Taylor & Francis
Taylor & Francis Group

Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: http://www.tandfonline.com/loi/ubes20

Comment on "HAR Inference: Recommendations for Practice" by E. Lazarus, D. J. Lewis, J. H. Stock and M. W. Watson

Ulrich K. Müller

To cite this article: Ulrich K. Müller (2018) Comment on "HAR Inference: Recommendations for Practice" by E. Lazarus, D. J. Lewis, J. H. Stock and M. W. Watson, Journal of Business & Economic Statistics, 36:4, 563-564, DOI: 10.1080/07350015.2018.1497502

To link to this article: https://doi.org/10.1080/07350015.2018.1497502

Published online: 02 Nov 2018.

Comment

Ulrich K. Müller
Department of Economics, Princeton University, Princeton, NJ 08544 (*umueller@princeton.edu*)

1. INTRODUCTION

The article applies the framework of Lazarus, Lewis, and Stock (2017) to suggest a specific recipe for heteroscedasticity and autocorrelation robust (HAR) inference. In many ways, the ultimate recommendations echo what I suggested more than a decade ago in Müller (2004) based on different set of arguments. I continue to think that the equal-weighted cosine (EWC) projection approach, with the resulting convenient student-t and Hotelling-T limit distributions, is a practically useful approach that should be more widely adopted.

In this comment, I want to make two points. First, I want to briefly investigate whether the authors' focus on the EWC form of the t-statistic is substantively restrictive. Specifically, are there functions of the same underlying data transformations that result in a more favorable size and power trade-off? At least for inference about the mean of a Gaussian AR(1) with $T = 200$ observations and coefficient $\rho \leq \bar{\rho} = 0.7$, the answer turns out to be: "no," further strengthening the case for the suggested procedures.

My second point takes issue with the suggested default rule of using $\nu = 0.4T^{2/3}$ cosine terms. In particular, this rule leads to arbitrarily large values of ν in large enough samples. But the marginal benefit of larger power quickly decreases for ν large, while the cost of less robustness remains potentially unbounded as soon as one considers more strongly persistent underlying processes. In my mind, this is an unattractive trade-off, since in most applications, there is no compelling reason why values of ρ larger than 0.7 are entirely impossible. I would therefore suggest a modification of the default rule to, say

$$\nu = \min(0.4T^{2/3}, 20)$$

ensuring that, if it is cheap to do so in terms of power, the resulting inference becomes robust to a wider range of persistence patterns.

2. CHOICE OF TEST STATISTIC

Consider the issue in a small-sample setting. Let $\{z_t\}_{t=1}^T$ be a Gaussian AR(1) with coefficient ρ and mean μ, so that with $Z_0 = T^{-1}\sum_{t=1}^T z_t$ and $Z_j = T^{-1}\sum_{t=1}^T \sqrt{2}\cos(\pi j(t-1/2)/T)z_t$, we have

$$Z = (Z_0, \ldots, Z_\nu)' \sim \mathcal{N}(\mu\iota_1, \Sigma_0(\rho))$$

with $\iota_1 = (1, 0, \ldots, 0)'$. We want to test the null hypothesis $H_0 : \mu = 0$ against the two-sided alternative. Note that the EWC t-statistic

$$t_{\text{EWC}} = \frac{Z_0}{\sqrt{\nu^{-1}\sum_{j=1}^{\nu} Z_j^2}}$$

is a function of Z. In the derivations of Lazarus, Lewis, and Stock (2017), it is an exogenous constraint not to consider alternative functions of Z as test statistics. This raises the question whether it might be possible, for given ν and thus Z, to find tests with a more advantageous size and power profile by considering alternative functions of Z.

To further simplify the problem, consider tests that maximize weighted average power, with a weight function on μ equal to $\mathcal{N}(0, \kappa^2/T)$. The scalar κ governs whether the weighting function emphasizes closer or more distant alternatives. For given κ, maximizing weighted average is then equivalent to maximizing power against the alternative H_1 (see Müller (2014) for additional details), where

$$H_0 : Z \sim \mathcal{N}(0, \Sigma_0(\rho)) \text{ against } H_1 : Z \sim \mathcal{N}(0, \Sigma_1(\rho)) \quad (1)$$

with $\Sigma_1(\rho) = \Sigma_0(\rho) + \kappa^2\iota_1\iota_1'/T$. The problem has thus become one of inference about the covariance matrix of a normal vector. Furthermore, consider tests that are scale invariant, that is, the decision of a test to reject (or not) remains unchanged when Z is multiplied by any nonzero scalar constant. Under this constraint, the innovation variance of the Gaussian AR(1) can be normalized to unity without loss of generality, so that $\Sigma_0(\rho)$ and $\Sigma_1(\rho)$ are solely functions of ρ. Both hypotheses in (1) are still composite, though, since they do not specify the value of ρ. For given values ρ_0 and ρ_1 under H_0 and H_1, respectively, the best level-α scale invariant test is of the form (see Müller 2014)

$$\frac{Z'\Sigma_0(\rho_0)^{-1}Z}{Z'\Sigma_1(\rho_1)^{-1}Z} > \text{cv} \quad (2)$$

with cv chosen such that the null rejection probability is equal to α.

Now let $\alpha_{\text{EWC}}(\rho)$ be the null rejection probability of the nominal level-α test based on t_{EWC} with student-t critical value cv_t, $\alpha_{\text{EWC}}(\rho) = E[t_{\text{EWC}}^2 > \text{cv}_t^2]$ with $Z \sim \mathcal{N}(0, \Sigma_0(\rho))$. We seek scale invariant tests that are functions of Z with null rejection bounded above by $\alpha_{\text{EWC}}(\rho)$ uniformly in ρ, but with potentially larger power against a specific alternative $H_1^* : Z \sim \mathcal{N}(0, \Sigma_1(\rho_1))$. Any such test must in particular have null rejection probability no larger than $\alpha_{\text{EWC}}(\rho)$ for

Journal of Business & Economic Statistics
October 2018, Vol. 36, No. 4
DOI: 10.1080/07350015.2018.1497502

$\rho = \rho_0 = \rho_1$—this amounts to imposing the uniform bound $\alpha_{\rm EWC}(\rho)$ only at the single value $\rho = \rho_1$. For $\rho_1 = 0$ (so that power is maximized in the iid model), the orthogonality of the Type II cosine transform implies that $\Sigma_0(0) = T^{-1}I_{\nu+1}$, and $\Sigma_1(0) = T^{-1}\mathrm{diag}(1+\kappa^2, I_\nu)$. But with this special structure on the quadratic forms in (2), the event in (2) can exactly be rewritten as $t^2_{\rm EWC} > \mathrm{cv}^2_t$, so that the best test that only imposes the $\alpha_{\rm EWC}(\rho)$-bound at $\rho = \rho_1$ equals $t_{\rm EWC}$. This immediately implies that the best test under the original uniform null rejection bound is also equal to $t_{\rm EWC}$, since it trivially satisfies the uniform bound. This argument goes through for any κ, so $t_{\rm EWC}$ is uniformly best in that sense.

For $T = 200$ and $0 < \rho_1 < 0.7$, $\Sigma_0(\rho_1)$ is still nearly diagonal, and the variation of the diagonal elements remains small. The rejection region of the test that imposes the null rejection probability bound only at $\rho_0 = \rho_1$ is thus still numerically very close to the rejection region $t^2_{\rm EWC} > \mathrm{cv}^2_t$, severely limiting the scope of any substantial improvement. A more detailed numerical analysis shows that for $\kappa^2 = 10\mathrm{tr}\Sigma_0(\rho)/(\nu+1)$, which induces approximately 50% weighted average power, the potential gain in power of 5% level tests is smaller than 0.1 percentage points for $t_{\rm EWC}$ with the suggested default choice of $\nu = 0.4T^{2/3} = 14$. Thus, for all practical purposes, one cannot improve over the performance of $t_{\rm EWC}$ by considering alternative functions of the underlying transformation Z.

3. CHOICE OF ν

The choice of ν represents a classic trade-off: small values of ν lead to robust tests, but with low power, while choosing ν large leads to powerful tests with potentially large size distortions. Under asymptotics where ν is fixed as $T \to \infty$, and weakly dependent data, the large-sample distribution of $t_{\rm EWC}$ is the same as the small-sample distribution of the usual t-statistic for inference about the mean of $\nu + 1$ iid Gaussian variates, under the null and local alternatives. Also, the large-sample distribution of the (in general infeasible) test that plugs-in the true value of the long-run variance is equal to the small-sample distribution of the usual z-test (known variance test) for inference about the mean of $\nu + 1$ iid Gaussian variates, under the null and local alternatives. These convergences hold jointly, so that the loss in local asymptotic power of $t_{\rm EWC}$ relative to the test with known long-run variance corresponds to the small-sample loss in power of the t-test relative to the z-test. Figure 1 plots these losses, maximized over the alternative. As ν becomes larger, the largest loss in power becomes small in absolute terms, and, what is more, the marginal gain in power (reduction of the maximal power loss) of increasing ν further quickly become very small.

The default rule suggested by the authors nevertheless picks large values of ν in large datasets. For instance, if $T = 1000$, then $\nu = 0.4T^{2/3} = 40$. This is because in their analysis, the persistence is assumed to be at most moderate and sample size-independent, $\rho \le \bar{\rho} = 0.7$. With ρ at most equal to 0.7, size distortions are very small even for fairly large ν if $T = 1000$, so that their criterion makes it relatively optimal to try to extract the last bit of power.

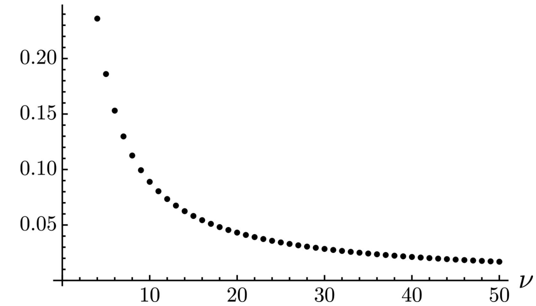

Figure 1. Maximal loss in power of 5% t-test with ν degrees of freedom relative to z-test.

Table 1. Null rejection probabilities of $t_{\rm EWC}$ with $T = 1000$

$\rho =$	0.7	0.8	0.9	0.95
$t_{\rm EWC}$ with $\nu = 20$	0.051	0.053	0.063	0.097
$t_{\rm EWC}$ with $\nu = 40$	0.055	0.062	0.094	0.176

NOTE: Table contains null rejection probabilities of nominal 5% level tests for mean of Gaussian AR(1) with coefficient ρ and $T = 1000$.

But this is only sensible if one is quite certain that stronger persistence is impossible. Table 1 reports null rejection probabilities of $t_{\rm EWC}$ with 20 and 40 degrees of freedom, respectively, of nominal 5% level tests about the mean of a Gaussian AR(1) with $T = 1000$ observations. As a point of reference, the largest power difference of these two tests under iid data is a mere 2.2 percentage points. It seems to me that unless there are application-specific reasons why stronger persistence can be entirely ruled out, $t_{\rm EWC}$ with $\nu = 20$ strikes a more attractive balance between size and power.

More generally, I am sympathetic to the idea that in larger samples, one does not want to forgo much power due to imperfect long-run variance estimation, leading to a default rule that has ν increasing in T. But one might also want to use the relative abundance of data to buy more insurance against forms of persistence that are stronger than maybe initially envisaged, especially if doing so is cheap in terms of power losses. A maybe reasonable compromise along those lines is the rule $\nu = \min(0.4T^{2/3}, 20)$ mentioned in the introduction.

[Received June 2018.]

REFERENCES

Lazarus, E., Lewis, D. J., and Stock, J. H. (2017), "The Size-Power Tradeoff in HAR Inference," manuscript, Harvard University. [563]

Müller, U. K. (2004), "A Theory of Robust Long-Run Variance Estimation," Working Paper, Princeton University. [563]

——— (2014), "HAC Corrections for Strongly Autocorrelated Time Series," *Journal of Business and Economic Statistics*, 32, 311–322. [563]