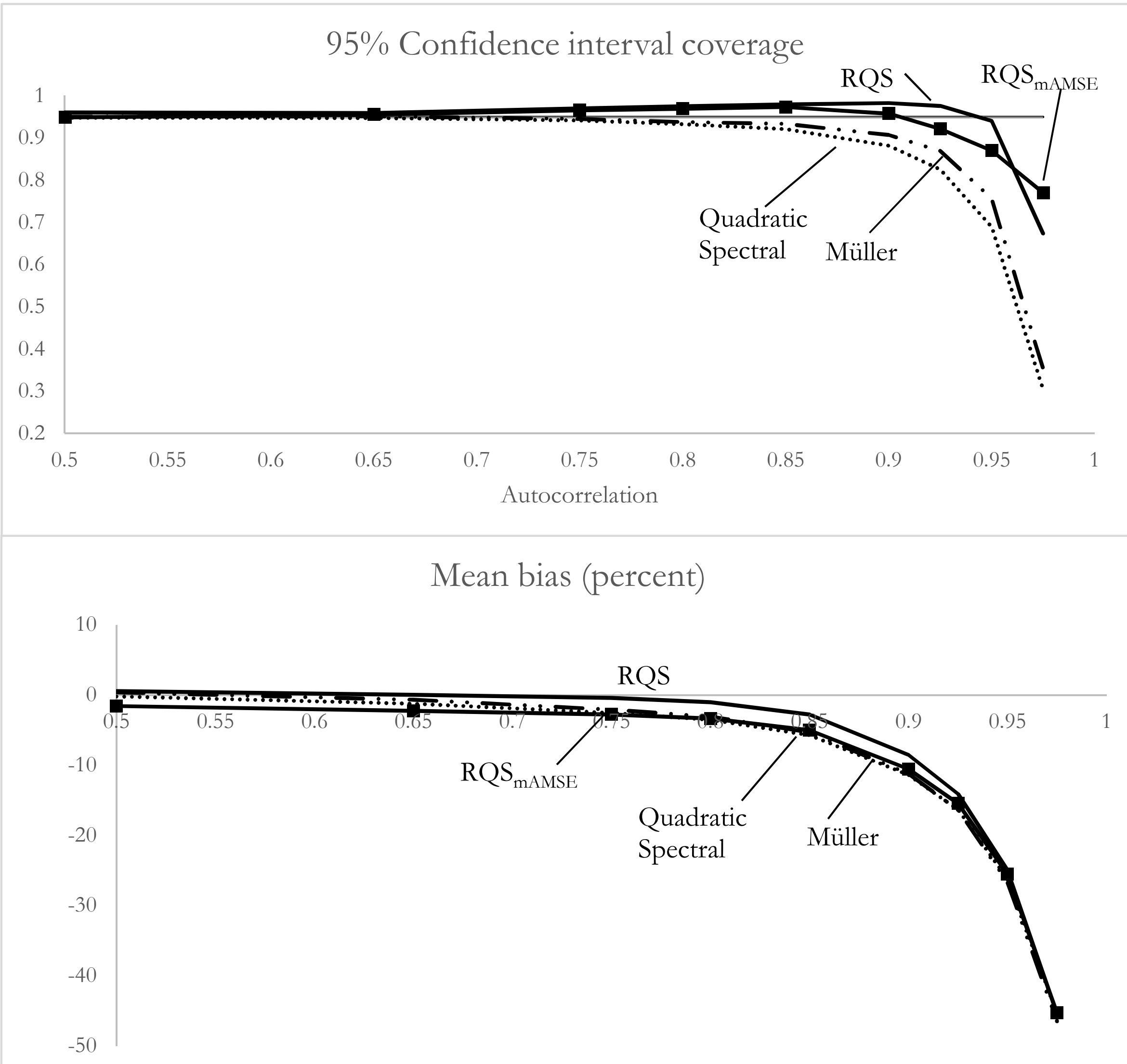
Notes: Spectra for calibrations of the five models used in the simulations. The spectra are all normalized so that their maximum value is 1. The long-run risk model is calibrated with Bansal and Yaron's (2004) original parameters. The long-memory model sets φ=0.4. The AR(1) model uses an autocorrelation of 0.9. The Markov switching model is calibrated to that the transition probability for the trend, x, is 3 percent per quarter. The notched spectrum model has R=2.

Figure A3. Bias/variance tradeoffs in an AR(1)



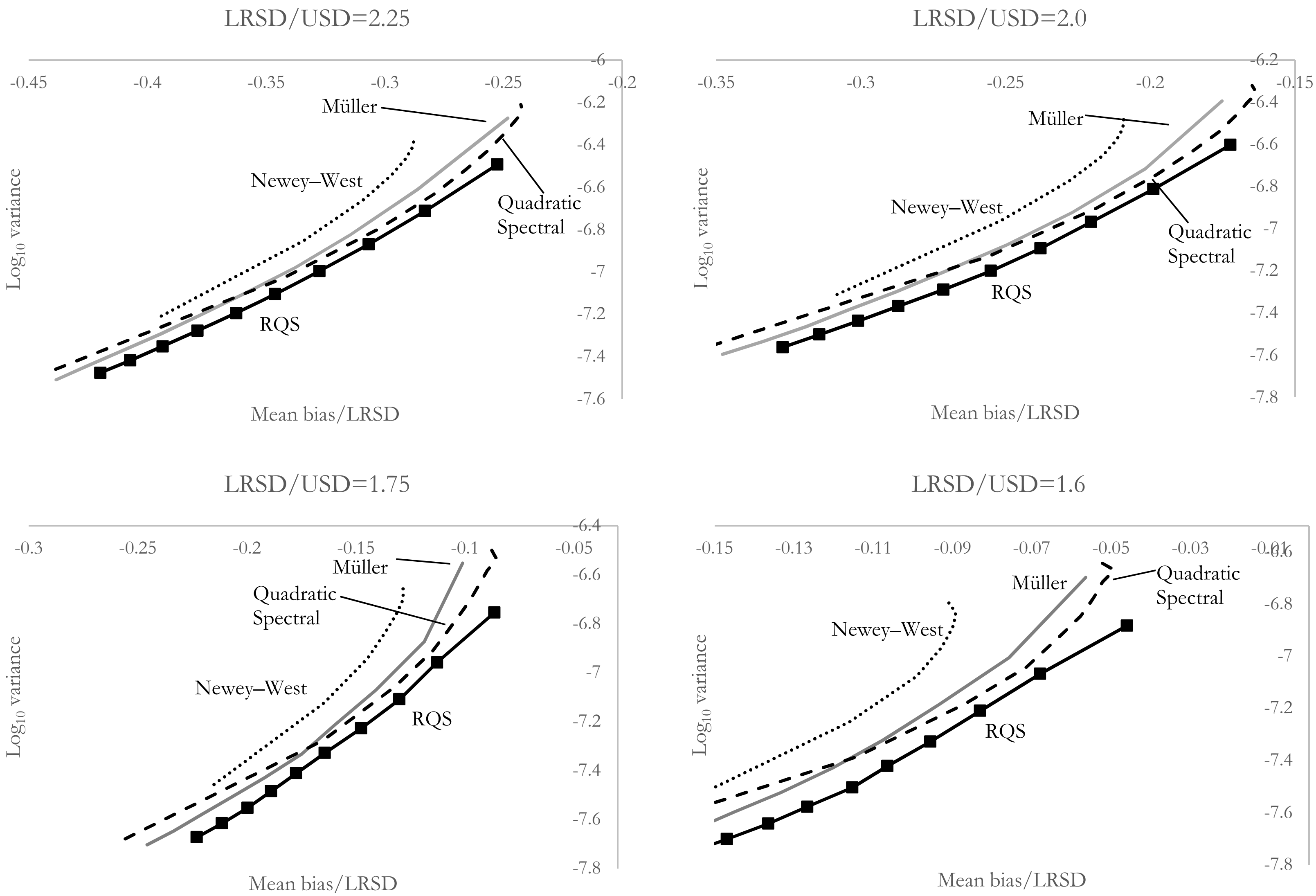
Notes: Each panel plots the bias/variance tradeoff for the four estimators of the LRSD. The lines result from varying the bandwidth of each estimator. The results are from simulations of an AR(1) model with different autocorrelations in each panel.

Figure A4. Performance in an AR(1) model



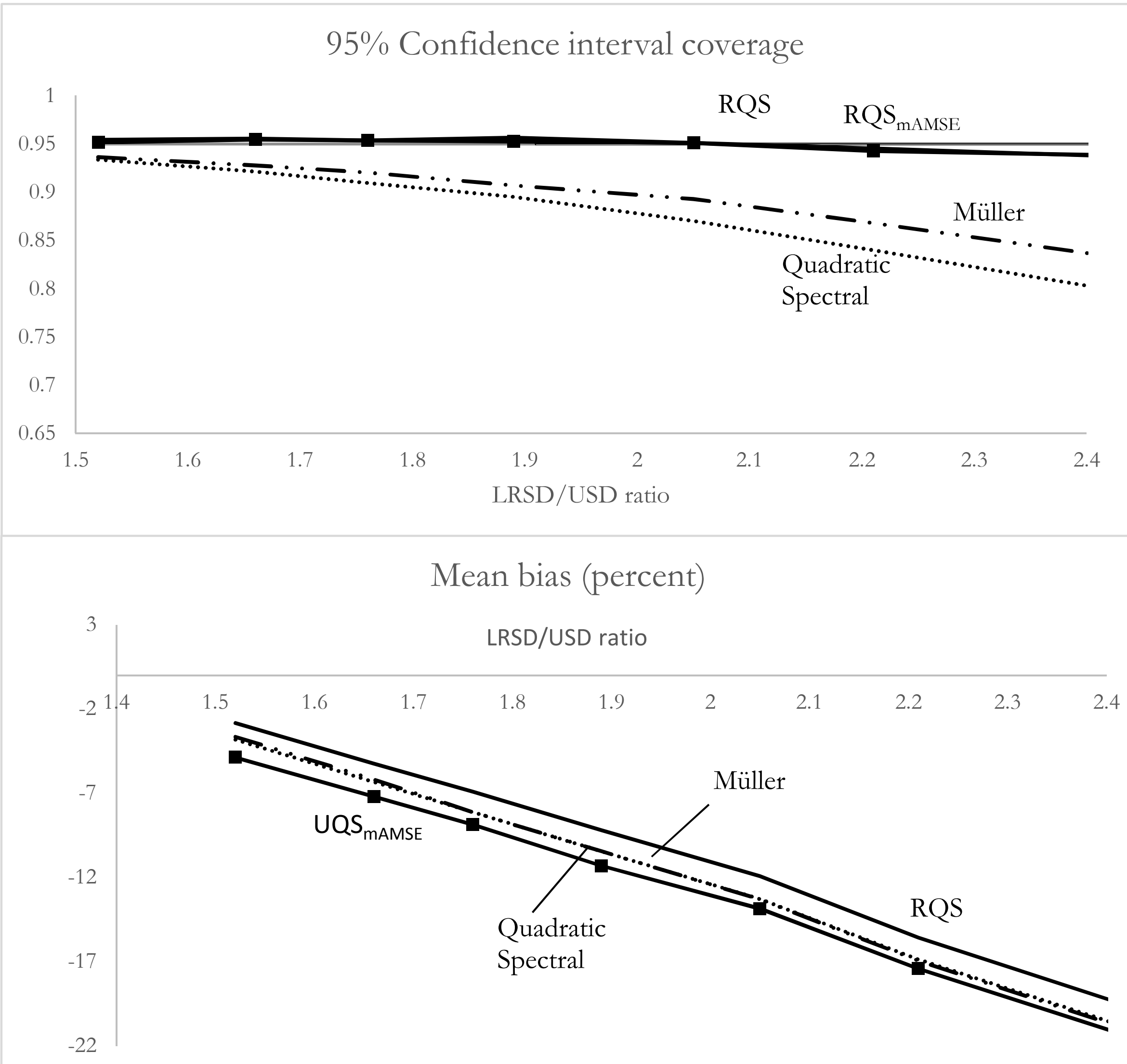
Notes: Confidence interval coverage and mean bias for four LRSD estimators. The bandwidths are set as in figure 3. The confidence interval coverage is for a one-sided (on the high side) confidence interval. The x-axis gives the autoregressive coefficient.

Figure A5. Bias/variance tradeoffs for long memory



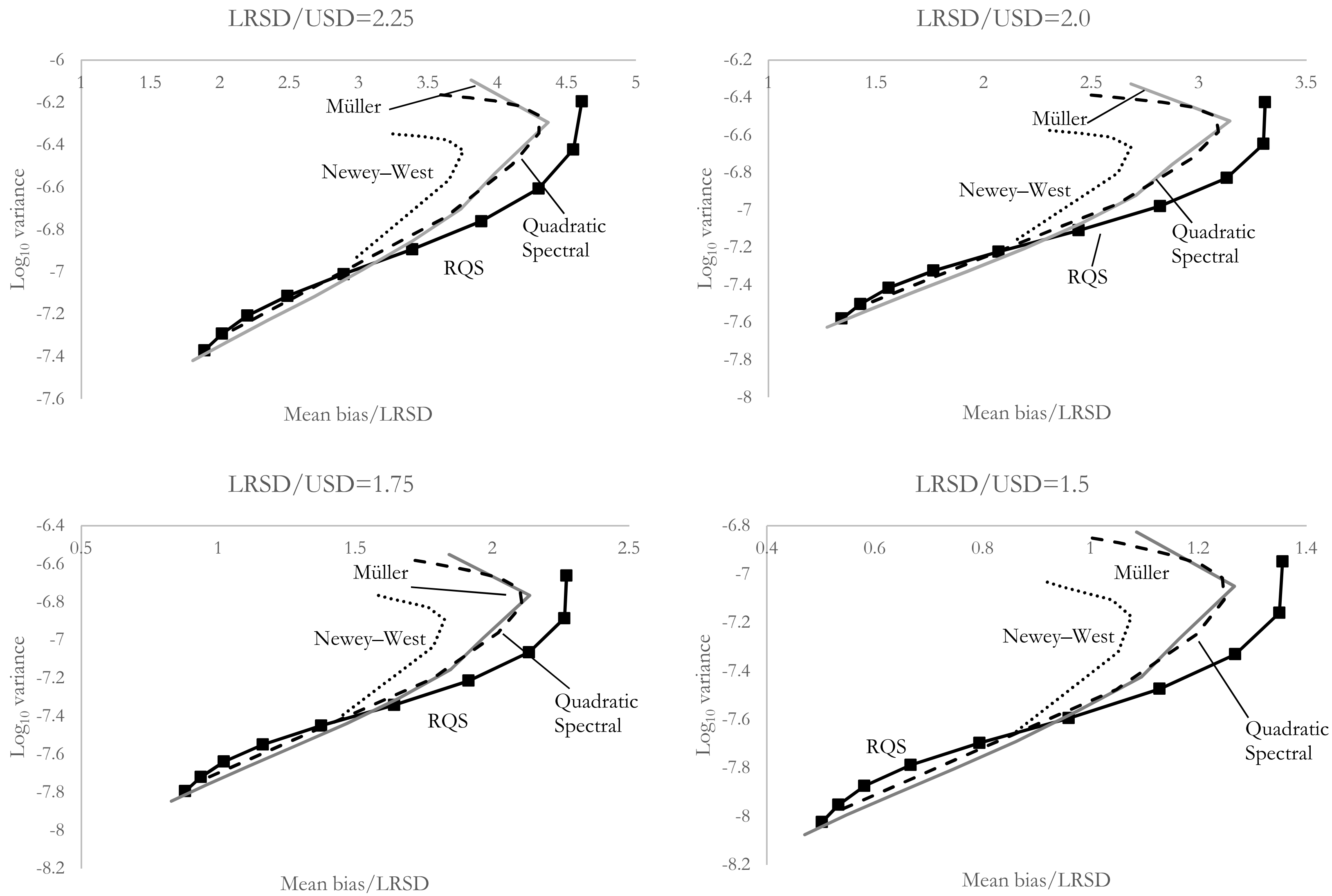
Notes: Each panel plots the bias/variance tradeoff for the four estimators of the LRSD. The lines result from varying the bandwidth of each estimator. The results are from simulations of the long-memory model allowing the long-run/unconditional standard deviation ratio to vary by changing the parameter ϕ.

Figure A6. Long memory simulations



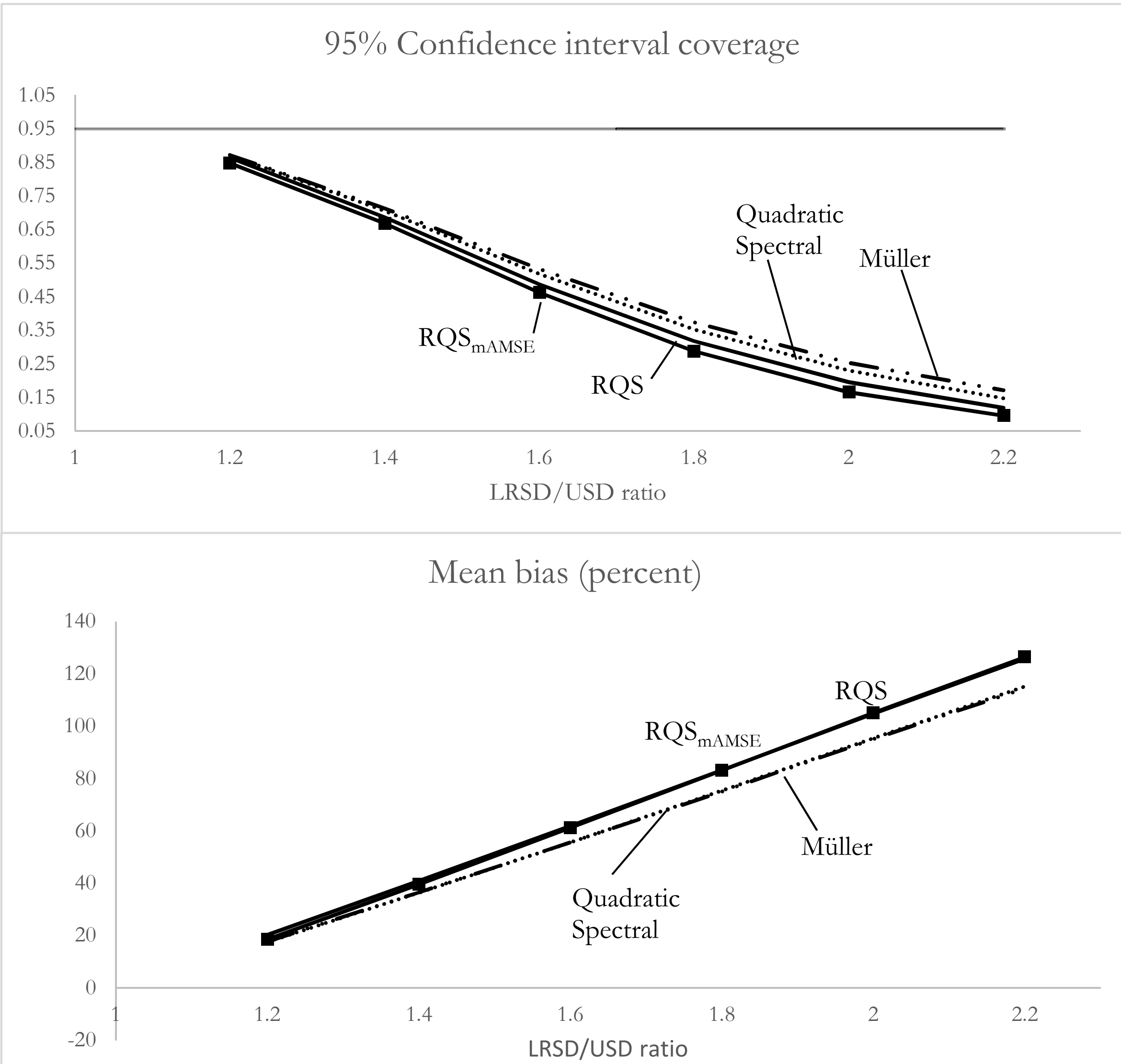
Notes: Confidence interval coverage and mean bias for four LRSD estimators. The bandwidths are chosen as in figure 3. The confidence interval coverage is for a one-sided (on the high side) confidence interval.

Figure A7. Bias/variance tradeoffs for notched spectrum



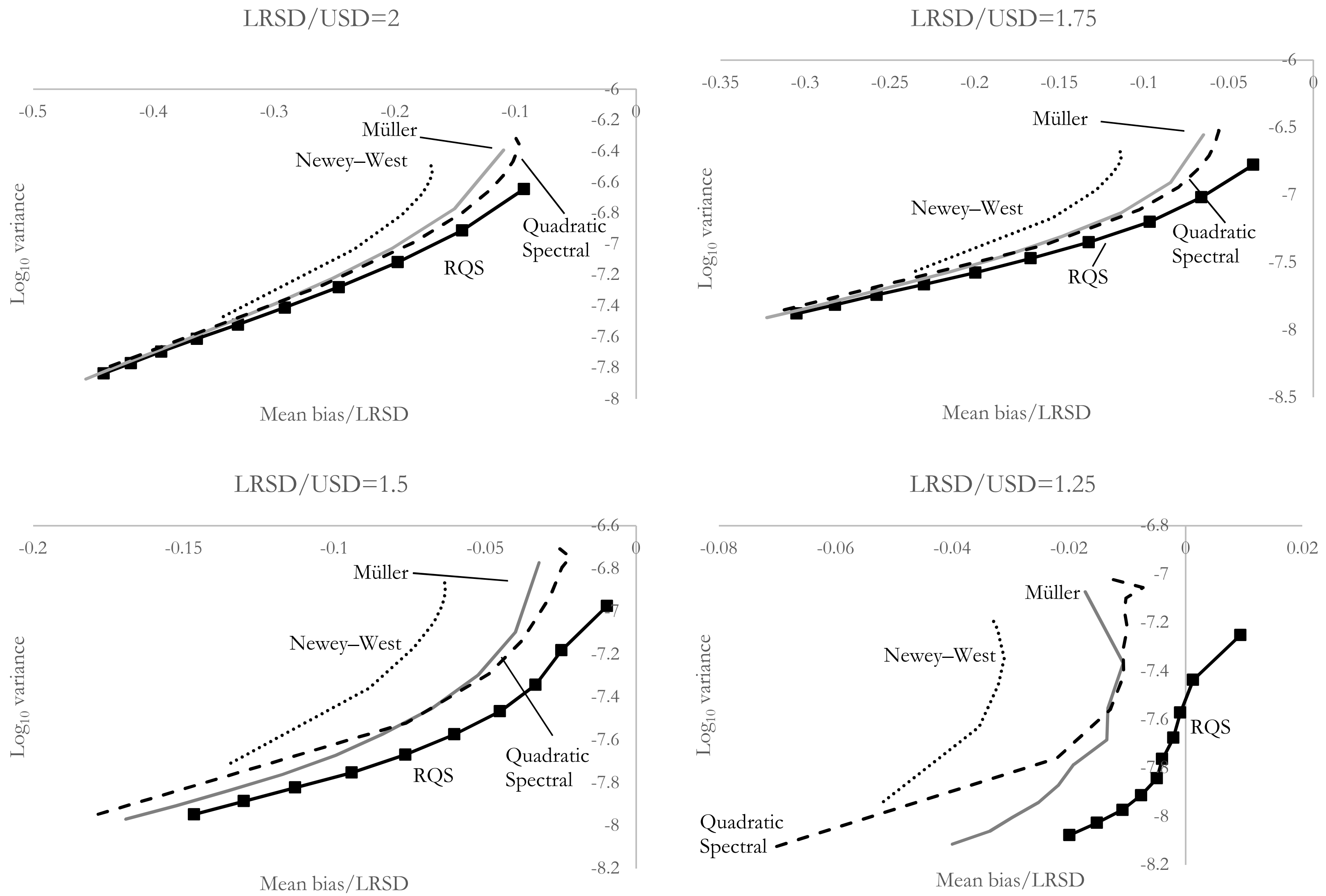
Notes: Each panel plots the bias/variance tradeoff for the four estimators of the LRSD. The lines result from varying the bandwidth of each estimator. The results are from simulations of the notched spectrum model allowing the long-run/unconditional standard deviation ratio to vary by changing the parameter R.

Figure A8. Notched spectrum simulations



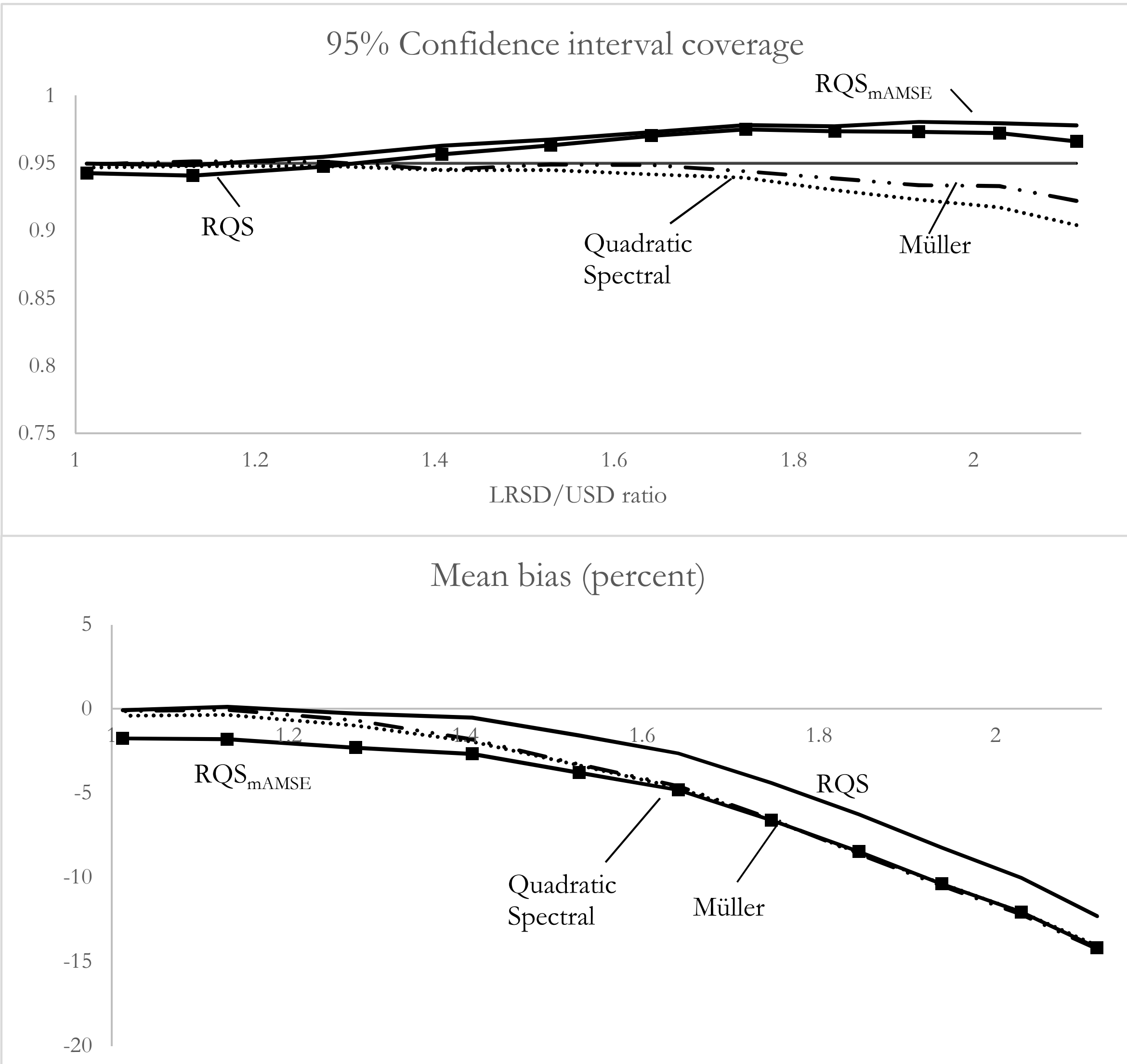
Notes: **Two-sided** confidence interval coverage and mean bias for four LRSD estimators. The bandwidths are chosenas in figure 3. The confidence interval coverage is for a two-sided confidence interval.

Figure A9. Bias/variance tradeoffs for Markov switching model



Notes: Each panel plots the bias/variance tradeoff for the four estimators of the LRSD. The lines result from varying the bandwidth of each estimator. The results are from simulations of the markov switching model allowing the long-run/unconditional standard deviation ratio to vary by changing the transition probability.

Figure A10. Markov switching model simulations



Notes: Confidence interval coverage and mean bias for four LRSD estimators. The bandwidths are chosen as in figure 3. The confidence interval coverage is for a one-sided (on the high side) confidence interval.

Table A1. Simulation results

	Mueller			Mueller (Infeasible)			QS			RQS		
	# of KLT points			# of KLT points			Lags/sample size			Minimum cycle length		
	1	2	8	1	2	8	0.621	0.13	0.064	8	12	16
AR(1) simulation												
Bias												
$\varrho=0$	0.04	0.05	0.04	0.07	0.06	0.07	0.05	0.03	0.01	0.02	0.02	0.03
$\varrho=-0.7$	0.12	0.11	0.12	0.11	0.11	0.12	0.14	0.07	0.07	0.05	0.06	0.06
$\varrho=0.7$	-0.01	-0.03	-0.15	0.05	0.03	-0.07	-0.02	-0.12	-0.29	-0.21	-0.11	-0.06
$\varrho=0.9$	-0.16	-0.24	-0.59	0.00	-0.07	-0.47	-0.69	-0.59	-0.73	-0.67	-0.55	-0.46
Root mean squared error												
$\varrho=0$	1.49	1.05	0.52	1.52	1.07	0.55	1.21	0.54	0.37	0.46	0.56	0.66
$\varrho=-0.7$	1.58	1.12	0.58	1.59	1.11	0.58	1.28	0.56	0.39	0.49	0.59	0.69
$\varrho=0.7$	1.39	0.98	0.45	1.47	1.01	0.47	1.13	0.50	0.41	0.39	0.47	0.57
$\varrho=0.9$	1.19	0.79	0.63	1.41	0.93	0.56	0.78	0.65	0.74	0.69	0.61	0.58
Coverage rate of 90% CI												
$\varrho=0$	0.90	0.89	0.89	0.90	0.89	0.88	0.90	0.90	0.90	0.90	0.89	0.90
$\varrho=-0.7$	0.89	0.89	0.88	0.89	0.89	0.87	0.90	0.90	0.89	0.88	0.88	0.88
$\varrho=0.7$	0.90	0.90	0.90	0.89	0.89	0.90	0.90	0.89	0.76	0.93	0.95	0.94
$\varrho=0.9$	0.91	0.92	0.54	0.90	0.90	0.70	0.90	0.64	0.10	0.36	0.79	0.94
MA(1) simulation												
Bias												
$\theta=-0.7$	-0.01	-0.02	-0.01	0.01	0.01	0.02	-0.01	-0.01	-0.03	0.00	0.00	0.00
$\theta=0.5$	0.08	0.09	0.13	0.06	0.06	0.11	0.10	0.09	0.18	0.05	0.04	0.05
$\theta=0.7$	0.33	0.35	0.51	0.20	0.20	0.36	0.39	0.34	0.72	0.23	0.19	0.18
$\theta=0.9$	3.72	3.81	5.83	1.89	1.92	3.71	4.28	3.87	8.47	3.79	2.78	2.40
Root mean squared error												
$\theta=-0.7$	1.40	0.99	0.49	1.44	1.02	0.52	1.14	0.53	0.36	0.44	0.54	0.64
$\theta=0.5$	1.52	1.10	0.59	1.48	1.06	0.58	1.24	0.57	0.45	0.55	0.62	0.70
$\theta=0.7$	1.93	1.39	0.94	1.71	1.22	0.79	1.56	0.74	0.93	0.89	0.85	0.88
$\theta=0.9$	7.51	6.09	7.49	4.47	3.88	5.04	6.74	4.77	9.20	6.45	4.97	4.50
Coverage rate of 90% CI												
$\theta=-0.7$	0.90	0.90	0.90	0.90	0.90	0.89	0.90	0.90	0.90	0.91	0.90	0.90
$\theta=0.5$	0.89	0.89	0.87	0.90	0.89	0.87	0.90	0.90	0.86	0.81	0.85	0.87
$\theta=0.7$	0.87	0.85	0.74	0.88	0.88	0.80	0.89	0.83	0.50	0.58	0.72	0.78
$\theta=0.9$	0.60	0.45	0.07	0.72	0.65	0.16	0.46	0.10	0.00	0.08	0.15	0.23

Notes: Results from simulations of AR(1) and MA(1) models with 100 observations. In all cases, the true LRSD is 1. The infeasible Müller estimator assumes that the mean of the process is known so that it is not demeaned. The confidence interval for the QS estimator is based on the fixed-b asymptotics of Kiefer and Vogelsang (2005). The point estimate for the QS estimator includes the bias correction implied by the fixed-b distribution (dividing by the mean of that distribution to give an unbiased estimator).