Testing for ARCH in the Presence of a Possibly Misspecified Conditional Mean

Robin L. Lumsdaine* Serena Ng†

August 27, 1998

Abstract

Ever since the development of the Autoregressive Conditional Heteroskedasticity (ARCH) model (Engle [1982]), testing for the presence of ARCH has become a routine diagnostic. One popular method of testing for ARCH is T times the R^2 from a regression of the squared residuals on p of its lags. This test has been shown to have a lagrange multiplier interpretation and is asymptotically distributed as a $\chi^2(p)$ random variable. Underlying this test is the assumption of a correctly specified conditional mean. In this paper, we consider the properties of the ARCH test when there is a possibly misspecified conditional mean. Examples of misspecification include omitted variables, structural change, and parameter instability. We show that, in general, misspecification will lead to overrejection of the null hypothesis of conditional homoskedasticity. We demonstrate the use of recursive residuals to improve the fit of a first stage conditional mean regression. We illustrate these results via Monte Carlo simulations and consider two empirical examples.

KEYWORDS: Autoregressive Conditional Heteroskedasticity, Model Misspecification, Recursive Residuals, Lagrange Multiplier Test
JEL Classification Numbers: C22, C52

*Department of Economics, Box B, Brown University, Providence, RI 02912, (401)863-1978 (w), (401)863-1970 (fax), Robin_Lumsdaine@brown.edu

†Department of Economics, Boston College, Chestnut Hill, MA 02467, (617)552-2182 (w), (617)552-2308(fax), Serena.Ng @bc.edu.

We would like to thank to Rob Engle, Gloria Gonzalez-Rivera, Alastair Hall, Bruce Hansen, Jim Powell, Tom Rothenberg, Paul Ruud, Jim Stock, Timo Teräsvirta, Aman Ullah, Dick van Dijk, Herman van Dijk, and Mark Watson for helpful comments and discussions. Lumsdaine thanks the National Fellows Program, Hoover Institution, Stanford University and the Institute for Policy Analysis, University of Toronto. Ng gratefully acknowledges grants from SSHRC (Canada) and FCAR (Québec).

1 Introduction

Ever since the development of the Autoregressive Conditional Heteroskedasticity (ARCH) model (Engle [1982]), testing for the presence of ARCH has become a routine diagnostic. One popular method of testing for ARCH is T times the R^2 from a regression of the squared residuals on a constant and p of its lags. This test inherently assumes that the conditional mean is correctly specified.

Tests of parameter instability have likewise become a central focus of applied researchers. Failure to model the conditional mean correctly can lead to erroneous inference. An example of this is given in Perron [1989]; here failure to model a break in the trend function will lead to conclusions of excess persistence in many macroeconomic time series.

Failure to account for parameter instability in either the first or second moments is a specific example of the general notion of model misspecification. Robinson [1991], Wooldridge [1991a,b], and Hansen [1992], derive ways of "robustifying" tests to allow for possible misspecification of higher-order moments. In these cases, the tests of interest can be modified to correct for the misspecification of a higher moment by means of a bias correction which can be shown to be distributed as a χ^2 random variable. For example, Wooldridge [1991a] shows that valid inference regarding the conditional mean parameters is still possible in the presence of conditional variance misspecification. In general, tests about the k-th moment of the residuals implicitly assume correct specification of lower moments. Analogous "robustification" of tests about higher moments in the presence of possible lower-order moment misspecification is more difficult, though some authors have employed nonparametric methods to consider hypotheses about unconditional moments in this situation.[1]

A number of authors have documented cases where rejection of the null of *conditional* homoskedasticity may be a result of model misspecification. Bera, Higgins, and Lee [1992], citing Engle, Hendry, and Trumble [1985], note that "the presence of autocorrelation can readily be mistaken for ARCH when, in fact, no ARCH is present." Besides autocorrelation, ARCH is often detected in models with nonlinear dynamics. Bera and Higgins [1997] compare ARCH and bilinear processes, noting that the latter have "an unconditional moment structure very similar to ARCH and hence may be easily mistaken for ARCH" (see also Weiss [1986] and Tong [1990]). Via Monte Carlo simulation, Giles, Giles and Wong [1993] consider a number of tests for ARCH and GARCH in the case of an omitted regressor and find the null hypothesis of conditional homoskedasticity

[1] e.g., Lee [1992] considers tests for unconditional heteroskedasticity in the presence of conditional mean misspecification; Rilstone [1992] and Whang [1998] develop tests for normality based on third and fourth moments, allowing lower moments to take on a general (nonspecified) form.

tends to be rejected too often.

In this paper, we consider testing for ARCH in the general context of a possibly misspecified conditional mean. We assume that correct inference regarding the conditional variance is of primary interest to the researcher (as in many examples in finance). Section 2 first formalizes the problem of testing for ARCH when the conditional mean is possibly misspecified. We then discuss possibilities for guarding against misspecification in the mean function. One solution is to exploit the information available from the first step estimation to further improve the estimation of the conditional mean. This amounts to treating the estimation of the conditional mean as a first stage ancillary regression. Section 3 investigates this "robustified" ARCH test. Section 4 examines the properties of the statistic via Monte Carlo simulation. Section 5 contains two empirical examples. Section 6 concludes.

2 Testing for ARCH when the Conditional Mean is Misspecified

Suppose we have data $\{y_t : t = 1, \ldots, T\}$, with data generating process

$$y_t = m_t + \epsilon_t, \tag{1}$$

where $m_t = \mathrm{E}(y_t | I_t^*)$ is an (unknown) function of data and parameters (possibly nonlinear), I_t^* is Borel-measurable with respect to y_t, and ϵ_t has mean zero, conditional variance

$$h_t = \widetilde{\alpha}_0 + \widetilde{\alpha}_1 \epsilon_{t-1}^2 + \ldots + \widetilde{\alpha}_p \epsilon_{t-p}^2,$$

and unconditional variance $E(\epsilon_t^2) = \sigma^2$. We assume m_t is nonconstant and that the data are stationary.[2] The hypothesis of conditional homoskedasticity involves determining whether $\mathrm{E}(\epsilon_t^2 | I_{t-1}^*)$ is (a time-invariant) constant. Because m_t and I_t^* are unknown, we estimate the following model:

$$y_t = \mu_t + u_t \tag{2}$$

where $\mu_t \in I_t$ is a known function of data and parameters and I_t is the information set of the econometrician. Denote the estimated residuals by

$$\widehat{u}_t = (m_t - \widehat{\mu}_t) + \epsilon_t \equiv \eta_t + \epsilon_t. \tag{3}$$

where $\widehat{\mu}_t$ is the estimate of the conditional mean μ_t. If the mean function is correctly specified in the sense that $\mu_t \equiv m_t$ with $I_t \equiv I_t^*$, then $\widehat{\mu}_t \xrightarrow{a.s.} m_t$ under mild regularity conditions. It follows that $u_t \equiv \epsilon_t$, and $\widehat{u}_t = \epsilon_t + o_p(1)$. It is then obvious that testing hypotheses about $\mathrm{E}(\widehat{u}_t^2 | I_{t-1}^*)$

[2] The stationarity assumption is not necessary but is useful in deriving the argument below. In the simulations, we also consider the case where y_t is $I(1)$.

is asymptotically equivalent to testing $\mathrm{E}(\epsilon_t^2|I_{t-1})$. In this case, the estimated residuals $\widehat{u}_t$ are appropriate for testing the null hypothesis of interest. The standard ARCH test thus involves regressing $\widehat{u}_t^2$ on p of its lags, that is

$$\widehat{u}_t^2 = \alpha_0 + \alpha_1 \widehat{u}_{t-1}^2 + \cdots + \alpha_p \widehat{u}_{t-p}^2 + \nu_t, \tag{4}$$

and testing the hypothesis that $\alpha_i = 0, i = 1, \ldots, p$. Engle [1982] shows that if ϵ_t is conditionally normal, then TR^2 from this regression is asymptotically equivalent to a Lagrange Multiplier test and is distributed asymptotically as a $\chi^2(p)$ random variable under the null hypothesis. More formally, the statistic is

$$LM = \frac{T\widehat{f}'\widehat{Z}(\widehat{Z}'\widehat{Z})^{-1}\widehat{Z}'\widehat{f}}{\widehat{f}'\widehat{f}},$$

where $\widehat{f}' = (\widehat{f}_1, \ldots, \widehat{f}_T)$, $\widehat{f}_t = (\widehat{u}_t^2/\widehat{\sigma}^2 - 1)$, $\widehat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\widehat{u}_t^2$, and $\widehat{Z}' = (\widehat{z}_1, \ldots, \widehat{z}_T)$, $\widehat{z}_t = (1, \widehat{u}_{t-1}^2, \ldots, \widehat{u}_{t-p}^2)'$. Weiss [1986] extended this result to cases where ϵ_t has finite fourth moments.

Consider now the case when the model is not correctly specified in a way that $\eta_t \equiv m_t - \widehat{\mu}_t \overset{p}{\not\to} 0$ with $\widehat{u}_t = \epsilon_t + \eta_t$.[3] It follows that

$$\widehat{u}_t^2 = (m_t - \widehat{\mu}_t)^2 + 2\epsilon_t(m_t - \widehat{\mu}_t) + \epsilon_t^2 \equiv \xi_t + \epsilon_t^2. \tag{5}$$

The conditional second moment of the squared estimated residuals can be separated into two pieces: a part that is related to the conditional mean $\mathrm{E}[\xi_t|I_{t-1}]$, and a conditional variance part $\mathrm{E}[\epsilon_t^2|I_{t-1}]$. The conditional and unconditional means of $\widehat{u}_t^2$ generally will differ from those of ϵ_t^2 if $m_t - \widehat{\mu}_t \overset{p}{\not\to} 0$.[4] Thus the LM test implicitly tests the constancy of ξ_t and ϵ_t^2 rather than just ϵ_t^2. Put somewhat differently, the ARCH regression is an autoregression with errors-in-variables when $m_t - \widehat{\mu}_t \overset{p}{\not\to} 0$.

Model misspecification can arise if the functional form and/or the conditioning information set is misspecified. For linear dynamic models, notable examples are omitted shifts in the trend function, selecting a lag length in an autoregression that is lower than the true order, failure to account for parameter instability, residual autocorrelation, and omitted time series variables. Of note is that such misspecifications in dynamic models often result in serial correlation in $\eta_t = m_t - \widehat{\mu}_t$. It is easy to show that the sum of two stationary processes will be serially correlated if at least one series is serially correlated. As shown in Granger and Morris (1976), the resulting (summed) process will have ARMA like autocovariance properties. Thus, for model misspecification of the types likely to arise in practice, $\widehat{u}_t = \epsilon_t + \eta_t$ will, in general, be serially correlated insofar as η_t is serially correlated.

[3] Throughout the remainder of the paper, the term 'correctly specified' will refer to this definition.

[4] In the special case that $(m_t - \widehat{\mu}_t) \in I_{t-1}$, then $\epsilon_t(m_t - \widehat{\mu}_t)$ has a conditional mean of zero by the law of iterated expectations, and so $\mathrm{E}(\widehat{u}_t^2|I_{t-1}) > \mathrm{E}(\epsilon_t^2|I_{t-1})$. In general, however, the conditional second moment of (5) will involve both squared terms and the cross-product term.

To see how misspecification in the conditional mean affects the LM test, consider the test for ARCH(1). If we could observe the ϵ_t^2, we would run the following regression

$$\epsilon_t^2 = \widetilde{\alpha}_0 + \widetilde{\alpha}_1 \epsilon_{t-1}^2 + \widetilde{\nu}_t$$

to test for ARCH(1). However, because we only have $\widehat{u}_t^2$, we inadvertently run the regression

$$\widehat{u}_t^2 = \alpha_0 + \alpha_1 \widehat{u}_{t-1}^2 + \nu_t.$$

If ϵ_t^2 is indeed serially uncorrelated with $\widetilde{\alpha}_1 = 0$, the ARCH test based upon $\widehat{u}_t^2$ will yield correct inference only if $\widehat{u}_t^2$ is also not serially correlated. But as discussed earlier, $\widehat{u}_t$ will, in general, be serially correlated when the conditional mean of the dynamic model is misspecified. Since the square of a serially correlated process is itself serially correlated, $\widehat{u}_t^2$ will be serially correlated in most instances even if ϵ_t^2 is not.[5]

As an example, suppose y_t is generated by $z_t\gamma + \epsilon_t$, ϵ_t is conditionally and unconditionally homoskedastic, $z_t - \mu_z = \phi(z_{t-1} - \mu_z) + \epsilon_t^z$, and ϵ_t^z is a Gaussian error (with variance σ_z^2) that is uncorrelated with ϵ_t. Without loss of generality, let $\gamma = 1$. Consider the test for ARCH(1) when the researcher does not control for z_t and simply uses y_t^2 to test for ARCH(1). That is, we omit an AR(1) regressor when estimating the conditional mean, and $\eta_t = z_t$ in this example. As shown in the Appendix, z_t^2 is then an ARMA(2,1) process and can thus be represented as $z_t^2 = \mu_{z^2} + \Phi(L)\omega_t$, where ω_t is a white noise process, $\Phi(L) = \sum_{i=0}^{\infty} \Phi_i L^i$, Φ_is are functions of ϕ, μ_z, and σ_z^2. Most importantly, $\Phi_i = 0$ only if $\phi = 0$. Then $\xi_t = z_t^2 + 2z_t\epsilon_t = \mu_{z^2} + \Phi(L)\omega_t + 2z_t e_t$. Consider the least squares estimate of α_1. Let $\widetilde{u}_t^2 = \widehat{u}_t^2 - \frac{1}{T}\sum_{t=1}^{T} \widehat{u}_t^2$. We have

$$\widehat{\alpha}_1 \;=\; \frac{\sum_{t=2}^{T} \widetilde{u}_{t-1}^2 \widetilde{u}_t^2}{\sum_{t=1}^{T} (\widetilde{u}_{t-1}^2)^2} \xrightarrow{p} \frac{\sum_{i=0}^{\infty} \Phi_i \Phi_{i+1}}{\sum_{i=0}^{\infty} \Phi_i^2}.$$

Thus, $\widehat{\alpha}_1$ does not consistently estimate $\widetilde{\alpha}_1 = 0$ unless $\Phi_i = 0, \forall i > 1$, which in turn requires that $\phi = 0$.[6] Given the probability limit of $\widehat{\alpha}_1$, the R^2 of the ARCH(1) regression (which, in this simple example is just $\widehat{\alpha}_1^2$) will be biased upwards. Evidently, the autocorrelations at longer lags of $\widehat{u}_t^2$ are also non-zero in this example. As discussed in Granger and Teräsvirta (1993), the LM test can be seen as a Box-Ljung statistic for testing the significance of the autocorrelation coefficients corresponding to $\widehat{u}_t^2$. Thus, evidence for higher order ARCH effects could be found significant.[7]

---

[5] For Gaussian processes, this result follows from Granger and Newbold (1976). For non-Gaussian processes, serial correlation in the squared process arises even under mild moment conditions.

[6] In particular, if we denote the resulting ARMA(2,1) representation as $(1 - \phi_1 L - \phi_2 L^2)z_t^2 = (1 + \theta L)\omega_t$, with $\phi_1 = \rho + \rho^2, \phi_2 = -\rho^2$, then $\Phi(L)$ is given by $\Phi_0 = 1$, $\Phi_1 = \phi_1 + \theta$, and $\Phi_i = \phi_1 \Phi_{i-1} + \phi_2 \Phi_{i-2}$, for $i \geq 2$. These calculations are available from the authors on request.

[7] This suggests model selection methods such as the AIC and BIC will likely conclude in favor of an higher order ARCH process. In a similar example, Ng and Vogelsang [1996] find the AIC overparameterizes a VAR when shifts in the mean function are omitted.

Although the probability limit of $\widehat{\alpha}_1$ will depend on the exact nature of the model misspecification, this example shows that serial correlation in $m_t - \widehat{\mu}_t$ is the root of the problem as far as inference is concerned.[8]

3 Robustifying the ARCH Test

If the existence of ARCH is of primary importance to the applied researcher (as in the case, for example, of options pricing), and the problem with a misspecified conditional mean is $\widehat{u}_t^2 = \epsilon_t^2 + \xi_t$, it is evident that any solution to this errors-in-variables problem must involving removing or minimizing the ξ_t piece. Otherwise, persistence in $\widehat{u}_t^2$ arising from model misspecification will continue to be misinterpreted as evidence of ARCH. But ξ_t originates from $\eta_t = m_t - \widehat{\mu}_t$. It is then clear that our goal should be to obtain as accurate an approximation to m_t as possible (in an L^2 sense).

We propose two ways of guarding the LM test against misspecification of the regression function; both aim at maximizing the fit from the first stage (ancillary) regression (2). Both are motivated by the fact that projecting y_t onto the sigma-algebra generated by $\{y_{t-1}, y_{t-2}, \ldots, X_{t-1}\}$ is valid and will approximate m_t no worse (and possibly better) than a projection of y_t onto the sigma-algebra generated by $\{y_{t-1}, y_{t-2}, \ldots\}$ only, provided X_{t-1} is in the econometrician's information set. Identifying the correct conditioning information set (i.e., X_{t-1}) is an important part of this exercise.

In what follows, we assume that the (possibly misspecified) conditional mean is given by a linear regression model,

$$y_t = z_t'\gamma + u_t.$$

The first approach we consider is to include additional lags in the estimated model. The motivation comes from the fact that we can approximate the serial correlation in u_t by $A(L)u_t = e_t$, where $A(L)$ is the induced polynomial in the lag operator and e_t is white noise. Inverting $A(L)$ to solve for u_t immediately yields a specification for y_t in lags of y_t and z_t with a lag order that is higher than the true order. We refer to this as the "naive" approach, in that it is perhaps the solution that one's intuition would first suggest.

Remember, however, that η_t (and hence ξ_t) is unobserved. Thus any Wold representation is merely an approximation to the true autocorrelation function of η_t, which could be highly nonlinear. The second approach, therefore, is to approximate it by functions of the recursive residuals defined

[8] Van Dijk, Franses, and Lucas [1996] compute the numerator of the score for α_1 in the case of additive outliers in and AR(1) model. These authors showed that the presence of outliers also introduces non-centrality to the χ^2 distribution.

in Brown, Durbin, and Evans [1975]. Our motivation is that any unobserved nonlinearlities will be manifested in the recursive residuals; this is supported by discussion in Kianifard and Swallow [1996], who also demonstrate that among many standard tests for model misspecification, use of recursive residuals (rather than standard OLS residuals) increases the power of such tests.

More precisely, we suggest a two-step estimation procedure. Step 1 is to start from the $k+1^{th}$ observation for some predetermined k and perform recursive estimation of y_t on z_t over the remaining $T-k$ observations. This leads to a set of $\widehat{\gamma}_t$ and a set of recursive residuals $\widehat{w}_t = y_t - z_t' \widehat{\gamma}_{t-1}$. These recursive residuals contain the information used to update $\widehat{\gamma}_t$ from $\widehat{\gamma}_{t-1}$ and cannot be predicted by the regression model given information at time $t-1$. They are serially uncorrelated by construction if the model is correctly specified, but when the model is misspecified, $\widehat{w}_t$ will contain information about the true conditional mean not captured by the regression function. Step 2 is to estimate

$$y_t = z_t'\gamma + g(\widehat{w}_{t-1}) + v_t, \tag{6}$$

where $g(\widehat{w}_{t-1})$ is a (possibly non-linear) function of the recursive residuals $\widehat{w}_{t-1}$. Then use $\widehat{v}_t^2$ (the square of the residuals from estimation of (6)) to test for ARCH effects.

It should be evident that minimizing $(m_t - \widehat{\mu}_t)$ is equivalent to making v_t closer to zero, so that the objective of (6) is to improve the estimate of the conditional mean using the information available. In this framework, this is provided by $g(\widehat{w}_{t-1})$. Another way to think about the role of $g(\widehat{w}_{t-1})$ is that it attempts to orthogonalize u_t so that the conditional mean of the resulting regression error v_t shrinks towards zero.

The recursive residuals are appealing not just because they are easy to compute, but because $\widehat{w}_{t-1} \in I_{t-1}$ and hence is in the econometrician's information set at time t. This is why $\widehat{w}_{t-1}$ is used in (6) at time t rather than $\widehat{w}_t$. Note that for this same reason, the use of the OLS residuals is invalid. In non-time series situations, one would have considered using a semiparametric estimate of the conditional mean (see, for example, Robinson [1988]), using a two-sided window with triangular weights. Because the hypothesis of ARCH involves conditional moments, only information available at time $t-1$ can be used. Thus leads of the recursive residuals are also not valid.

To make the two-step procedure operational, it remains to specify g. Given that the objective of the exercise is to guard against misspecification in functional form and the conditioning information set, the natural candidate is to make g as flexible a function of the recursive residuals as possible. One simple alternative is to use a polynomial in the recursive residuals, i.e. $g(\widehat{w}_{t-1}) = \sum_{i=1}^{m} \beta_i \widehat{w}_{t-1}^i$ for a series expansion of length m in $\widehat{w}_{t-1}$. This is appealing because polynomials have a nonparametric interpretation. Furthermore, significance of $\widehat{\beta}_i$ can be interpreted as a diagnostic for misspecification in the conditional mean.

We additionally propose use of the cumulated sum of the recursive residuals with $g(\widehat{w}_{t-1}) = \beta_1(\sum_{i=1}^{t-1} \widehat{w}_i)$. Many authors (e.g., Harvey [1990]) have suggested the use of cusum and cusumsq tests to detect potential nonlinearity. However, in the present context, the cusum of $\widehat{w}_{t-1}$ could be especially useful in cases of omitted mean or trend shifts and when the data have unit roots. This is because misspecification errors are cumulative when the data are I(1), and the cusum of the recursive residuals is the same order in probability as the error causing misspecification. The use of cusum of $\widehat{w}_{t-1}$ nevertheless requires caution as the partial summed series is no longer stationary.

Another alternative, similar to the cusum but which retains stationarity of the $g(\widehat{w}_{t-1})$ is a one-sided sum of the past recursive residuals over a fixed length. This is analogous to a non-parametric estimation of $g(\widehat{w}_{t-1})$ using a flat kernel with a truncated bandwidth, and is more common in the literature. A further possibility is to use varying instead of fixed kernel weights. Hong and Shehadeh [1996] propose tests using declining weights in non-parametric tests for ARCH and show that the performance of these tests is better than the standard LM test. These alternatives are not considered here but are worthy of further investigation.

It should be clear that in contexts when the nature of the misspecification involves contemporaneous variables, use of $g(\widehat{w}_{t-1})$ will not be optimal since they will only provide information with a one-period lag. But the contemporaneous variable must not have been considered by the econometrician in the first place, so that while $g(\widehat{w}_{t-1})$ is a suboptimal improvement, it still can do no worse than simply using z_t as the regressor. Clearly, if the econometrician has some idea as to the type of misspecification, the estimated model should incorporate the information a priori. Augmenting the mean regression with functions of $\widehat{w}_{t-1}$ as discussed above is only a guard against model misspecification. It should also be clear that the robustified test statistic will have a *central* χ^2 distribution only if m_t can be estimated consistently. Nevertheless, if our robust procedures succeed in reducing the deviation of the finite sample distribution of the test from the χ_p^2 distribution, size distortions will be reduced relative to the standard test. In effect, including functions of the recursive residuals serves several precautionary purposes: first, there will be cases when the exact form of model misspecification is unknown; and second, when the model is misspecified in more than one way (such as the case where there is a structural change and a few large outliers, for example). As we will see in the simulations, inclusion of $g(\widehat{w}_{t-1})$ alleviates the size distortion of the TR^2 test associated with model misspecification without a substantial loss of power;

As an example of when we believe our two-step ARCH test will be useful, consider the data shown in Figure 1. We initially estimate an AR(1); from the figure, it is not entirely obvious that the model is misspecified. Estimating an AR(1) with constant autoregressive coefficient, a test of the residuals for ARCH(1) rejects the null hypothesis of no conditional heteroskedasticity; the value of the test statistic is 5.07. Inclusion of the recursive residuals and their squares results in a reduction of the test statistic to 2.81, which is no longer statistically significant; adding the third and fourth powers further reduces the statistic to 2.22. Alternatively, adding in the cusum of the recursive residuals immediately reduces the statistic to 1.34. Using both the cusum and the first two terms of the polynomial reduces the statistic to 0.91.

Is there ARCH in the data? In fact, the DGP that generates this data, detailed in Example 2 to follow, is an AR(1) with autoregressive coefficient equal to 0.3, but it contains a mean shift half way through the sample that cannot be immediately visualized. The detection of ARCH here is therefore spurious; rejection of the standard ARCH test is due to failure to account for the shifted mean. The optimal procedure in this example is obviously to remove the break. Indeed, a test for structural change in the mean strongly rejects a null hypothesis of no break. If we include an indicator variable in the regression to allow for a mean shift at the *true* break date, the value of the test statistic for ARCH is 0.48. Even if we did not know the true break date, the break fraction can be consistently estimated under mild regularity conditions[9] and imposing the estimated break date will also restore the asymptotic convergence of $\widehat{\mu}_t$ to m_t. Using the procedure developed in Banerjee, Lumsdaine, and Stock [1992] to control for the mean shift, for example, the LM test is 1.67, still well below the rejection region. Thus, if we suspect a break, we should test and remove it prior to testing for ARCH.[10]

However, failing to diagnose the break may not be altogether damaging if we consider the robust procedures. Figure 2 shows a plot of the recursive residuals from the first step regression, along with corresponding standard error bands, where $g(\widehat{w}_{t-1}) = 0$, $\gamma_1 \widehat{w}_{t-1} + \gamma_2 \widehat{w}_{t-1}^2$, and $\gamma_1 \widehat{w}_{t-1} + \gamma_2 \widehat{w}_{t-1}^2 + \gamma_3 \sum_{i=1}^{t-1} \widehat{w}_i$, respectively. It is clear from the plot that the addition of the cusum term results in recursive residuals that are much closer to zero. Most importantly, either specification for $g(\widehat{w}_{t-1})$ will shift inference from rejection to non-rejection, illustrating that even "non-optimal" functions will reduce the difference $m_t - \widehat{\mu}_t$. This underscores the importance of accounting for model misspecification and how consistently estimating the conditional mean is fundamental to inference regarding the ARCH test.

4 Monte Carlo Evidence

In each of the simulations, we consider five LM tests which differ in the estimation of the conditional mean. These are specified as:

[9] See, e.g., Banerjee, Lumsdaine, and Stock [1992] or Zivot and Andrews [1992].

[10] Similarly, Lamoureux and Lastrapes [1990] include equally spaced dummy variables in the ARCH equation to absorb potential nonlinearity and demonstrate that the evidence for ARCH is weakened in their empirical example.

$$
\begin{aligned}
&(LM) && y_t && = \gamma_0 + \gamma_1 y_{t-1} + u_t, && (7a)\\
&(LM - naive) && y_t && = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \gamma_3 y_{t-3} + \gamma_4 y_{t-4} + v_{2t}, && (7b)\\
&(LMa) && y_t && = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 \widehat{w}_{t-1} + \gamma_3 \widehat{w}_{t-1}^2 + v_{3t}, && (7c)\\
&(LMb) && y_t && = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 \textstyle\sum_{i=1}^{t-1} \widehat{w}_i + v_{4t}, && (7d)\\
&(LMc) && y_t && = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 \widehat{w}_{t-1} + \gamma_3 \widehat{w}_{t-1}^2 + \gamma_4 \textstyle\sum_{i=1}^{t-1} \widehat{w}_i + v_{5t}, && (7e)
\end{aligned}
$$

where $\widehat{w}_t$ are the recursive residuals and computed as discussed in Section 3. The five LM tests are based on $\widehat{u}_t$ and $\widehat{v}_{it}^2$, $i = 2, \ldots 5$. The sample size is $T = 200$. Random numbers are generated using the rndn( ) function in Gauss V 3.27 with seed=99. The simulations are performed on a Pentium 300 Mhz personal computer running Windows NT 4.0. Throughout the 10,000 simulations, $\alpha_0 = .5$ and we use a nominal size of 5% and thus the standard error of the Monte Carlo simulations is .0022.

Example 1. Slope Shift

The data are generated according to the following:

$$
y_t \;= a_t y_{t-1} + \epsilon_t \qquad\qquad \epsilon_t \sim N(0,1)
$$

$$
a_t \quad = \begin{cases} a_1, & \text{for } t \leq \frac{1}{2}T \\ a_2, & \text{for } t > \frac{1}{2}T \end{cases}
$$

where T is the sample size. That is, the data generating process has a break in the coefficient on the lagged dependent variable halfway through the sample. Note here that ϵ_t is conditionally homoskedastic but that y_t is conditionally heteroskedastic due to the mean shift. The results are in Table 1. Because the null hypothesis of no ARCH is true, this table documents the size of the test. The first panel of the table gives the percentage of rejections for the standard LM statistic for a variety of values of a_1 and a_2. Along the diagonal of the matrix (when $a_1 = a_2$), the model is correctly specified. Here the size of the test is approximately equal to its nominal level. The off-diagonal elements detail the extent of size distortion that occurs due to misspecification of the conditional mean. As we saw in section 2, the test for ARCH is contaminated by the misspecification of the conditional mean. It is evident that in many cases, the size distortions are large. When $|a_1 - a_2| > 0.6$, the size distortion is more than twice its level.

The next panel of table 1 shows the improvement when the conditioning variables are naively expanded to include longer lags. In this case, overparametrizing the model has desirable effects because lags of y_t approximate the serial correlation in $\widehat{u}_t$. In panel three, the information set is expanded to include recursive residuals and their squares. For the most extreme cases (where $|a_1 - a_2| > 0.6$), the improvement is greater than 50%. Even greater improvement in test performance results from inclusion of the cumulative sum of the recursive residuals, as shown in the fifth panel of table 1. The cusum by itself, however, does not greatly reduce the size distortions (panel 4).

Example 2. Mean Shift

In this example, the data are generated according to the following:

$$y_t = \begin{cases} a_1 y_{t-1} + \epsilon_t, & \text{for } t \leq \frac{1}{2}T \qquad \epsilon_t \sim N(0,1) \\ \mu + a_1 y_{t-1} + \epsilon_t, & \text{for } t > \frac{1}{2}T \end{cases}$$

where T is the sample size. That is, the data generating process has a shift in the mean halfway through the sample. The four panels in Table 2 mirror those in Table 1, although here the first column (under the 0.0 heading) corresponds to the case of no structural change. In this first column, the standard ARCH test seems to be slightly undersized. For the most part, the standard ARCH test is fairly robust to mean-shift misspecification, except when the shift is very large (2 times the standard deviation) or when the autoregressive parameter displays a lot of persistence. The naive approach of including more lags alleviates size distortions in many cases but is still inadequate for extremely large breaks. The third panel contains rejection frequencies of the modified ARCH test, where the first stage residuals are generated from a regression of y_t on one lag, the lagged recursive residuals and their squares. The size distortions that were evident in panel 1 for high values of μ and a_1 are significantly reduced. Addition of the cusum function reduces the distortions in the unit root case substantially.

Example 3. Additive Outlier

This example is similar to the one considered in van Dijk, Franses, and Lucas [1996]. The data are generated according to

$$\begin{aligned} y_t &= a_1 y_{t-1} + \epsilon_t + \psi_t & \epsilon_t &\sim N(0,\,1) \\ \psi_t &\sim N(0, \sigma_2^2) & \text{for } \tfrac{T}{2} - 1 &\leq t \leq \tfrac{T}{2} + 1 \end{aligned}$$

so that the data generating process experiences an outlier in the middle three periods of the sample (for $T = 200$). The results are in Table 3. The standard ARCH test has size approximately equal to its level when the variance of the outlier (i.e.σ_2^2) is small. When σ_2^2 exceeds the variance of ϵ, however, the standard ARCH test rejects the null hypothesis too frequently. There is also some evidence that this is exacerbated by higher levels of persistence (as given by higher a_1). When the information set includes the recursive residuals and their squares, the size distortions are not appreciably smaller when the outlier variance is large (as seen in panel 2 of Table 3). The inclusion of the cusum function reduces the distortions by up to 40%. It is not surprising that the proposed functions do not improve the size of the ARCH test as substantially as they did in the previous two examples. In this example, the misspecification is contemporaneous, that is, it is unpredictable at previous periods. Therefore, functions of lagged information will not completely robustify the ARCH test.

Other Examples

Table 4 presents three additional simulations. The first panel shows results from ignoring an MA(1) error in the data generating process. That is, the DGP is $y = C + e_t$, where e_t is MA(1) with moving average coefficient θ. From the first row of this panel, we see that failure to account for an MA(1) error results in large size distortions for the ARCH test, especially when the MA root is large. Inclusion of the second order polynomial in the recursive residuals (lines 3 and 5) virtually eliminates these distortions, while the cusum (line 4) has little effect. The cusum result is not surprising, since autocorrelations of an MA(1) process are zero beyond the first lag. Thus we would not expect lags of standardized residuals (beyond the first lag) to have any predictive power. Inclusion of the lagged recursive residual, however, is analogous to performing a GLS correction. Similarly, the "naive" approach of adding four lags of the dependent variable removes much of the distortion; this is because the example involves unaccounted for serial correlation and adding additional lags alleviates some of the bias induced by this type of misspecification.

The second panel of Table 4 explores the effects of omitted variables bias, when the omitted variable is serially correlated. That is, the data generating process is $y_t = C + x_t + \epsilon_t$ where $x_t = \rho x_{t-1} + \nu_t$ but the regression is $y_t = C + u_t$ (this is the example considered in section 2). For highly persistent values of ρ, the size distortion is very large because the variance of x_t, ($\frac{\sigma_\nu^2}{1-\rho^2}$), is large relative to that of e_t. Thus, as in the example with the additive outlier, size distortions are large when the variance from the specification error is large. In the extreme case when $\rho = 1$, x_t is an integrated process and the ARCH test rejects 100% of the time. This is consistent with the example presented in Section 2 which shows that the extent of serial correlation in the squared residuals is determined by the degree of serial correlation in η_t. In these highly persistent cases ($\rho = 0.9$ and $\rho = 1$), the size distortions are substantially reduced when the recursive residuals and their squares are included because they provide a GLS-type correction. However, the cusum provides almost no improvement because the cusum series is doubly integrated and its inclusion leaves the regression unbalanced.

The third panel of Table 4 considers the common question of trend-stationarity versus difference-stationarity. In particular, suppose the data generating process is $y_t = d_0 + d_1 t + a y_{t-1} + \epsilon_t$, with $|a| < 1$, so that y_t is trend stationary but we estimate $\Delta y_t = c + u_t$; that is, we accidentally first difference the data and thus u_t is over-differenced. In the simulation, we choose $d_0 = 0$ and $d_1 = 0.1$. The size distortions of the ARCH test are quite severe when a is small. In the extreme when $a = 0$, $\widehat{u}_t$ has a non-invertible moving-average component. As a increases, the ARCH test has size approximately equal to its level. This is because a near unit root is well approximated by a unit root process in finite samples, and first-differencing appears to have little implication for the ARCH test.

Size-adjusted Power of the Tests

Table 5 considers the power of the tests when the alternative of ARCH is true. The data are generated as in example 1 but the errors ϵ_t are conditionally normal with mean 0 and conditional variance h_t. We assume for this example that h_t is an ARCH(1), that is $h_t = \omega + \alpha_0 \epsilon_{t-1}^2$, where $\alpha_0 = 0.5$ and $\omega = 1$. There appears to be a tradeoff between the number of regressors included in the first stage regression and the power of the tests. While increasing the information set only can improve the extent to which we fit m_t, beyond a certain point, the marginal improvement is minimal. Thus it is important to determine the optimal set of projection instruments.[11]

5 Empirical Examples

Our first example considers the daily S&P500 returns from Bera and Higgins [1997]. Following Bera and Higgins, we estimate an AR(1) for the conditional mean equation and use the residuals to conduct a standard ARCH(1) test; the value of the test statistic is 5.39, rejecting the null hypothesis of conditional homoskedasticity. If we additionally include the recursive residuals and their squares in the conditional mean equation, however, the resulting ARCH(1) test statistic is 0.022, well below the corresponding critical value. Further tests for structural change indicate strong evidence in favor of these types of nonlinearity. Bera and Higgins compare two competing nonlinear models and find that neither can be rejected, supporting the notion that this series exhibits nonlinearity. They argue that a GARCH(1,1) model is preferred to a bilinear specification (which allows for nonlinearity in the conditional mean). Our results suggest that perhaps some other nonlinear model would be preferred to the GARCH specification; accounting for this nonlinearity weakens the evidence in favor of conditional heteroskedasticity.

Our second example is meant to show that the modifications we propose do not substantially reduce the power of the test if in fact there is ARCH. The example uses data from Engle-Bollerslev [1986] (the weekly US$/Swiss Franc exchange rate, originally from Diebold and Nerlove [1989]); the data are in log first differences. The data are tested for ARCH(4), with one lag included in the

[11] Similar conclusions regarding power hold for the other examples considered – these are available from the authors on request.

conditional mean equation. The value of the TR^2 test is 50.21. Under the null hypothesis of no ARCH, this statistic is distributed $\chi^2(4)$; this hypothesis is strongly rejected. Using the "naive" approach of including four lags in the conditional mean equation to control for some apparent serial correlation, the statistic is reduced to 49.95. Including the recursive residuals and their squares reduces the statistic further, to 37.90. Thus while the evidence for conditional heteroskedasticity is diminished, it is still apparent in this dataset. An informal examination of a plot of the pre-differenced data suggests the possible presence of a structural break in trend. As discussed earlier, if the nature of the misspecification is known, it is possible to improve the size properties of the TR^2 test to account for such misspecification; a formal test of the null hypothesis of no structural break using a recursive Quandt likelihood ratio statistic with four lags (Banerjee, Lumsdaine, Stock [1992]) rejects this hypothesis, estimating a break in November 1977. We recompute the ARCH test controlling for a break at this estimated date (i.e., we regress the data on a constant and an indicator variable equal to zero before the date and 1 afterwards; we use the residuals from this first stage regression to perform the ARCH test); the value of the statistic decreases to 47.47 but is still highly significant.

6 Conclusions

This paper has emphasized the importance of testing for model misspecification in lower moments before conducting routine diagnostic tests regarding conditional heteroskedasticity. We have shown that failure to account for conditional mean misspecification can produce spurious results and lead to overrejection of the null hypothesis of conditional homoskedasticity when testing for ARCH(1). This intuition extends more generally to LM-based tests for higher order ARCH. We have proposed a method for adjusting the standard ARCH test to allow for possible misspecification of unknown form. This method has been shown in simulations to reduce the size distortions of the ARCH(1) test by a substantial amount. In addition, we consider two empirical examples.

Appendix

If X_t is AR(1), then X_t^2 is an ARMA(2,1):

Suppose a random variable X_t is an AR(1) with Gaussian errors e_t (mean zero, variance σ_e^2).

$$X_t - \mu = \phi(X_{t-1} - \mu) + e_t.$$

Let $\sigma^2 = \sigma_e^2/(1-\phi^2)$ denote the variance of X_t. Then $\zeta_t = (X_t - \mu)/\sigma$, is a mean zero, unit variance AR(1) process with $cov(\zeta_t, \zeta_{t-\tau}) = \rho^\tau$ for $\tau > 0$. Consider

$$\begin{aligned} Y_t &= X_t^2 = (\mu + \sigma\zeta_t)^2 = \mu^2 + 2\mu\sigma\zeta_t + \sigma^2\zeta_t^2, \\ E(Y_t) &= \mu^2 + \sigma^2, \quad \text{and} \\ Y_t - E(Y_t) &= 2\mu\sigma\zeta_t + \sigma^2(\zeta_t^2 - 1). \end{aligned}$$

Then

$$\begin{aligned} cov(Y_t, Y_{t-\tau}) &= E[4\mu^2\sigma^2\zeta_t\zeta_{t-\tau} + \sigma^4(\zeta_t^2 - 1)(\zeta_{t-\tau}^2 - 1) + 2\mu\sigma^3\zeta_t(\zeta_{t-\tau}^2 - 1) + 2\mu\sigma^3\zeta_{t-\tau}(\zeta_t^2 - 1)] \\ &= 4\mu^2\sigma^2\rho^\tau + 2\sigma^4\rho^{2\tau}. \end{aligned}$$

To show that Y_t has the same autocovariance structure as an ARMA(2,1) process, it suffices to show that ζ_t^2 can be written as the sum of two AR(1) processes (see, e.g., Hamilton, page 108). As in Granger and Newbold (1976), consider the following two AR processes:

$$\omega_{1t} = \rho\omega_{1t-1} + \nu_{1t}$$

$$\omega_{2t} = \rho^2\omega_{2t-1} + \nu_{2t}$$

where ν_{1t} and ν_{2t} are Gaussian and uncorrelated with each other, with variance $\sigma_{\nu_1}^2 = 4\mu^2\sigma^2(1-\rho^2)$ and $\sigma_{\nu_2}^2 = 2\sigma^4(1-\rho^4)$, respectively. Let $\Gamma_i(\tau)$ denote the τ-th autocovariance for ω_{it}, $i = 1, 2$. Consider ω_{1t}. Then

$$\Gamma_{\omega_1}(\tau) = \rho^\tau\Gamma_{\omega_1}(0)$$

where $\Gamma_{\omega_1}(0) = \frac{\sigma_{\nu_1}^2}{1-\rho^2}$. Therefore $\Gamma_{\omega_1}(\tau) = 4\rho^\tau\mu^2\sigma^2$. Similarly, consider ω_{2t}. Then

$$\Gamma_{\omega_2}(\tau) = \rho^{2\tau}\Gamma_{\omega_2}(0)$$

where $\Gamma_{\omega_2}(0) = \frac{\sigma_{\nu_2}^2}{1-\rho^4}$. Therefore $\Gamma_{\omega_2}(\tau) = 2\rho^{2\tau}\sigma^4$. Because ω_{1t} and ω_{2t} are uncorrelated covariance stationary AR(1) processes, the autocovariance generating function of their sum (i.e., ζ_t^2) is equal to the sum of their individual autocovariance generating functions. By construction, this is $4\mu^2\sigma^2\rho^\tau + 2\sigma^4\rho^{2\tau}$. ■

REFERENCES

Banerjee, Anindya, Robin L. Lumsdaine, and James H. Stock [1992], "Recursive and Sequential Tests of the Unit Root and Trend Break Hypotheses: Theory and International Evidence," *Journal of Business and Economic Statistics* 10, 271-288.

Bera, Anil K., and Matthew L. Higgins [1997], "ARCH and Bilinearity as Competing Models for Nonlinear Dependence," *Journal of Business and Economic Statistics* 15(1), 43-50.

Bera, Anil K., Matthew L. Higgins, and Sangkyu Lee [1992], "Interaction Between Autocorrelation and Conditional Heteroskedasticity: A Random-Coefficient Approach," *Journal of Business and Economic Statistics* 10(2), 133-142.

Bera, Anil K., and Sangkyu Lee [1993], "Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis," *Review of Economic Studies* 60, 229-240.

Brown, R.L., J. Durbin, and J.M. Evans [1975], "Techniques for Testing the Constancy of Regression Relationship over Time," *Journal of the Royal Statistical Society, Series B* , Ser. B, 149-192.

Diebold, Francis X., and Marc Nerlove [1989], "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics* 4(1), 1-21.

Engle, Robert F. [1982], "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50(4), 987-1007.

Engle, Robert F., David F. Hendry, and David Trumble [1985], "Small Sample Properties of ARCH Estimators and Tests," *Canadian Journal of Economics* 18, 66-93.

Giles, David E.A., Judith A. Giles, and Jason K. Wong [1993], "Testing for ARCH-GARCH Errors in a Mis-Specified Regression," *Computational Statistics* 8, 109-126.

Granger, C.W.J., and M.J. Morris [1976], "Time Series Modelling and Interpretation," *Journal of the Royal Statistical Society, Series B* A 139, 246-257.

Granger, C.W.J., and P. Newbold [1976], "Forecasting Transformed Series," *Journal of the Royal Statistical Society, Series B* B, 38, 189-203.

Granger, Clive W.J., and Timo Teräsvirta [1993]. *Modelling Nonlinear Economic Relationships.* New York: Oxford University Press, Inc.

Hansen, Bruce E. [1992], "Testing for Parameter Instability in Linear Models," *Journal of Policy Modelling* 14(4), 517-533.

Harvey, Andrew C. [1990]. *The Econometric Analysis of Time Series, 2nd edition.* Cambridge, MA: MIT Press.

Harvey, Andrew C., and P. Collier [1977], "Testing for Function Misspecification in Regression Analysis," *Journal of Econometrics* 6, 103-119.

Harvey, Andrew C., and G.D.A. Phillips [1974], "A Comparison of the Power of Some Tests of

Heteroskedasticity in the General Linear Model," *Journal of Econometrics* 2, 307-316.

Hong, Yongmiao, and Ramsey D. Shehadeh [1996], "A New Test for ARCH Effects and its Finite Sample Performance," unpublished manuscript, Cornell University.

Kianifard, Farid, and William H. Swallow [1996], "A Review of the Development and Application of Recursive Residuals in Linear Models," *Journal of the American Statistical Association* 91(433), 391-400.

Lamoureux, Christopher G., and William D. Lastrapes [1990], "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business and Economic Statistics* 8, 225-234.

Lee, Byung-Joo [1992], "A Heteroskedasticity Test Robust to Conditional Mean Misspecification," *Econometrica* 60(1), 159-171.

Ng, Serena, and Timothy J. Vogelsang [1996], "Analysis of Vector Autoregressions in the Presence of Shifts in Mean," manuscript, Cornell University.

Perron, Pierre [1989], "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica* 57, 1361-1401.

Rilstone, Paul [1992], "A Simple Bera-Jarque Normality Test for Nonparametric Residuals," *Econometric Reviews* 11(3), 355-365.

Robinson, Peter M. [1988], "Root-N Consistent Semiparametric Regression," *Econometrica* 56(4), 931-954.

Robinson, Peter M. [1991], "Testing for Strong Serial Correlation and Dynamic Conditional Heteroskedasticity in Multiple Regression," *Journal of Econometrics* 47, 67-84.

Tong, Howell [1990]. *Nonlinear Time Series.* Oxford: Oxford Science Publications.

van Dijk, Dick, Philip Hans Franses, and André Lucas [1996], "Testing for ARCH in the Presence of Additive Outliers," unpublished manuscript, Erasmus University.

Weiss, Andrew A. [1986], "Asymptotic Theory for ARCH Models: Estimation and Testing," *Econometric Theory* 2, 107-131.

Whang, Y-J. [1998], "A Test of Normality Using Nonparametric Residuals," forthcoming, *Econometric Reviews* .

Wooldridge, Jeffrey M. [1991a], "On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances," *Journal of Econometrics* 47, 5-46.

———[1991b], "Specification Testing and Quasi-Maximum Likelihood Estimation," *Journal of Econometrics* 48, 29-55.

Zivot, Eric, and Donald W.K Andrews [1992], "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business and Economic Statistics* 10, 251-270.

Figure 1: Is there ARCH in this simulated series?

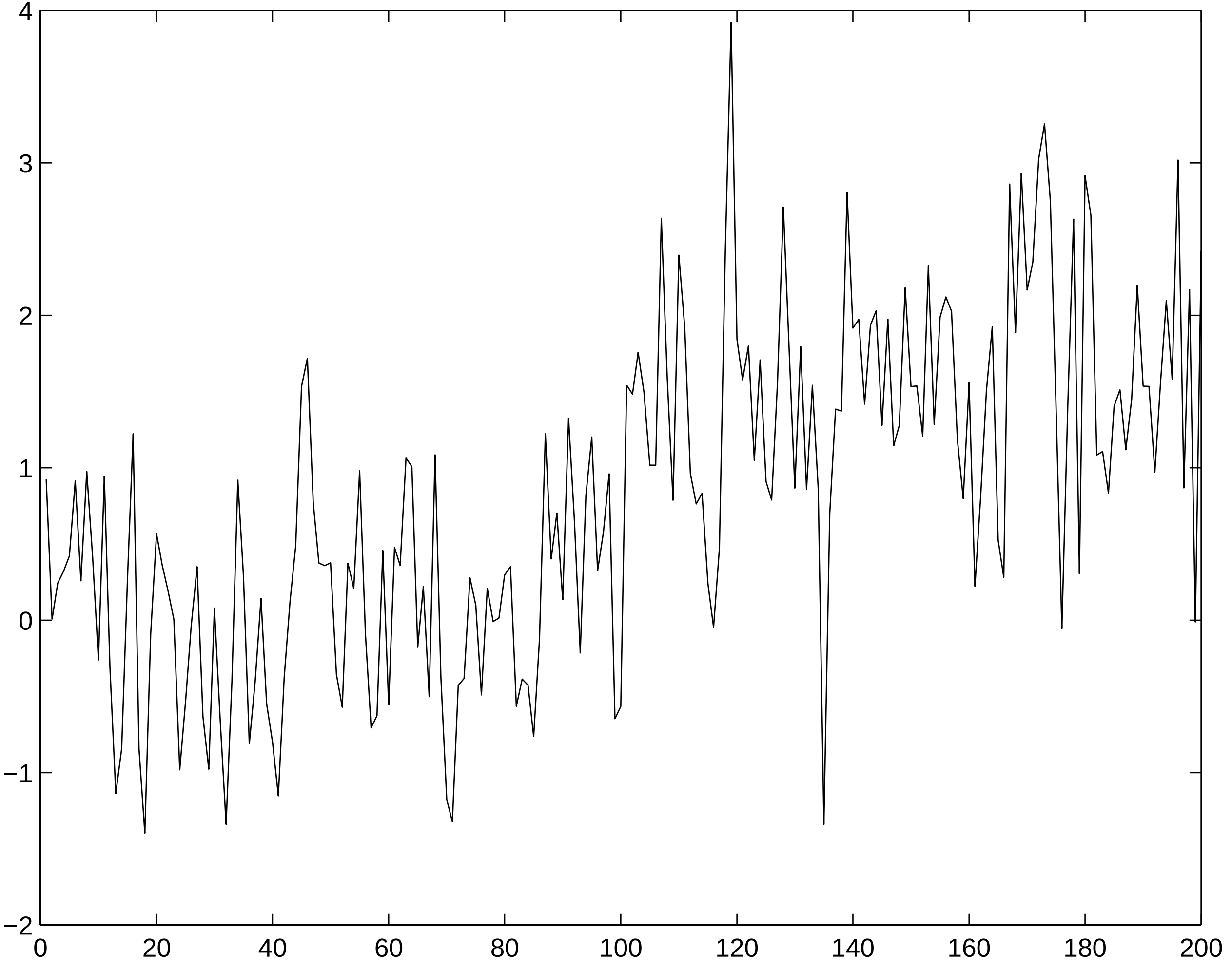

Figure 2: Plot of Recursive Residuals

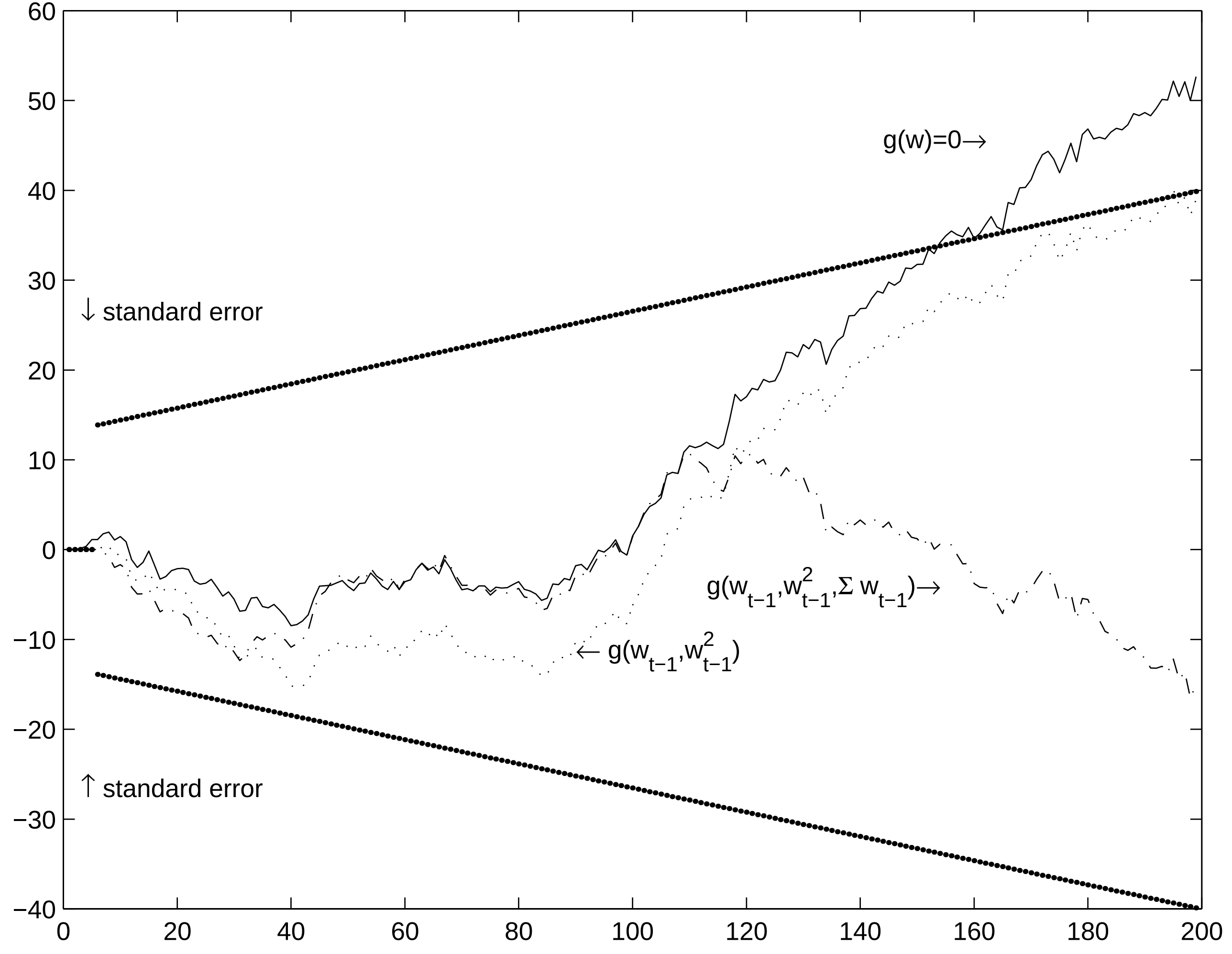

Table 1: Size of the Tests: Nominal Size=5%

$$\begin{aligned} y_t &= a_1 y_{t-1} + e_t \quad & t <= 100 \quad & e_t \sim N(0,1) \\ y_t &= a_2 y_{t-1} + e_t \quad & 101 \leq t \leq 200 & \end{aligned}$$

Standard LM test:

a_1 \ a_2	0.1	0.3	0.5	0.7	0.9	1.0
0.1	0.0417	0.0479	0.0736	0.1702	0.3427	0.5111
0.3	0.0446	0.0413	0.0447	0.0731	0.1498	0.2341
0.5	0.0781	0.0455	0.0385	0.0457	0.0712	0.0985
0.7	0.1712	0.0676	0.0448	0.0426	0.0482	0.0543
0.9	0.3495	0.1534	0.0712	0.0493	0.0460	0.0456
1.0	0.1690	0.2235	0.4067	0.3540	0.1038	0.0410

"Naive" LM:

0.1	0.0413	0.0425	0.0649	0.1273	0.1742	0.1686
0.3	0.0445	0.0383	0.0460	0.0619	0.0943	0.0985
0.5	0.0687	0.0430	0.0405	0.0439	0.0649	0.0644
0.7	0.1249	0.0668	0.0435	0.0397	0.0424	0.0486
0.9	0.1807	0.0988	0.0605	0.0432	0.0394	0.0404
1.0	0.1531	0.2172	0.3064	0.2876	0.0633	0.0433

Modified LM(a):

0.1	0.0395	0.0411	0.0581	0.1285	0.2402	0.2801
0.3	0.0397	0.0402	0.0395	0.0538	0.0989	0.1221
0.5	0.0581	0.0374	0.0371	0.0396	0.0516	0.0641
0.7	0.1047	0.0516	0.0388	0.0411	0.0422	0.0420
0.9	0.1497	0.0816	0.0498	0.0398	0.0400	0.0388
1.0	0.0531	0.0554	0.0617	0.0769	0.0494	0.0383

Modified LM(b):

0.1	0.0417	0.0461	0.0739	0.1681	0.3428	0.4415
0.3	0.0457	0.0422	0.0450	0.0710	0.1509	0.1986
0.5	0.0772	0.0474	0.0390	0.0450	0.0704	0.0887
0.7	0.1670	0.0663	0.0431	0.0430	0.0474	0.0497
0.9	0.3313	0.1432	0.0723	0.0450	0.0454	0.0447
1.0	0.1635	0.2283	0.3968	0.3394	0.0917	0.0431

Modified LM(c):

0.1	0.0412	0.0408	0.0572	0.1242	0.2353	0.2697
0.3	0.0370	0.0396	0.0386	0.0537	0.0975	0.1167
0.5	0.0583	0.0369	0.0386	0.0382	0.0515	0.0617
0.7	0.1036	0.0500	0.0365	0.0400	0.0424	0.0418
0.9	0.1537	0.0804	0.0494	0.0386	0.0408	0.0379
1.0	0.0549	0.0558	0.0598	0.0737	0.0492	0.0399

Notes: The simulations are based on 200 observations and 10,000 replications. Each element of the table gives the rejection frequencies of TR^2 from an autoregression of the squared residuals onto a constant and one lag, where the residuals are obtained via estimation of the following five models (equations 7a-e in the text): Standard LM uses a first order autoregression, "Naive" LM is a fourth order autoregression; LM(a), LM(b), and LM(c) are based on estimated models of the form $y_t = \gamma_0 + \gamma_1 y_{t-1} + g(\widehat{w}_{t-1}) + v_t$, where $g(\widehat{w}_{t-1}) = \widehat{w}_{t-1} + \widehat{w}^2_{t-1}$, $\sum_{i=t}^{t-1} \widehat{w}_i$, and $\widehat{w}_{t-1} + \widehat{w}^2_{t-1} + \sum_{i=t}^{t-1} \widehat{w}_i$, for (a), (b), and (c), respectively, and $\widehat{w}_{t-1}$ are the lagged recursive residuals, computed from the first order autoregression (used in the standard LM).

Table 2: Size of the Tests: Nominal Size=5%

$$y_t = a_1 y_{t-1} + e_t \qquad t <= 100$$
$$y_t = \mu + a_1 y_{t-1} + e_t \qquad 101 \leq t \leq 200$$

LM test:

a_1 \ μ	0.0	0.5	1.0	1.5	2.0	2.5
0.1	0.0417	0.0413	0.0426	0.0567	0.1102	0.1781
0.3	0.0388	0.0420	0.0410	0.0651	0.1188	0.1619
0.5	0.0406	0.0422	0.0492	0.0688	0.1087	0.1652
0.7	0.0423	0.0413	0.0489	0.0699	0.1350	0.2939
0.9	0.0387	0.0454	0.0519	0.1396	0.4184	0.7718
1.0	0.0413	0.0432	0.0727	0.2269	0.6180	0.9264

“Naive” LM:

0.1	0.0413	0.0391	0.0393	0.0491	0.0698	0.1078
0.3	0.0380	0.0402	0.0411	0.0496	0.0774	0.1401
0.5	0.0375	0.0408	0.0422	0.0568	0.1021	0.2068
0.7	0.0407	0.0437	0.0451	0.0681	0.1654	0.3362
0.9	0.0395	0.0414	0.0471	0.1004	0.2476	0.4340
1.0	0.0424	0.0391	0.0420	0.0449	0.0534	0.0763

Modified LM(a):

0.1	0.0395	0.0393	0.0462	0.0612	0.0803	0.1142
0.3	0.0370	0.0395	0.0401	0.0500	0.0732	0.1133
0.5	0.0394	0.0366	0.0401	0.0493	0.0778	0.1235
0.7	0.0391	0.0374	0.0426	0.0562	0.0820	0.1373
0.9	0.0370	0.0407	0.0421	0.0554	0.0874	0.1430
1.0	0.0384	0.0379	0.0353	0.0739	0.3688	0.8241

Modified LM(b):

0.1	0.0417	0.0435	0.0425	0.0513	0.0776	0.1190
0.3	0.0398	0.0420	0.0435	0.0559	0.0970	0.1577
0.5	0.0416	0.0429	0.0469	0.0594	0.1131	0.2038
0.7	0.0419	0.0405	0.0461	0.0743	0.1312	0.2530
0.9	0.0398	0.0433	0.0493	0.1180	0.3356	0.6264
1.0	0.0422	0.0432	0.0742	0.2025	0.4220	0.6124

Modified LM(c):

0.1	0.0412	0.0402	0.0356	0.0388	0.0511	0.0877
0.3	0.0365	0.0402	0.0362	0.0413	0.0661	0.1034
0.5	0.0387	0.0357	0.0399	0.0449	0.0743	0.1200
0.7	0.0392	0.0375	0.0383	0.0545	0.0730	0.1145
0.9	0.0385	0.0405	0.0397	0.0519	0.0819	0.1185
1.0	0.0393	0.0401	0.0399	0.0553	0.0874	0.1313

See notes to Table 1.

Table 3: Size of the Tests: Nominal Size=5%

$$y_t = a_1 y_{t-1} + e_t + \psi_t$$
$$\psi_t \sim N(0, \sigma_2^2) \quad 101 \leq t \leq 103 \qquad \text{and 0 otherwise.}$$

LM test:

σ_2 / a_1	0.0	1.0	2.5	5.0	7.5	10.0
0.1	0.0422	0.0455	0.1269	0.4139	0.4805	0.4904
0.3	0.0379	0.0449	0.1127	0.4200	0.4919	0.4971
0.5	0.0395	0.0415	0.1222	0.4261	0.5057	0.5095
0.7	0.0419	0.0481	0.1221	0.4330	0.5246	0.5342
0.9	0.0409	0.0502	0.1251	0.4542	0.5377	0.5625
1.0	0.0423	0.0439	0.1369	0.4628	0.5532	0.5740

"Naive" LM:

0.1	0.0382	0.0432	0.1206	0.4106	0.4773	0.4992
0.3	0.0427	0.0473	0.1190	0.4016	0.4728	0.4921
0.5	0.0403	0.0441	0.1180	0.4141	0.4826	0.4926
0.7	0.0428	0.0470	0.1162	0.4091	0.4739	0.4883
0.9	0.0454	0.0426	0.1168	0.4065	0.4793	0.4973
1.0	0.0418	0.0455	0.1184	0.4065	0.4768	0.4836

Modified LM(a):

0.1	0.0439	0.0463	0.1267	0.4122	0.4786	0.4884
0.3	0.0391	0.0436	0.1143	0.4172	0.4897	0.4944
0.5	0.0402	0.0431	0.1214	0.4223	0.5021	0.5069
0.7	0.0413	0.0481	0.1233	0.4313	0.5212	0.5322
0.9	0.0405	0.0505	0.1245	0.4503	0.5368	0.5615
1.0	0.0437	0.0448	0.1334	0.4595	0.5477	0.5662

Modified LM(b):

0.1	0.0370	0.0420	0.0882	0.2510	0.2827	0.2866
0.3	0.0367	0.0395	0.0781	0.2512	0.2894	0.2875
0.5	0.0385	0.0372	0.0805	0.2587	0.2874	0.2865
0.7	0.0373	0.0386	0.0806	0.2548	0.3049	0.3014
0.9	0.0373	0.0439	0.0802	0.2688	0.3078	0.3234
1.0	0.0383	0.0389	0.0810	0.2805	0.3223	0.3374

Modified LM(c):

0.1	0.0379	0.0435	0.0884	0.2508	0.2809	0.2847
0.3	0.0373	0.0384	0.0782	0.2507	0.2887	0.2864
0.5	0.0384	0.0384	0.0816	0.2565	0.2877	0.2858
0.7	0.0371	0.0379	0.0791	0.2553	0.3023	0.3012
0.9	0.0369	0.0429	0.0810	0.2679	0.3079	0.3226
1.0	0.0385	0.0391	0.0791	0.2790	0.3186	0.3371

See notes to Table 1.

Table 4

MA(1) error:

θ	0.0	0.1	0.3	0.5	0.7	0.9	1.0
LM	0.0432	0.0505	0.1734	0.5126	0.7668	0.8509	0.8552
"Naive"	0.0412	0.0395	0.0379	0.0396	0.0403	0.0447	0.0489
LM(a)	0.0389	0.0380	0.0386	0.0447	0.0569	0.0689	0.0735
LM(b)	0.0431	0.0491	0.1692	0.4964	0.7531	0.8406	0.8449
LM(c)	0.0395	0.0384	0.0392	0.0424	0.0562	0.0678	0.0719

AR(1) regressor omitted:

ρ	0.1	0.3	0.5	0.7	0.9	1.0
LM	0.0454	0.0400	0.0408	0.0614	0.3971	1.0000
"Naive"	0.0423	0.0438	0.0459	0.0677	0.4000	1.0000
LM(a)	0.0433	0.0376	0.0331	0.0380	0.0446	0.0942
LM(b)	0.0455	0.0424	0.0404	0.0607	0.3263	1.0000
LM(c)	0.0433	0.0387	0.0331	0.0395	0.0387	0.5151

Difference a Trend-stationary Series:

α	0.1	0.3	0.5	0.7	0.9	1.0
LM	0.7224	0.3591	0.1457	0.0686	0.0451	0.1689
"Naive"	0.7247	0.3584	0.1484	0.0702	0.0438	0.1660
LM(a)	0.0584	0.0490	0.0491	0.0459	0.0400	0.0791
LM(b)	0.1297	0.0834	0.0588	0.0468	0.0440	0.0420
LM(c)	0.0532	0.0463	0.0411	0.0407	0.0396	0.0373

See notes to Table 1.

Table 5: Size-adjusted Power of the Tests:

$$\begin{aligned} y_t &= a_1 y_{t-1} + e_t \qquad t <= 100 \\ y_t &= a_2 y_{t-1} + e_t \qquad 101 \le t \le 200 \end{aligned}$$

LM test:

a_2 / a_1	0.1	0.3	0.5	0.7	0.9	1.0
0.1	0.9729	0.9732	0.9599	0.8853	0.6958	0.5256
0.3	0.9726	0.9765	0.9737	0.9571	0.8534	0.6934
0.5	0.9586	0.9736	0.9765	0.9700	0.9433	0.8745
0.7	0.8931	0.9620	0.9726	0.9725	0.9678	0.9467
0.9	0.6734	0.8508	0.9368	0.9668	0.9792	0.9690
1.0	0.2936	0.3665	0.3390	0.3667	0.9451	0.9745

"Naive" LM:

0.1	0.9715	0.9696	0.9570	0.9008	0.7807	0.6867
0.3	0.9733	0.9720	0.9659	0.9538	0.8897	0.8109
0.5	0.9563	0.9711	0.9725	0.9717	0.9405	0.8950
0.7	0.9130	0.9511	0.9671	0.9714	0.9663	0.9385
0.9	0.7566	0.8711	0.9380	0.9660	0.9706	0.9644
1.0	0.3167	0.3478	0.4082	0.5728	0.9471	0.9703

Modified LM(a):

0.1	0.9252	0.9254	0.9024	0.7830	0.6187	0.5169
0.3	0.9325	0.9305	0.9265	0.8996	0.7825	0.7048
0.5	0.8978	0.9266	0.9280	0.9190	0.8799	0.8257
0.7	0.8016	0.9036	0.9246	0.9240	0.9126	0.8836
0.9	0.6338	0.7691	0.8663	0.9130	0.9302	0.9114
1.0	0.2638	0.3470	0.4790	0.6544	0.8902	0.9230

Modified LM(b):

0.1	0.9531	0.9497	0.9236	0.8035	0.5825	0.4385
0.3	0.9514	0.9571	0.9500	0.9189	0.7664	0.6194
0.5	0.9156	0.9504	0.9570	0.9450	0.8987	0.8165
0.7	0.8188	0.9252	0.9521	0.9564	0.9445	0.9144
0.9	0.5726	0.7753	0.8937	0.9467	0.9568	0.9428
1.0	0.2235	0.2767	0.2210	0.1870	0.8973	0.9522

Modified LM(c):

0.1	0.9192	0.9200	0.9010	0.7848	0.6173	0.5194
0.3	0.9282	0.9258	0.9227	0.8956	0.7812	0.7074
0.5	0.8923	0.9236	0.9240	0.9162	0.8742	0.8221
0.7	0.8014	0.9014	0.9219	0.9212	0.9088	0.8774
0.9	0.6315	0.7666	0.8628	0.9095	0.9250	0.9076
1.0	0.2635	0.3461	0.4787	0.6586	0.8851	0.9121

Notes: $e_t \sim N(0, h_t)$, where $h_t = 1.0 + 0.5e_{t-1}^2$, as discussed in the text. Also see notes to Table 1.