Heteroskedasticity-Autocorrelation Robust Standard Errors Using the Bartlett Kernel Without Truncation

Nicholas M. Kiefer*   Timothy J. Vogelsang†‡

September, 2000; Revised February, 2001

Abstract

In this paper we analyze heteroskedasticity-autocorrelation (HAC) robust tests constructed using the Bartlett kernel without truncation. We show that while such an HAC estimator is not consistent, asymptotically valid testing is still possible. We show that tests using the Bartlett kernel without truncation are exactly equivalent to recent HAC robust tests proposed by Kiefer, Vogelsang and Bunzel (2000, Econometrica, 68, pp 695-714).

Keywords: HAC estimators, KVB statistic inference, robust testing, autocorrelation, covariance matrix estimation, truncation lag, automatic bandwidth.

*School of Economics and Management, CAF, CDME and CLS, University of Aarhus, Denmark and Departments of Economics and Statistical Science, Cornell University, Uris Hall, Ithaca, N.Y. 14583-7601 Tel. (607) 255-6315 Fax. (607) 255-2818. Email: nmk1@cornell.edu

†(Corresponding Author) Departments of Economics and Statistical Sciences, Cornell University Uris Hall, Ithaca, N.Y. 14583-7601 Tel. (607) 255-5108 Fax. (607) 255-2818. Email: tjv2@cornell.edu

‡We thank Whitney Newey for stimulating conversations that inspired this paper. We thank an editor and two anonymous referees for helpful comments.

1 Introduction

Over the past 15 years an important literature has emerged in econometrics on methods for consistently estimating asymptotic covariance matrices of parameters estimates in models with heteroskedasticity and autocorrelation of unknown form. Asymptotic theory for heteroskedasticity-autocorrelation (HAC) consistent variance estimators has developed rapidly with the literature primarily focused on the class of nonparametric estimators derived from the spectral analysis literature. See Kiefer, Vogelsang and Bunzel (2000) for relevant references. The nonparametric class of HAC estimators are zero frequency spectral density estimators that are weighted sums of sample autocovariances. The weights are determined by a kernel function and truncation lag or "bandwidth". Regarding the bandwidth, consistency of the HAC estimator only requires that the bandwidth increase with the sample size but at a slower rate. Unfortunately, this asymptotic theory provides little guidance for kernel or bandwidth selection in finite samples because any choice of bandwidth for a particular finite sample can be made consistent with any rate of growth. This well known and old problem has led to the development of data dependent methods for choosing bandwidths. But, data dependent bandwidths do not provide complete solutions because they require the choice of an approximate parametric model for the autocorrelation, Andrews (1991), or initial nonparametric estimates that require their own bandwidth choice, Newey and West (1994).

In this paper we begin with a different approach to the problem of choosing the kernel and bandwidth. Rather than focus on asymptotic variances and their consistent estimation, we take a finite sample perspective and focus on exact variances. Our approach immediately suggests that the Bartlett kernel without truncation (bandwidth equal to the sample size) provides the natural weights. Even though this HAC estimator is inconsistent, we show that valid testing is possible nonetheless. We also prove that tests using the Bartlett kernel without truncation are exactly equivalent to recent HAC robust tests proposed by Kiefer et al. (2000).

2 The Bartlett Kernel Without Truncation

For clarity, we focus on the simple linear regression model

$$y_t = x_t'\beta + u_t \qquad t = 1, 2, ..., T,$$

where β and x_t are $k \times 1$ vectors, u_t is autocorrelated and possibly conditionally heteroskedastic, and $E(u_t|x_t) = 0$. This last condition rules out lagged dependent variables but can be dropped by doing the analysis in the context of instrumental variable estimation. See Vogelsang (2000).

The focus is testing linear hypotheses about β. We consider the ordinary least squares (OLS) estimator, $\hat{\beta} = \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sum_{t=1}^{T} x_t y_t$. Define $v_t = x_t u_t$. Using standard calculations we can write the normalized estimator as $\sqrt{T}\left(\hat{\beta} - \beta\right) = \left(T^{-1} \sum_{t=1}^{T} x_t x_t'\right)^{-1} T^{-1/2} \sum_{t=1}^{T} v_t$. Because the variance of $T^{-1/2} \sum_{t=1}^{T} v_t$ is crucial for constructing tests about β, it is useful to consider the exact variance,

$$var\left(T^{-1/2} \sum_{t=1}^{T} v_t\right) = \Gamma_0 + \sum_{j=1}^{T-1} \left(1 - \frac{j}{T}\right)\left(\Gamma_j + \Gamma_j'\right), \qquad (1)$$

where $\Gamma_j = cov\left(v_t v_{t-j}'\right)$. Notice that the weights in (1) are the Bartlett kernel weights without truncation. The standard approach is to take the limit of (1) as $T \to \infty$ which gives $\Omega = \Lambda\Lambda' = \Gamma_0 + \sum_{j=1}^{\infty}\left(\Gamma_j + \Gamma_j'\right)$. Then, Ω is estimated using a HAC estimator.

From the finite sample perspective, it makes more sense to estimate (1) directly which is what would be done if an exact test for β were feasible. If the Γ_j's are replaced with sample analogs, $\hat\Gamma_j = T^{-1}\sum_{t=j+1}^{T}\hat v_t \hat v_{t-j}'$ for $j \geq 0$, $\hat\Gamma_j = \hat\Gamma_{-j}'$ for $j < 0$ where $\hat v_t = x_t\left(y_t - x_t'\hat\beta\right)$, then $\hat\Omega = \hat\Gamma_0 + \sum_{j=1}^{T-1}\left(1 - \frac{j}{T}\right)\left(\hat\Gamma_j + \hat\Gamma_j'\right)$ is obtained. $\hat\Omega$ is the Bartlett kernel (i.e. Newey and West (1987)) HAC estimator without truncation. Because there is no truncation, all information (some of it noisy) in the data regarding the Γ_j's is used. Although $\hat\Omega$ is not a consistent estimator of Ω, valid tests can still be obtained because $\hat\Omega$ is asymptotically proportional to Ω as we now show.

3 Asymptotics and Inference

Let $W_k(r)$ denote a k-vector of independent standard Wiener processes, and define $B_k(r) = W_k(r) - rW_k(1)$. Let $[rT]$ denote the integer part of rT where $r \in [0,1]$. We use $\Rightarrow$ to denote weak convergence. The following assumption is sufficient for our results:

Assumption 1 $T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \Lambda W_k(r)$, $p\lim T^{-1}\sum_{t=1}^{[rT]} x_t x_t' = rQ$ uniformly in r with $\det(Q) > 0$.

Define $\hat S_t = \sum_{j=1}^{t}\hat v_j$. We derive the asymptotic distribution of $\hat\Omega$ by showing that $\hat\Omega = 2\hat C$ where $\hat C = T^{-2}\sum_{t=1}^{T}\hat S_t \hat S_t'$. Making use of the identity $\sum_{j=1}^{T} a_j b_j = \sum_{j=1}^{T-1}\left(a_j - a_{j+1}\right)\left(\sum_{i=1}^{j} b_i\right) + a_T \sum_{i=1}^{T} b_i$ and the fact that $\hat S_T = 0$ by the normal equations for OLS, simple algebra gives

$$\hat\Omega = T^{-1}\sum_{i=1}\sum_{j=1}\hat v_i\left(1 - \frac{|i-j|}{T}\right)\hat v_j' = T^{-1}\sum_{i=1}\hat v_i\sum_{j=1}\left(1 - \frac{|i-j|}{T}\right)\hat v_j' = T^{-1}\sum_{i=1}\hat v_i\sum_{j=1}\frac{(-1)^{-(i-j)}}{T}\hat S_j'$$

$$= T^{-1}\sum_{j=1}\left(\sum_{i=1}\hat v_i\frac{(-1)^{-(i-j)}}{T}\right)\hat S_j' = T^{-1}\sum_{j=1}\left(\sum_{i=1}\hat S_i 1(i=j)\frac{2}{T}\right)\hat S_j' = 2T^{-2}\sum_{j=1}\hat S_j \hat S_j' = 2\hat C. \quad (2)$$

Using (2) it directly follows from Kiefer et al. (2000) and Assumption 1 that

$$\hat\Omega \Rightarrow 2\Lambda\int_0^1 B_k(r)B_k(r)'dr\Lambda'$$

as $T \to \infty$. Note that $\hat\Omega$ is asymptotically proportional to Ω through $\Lambda\Lambda'$ and otherwise only depends on known random variables.

Consider testing the null hypothesis $H_0 : R\beta = r$ against the alternative hypothesis $H_1 : R\beta \neq r$ where R is a $q \times k$ matrix of constants with rank q and r is a $q \times 1$ vector of constants. Under H_0,

$$\sqrt{T}\left(R\hat\beta - r\right) = R\left(T^{-1}\sum_{t=1}^{T} x_t x_t'\right)^{-1} T^{-1/2}\sum_{t=1}^{T} v_t \Rightarrow RQ^{-1}\Lambda W_k(1).$$

Because $W_k(1)$ is a $k-$vector of standard normal random variables, $RQ^{-1}\Lambda W_k(1)$ is a $q-$vector of linearly independent normal random variables with variance matrix $RQ^{-1}\Omega Q^{-1}R'$. A Wald statistic could be constructed using a consistent estimator of $RQ^{-1}\Omega Q^{-1}R'$. Instead, we consider the inconsistent estimator, $R\hat Q^{-1}\hat\Omega\hat Q^{-1}R'$, where $\hat Q = T^{-1}\sum_{t=1}^{T} x_t x_t'$ giving the statistic

$$F^* = T\left(R\hat\beta - r\right)'\left[R\hat Q^{-1}\hat\Omega\hat Q^{-1}R'\right]^{-1}\left(R\hat\beta - r\right)/q.$$

In the case where $q = 1$ so that one restriction is being tested, we can construct a $t-$type statistic

$$t^* = \sqrt{T}\left(R\hat{\beta} - r\right) / \sqrt{R\hat{Q}^{-1}\hat{\Omega}\hat{Q}^{-1}R'}.$$

It follows from (2) and Theorem 1 of Kiefer et al. (2000) that under Assumption 1 as $T \to \infty$,

$$F^* \Rightarrow W_q(1)'\left(2\int_0^1 B_q(r)B_q(r)'dr\right)^{-1} W_q(1)/q, \qquad t^* \Rightarrow W_1(1)/\sqrt{2\int_0^1 B_1(r)^2 dr}.$$

Critical values for the asymptotic distribution of t^* have been obtained analytically by Abadir and Paruolo (1997) and are tabulated in Table I for convenience. Asymptotic critical values for F^* for $q = 1, 2, 3, ..., 30$ can be obtained by multiplying by 0.5 the critical values tabulated by Kiefer et al. (2000) in their Table II.

4 Directions for Future Research

Given that the Bartlett kernel without truncation delivers promising HAC robust tests, it is logical to ask how tests using other kernels without truncation behave and perform. Kiefer and Vogelsang (2000) analyze the case of a general kernel, $k(x)$, with continuous second derivative, $k''(x)$, and show that under Assumption 1, $\hat{\Omega} = \hat{\Gamma}_0 + \sum_{j=1}^{T-1} k(j/T)\left(\hat{\Gamma}_j + \hat{\Gamma}_j'\right) \Rightarrow -\Lambda\int_0^1\int_0^1 k''(r-s)B_k(r)B_k(s)'drds\Lambda'$ as $T \to \infty$. Because asymptotic proportionality to Ω is obtained for general kernels, valid asymptotic testing is possible using the class of kernel HAC estimators with bandwidth equal to sample size. A preliminary analysis on this class of tests is given in Kiefer and Vogelsang (2000).

References

Abadir, K. M. and Paruolo, P. (1997), Two Mixed Normal Densities from Cointegration Analysis, *Econometrica* **65**, 671–680.

Andrews, D. W. K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* **59**, 817–854.

Kiefer, N. M. and Vogelsang, T. J. (2000), A New Approach to the Asymptotics of Heteroskedasticity-Autocorrelation Robust Testing, mimeo, Department of Economics, Cornell University.

Kiefer, N. M., Vogelsang, T. J. and Bunzel, H. (2000), Simple Robust Testing of Regression Hypotheses, *Econometrica* **68**, 695–714.

Newey, W. K. and West, K. D. (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* **55**, 703–708.

Newey, W. K. and West, K. D. (1994), Automatic Lag Selection in Covariance Estimation, *Review of Economic Studies* **61**, 631–654.

Vogelsang, T. J. (2000), Testing in GMM Models Without Truncation, mimeo, Department of Economics, Cornell University.

Table I: Asymptotic Critical values of t^*

1.0%	2.5%	5.0%	10.0%	50.0%	90.0%	95%	97.5%	99.0%
-6.090	-4.771	-3.764	-2.740	0.000	2.740	3.764	4.771	6.090

Source: Line 1 of Table I from Abadir and Paruolo (1997, p. 677) scaled by $1/\sqrt{2}$.