Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form

Sílvia Gonçalves
(CRDE, CIRANO and Université de Montréal)

Lutz Kilian
(University of Michigan, European Central Bank and CEPR)

Discussion paper 26/02
Economic Research Centre
of the Deutsche Bundesbank

November 2002

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 95 66-1
Telex within Germany 4 1 227, telex from abroad 4 14 431, fax +49 69 5 60 10 71

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax No. +49 69 95 66-30 77

ISBN 3–935821–35–2

Contents

List of Tables

List of Figures

Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form*

Abstract

Conditional heteroskedasticity is an important feature of many macroeconomic and financial time series. Standard residual-based bootstrap procedures for dynamic regression models treat the regression error as i.i.d. These procedures are invalid in the presence of conditional heteroskedasticity. We establish the asymptotic validity of three easy-to-implement alternative bootstrap proposals for stationary autoregressive processes with m.d.s. errors subject to possible conditional heteroskedasticity of unknown form. These proposals are the fixed-design wild bootstrap, the recursive-design wild bootstrap and the pairwise bootstrap. In a simulation study all three procedures tend to be more accurate in small samples than the conventional large-sample approximation based on robust standard errors. In contrast, standard residual-based bootstrap methods for models with i.i.d. errors may be very inaccurate if the i.i.d. assumption is violated. We conclude that in many empirical applications the proposed robust bootstrap procedures should routinely replace conventional bootstrap procedures based on the i.i.d. error assumption.

Zusammenfassung

Bedingte Heteroskedastizität ist eine wichtige Eigenschaft von vielen Daten über Finanzmärkte und die Makroökonomie. Standard bootstrap Verfahren für dynamische Regressionsmodelle behandeln die Residuen der Regression als i. i. d. Bei bedingter Heteroskedastizität sind diese Prozeduren nicht angemessen. Wir zeigen die asymptotische Gültigkeit von 3 alternativen bootstrap Methoden für stationäre autoregressive Prozesse mit m. d. s. Fehler, die eine bedingte Heteroskedastizität unbekannter Form aufweisen. Es geht dabei um ein fixed-design wild bootstrap, den recursive-design wild bootstrap und den paarweisen bootstrap. In einer Simulationsstudie erscheinen alle 3 Prozeduren in kleinen Stichproben angewandt genauer als die konventionellen Approximationen, die auf robusten Standardfehlern basieren. Diese letztgenannten Methoden können dagegen sehr ungenau sein, wenn die i. i. d. Annahme nicht gilt. Wir schließen daraus, dass bei vielen empirischen Anwendungen die robusten bootstrap Verfahren, die hier vorgestellt werden und leicht zu implementieren sind, die üblichen bootstrap Verfahren ersetzen sollten.

JEL:	C15, C22, C52
Keywords:	wild bootstrap; pairwise bootstrap; robust inference;
Garch:	stochastic volatility.

---

* We thank Atushi Inoue, Simone Manganelli, Nour Meddahi, Michael Wolf and Jonathan Wright for helpful discussions. The views expressed in this paper do not necessarily reflect the opinion of the ECB or its staff.

Sílvia Gonçalves, Département de sciences économiques, Université de Montréal, Canada. E-mail: silivia.goncalves@umontreal.ca.
Lutz Kilian, European Central Bank, DG Research, Kaiserstraße 29, D-60311 Frankfurt a.M., Germany. E-mail: lkilian@umich.edu.

1. Introduction

It is well known that there is evidence of conditional heteroskedasticity in the residuals of many estimated dynamic regression models in finance and in macroeconomics. This evidence is particularly strong for regressions involving monthly, weekly and daily data. Standard residual-based bootstrap methods of inference for autoregressions treat the error term as independent and identically distributed (i.i.d.) and are invalidated by conditional heteroskedasticity. In this paper, we analyze two main proposals for dealing with conditional heteroskedasticity of unknown form in autoregressions.

The first proposal is very easy to implement and involves an application of the wild bootstrap (WB) to the residuals of the dynamic regression model. The WB method allows for regression errors that follow martingale difference sequences (m.d.s.) with possible conditional heteroskedasticity. We investigate both the fixed-design and the recursive-design implementation of the WB for autoregressions. We prove their first-order asymptotic validity for the autoregressive parameters (and smooth functions thereof) under fairly general conditions including, for example, many stationary ARCH, GARCH and stochastic volatility error processes.

There are several fundamental differences between this paper and earlier work on the WB in regression models. First, existing theoretical work has largely focused on the classical linear regression model (see Davidson and Flachaire 2000). Second, Davidson and Flachaire (2000) establish the validity of the WB in the presence of unconditional heteroskedasticity in cross-sections, whereas we focus on conditional heteroskedasticity in time series. Third, much of the earlier work has focused on bootstrapping models restricted under the null hypothesis of a test, whereas we focus on the construction of bootstrap confidence intervals from unrestricted regression models (see Davidson and Flachaire 2000, Godfrey and Orme 2001).

The work most closely related to ours is Kreiss (1997). Kreiss established the asymptotic validity of a fixed-design WB for stationary autoregressions with known finite lag order when the error term exhibits a specific form of conditional heteroskedasticity. We provide a generalization of this result to m.d.s. errors with possible conditional heteroskedasticity of unknown form. Our results cover as special cases the N-GARCH, t-GARCH and asymmetric GARCH models, as well as stochastic volatility models.

Kreiss (1997) also proposed a recursive-design WB, under the name of "modified wild bootstrap", but he did not establish the consistency of this bootstrap proposal for autoregressive processes with conditional heteroskedasticity. We prove the first-order asymptotic validity of the recursive-design WB for finite-order autoregressions with m.d.s. errors subject to possible conditional heteroskedasticity of unknown form. The proof holds under slightly stronger assumptions than the proof for the fixed-design WB.

Tentative simulation evidence shows that the recursive-design WB scheme works well in small samples for a wide range of models of conditional heteroskedasticity. In contrast, conventional residual-based resampling schemes based on the i.i.d. assumption may be very inaccurate in the presence of conditional heteroskedasticity. Moreover, the recursive-design WB method works equally well in the i.i.d. error case. The recursive-design WB method is typically more accurate in small samples than the fixed-design WB method. It also tends to be more accurate than the Gaussian large-sample approximation based on robust standard errors.

The second proposal for dealing with conditional heteroskedasticity of unknown form involves the pairwise resampling of the observations. This method was originally suggested by Freedman (1981) for cross-sectional models. We establish the asymptotic validity of this method in the autoregressive context and compare its performance to that of the fixed-design and of the recursive-design WB. The pairwise bootstrap is less efficient than the residual-based WB, but - like the fixed-design WB - it remains valid for a broader range of GARCH processes than the recursive-design WB, including EGARCH, AGARCH and GJR-GARCH processes, which have been proposed specifically to capture asymmetric responses to shocks in asset returns (see, e.g., Engle and Ng (1993) for a review). We find in Monte Carlo simulations that the pairwise bootstrap is typically more accurate than the fixed-design WB method, but in small samples tends to be somewhat less accurate than the recursive-design WB when the data are persistent. For large samples these differences vanish, and the pairwise bootstrap is as accurate as the recursive-design WB.

The theoretical and simulation results in this paper suggest that no single method of dealing with conditional heteroskedasticity of unknown form will be optimal in all cases. We conclude that the recursive-design WB should replace conventional recursive-design i.i.d. bootstrap methods in many standard applications in empirical macroeconomics. This method performs equally well, whether the error term is i.i.d. or conditionally heteroskedastic, but it lacks a theoretical justification for some forms of GARCH that have figured prominently in the literature on high-frequency returns. When sample sizes are at least moderately large and the possibility of asymmetric forms of GARCH is a practical concern, the pairwise bootstrap provides a suitable alternative.

A third proposal for dealing with conditional heteroskedasticity of unknown form is the resampling of blocks of autoregressive residuals (see, e.g., Berkowitz, Birgean and Kilian 2000). No formal theoretical results exist that would justify such a bootstrap proposal. We do not consider this proposal for two reasons. First, in the context of a well-specified parametric model this proposal involves a loss of efficiency relative to the WB because it allows for serial correlation in the error term in addition to conditional heteroskedasticity. Second, the residual-based block bootstrap requires the choice of an additional tuning parameter in the form of the block size. In practice, results may be sensitive to the choice of block size. Although there are data-dependent rules for block size selection, these procedures are very computationally intensive and little is known about their accuracy in small samples. In contrast, the methods we propose are no more computationally burdensome than the standard residual-based algorithm and very easy to implement.

The paper is organized as follows. In section 2 we provide some empirical and theoretical motivation for the use of the m.d.s. assumption in resampling and highlight the limitations of existing bootstrap and asymptotic methods of inference for dynamic regression models such as autoregressions. In section 3 we describe the bootstrap algorithms and state our main theoretical results. Details of the proofs are relegated to the appendix. In section 4, we provide some tentative simulation evidence for the small-sample performance of alternative bootstrap proposals. We conclude in section 5.

2. Evidence Against the Assumption of i.i.d. Errors

Standard residual-based bootstrap methods of inference for dynamic regression models treat the error term as i.i.d. The i.i.d. assumption does not follow naturally from economic models. Nevertheless, in many cases it has proved convenient for theoretical purposes to treat the error term of dynamic regression models as i.i.d. This would be of little concern if actual data were well represented by models with i.i.d. errors. Unfortunately, this is not the case in many empirical studies. Two illustrative examples are asset return regressions in empirical finance and autoregressions in empirical macroeconomics.

Dating back to work by Fama and French (1988), there has been great interest in testing the null hypothesis of uncorrelated stock returns. It is common to use *nonparametric* bootstrap tests of this hypothesis that impose the much stronger assumption of i.i.d. returns (see, e.g., Goetzmann and Jorion 1993). Figure 1a shows clear evidence of volatility clustering in monthly value-weighted CRSP returns for 1927.1-2000.12 that invalidates that assumption. This conclusion is also supported by a formal LM test of the null of conditional homoskedasticity in Table 1 (see Engle 1982). A related problem arises in the international finance literature. The random walk hypothesis due to Meese and Rogoff (1983) implies that changes in exchange rates should be unpredictable. It is standard to employ bootstrap tests of this hypothesis. In actuality, however, these tests impose the much more stringent assumption of i.i.d. returns (see Mark 1995, Kilian 1999). The evidence in Figure 1b and Table 1 (based on the DM-U.S. dollar exchange rate for 1973.1-2001.10) suggests that this assumption is highly questionable, at least for exchange returns at monthly or higher frequency.

An alternative approach in empirical finance involves the use of finite-sample critical values based on fitted VAR models for returns and a set of additional predictors. This approach may be interpreted as a *parametric* bootstrap approach. Often, however, these VAR models ignore evidence of conditional heteroskedasticity in the VAR errors (see e.g., Goetzmann and Jorion 1995). In principle, we may modify the bootstrap approach by postulating a parametric model of conditional heteroskedasticity. For example, Hodrick (1992) and Bekaert and Hodrick (2001) postulate a VAR model with conditionally Gaussian GARCH(1,1) errors. Similarly, Lamoureux and Lastrapes (1990) augment the return regression by a parametric GARCH(1,1) model. This approach is unlikely to solve the problem. Even in the unlikely case that we could agree that the class of GARCH models is appropriate for a given data set, in practice the precise form of the GARCH model will be unknown and different specifications may yield different results (see Wolf 2000). The same holds for the class of stochastic volatility models. This fact points to the need for a nonparametric treatment of conditional heteroskedasticity in dynamic regression models.

This need is reinforced by the fact that it is exceedingly difficult to obtain reliable numerical estimates of multivariate GARCH models. In practice, researchers often impose additional ad hoc restrictions on the covariance structure of the model (see, e.g., Bollerslev, Engle and Wooldridge 1988, Bollerslev 1990, Bekaert et. al. 1997). These restrictions have no theoretical justification (also see Ledoit, Santa-Clara and Wolf 2001). Finally, we note that even with such restrictions it seems next to impossible to model conditional heteroskedasticity in high-dimensional VAR models unless the sample size is very large. This problem is most apparent in macroeconomic applications with many variables.

Whereas the failure of the i.i.d. assumption is well-documented in empirical finance, it is less well known that many monthly macroeconomic variables also exhibit strong conditional heteroskedasticity. The workhorse model of empirical macroeconomics is the linear autoregression. Figure 2 plots the squared residuals of six univariate monthly autoregressive models (for the growth rate of industrial output, M1 growth, CPI inflation, the real 3-month T-Bill rate, the nominal Federal Funds rate and the percent change in the price of oil). The data source is FRED, the sample period 1959.1-2001.8, and the lag orders of the AR models have been selected by the AIC. Figure 2 shows strong evidence of departures from conditional homoskedasticity. Formal LM tests of the null hypothesis of no ARCH in Table 1 also provide overwhelming evidence against the i.i.d. assumption. The evidence in Table 1 is important because many methods of inference developed for smooth functions of autoregressive parameters (such as impulse responses) do not allow for conditional heteroskedasticity. For example, standard residual-based bootstrap methods for autoregressions rely on the i.i.d. error assumption and are invalid in the presence of conditional heteroskedasticity, as we will show in the next section. Similarly, the grid bootstrap of Hansen (1999) is based on the assumption of an autoregression with i.i.d. errors.

It may seem that standard asymptotic methods would be less restrictive, but this is not necessarily the case. For example, the closed-form solutions for the asymptotic normal approximation proposed by Lütkepohl (1990) also rely on the assumption of conditional homoskedasticity. They are based on least-squares estimates of the variance of the estimator that are inconsistent in the presence of conditional heteroskedasticity. Similarly, Wright's (2000) local-to-unity intervals for AR(p) impulse responses rely on the assumption of i.i.d. innovations. Although these methods could presumably be modified to allow for conditional heteroskedasticity, current implementations of these methods are invalid in the presence of conditional heteroskedasticity. Other papers make the even stronger assumption of Gaussian i.i.d. errors, including Wright (2001), Andrews (1993) and Andrews and Chen (1994). Although the latter two papers provide some simulation evidence that their method is fairly robust to non-Gaussian i.i.d. innovations, they do not consider conditionally heteroskedastic errors. Finally, although this paper does not cover the Bayesian approach, it should be noted that the popular Bayesian Monte Carlo integration method for forming Bayesian error bands for VAR impulse responses also assumes that the VAR innovations are i.i.d. (see Doan 1990, Sims and Zha 1999).

In this paper we study several easy-to-implement bootstrap methods that allow inference in autoregressions with possible conditional heteroskedasticity of unknown form. Unlike the standard residual-based bootstrap for models with i.i.d. innovations these bootstrap methods remain valid under the much weaker assumption of m.d.s. innovations, and they do not require the researcher to take a stand on the existence or specific form of conditional heteroskedasticity. For expository purposes we focus on univariate autoregressive models. Analogous results for the multivariate case are possible at the cost of additional notation.

3. Theory

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_t\}$ a sequence of increasing σ-fields of $\mathcal{F}$. The sequence of martingale differences $\{\varepsilon_t,\, t \in \mathbb{Z}\}$ is defined on $(\Omega, \mathcal{F}, P)$, where each ε_t is assumed to be measurable with respect to $\mathcal{F}_t$. We observe a sample of data $\{y_{-p+1}, \ldots, y_0, y_1, \ldots, y_n\}$ from the following data generating process for the time series y_t,

$$\phi\,(L)\, y_t = \varepsilon_t, \tag{3.1}$$

where $\phi\,(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$, $\phi_p \neq 0$, is assumed to have all roots outside the unit circle. $\phi = \left(\phi_1, \ldots, \phi_p\right)'$ is the parameter of interest, which we estimate by ordinary least squares (OLS) using observations 1 through n:

$$\hat{\phi}_n = \left(n^{-1} \sum_{t=1}^{n} Y_{t-1} Y_{t-1}'\right)^{-1} n^{-1} \sum_{t=1}^{n} Y_{t-1} y_t,$$

where $Y_{t-1} = (y_{t-1}, \ldots, y_{t-p})'$. In this paper we focus on bootstrap confidence intervals for ϕ that are robust to the presence of conditional heteroskedasticity of unknown form in the innovations $\{\varepsilon_t\}$. More specifically, we assume the following condition:

Assumption A

(i) $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$, almost surely, where $\mathcal{F}_{t-1} = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)$ is the σ-field generated by $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$.

(ii) $E(\varepsilon_t^2) = \sigma^2 < \infty$.

(iii) $\lim_{n\to\infty} n^{-1} \sum_{t=1}^{n} E(\varepsilon_t^2|\mathcal{F}_{t-1}) = \sigma^2 > 0$ in probability.

(iv) $E(\varepsilon_t^2 \varepsilon_{t-r} \varepsilon_{t-s}) = \sigma^4 \tau_{r,s}$ is uniformly bounded for all $t, r \geq 1$, $s \geq 1$; $\tau_{r,r} > \underline{\alpha}$ for some $\underline{\alpha} > 0$ for all r.

(v) $\lim_{n\to\infty} n^{-1} \sum_{t=1}^{n} \varepsilon_{t-r} \varepsilon_{t-s} E(\varepsilon_t^2|\mathcal{F}_{t-1}) = \sigma^4 \tau_{r,s}$ in probability for any $r \geq 1$, $s \geq 1$.

(vi) $E|\varepsilon_t|^{4r}$ is uniformly bounded, for some $r > 1$.

Assumption A replaces the usual i.i.d. assumption on the errors $\{\varepsilon_t\}$ by the broader martingale difference sequence assumption. In particular, Assumption A does not impose conditional homoskedasticity on the sequence $\{\varepsilon_t\}$, which need not be strictly stationary (although it is covariance stationary). Assumption A covers a variety of conditionally heteroskedastic models such as ARCH, GARCH, EGARCH and stochastic volatility models (see, e.g. Deo (2000), who shows that a stronger version of Assumption A is satisfied for stochastic volatility and GARCH models). Assumptions (iv) and (v) restrict the fourth order cumulants of ε_t.

The following theorem gives the asymptotic distribution of the OLS estimator $\hat{\phi}_n$ for the parameter vector ϕ under the martingale difference sequence Assumption A. This result could be obtained as a special case of Kuersteiner's (2001) Theorem 3.4, which gives the asymptotic distribution of efficient instrumental variables estimators in the context of ARMA models with martingale difference sequence errors. In particular, in addition to the martingale difference sequence assumption, his Assumption A1 assumes $\{\varepsilon_t\}$ to be stationary ergodic, and it imposes a summability condition on the fourth order cumulants. Here, we use Assumption A, which relaxes the stationarity and ergodicity assumptions and the summability condition. We use Kuersteiner's (2001) notation to characterize the asymptotic covariance matrix of $\hat{\phi}_n$. Using $\phi^{-1}(L) = \sum_{j=0}^{\infty} \psi_j L^j$, we let $b_j = (\psi_{j-1}, \ldots, \psi_{j-p})'$ with $\psi_0 = 1$ and $\psi_j = 0$ for $j < 0$. The coefficients ψ_j satisfy the recursion $\psi_s - \phi_1 \psi_{s-1} - \ldots - \phi_p \psi_{s-p} = 0$ for all $s > 0$ and $\psi_0 = 1$. We let $\Rightarrow$ denote convergence in distribution throughout.

Theorem 3.1. *Under Assumption A,* $\sqrt{n}\left(\hat{\phi}_n - \phi\right) \Rightarrow N(0, C)$, *where*

$$
\begin{aligned}
C &= A^{-1} B A^{-1}, \\
A &= \sigma^2 \sum_{j=1}^{\infty} b_j b_j' = \sigma^2 \left[\sum_{j=0}^{\infty} \psi_j \psi_{j+|k-l|} \right]_{k,l=1,\ldots,p} \\
B &= \sigma^4 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} b_i b_j' \tau_{i,j} = \sigma^4 \left[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \tau_{l+i,k+j} \right]_{k,l=1,\ldots,p}.
\end{aligned}
$$

The asymptotic covariance of $\hat{\phi}_n$ is of the traditional "sandwich" form, where $A = E\left(n^{-1} \sum_{t=1}^{n} Y_{t-1} Y_{t-1}'\right)$ and $B = Var\left(n^{-1/2} \sum_{t=1}^{n} Y_{t-1} \varepsilon_t\right)$. Under conditional homoskedasticity, we obtain simplified expressions for A and B. In particular, by application of the law of iterated expectations, we have that $\tau_{i,i} \equiv \sigma^{-4} E\left(\varepsilon_t^2 \varepsilon_{t-i}^2\right) = \sigma^{-4} E\left(\varepsilon_{t-i}^2 E\left(\varepsilon_t^2 | F_{t-1}\right)\right) = \sigma^{-4} E\left(\varepsilon_{t-i}^2 \sigma^2\right) = 1$ for all $i = 1, 2, \ldots$. Similarly, we can show that $\tau_{i,j} = 0$ for all $i \neq j$. Thus, for instance in the AR(1) case, the asymptotic variance of $\hat{\phi}_n = \hat{\phi}_{1n}$ simplifies to $C = \left(\sigma^2 \sum_{i=0}^{\infty} \psi_i^2\right)^{-2} \left(\sigma^4 \sum_{i=0}^{\infty} \psi_i^2\right) = 1 - \phi_1^2$.

The validity of any bootstrap method in the context of autoregressions with conditional heteroskedasticity depends crucially on the ability of the bootstrap to estimate consistently the asymptotic covariance matrix C. The standard residual-based bootstrap method fails to do so by not correctly mimicking the behavior of the fourth order cumulants of ε_t in the conditionally heteroskedastic case, as we now show. Let $\hat{\varepsilon}_t^*$ be resampled with replacement from the centered residuals. The standard residual-based bootstrap builds y_t^* recursively from $\hat{\varepsilon}_t^*$ according to

$$
y_t^* = Y_{t-1}^{*\prime} \hat{\phi}_n + \hat{\varepsilon}_t^*, \; t = 1, \ldots, n,
$$

where $Y_{t-1}^* = \left(y_{t-1}^*, \ldots, y_{t-p}^*\right)'$, given some initial conditions. The bootstrap analogues of A and B are $A_n^* = n^{-1} \sum_{t=1}^{n} E^*\left(Y_{t-1}^* Y_{t-1}^{*\prime}\right)$ and $B_n^* = Var^*\left(n^{-1/2} \sum_{t=1}^{n} Y_{t-1}^* \hat{\varepsilon}_t^*\right)$, respectively. Because $\hat{\varepsilon}_t^*$ is i.i.d.

$\left(0, \hat{\sigma}^2\right)$, where $\hat{\sigma}^2 = n^{-1} \sum_{t=1}^{n} \left(\hat{\varepsilon}_t - \overline{\hat{\varepsilon}}\right)^2$, $\hat{\varepsilon}_t^*$ and Y_{t-1}^* are (conditionally) independent, and

$$\begin{aligned} B_n^* &= n^{-1} \sum_{t=1}^{n} E^* \left(Y_{t-1}^* Y_{t-1}^{*\prime} \hat{\varepsilon}_t^{*2}\right) = n^{-1} \sum_{t=1}^{n} E^* \left(Y_{t-1}^* Y_{t-1}^{*\prime}\right) E^* \left(\hat{\varepsilon}_t^{*2}\right) \\ &= \hat{\sigma}^2 A_n^*. \end{aligned}$$

Thus, the bootstrap analogue of C, $C_n^* \equiv A_n^{*-1} B_n^* A_n^{*-1} = \hat{\sigma}^2 A_n^{*-1}$, converges in probability to $\sigma^2 A^{-1}$, implying that the limiting distribution of the recursive i.i.d. bootstrap is $N\left(0, \sigma^2 A^{-1}\right)$. As Theorem 3.1 above shows, $\sigma^2 A^{-1}$ is *not* the correct asymptotic covariance matrix of $\hat{\phi}_n$ without further conditions, e.g., that ε_t is conditionally homoskedastic. In the general conditionally heteroskedastic case, B depends on $\sigma^4 \tau_{i,j}$. The recursive i.i.d. bootstrap implies $E^* \left(\hat{\varepsilon}_{t-i}^* \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^{*2}\right) = \hat{\sigma}^4$ when $i = j$ and zero otherwise, and thus implicitly sets $\tau_{i,j} = 1$ for $i = j$ and 0 for $i \neq j$.

Given the failure of the standard-residual based bootstrap, we are interested in establishing the first-order asymptotic validity of three alternative bootstrap methods in this environment. Two of the bootstrap methods we study rely on an application of the wild bootstrap (WB). The WB has been originally developed by Wu (1986), Liu (1988) and Mammen (1993) in the context of static linear regression models with (unconditionally) heteroskedastic errors. We consider both a recursive-design and a fixed-design version of the WB. The third method is a natural generalization of the pairwise bootstrap for linear regression first suggested by Freedman (1981) for cross sectional data.

As we will see next, the recursive-design WB requires a strengthening of Assumption A in order to ensure convergence towards the correct asymptotic covariance matrix C. In contrast, the fixed-design WB and the pairwise bootstrap are valid under the more general Assumption A.

Recursive-design wild bootstrap

The recursive-design WB is a simple modification of the usual recursive-design bootstrap method for autoregressions (see e.g. Bose, 1988) which consists of replacing Efron's i.i.d. bootstrap by the wild bootstrap when bootstrapping the errors of the AR model. More specifically, the recursive-design WB bootstrap generates a pseudo time series $\{y_t^*\}$ according to the autoregressive process:

$$y_t^* = \hat{\phi}_{1n} y_{t-1}^* + \hat{\phi}_{2n} y_{t-2}^* + \ldots + \hat{\phi}_{pn} y_{t-p}^* + \hat{\varepsilon}_t^*,\ t = 1, \ldots, n,$$

where $\hat{\varepsilon}_t^* = \hat{\varepsilon}_t \eta_t$, with $\hat{\varepsilon}_t = \hat{\phi}_n(L) y_t$, and where η_t is an i.i.d. sequence with mean zero and variance one such that $E^* |\eta_t|^4 \leq \Delta < \infty$. We let $y_t^* = 0$ for all $t \leq 0$. Kreiss (1997) suggested this method in the context of autoregressive models with i.i.d. errors, but did not investigate its theoretical justification in more general models. Here, we will provide conditions for the asymptotic validity of the recursive-design WB proposal for finite-order autoregressive processes with possibly conditionally heteroskedastic errors. To show this result we need to strengthen Assumption A as follows:

Assumption A′

(iv′) $E\left(\varepsilon_t^2 \varepsilon_{t-r} \varepsilon_{t-s}\right) = 0$ for all $r \neq s$, for all t, $r \geq 1$, $s \geq 1$.

(vi′) $E |\varepsilon_t|^{4r}$ is uniformly bounded for some $r \geq 2$ and for all t.

Assumption A′ restricts the class of conditionally heteroskedastic autoregressive models in two dimensions. First, Assumption A′ (iv′) requires the product moments of $\{\varepsilon_t\}$ up to order four to behave as those of an independent series. Milhøj (1985) shows that this assumption is satisfied for the ARCH(p) model with innovations having a symmetric distribution. Bollerslev(1986) and He and Teräsvirta (1999) extend the argument to the GARCH(p, q) case. In addition, Deo (2000) shows that this assumption is satisfied by certain stochastic volatility models. Nevertheless, Assumption A′ (iv′) excludes some non-symmetric parametric models such as asymmetric EGARCH. Second, we now require the existence of at least eight moments for the martingale difference sequence $\{\varepsilon_t\}$ as opposed to only $4r$ moments, for some $r > 1$, as in Assumption A. A similar moment condition was used by Kreiss (1997) in his Theorem 4.3, which shows the validity of the recursive-design WB for possibly infinite-order AR processes with i.i.d. innovations.

The strengthening of Assumption A is crucial to showing the asymptotic validity of the recursive-design WB in the martingale difference context. In particular, conditional on the data, and given the independence of $\{\eta_t\}$, $\left\{Y_{t-1}^* \hat{\varepsilon}_t^*, \mathcal{F}_t^*\right\}$ can be shown to be a vector m.d.s., where $\mathcal{F}_t^* = \sigma\left(\eta_t, \eta_{t-1}, \ldots, \eta_1\right)$. We use Assumption A′ (vi′) to ensure convergence of $n^{-1} \sum_{t=1}^n Y_{t-1}^* Y_{t-1}^{*\prime} \hat{\varepsilon}_t^{*2}$ to $B_n^* \equiv Var^*\left(n^{-1/2} \sum_{t=1}^n Y_{t-1}^* \varepsilon_t^*\right)$, thus verifying one of the conditions of the CLT for m.d.s. Assumption A′ (iv′) ensures convergence of the recursive-design WB variance B_n^* to the correct limiting variance of $n^{-1/2} \sum_{t=1}^n Y_{t-1} \varepsilon_t$. More specifically, letting $Y_{t-1}^{*} \equiv \sum_{j=1}^{t-1} \hat{b}_j \hat{\varepsilon}_{t-j}^{*}$ with $\hat{b}_j \equiv \left(\hat{\psi}_{j-1}, \ldots, \hat{\psi}_{j-p}\right)'$, $\hat{\psi}_0 = 1$ and $\hat{\psi}_j = 0$ for $j < 0$, it follows by direct evaluation that

$$B_n^{*} = n^{-1} \sum_{t=1}^{n} \sum_{j=1}^{t-1} \sum_{i=1}^{t-1} \hat{b}_j \hat{b}_i' E^{*} \left(\hat{\varepsilon}_{t-j}^{*} \hat{\varepsilon}_{t-i}^{*} \hat{\varepsilon}_t^{*2}\right),$$

where $E^{*}\left(\hat{\varepsilon}_{t-j}^{*} \hat{\varepsilon}_{t-i}^{*} \hat{\varepsilon}_t^{*2}\right) = \hat{\varepsilon}_{t-i}^{2} \hat{\varepsilon}_t^{2}$ for $i = j$ and zero otherwise. We can rewrite B_n^{*} as $\sum_{j=1}^{n-1} \hat{b}_j \hat{b}_j' n^{-1}$ $\sum_{t=1+j}^{n} \hat{\varepsilon}_t^2 \hat{\varepsilon}_{t-j}^2$, which converges in probability to $\tilde{B} \equiv \sum_{j=1}^{\infty} b_j b_j' \sigma^4 \tau_{jj}$ under Assumption A. Without Assumption A$'$ (iv$'$) an asymptotic bias term appears in the estimation of $B \equiv \sigma^4 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} b_i b_j' \tau_{i,j}$, which is equal to $-\sigma^4 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} b_i b_j' \tau_{i,j}$ for all $i \neq j$. Assumption A$'$ (iv$'$) sets $\tau_{i,j}$ equal to zero for $i \neq j$, and thus ensures that the recursive-design WB consistently estimates B.

Theorem 3.2 formally states the asymptotic validity of the recursive-design WB for finite-order autoregressions with heteroskedastic errors. Let $\hat{\phi}_n^{*}$ denote the recursive-design WB OLS estimator, i.e. $\hat{\phi}_n^{*} = \left(n^{-1} \sum_{t=1}^{n} Y_{t-1}^{*} Y_{t-1}^{*\prime}\right)^{-1} n^{-1} \sum_{t=1}^{n} Y_{t-1}^{*} y_t^{*}$.

Theorem 3.2. *Under Assumption A strengthened by Assumption A$'$ (iv$'$) and (vi$'$), it follows that*

$$\sup_{x \in \mathbb{R}^p} \left| P^{*}\left(\sqrt{n}\left(\hat{\phi}_n^{*} - \hat{\phi}_n\right) \leq x\right) - P\left(\sqrt{n}\left(\hat{\phi}_n - \phi\right) \leq x\right)\right| \xrightarrow{P} 0,$$

where P^{*} denotes the probability measure induced by the recursive-design WB.

Fixed-design wild bootstrap

The fixed-design WB generates $\{y_t^{*}\}_{t=1}^{n}$ according to the equation

$$y_t^{*} = \hat{\phi}_{1n} y_{t-1} + \hat{\phi}_{2n} y_{t-2} + \ldots + \hat{\phi}_{n,p} y_{t-p} + \hat{\varepsilon}_t^{*}, \; t = 1, \ldots, n, \tag{3.2}$$

where $\hat{\varepsilon}_t^{*} = \hat{\varepsilon}_t \eta_t$, $\hat{\varepsilon}_t = \hat{\phi}_n(L) y_t$, and where η_t is an i.i.d. sequence with mean zero and variance one such that $E^{*} |\eta_t|^{2r} \leq \Delta < \infty$. The bootstrap estimator is $\hat{\phi}_n^{*} = \left(n^{-1} \sum_{t=1}^{n} Y_{t-1} Y_{t-1}'\right)^{-1} n^{-1} \sum_{t=1}^{n} Y_{t-1} y_t^{*}$. The fixed-design WB corresponds to a regression-type bootstrap method in that (3.2) is a fixed-design regression model, conditional on the original sample. The fixed-design WB was suggested by Kreiss (1997). Kreiss' (1997) Theorem 4.2 provides the first-order asymptotic validity of the fixed-design WB for finite-order autoregressions with conditional heteroskedasticity of a specific form. More specifically, he assumes a data generating process of the form $y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sigma(y_{t-1}) v_t$, where v_t is i.i.d. $(0,1)$

with finite fourth moment. The i.i.d. assumption on the rescaled innovations v_t is violated if for instance the conditional moments of v_t depend on past observations. We prove the first-order asymptotic validity of the fixed-design WB of Kreiss (1997) under a broader set of regularity conditions, namely Assumption A.

Theorem 3.3. *Under Assumption A,*

$$\sup_{x \in \mathbb{R}^p} \left| P^* \left(\sqrt{n} \left(\hat{\phi}_n^* - \hat{\phi}_n \right) \leq x \right) - P \left(\sqrt{n} \left(\hat{\phi}_n - \phi \right) \leq x \right) \right| \xrightarrow{P} 0,$$

where P^* denotes the probability measure induced by the fixed-design WB.

In contrast to the recursive-design WB, the ability of the fixed-design WB to consistently estimate the variance, and hence the limiting distribution, of $\hat{\phi}_n$ does not require a strengthening of Assumption A. Specifically, the variance of the limiting conditional bootstrap distribution of $\hat{\phi}_n^*$ is given by $A_n^{*-1} B_n^* A_n^{*-1}$, where $A_n^* = n^{-1} \sum_{t=1}^n Y_{t-1} Y_{t-1}'$ and $B_n^* \equiv Var^* \left(n^{-1/2} \sum_{t=1}^n Y_{t-1} \hat{\varepsilon}_t^* \right) = n^{-1} \sum_{t=1}^n Y_{t-1} Y_{t-1}' \hat{\varepsilon}_t^2$. Under Assumption A one can show that $A_n^* \xrightarrow{P} A$ and $B_n^* \xrightarrow{P} B$, thus ensuring that $A_n^{*-1} B_n^* A_n^{*-1} \xrightarrow{P} A^{-1} B A^{-1} \equiv C$.

Pairwise bootstrap

Another bootstrap method that captures the presence of conditional heteroskedasticity in autoregressive models consists of bootstrapping "pairs", or tuples, of the dependent and explanatory variables in the autoregression. This method is an extension of Freedman's (1981) bootstrap method for the correlation model to the autoregressive context. In the AR(p) model, it amounts to resampling with replacement from the set of tuples $\left(y_t, Y_{t-1}' \right) = (y_t, y_{t-1}, \ldots, y_{t-p})$, $t = 1, \ldots, n$. Let $\left\{ \left(y_t^*, Y_{t-1}^{*\prime} \right) = \left(y_t^*, y_{t-1}^*, \ldots, y_{t-p}^* \right), t = 1, \ldots, n \right\}$ be an i.i.d. resample from this set. Then the pairwise bootstrap estimator is defined by $\hat{\phi}_n^* = \left(n^{-1} \sum_{t=1}^n Y_{t-1}^* Y_{t-1}^{*\prime} \right)^{-1} n^{-1} \sum_{t=1}^n Y_{t-1}^* y_t^*$. The bootstrap analogue of ϕ is $\hat{\phi}_n$ since $\hat{\phi}_n$ is the parameter value that minimizes $E^* \left[\left(y_t^* - \phi_1 y_{t-1}^* - \ldots - \phi_p y_{t-p}^* \right)^2 \right]$. The following theorem establishes the asymptotic validity of the pairwise bootstrap for the AR(p) process with m.d.s. errors satisfying Assumption A.

Theorem 3.4. *Under Assumption A, it follows that*

$$\sup_{x\in\mathbb{R}^p}\left|P^*\left(\sqrt{n}\left(\hat{\phi}_n^*-\hat{\phi}_n\right)\le x\right)-P\left(\sqrt{n}\left(\hat{\phi}_n-\phi\right)\le x\right)\right|\xrightarrow{P}0,$$

where P^* denotes the probability measure induced by the pairwise bootstrap.

Asymptotic validity of bootstrapping the studentized slope parameter

Corollary 3.1 below establishes the asymptotic validity of bootstrapping the *t-statistic* for the elements of ϕ. To conserve space, we let $\hat{\phi}_n^*$ denote the OLS estimator of ϕ obtained under any of the three bootstrap resampling schemes studied above. Similarly, we use $(y_t^*, Y_{t-1}^{*\prime})$ to denote bootstrap data in general. In particular, we implicitly set $Y_{t-1}^* = Y_{t-1}$ for the fixed-design WB.

For a typical element ϕ_j a bootstrap percentile-t confidence interval is based on $t_{\hat{\phi}_{jn}^*}=\frac{\sqrt{n}(\hat{\phi}_{jn}^*-\hat{\phi}_{jn})}{\sqrt{\hat{C}_{n,jj}^*}}$, the bootstrap analogue of the t-statistic $t_{\hat{\phi}_{jn}}=\frac{\sqrt{n}(\hat{\phi}_{jn}-\phi_j)}{\sqrt{\hat{C}_{n,jj}}}$. In the context of (conditional) heteroskedasticity, $\hat{C}_{n,jj}$ and $\hat{C}_{n,jj}^*$ are the heteroskedasticity-consistent variance estimators evaluated on the original and on the bootstrap data, respectively. Specifically, for the bootstrap t-statistic let

$$\begin{aligned}\hat{C}_n^* &= \hat{A}_n^{*-1}\hat{B}_n^*\hat{A}_n^{*-1}, \text{ with}\\ \hat{A}_n^* &= n^{-1}\sum_{t=1}^n Y_{t-1}^*Y_{t-1}^{*\prime} \text{ and } \hat{B}_n^* = n^{-1}\sum_{t=1}^n Y_{t-1}^*Y_{t-1}^{*\prime}\widetilde{\varepsilon}_t^{*2},\end{aligned}$$

where $\widetilde{\varepsilon}_t^* = y_t^* - \hat{\phi}_n^{*\prime}Y_{t-1}^*$ are the bootstrap residuals.

Corollary 3.1. *Assume Assumption A holds. Then, for the fixed-design WB and the pairwise bootstrap, it follows that*

$$\sup_{x\in\mathbb{R}}\left|P^*\left(t_{\hat{\phi}_{jn}^*}\le x\right)-P\left(t_{\hat{\phi}_{jn}}\le x\right)\right|\xrightarrow{P}0,\quad j=1,\ldots,p.$$

If Assumption A is strengthened by Assumption A′ (iv′) and (vi′), then the above result also holds for the recursive-design WB.

4. Simulation Evidence

In this section, we study the accuracy of the bootstrap approximation proposed in section 3 for sample sizes of interest in applied work. We focus on the AR(1) model as the leading example of an autoregressive process. The DGP is $y_t = \phi_1 y_{t-1} + \varepsilon_t$ with $\phi_1 \in \{0, 0.9\}$. In our simulation study we allow for GARCH(1,1) errors of the form $\varepsilon_t = \sqrt{h_t} v_t$, where v_t is i.i.d. $N(0,1)$ and $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$, $t = 1, \ldots, n$. We normalize the unconditional variance of ε_t to one. In addition to conditional N(0,1) innovations we also consider GARCH models with conditional t_5-errors (suitably normalized to have unit variance). For $\beta = 0$ this model reduces to an ARCH(1) model. For $\alpha = 0$ and $\beta = 0$ the error sequence reduces to a sequence of (possibly non-Gaussian) i.i.d errors. We allow for varying degrees of volatility persistence modeled as GARCH processes with $\alpha + \beta \in \{0, 0.9, 0.99\}$. In addition, we consider AR(1) models with exponential GARCH errors (EGARCH), asymmetric GARCH errors (AGARCH) and with the GJR-GARCH errors proposed by Glosten, Jaganathan and Runkle (1993). Our parameter settings are based on Engle and Ng (1993). Note that many of these processes are not covered by either the conventional asymptotic theory or by the asymptotic theory for the bootstrap. In particular, the assumption of a finite fourth moment may be violated for some parameter settings. Nevertheless, it is important to investigate the robustness of these methods to such departures from our assumptions.

Finally, we also consider the stochastic volatility model $\varepsilon_t = v_t \exp(h_t)$ with $h_t = \lambda h_{t-1} + 0.5 u_t$, where $|\lambda| < 1$ and (u_t, v_t) is a sequence of independent bivariate normal random variables with zero mean and covariance matrix $diag(\sigma_u^2, 1)$. This model is a m.d.s. model and satisfies Assumption A. We follow Deo (2000) in postulating the values $(0.936, 0.424)$ and $(0.951, 0.314)$ for (λ, σ_u). These are values obtained by Shephard (1996) by fitting this stochastic volatility model to real exchange rate data.

We generate repeated trials of length $n = 120$ and $n = 240$ from these processes and conduct bootstrap inference based on the fitted AR(1) model for each trial. All fitted models include an intercept. The number of Monte Carlo trials is 1,000 with 1,000 bootstrap replications each. The fixed-design and recursive-design WB involve applying the WB to the residuals of the fitted model. Recall that the WB innovation is $\varepsilon_t^* = \hat{\varepsilon}_t \eta_t$, with $\hat{\varepsilon}_t = y_t - \hat{\phi}_{1n} y_{t-1}$, where η_t is an i.i.d. sequence with mean zero and variance one such that $E^* |\eta_t|^{4r} \leq \Delta < \infty$. In practice, there are several choices for η_t that satisfy these conditions. In the simulations we use $\eta_t \sim N(0,1)$. Our results are robust to alternative choices including the two-point distribution $\eta_t = -(\sqrt{5}-1)/2$ with probability $p = (\sqrt{5}+1)/(2\sqrt{5})$ and $\eta_t = (\sqrt{5}+1)/2$ with probability $1-p$, as proposed by Mammen (1993), and the two-point distribution $\eta_t = 1$ with probability 0.5 and $\eta_t = -1$ with probability 0.5, as proposed by Liu (1988).

We are interested in studying the coverage accuracy of nominal 90% symmetric percentile-t bootstrap confidence intervals for the slope parameter ϕ_1. We also considered equal-tailed percentile-t intervals, but found that symmetric percentile-t intervals in all cases were at least as accurate. Unlike the percentile interval, the construction of the bootstrap t-interval requires the use of an estimate of the standard error of $n^{1/2}(\widehat{\phi}^*_{1n} - \widehat{\phi}_{1n})$. We use the heteroskedasticity-robust estimator of the covariance proposed by Nicholls and Pagan (1983) based on work by Eicker (1963) and White (1980). We also experimented with several modified robust covariance estimators (see MacKinnon and White 1985, Chesher and Jewitt 1987, Davidson and Flachaire 2000). For our sample sizes, none of these estimators performed better than the basic estimator proposed by Nicholls and Pagan (1983). Finally, virtually identical results were obtained based on WB bootstrap standard error estimates. The latter approach involves a nested bootstrap loop and is not recommended for computational reasons. As a benchmark we also include the coverage rates of the Gaussian large-sample approximation based on Nicholls-Pagan robust standard errors.

We begin with a review of the simulation results for the stationary AR(1) model. Starting with the results for N-GARCH errors in Table 2 several broad tendencies emerge. First, the accuracy of the standard recursive-design bootstrap procedure based on i.i.d. resampling of residuals is high when the model errors are truly i.i.d., but can be very poor in the presence of N-GARCH. Second, conventional large-sample approximations based on robust standard errors are more accurate than the recursive-design i.i.d. bootstrap in the presence of N-GARCH, but less accurate for models with i.i.d. errors. In either case, their coverage rates may be substantially below the nominal level. Third, all three robust bootstrap methods are more accurate than the i.i.d. bootstrap or the conventional Gaussian approximation. Fourth, the recursive-design WB is always at least as accurate as the fixed-design WB and the pairwise resampling procedures, and its accuracy is very high for all variations of the DGP, including models with i.i.d. innovations. Specifically, for $n = 120$ and AR(1) models with high persistence, the accuracy of the recursive-design WB tends to be higher than for the pairwise bootstrap. For $n = 240$, these differences vanish and both methods are equally accurate. The fixed-design WB is typically less accurate than the recursive-design WB both for $n = 120$ and for $n = 240$, although the discrepancies diminish with the larger sample size.

The results for the AR(1) model with t_5-GARCH errors in Table 3 are qualitatively similar, except that the recursive-design i.i.d. bootstrap and the conventional Gaussian approximation are even less accurate than for N-GARCH processes. In Table 4 we explore a number of additional models of conditional heteroskedasticity that have been used primarily to model returns in empirical finance. The results for the stochastic volatility model are qualitatively the same as for N-GARCH and t-GARCH. For the other three models, we find that there is little to choose between the recursive-design WB and the pairwise bootstrap. Their accuracy for $n = 120$ and highly persistent data tends to be slightly below nominal coverage, but consistently higher than that of any alternative method. In all other cases both methods are highly accurate. Neither the i.i.d. bootstrap nor the conventional Gaussian approximation perform well. The high accuracy of the recursive-design WB even for EGARCH, AGARCH and GJR-GARCH error processes is surprising, given its lack of theoretical support for these DGPs. Apparently, the asymptotic inconsistency of the recursive-design WB method has little effect on its performance in small samples. Fortunately, applications in finance, for which such asymmetric volatility models have been developed, invariably involve large sample sizes, conditions under which pairwise resampling is just as accurate as the recursive-design WB and theoretically justified.

Given the computational costs of the simulation study, we have chosen to focus on a stylized autoregressive model, but have explored a wide range of conditionally heteroskedastic errors. Although our simulation results are necessarily tentative, they suggest that the recursive-design WB should replace conventional recursive design i.i.d. bootstrap methods in many standard applications. The pairwise bootstrap provides a suitable alternative when sample sizes are at least moderately large and the possibility of asymmetric forms of GARCH is a practical concern. Even for moderate sample sizes the accuracy of the pairwise bootstrap is slightly higher than that of the fixed-design bootstrap, which appears only suited for very large samples.

5. Concluding Remarks

The aim of the paper has been to extend the range of applications of autoregressive bootstrap methods in empirical finance and macroeconometrics. We documented widespread evidence of conditional heteroskedasticity not just in financial time series, but also in monthly macroeconomic data. We analyzed the theoretical properties of three bootstrap procedures for stationary autoregressions that are robust to conditional heteroskedasticity of unknown form: the fixed-design WB, the recursive-design WB and the pairwise bootstrap.

Throughout the paper, we established conditions for the first-order asymptotic validity of these bootstrap procedures. We did not attempt to address the issue of the existence of higher-order asymptotic refinements provided by the bootstrap approximation. Arguments aimed at proving asymptotic refinements require the existence of an Edgeworth expansion for the distribution of the estimator of interest. Establishing the existence of such an Edgeworth expansion is beyond the scope of this paper. Moreover, the quality of the finite-sample approximation provided by analytic Edgeworth expansions often is poor and less accurate than bootstrap approximations. Thus, Edgeworth expansions in general are imperfect guides to the relative accuracy of alternative bootstrap methods (see Härdle, Horowitz and Kreiss 2001). Indeed, preliminary simulation evidence indicates that wild bootstrap methods based on two-point distributions that may yield asymptotic refinements in our context tend to perform no better than - and in some cases worse than - the first-order accurate methods studied in this paper. Nevertheless, we found that the robust bootstrap approximation was typically more accurate in small samples than the usual first-order asymptotic approximation based on robust standard errors. Our simulation results also highlighted the dangers of incorrectly modelling the error term in dynamic regression models as i.i.d. We found that conventional residual-based bootstrap methods may be very inaccurate in the presence of conditional heteroskedasticity.

The theoretical and simulation results in this paper suggested that no single bootstrap method for dealing with conditional heteroskedasticity of unknown form will be optimal in all cases. We concluded that the recursive-design WB is well-suited for many applications in empirical macroeconomics. This method performs equally well, whether the error term is i.i.d. or conditionally heteroskedastic, but it lacks a theoretical justification for some forms of GARCH that have figured prominently in the literature on high-frequency returns. When the sample size is at least moderately large and asymmetric forms of GARCH are a practical concern, the pairwise bootstrap method provides a suitable alternative . The fixed-design WB has the same theoretical justification as the pairwise bootstrap for parametric models, but based on our simulation evidence appears only suited for very large samples.

There are several interesting extensions of the approach taken in this paper. One possible extension is the development of bootstrap methods for conditionally heteroskedastic stationary autoregressions of possibly infinite order. This extension is the subject of ongoing research. Another useful extension would be to establish the validity of the recursive-design WB for regression parameters in I(1) autoregressions that can be written in terms of zero mean stationary regressors, generalizing recent work by Inoue and Kilian (2002) on I(1) autoregressive models with i.i.d. errors. Yet another useful extension would be to establish the asymptotic validity of robust versions of the grid bootstrap of Hansen (1999). These extensions are nontrivial and left for future research.

Table 1. Approximate Finite-Sample P-Values of LM Test of No-ARCH(q) Hypothesis (in Percent)

Univariate AR Models

q	1	2	3	4	5
CRSP Returns	0.00	0.00	0.00	0.00	0.00
DM-U.S. Dollar Returns	1.25	5.99	8.28	1.15	1.18
Industrial Output Growth	1.58	2.40	3.28	1.61	1.47
M1 Growth	0.00	0.01	0.01	0.02	0.01
CPI Inflation	0.50	1.13	1.79	2.35	2.05
Real T-Bill Rate	0.08	0.18	0.29	0.37	0.34
Federal Funds Rate	3.37	0.45	0.71	0.94	0.90
Percent Change in Oil Price	2.39	3.77	5.25	4.60	6.44

SOURCE: Based on 20000 bootstrap replications under i.i.d. error null hypothesis. The LM test is based on Engle (1982). All data are monthly. The macroeconomic data have been filtered using an autoregressive approximation selected by the AIC. The returns are unfiltered.

Table 2. Coverage Rates of
Nominal 90% Symmetric Percentile-t Intervals for ϕ_1

AR(1)-N-GARCH Model

DGP:	$y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = h_t^{1/2} v_t, h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}, v_t \sim N(0,1)$								
					Recursive iid	Recursive WB	Fixed WB	Pairwise	Robust SE Gaussian
n	ϕ_1	$\alpha+\beta$	α	β					
120	0	0	0	0	0.92	0.91	0.91	0.91	0.90
		0.9	0.9	0	0.60	0.89	0.87	0.89	0.85
			0.7	0.2	0.64	0.89	0.88	0.90	0.87
			0.45	0.45	0.73	0.89	0.89	0.91	0.88
			0.2	0.7	0.84	0.90	0.90	0.90	0.88
		0.99	0.99	0	0.57	0.88	0.87	0.89	0.83
			0.79	0.2	0.60	0.88	0.87	0.91	0.84
			0.495	0.495	0.69	0.90	0.89	0.89	0.85
			0.2	0.79	0.82	0.91	0.90	0.90	0.89
	0.9	0	0	0	0.87	0.88	0.86	0.84	0.83
		0.9	0.9	0	0.75	0.89	0.86	0.87	0.83
			0.7	0.2	0.76	0.88	0.86	0.87	0.84
			0.45	0.45	0.79	0.88	0.86	0.88	0.85
			0.2	0.7	0.84	0.89	0.87	0.87	0.84
		0.99	0.99	0	0.73	0.89	0.87	0.88	0.84
			0.79	0.2	0.73	0.88	0.85	0.87	0.85
			0.495	0.495	0.77	0.88	0.85	0.87	0.83
			0.2	0.79	0.84	0.88	0.85	0.87	0.84
240	0	0	0	0	0.92	0.90	0.90	0.91	0.90
		0.9	0.9	0	0.56	0.88	0.87	0.90	0.86
			0.7	0.2	0.59	0.89	0.87	0.90	0.87
			0.45	0.45	0.69	0.88	0.87	0.91	0.88
			0.2	0.7	0.81	0.90	0.89	0.91	0.90
		0.99	0.99	0	0.51	0.88	0.86	0.89	0.84
			0.79	0.2	0.56	0.88	0.87	0.90	0.85
			0.495	0.495	0.64	0.89	0.88	0.91	0.88
			0.2	0.79	0.78	0.90	0.89	0.92	0.90
	0.9	0	0	0	0.89	0.89	0.87	0.87	0.86
		0.9	0.9	0	0.72	0.88	0.86	0.90	0.87
			0.7	0.2	0.72	0.88	0.86	0.89	0.87
			0.45	0.45	0.76	0.89	0.88	0.89	0.87
			0.2	0.7	0.83	0.89	0.87	0.88	0.86
		0.99	0.99	0	0.67	0.88	0.87	0.90	0.86
			0.79	0.2	0.67	0.89	0.85	0.89	0.85
			0.495	0.495	0.70	0.90	0.85	0.89	0.85
			0.2	0.79	0.81	0.90	0.88	0.88	0.87

SOURCE: 1000 Monte Carlo trials with 1000 bootstrap replications each. The regression model includes an intercept. The bootstrap algorithms are described in the text.

Table 3. Coverage Rates of
Nominal 90% Symmetric Percentile-t Intervals for ϕ_1

AR(1)-t_5-GARCH Model

DGP:	$y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = h_t^{1/2} v_t, h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}, v_t \sim t_5$								
					Recursive iid	Recursive WB	Fixed WB	Pairwise	Robust SE Gaussian
n	ϕ_1	$\alpha+\beta$	α	β					
120	0	0	0	0	0.91	0.91	0.88	0.90	0.88
		0.9	0.9	0	0.58	0.88	0.87	0.90	0.83
			0.7	0.2	0.62	0.89	0.86	0.90	0.83
			0.45	0.45	0.69	0.89	0.87	0.90	0.83
			0.2	0.7	0.81	0.91	0.87	0.90	0.85
		0.99	0.99	0	0.55	0.88	0.87	0.91	0.79
			0.79	0.2	0.58	0.89	0.86	0.89	0.81
			0.495	0.495	0.65	0.90	0.86	0.89	0.83
			0.2	0.79	0.79	0.90	0.88	0.91	0.85
	0.9	0	0	0	0.88	0.90	0.85	0.86	0.84
		0.9	0.9	0	0.75	0.90	0.85	0.89	0.82
			0.7	0.2	0.77	0.91	0.85	0.88	0.83
			0.45	0.45	0.79	0.90	0.86	0.87	0.83
			0.2	0.7	0.84	0.91	0.86	0.87	0.84
		0.99	0.99	0	0.73	0.91	0.85	0.89	0.81
			0.79	0.2	0.74	0.90	0.85	0.88	0.81
			0.495	0.495	0.75	0.89	0.86	0.88	0.83
			0.2	0.79	0.83	0.91	0.86	0.87	0.85
240	0	0	0	0	0.91	0.90	0.89	0.91	0.89
		0.9	0.9	0	0.49	0.88	0.87	0.90	0.85
			0.7	0.2	0.56	0.89	0.89	0.90	0.87
			0.45	0.45	0.67	0.90	0.90	0.90	0.88
			0.2	0.7	0.78	0.91	0.90	0.91	0.88
		0.99	0.99	0	0.46	0.88	0.87	0.90	0.83
			0.79	0.2	0.53	0.88	0.88	0.89	0.85
			0.495	0.495	0.61	0.90	0.89	0.89	0.86
			0.2	0.79	0.74	0.89	0.88	0.89	0.87
	0.9	0	0	0	0.90	0.89	0.86	0.89	0.85
		0.9	0.9	0	0.69	0.89	0.87	0.90	0.86
			0.7	0.2	0.71	0.90	0.88	0.90	0.86
			0.45	0.45	0.76	0.90	0.88	0.89	0.87
			0.2	0.7	0.82	0.89	0.88	0.89	0.87
		0.99	0.99	0	0.67	0.89	0.87	0.90	0.84
			0.79	0.2	0.67	0.89	0.87	0.90	0.84
			0.495	0.495	0.69	0.90	0.87	0.91	0.84
			0.2	0.79	0.79	0.89	0.87	0.90	0.85

SOURCE: See Table 2.

Table 4. Coverage Rates of
Nominal 90% Symmetric Percentile-t Intervals for ϕ_1

(a) AR(1)-EGARCH Model (Engle and Ng 1993)

DGP: $y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = {h_t}^{1/2} v_t, ln(h_t) = -0.23 + 0.9 ln(h_{t-1}) + 0.25[|v_{t-1}^2| - 0.3 v_{t-1}]$
$v_t \sim N(0,1)$

		Recursive iid	Recursive WB	Fixed WB	Pairwise	Robust SE Gaussian
n	ϕ_1					
120	0	0.72	0.88	0.88	0.89	0.86
	0.9	0.79	0.87	0.85	0.87	0.83
240	0	0.69	0.89	0.88	0.90	0.87
	0.9	0.76	0.91	0.89	0.90	0.88

(b) AR(1)-AGARCH Model (Engle 1990)

DGP: $y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = {h_t}^{1/2} v_t, h_t = 0.0216 + 0.6896 h_{t-1} + 0.3174[\varepsilon_{t-1} - 0.1108]^2$
$v_t \sim N(0,1)$

		Recursive iid	Recursive WB	Fixed WB	Pairwise	Robust SE Gaussian
n	ϕ_1					
120	0	0.73	0.89	0.88	0.89	0.87
	0.9	0.78	0.87	0.85	0.87	0.84
240	0	0.68	0.90	0.88	0.89	0.87
	0.9	0.73	0.90	0.89	0.88	0.87

(c) AR(1)-GJR GARCH Model (Glosten, Jaganathan and Runkle 1993)

DGP: $y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = {h_t}^{1/2} v_t, h_t = 0.005 + 0.7 h_{t-1} + 0.28[|\varepsilon_{t-1}| - 0.23 \varepsilon_{t-1}]^2$
$v_t \sim N(0,1)$

		Recursive iid	Recursive WB	Fixed WB	Pairwise	Robust SE Gaussian
n	ϕ_1					
120	0	0.75	0.89	0.88	0.89	0.86
	0.9	0.79	0.87	0.85	0.87	0.84
240	0	0.70	0.90	0.89	0.90	0.88
	0.9	0.75	0.91	0.90	0.89	0.87

(d) AR(1)-Stochastic Volatility Model (Shephard 1996)

DGP: $y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t = v_t exp(h_t), h_t = \lambda h_{t-1} + 0.5 u_t, (u_t, v_t) \sim N[0, diag(\sigma_u^2, 1)]$

				Recursive iid	Recursive WB	Fixed WB	Pairwise	Robust SE Gaussian
n	ϕ_1	λ	σ_u					
120	0	0.936	0.424	0.76	0.89	0.88	0.89	0.86
		0.951	0.314	0.81	0.89	0.89	0.89	0.87
	0.9	0.936	0.424	0.79	0.90	0.88	0.86	0.83
		0.951	0.314	0.82	0.89	0.88	0.86	0.83
240	0	0.936	0.424	0.73	0.88	0.87	0.91	0.89
		0.951	0.314	0.79	0.89	0.89	0.91	0.90
	0.9	0.936	0.424	0.80	0.88	0.87	0.90	0.88
		0.951	0.314	0.83	0.89	0.88	0.89	0.88

SOURCE: See Table 2.

Figure 5.1: Squared Returns

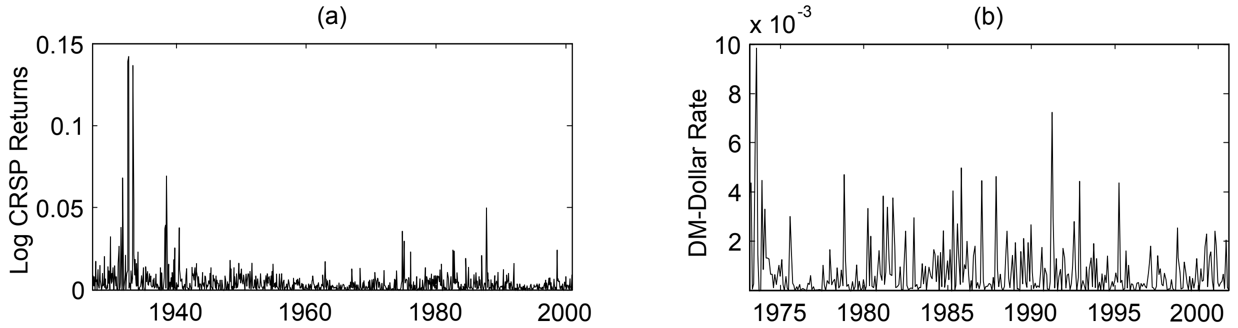
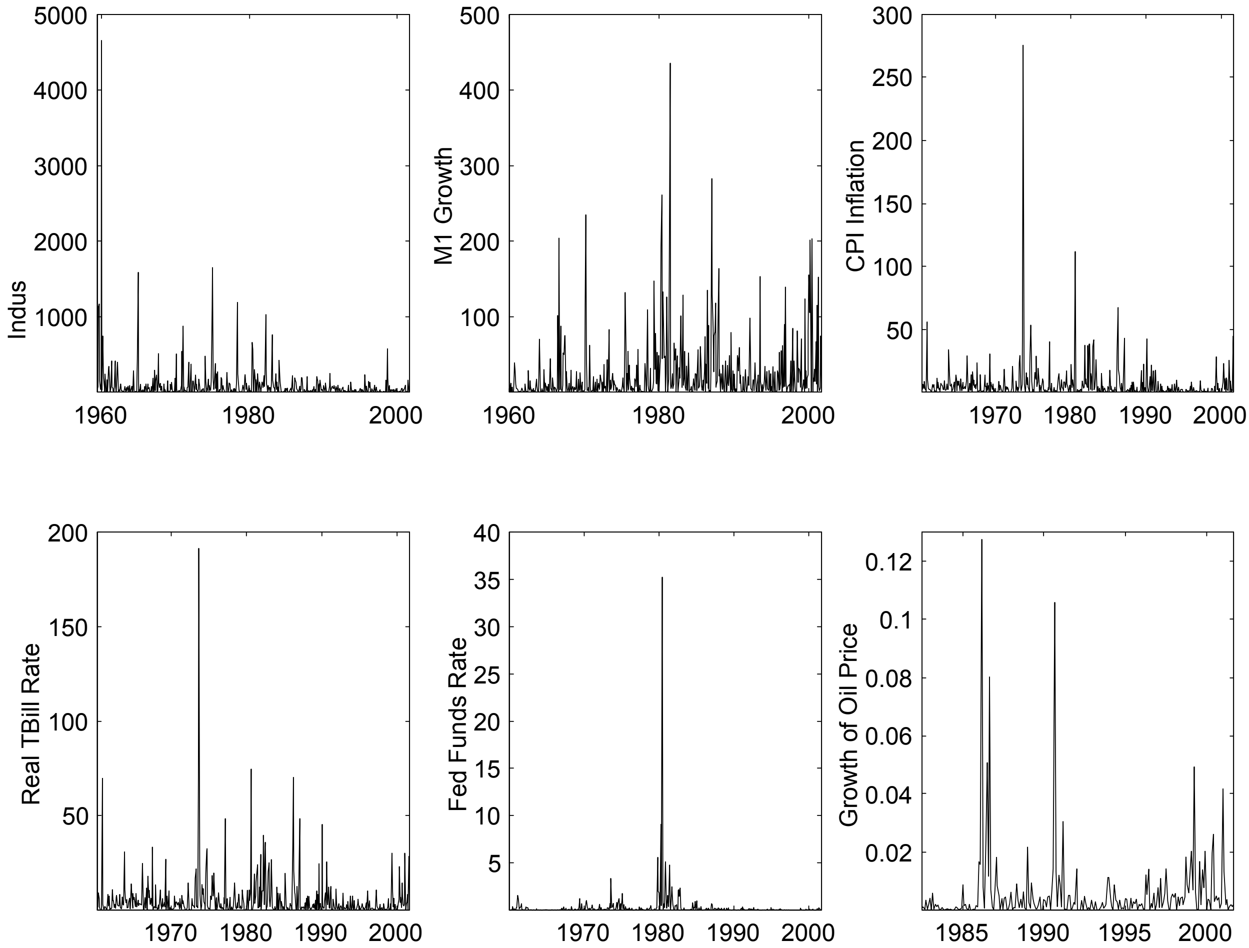


Figure 5.2: Squared Residuals of Autoregressions



A. Appendix

Throughout this Appendix, K denotes a generic constant independent of n. We use $u.i.$ to mean uniformly integrable. Given an $m \times n$ matrix A, let $\|A\| = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|$; for a $m \times 1$ vector a, let $|a| = \sum_{i=1}^{m} |a_i|$. For any $n \times n$ matrix A, $diag\,(a_{11}, \ldots, a_{nn})$ denotes a diagonal matrix with a_{ii}, $i = 1, \ldots, n$ in the main diagonal. Similarly, let $[a_{ij}]_{i,j=1,\ldots,n}$ denote a matrix A with typical element a_{ij}. For any bootstrap statistic T_n^* we write $T_n^* \xrightarrow{P^*} 0$ in probability when $\lim_{n\to\infty} P\left[P^*\left(|T_n^*| > \delta\right)\right] = 0$ for any $\delta > 0$, i.e. $P^*\left(|T_n^*| > \delta\right) = o_P(1)$. We write $T_n^* \Rightarrow^{d_{P^*}} D$, in probability, for any distribution D, when weak convergence under the bootstrap probability measure occurs in a set with probability converging to one. For simplicity, we omit the dependence on n of bootstrap estimators, e.g. $\hat{\varepsilon}_t^* \equiv \hat{\varepsilon}_{nt}^*$, $Y_t^* \equiv Y_{nt}^*$. Likewise, $\hat{\phi} \equiv \hat{\phi}_n$ throughout

The following CLT will be useful in proving results for the bootstrap (cf. White, 1999, p. 133; the Lindeberg condition there has been replaced by the stronger Lyapunov condition here):

Theorem A.1 (Martingale Difference Arrays CLT). *Let $\{Z_{nt}, \mathcal{F}_{nt}\}$ be a martingale difference array such that $\sigma_{nt}^2 = E\left(Z_{nt}^2\right)$, $\sigma_{nt}^2 \neq 0$, and define $\bar{Z}_n \equiv n^{-1}\sum_{t=1}^{n} Z_{nt}$ and $\bar{\sigma}_n^2 \equiv Var\left(\sqrt{n}\bar{Z}_n\right) = n^{-1}\sum_{t=1}^{n} \sigma_{nt}^2$. If*

1. *$n^{-1}\sum_{t=1}^{n} Z_{nt}^2/\bar{\sigma}_n^2 - 1 \xrightarrow{P} 0$, and*
2. *$\lim_{n\to\infty} \bar{\sigma}_n^{-2(1+\delta)} n^{-(1+\delta)} \sum_{t=1}^{n} E\left|Z_{nt}\right|^{2(1+\delta)} = 0$ for some $\delta > 0$,*

then $\sqrt{n}\bar{Z}_n/\bar{\sigma}_n \Rightarrow N(0,1)$.

The following Lemma generalizes Kuersteiner's (2001) Lemma A.1. Kuersteiner's Assumption A.1 is stronger than our Assumption A in that it assumes $\{\varepsilon_t\}$ is stationary ergodic, and in that it imposes a summability condition on the fourth order cumulants.

Lemma A.1. *Under Assumption A, for each $m \in \mathbb{N}$, m fixed, the vector*

$$n^{-1/2}\sum_{t=1}^{n} \left(\varepsilon_t\varepsilon_{t-1}, \ldots, \varepsilon_t\varepsilon_{t-m}\right)' \Rightarrow N\left(0, \Omega_m\right),$$

where $\Omega_m = \sigma^4\left[\tau_{r,s}\right]_{r,s=1,\ldots,m}$.

Lemmas A.2-A.5 are used to prove the asymptotic validity of the recursive-design WB (cf. Theorem 3.2). In these lemmas, $\hat{\varepsilon}_t^* = \hat{\varepsilon}_t\eta_t$, $t = 1, \ldots, n$, where $\hat{\varepsilon}_t = y_t - \hat{\phi}_n' Y_{t-1}$, and η_t is i.i.d. $(0,1)$ such that $E^*\left|\eta_t\right|^4 \leq \Delta < \infty$.

Lemma A.2. *Under Assumption A, for fixed $m \in \mathbb{N}$,*

(i) $n^{-1}\sum_{t=j+1}^{n}\hat{\varepsilon}_{t-j}^{*2} \xrightarrow{P^*} \sigma^2$, *in probability,* $j = 0, 1, \ldots, m$;

(ii) $n^{-1}\sum_{t=j+1}^{n}\hat{\varepsilon}_{t-j}^{*}\hat{\varepsilon}_{t}^{*} \xrightarrow{P^*} 0$, *in probability,* $j = 1, \ldots, m$.

If we strengthen Assumption A by A′ (vi′), then

(iii) $n^{-1}\sum_{t=\max(i,j)+1}^{n}\hat{\varepsilon}_{t}^{*2}\hat{\varepsilon}_{t-j}^{*}\hat{\varepsilon}_{t-i}^{*} \xrightarrow{P^*} \sigma^4\tau_{ij}1(i=j)$, *in probability,* $j, i = 1, \ldots, m$, *where* $1(i=j)$ *is 1 if* $i = j$, *and 0 otherwise.*

The following lemma is the WB analogue of Lemma A.1.

Lemma A.3. *Under Assumption A strengthened by A(vi′), for all fixed $m \in \mathbb{N}$,*

$$n^{-1/2}\sum_{t=m+1}^{n}\left(\hat{\varepsilon}_{t}^{*}\hat{\varepsilon}_{t-1}^{*}, \ldots, \hat{\varepsilon}_{t}^{*}\hat{\varepsilon}_{t-m}^{*}\right)' \Rightarrow^{d_{P^*}} N\left(0, \tilde{\Omega}_m\right),$$

in probability, where $\tilde{\Omega}_m \equiv \sigma^4 diag\left(\tau_{11}, \ldots, \tau_{mm}\right)$ *and* $\Rightarrow^{d_{P^*}}$ *denotes weak convergence under the bootstrap probability measure.*

Lemma A.4. *Suppose Assumption A holds. Then,* $n^{-1}\sum_{t=1}^{n}Y_{t-1}^{*}Y_{t-1}^{*\prime} \xrightarrow{P^*} A$, *in probability, where* $A \equiv \sigma^2\sum_{j=1}^{\infty}b_j b_j'$.

Lemma A.5. *Suppose Assumption A strengthened by A(vi′) holds. Then,*

$$n^{-1/2}\sum_{t=1}^{n}Y_{t-1}^{*}\hat{\varepsilon}_{t}^{*} \Rightarrow^{d_{P^*}} N\left(0, \tilde{B}\right),$$

in probability, where $\tilde{B} = \sum_{j=1}^{\infty}b_j b_j'\sigma^4\tau_{jj}$.

Proof of Theorem 3.1. We show that (i) $A_{1n} \equiv n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}' \xrightarrow{P} A$; and (ii) $A_{2n} \equiv n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\varepsilon_t \Rightarrow N(0, B)$. First, notice that for any stationary AR(p) process we have $y_t = \sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$, where $\{\psi_j\}$ satisfy the recursion $\psi_s - \phi_1\psi_{s-1} - \ldots - \phi_p\psi_{s-p} = 0$ with $\psi_0 = 1$ and $\psi_j = 0$ for $j < 0$, implying that $\sum_{j=0}^{\infty}j\left|\psi_j\right| < \infty$. We can write $Y_{t-1} = \left(\sum_{j=0}^{\infty}\psi_j\varepsilon_{t-1-j}, \ldots, \sum_{j=0}^{\infty}\psi_j\varepsilon_{t-p-j}\right)' = \sum_{j=1}^{\infty}b_j\varepsilon_{t-j}$ with $b_j = \left(\psi_{j-1}, \ldots, \psi_{j-p}\right)'$, where $\psi_{-j} = 0$ for all $j > 0$. Hence, by direct evaluation,

$$A \equiv E\left(Y_{t-1}Y_{t-1}'\right) = E\left[\left(\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}b_j b_i'\varepsilon_{t-j}\varepsilon_{t-i}\right)\right] = \sigma^2\sum_{j=1}^{\infty}b_j b_j' = \sigma^2\left[\sum_{j=0}^{\infty}\psi_j\psi_{j+|k-l|}\right]_{k,l=1,\ldots,p}$$

since $E\left(\varepsilon_{t-i}\varepsilon_{t-j}\right) = 0$ for $i \neq j$ under the m.d.s. assumption, and $\sum_{j=0}^{\infty}\left|\psi_j\psi_{j+|k-l|}\right| < \sum_{j=0}^{\infty}\left|\psi_j\right|\sum_{j=0}^{\infty}\left|\psi_{j+|k-l|}\right| < \infty$ for all k, l. To show (i), for fixed $m \in \mathbb{N}$, define $A_{1n}^{m} \equiv n^{-1}\sum_{t=1}^{n}Y_{t-1,m}Y_{t-1,m}'$, where $Y_{t-1,m} = \sum_{j=0}^{m} b_j \varepsilon_{t-j}$. It suffices to show: (a) $A_{1n}^m \to A_1^m \equiv \sigma^2 \sum_{j=1}^{m} b_j b_j'$ as $n \to \infty$, for each fixed m; (b) $A_1^m \to A$ as $m \to \infty$, and (c) $\lim_{m\to\infty} \limsup_{n\to\infty} P\left[\left|A_{1n} - A_{1n}^m\right| \geq \eta\right] = 0$ for all $\eta > 0$ (cf. Proposition 6.3.9 of Brockwell and Davis (BD) (1991), p.207). For (a), we have $A_{1n}^m = \sum_{j=0}^{m}\sum_{i=0}^{m} b_j b_i' n^{-1} \sum_{t=1}^{n} \varepsilon_{t-j}\varepsilon_{t-i}$. For fixed $i \neq j$ it follows that $n^{-1}\sum_{t=1}^{n} \varepsilon_{t-j}\varepsilon_{t-i} \xrightarrow{P} 0$ by Andrews' LLN (1988) for *u.i.* L_1-mixingales since $\{\varepsilon_{t-j}\varepsilon_{t-i}\}$ is a m.d.s. with $E\left|\varepsilon_{t-j}\varepsilon_{t-i}\right|^r \leq \left\|\varepsilon_{t-j}\right\|_{2r}^r \left\|\varepsilon_{t-i}\right\|_{2r}^r < \Delta^{2r} < \infty$ by Cauchy-Schwartz and Assumption A(vi). For fixed $i = j$, we can write $n^{-1}\sum_{t=1}^{n} \varepsilon_{t-j}^2 - \sigma^2 = n^{-1}\sum_{t=1}^{n} z_t + n^{-1}\sum_{t=1}^{n} E\left(\varepsilon_{t-j}^2|\mathcal{F}_{t-j-1}\right) - \sigma^2$, with $z_t = \varepsilon_{t-j}^2 - E\left(\varepsilon_{t-j}^2|\mathcal{F}_{t-j-1}\right)$. Since z_t can be shown to be an *u.i.* m.d.s, the first term goes to zero in probability by Andrews' LLN. The second term also vanishes in probability by Assumption A(iii). Thus, $n^{-1}\sum_{t=1}^{n} \varepsilon_{t-j}^2 - \sigma^2 \xrightarrow{P} 0$ for fixed j. It follows that $A_{1n}^m \xrightarrow{P} \sigma^2 \sum_{j=0}^{m} b_j b_j' \equiv A_1^m$, proving (a). (b) follows from the dominated convergence theorem, given that $\left\|\sum_{j=1}^{\infty} b_j b_j'\right\| = \sum_{j=1}^{\infty} |b_j|^2 < \infty$. To prove (c), note that

$$
\begin{aligned}
P\left[\left|A_{1n} - A_{1n}^m\right| \geq \eta\right] &\leq E\left|A_{1n} - A_{1n}^m\right| \\
&\leq 2\left(\sum_{j>m}^{\infty} |b_j|\right)\left(\sum_{j=1}^{\infty} |b_j|\right) n^{-1}\sum_{t=1}^{n} E\left|\varepsilon_{t-i}\varepsilon_{t-j}\right| \leq \left(\sum_{j>m}^{\infty} |b_j|\right) K \to 0 \text{ as } m \to \infty,
\end{aligned}
$$

since $E\left|\varepsilon_{t-i}\varepsilon_{t-j}\right| \leq \Delta$ for some $\Delta < \infty$, and since $\sum_{j=1}^{\infty} |b_j| < \infty$. Next, we prove (ii). We apply Proposition 6.3.9 of BD. Let $Z_t = Y_{t-1}\varepsilon_t \equiv \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}\varepsilon_t$. For fixed m, define $Z_t^m = Y_{t-1,m}\varepsilon_t = \sum_{j=0}^{m} b_j \varepsilon_{t-j}\varepsilon_t$, where $Y_{t-1,m}$ is as above. We first show $n^{-1/2}\sum_{t=1}^{n} Z_t^m \Rightarrow N(0, B_m)$, with $B_m = \sum_{j=0}^{m}\sum_{i=0}^{m} b_j b_i' \sigma^4 \tau_{ji}$. We have

$$
n^{-1/2}\sum_{t=1}^{n} Z_t^m = n^{-1/2}\sum_{t=1}^{n}\sum_{j=0}^{m} b_j \varepsilon_{t-j}\varepsilon_t = \sum_{j=0}^{m} b_j n^{-1/2}\sum_{t=1}^{n} \varepsilon_{t-j}\varepsilon_t \equiv \sum_{j=0}^{m} b_j \mathcal{X}_{nj}.
$$

By Lemma A.1 we have that $(\mathcal{X}_{n1}, \ldots, \mathcal{X}_{nm}) \Rightarrow N(0, \Omega_m)$. Thus, $\sum_{j=0}^{m} b_j \mathcal{X}_{nj} \Rightarrow N(0, B_m)$, with $B_m = b'\Omega_m b$, $b = (b_1', \ldots, b_m')'$. Since $\left\|\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} b_j b_i' \sigma^4 \tau_{ji}\right\| \leq \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} |b_j|\,|b_i|\,\sigma^4\,|\tau_{ji}| < \infty$, it follows that $B_m \to B \equiv \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} b_j b_i' \sigma^4 \tau_{ji}$ as $m \to \infty$. Finally, for any $\lambda \in \mathbb{R}^p$, consider

$$
\lim_{m\to\infty}\limsup_{n\to\infty} P\left[\left|n^{-1/2}\sum_{t=1}^{n} \lambda' Z_t - n^{-1/2}\sum_{t=1}^{n} \lambda' Z_t^m\right| \geq \eta\right] = \lim_{m\to\infty}\limsup_{n\to\infty} P\left[\left|n^{-1/2}\sum_{t=1}^{n}\sum_{j>m} \lambda' b_j \varepsilon_{t-j}\varepsilon_t\right| \geq \eta\right]
$$

$$
\leq \lim_{m\to\infty}\limsup_{n\to\infty} \frac{1}{n\eta^2} E\left(\left|\sum_{t=1}^{n}\sum_{j>m} \lambda' b_j \varepsilon_{t-j}\varepsilon_t\right|^2\right) = \lim_{m\to\infty}\limsup_{n\to\infty} \frac{1}{n\eta^2}\left(\sum_{t=1}^{n}\sum_{j>m}\sum_{i>m} \lambda' b_j b_i' \lambda \sigma^4 \tau_{ji}\right) = 0,
$$

where the inequality holds by Chebyshev's inequality, the second-to-last equality holds by the fact that $E\left(\varepsilon_{t-j}\varepsilon_t\varepsilon_{s-i}\varepsilon_s\right) = 0$ for $s \neq t$, and all i, j, and the last equality holds by the summability of $\{\psi_j\}$ and the fact that τ_{ji} are uniformly bounded. ■

Proof of Theorem 3.2. By Lemma A.4, $n^{-1}\sum_{t=1}^{n}Y_{t-1}^{*}Y_{t-1}^{*\prime}\xrightarrow{P^*}A$, in probability, whereas Lemma A.5 implies $n^{-1/2}\sum_{t=1}^{n}Y_{t-1}^{*}\hat{\varepsilon}_t^{*}\Rightarrow^{d_{P^*}}N\left(0,\tilde{B}\right)$, in probability. Since under Assumption A(iv′), $B=\tilde{B}$, the result follows by Polya's Theorem, given that the normal distribution is everywhere continuous. ■

Proof of Theorem 3.3 We need to show (a) $n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\xrightarrow{P}A$, and (b) $n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\hat{\varepsilon}_t^{*}\Rightarrow^{d_{P^*}}N\left(0,B\right)$ in probability. (a) was proven in Theorem 3.1. To show (b) note that

$$\begin{aligned}n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\hat{\varepsilon}_t^{*} &= n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\varepsilon_t\eta_t-n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\left(\varepsilon_t-\hat{\varepsilon}_t\right)\eta_t\\ &= n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\varepsilon_t\eta_t-n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\eta_t\sqrt{n}\left(\hat{\phi}-\phi\right)\equiv A_1^*-A_2^*.\end{aligned}$$

First, note that $A_2^*\xrightarrow{P^*}0$, in probability, since $\sqrt{n}\left(\hat{\phi}-\phi\right)=O_P\left(1\right)$ and $n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\eta_t\xrightarrow{P^*}0$, in probability. This follows by showing $E^*\left(n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\eta_t\right)=0$ and $Var^*\left(n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\eta_t\right)=n^{-2}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}Y_{t-1}Y_{t-1}^{\prime}\xrightarrow{P}0$, under Assumption A. We next show $A_1^*\Rightarrow^{d_{P^*}}N\left(0,B\right)$ in probability P, where $B=Var\left(n^{-1/2}\sum_{t=1}^{n}Y_{t-1}\varepsilon_t\right)=n^{-1}\sum_{t=1}^{n}E\left(Y_{t-1}Y_{t-1}^{\prime}\varepsilon_t^2\right)$. For any $\lambda\in\mathbb{R}^p$, $\lambda^{\prime}\lambda=1$, let $Z_t^*=\lambda^{\prime}Y_{t-1}\varepsilon_t\eta_t$. $\{Z_t^*\}$ is (conditionally) independent such that $E^*\left(n^{-1/2}\sum_{t=1}^{n}Z_t^*\right)=0$ and $Var^*\left(n^{-1/2}\sum_{t=1}^{n}Z_t^*\right)=\lambda^{\prime}n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\varepsilon_t^2\lambda$. We now apply Lyapunov's Theorem (e.g. Durrett, 1995, p.121). Let $\alpha_n^{*2}=\lambda^{\prime}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\varepsilon_t^2\lambda$. By arguments similar to Theorem 3.1, $n^{-1}\alpha_n^{*2}\xrightarrow{P}B$. If for some $r>1$

$$\alpha_n^{*-2r}\sum_{t=1}^{n}E^*\left|Z_t^*\right|^{2r}\xrightarrow{P}0 \tag{A.1}$$

then $\alpha_n^{*-1}\sum_{t=1}^{n}Z_t^*\Rightarrow^{d_{P^*}}N\left(0,1\right)$ in probability. By Slutsky's Theorem, it follows that $n^{-1/2}\sum_{t=1}^{n}Z_t^*\Rightarrow^{d_{P^*}}N\left(0,\lambda^{\prime}B\lambda\right)$. To show (A.1), note that the RHS can be written as

$$\left(\frac{\alpha_n^{*2}}{n}\right)^{-r}n^{-r}\sum_{t=1}^{n}\left|\lambda^{\prime}Y_{t-1}\varepsilon_t\right|^{2r}E^*\left|\eta_t\right|^{2r}.$$

Thus, it suffices to show that $E\left|n^{-r}\sum_{t=1}^{n}\left|\lambda^{\prime}Y_{t-1}\varepsilon_t\right|^{2r}E^*\left|\eta_t\right|^{2r}\right|\to 0$. Since $E^*\left|\eta_t\right|^{2r}\leq\Delta<\infty$, this holds provided $E\left|\lambda^{\prime}Y_{t-1}\varepsilon_t\right|^{2r}\leq\Delta<\infty$, which follows under Assumption A. ■

Proof of Theorem 3.4 Let $\hat{\varepsilon}_t=y_t-\hat{\phi}^{\prime}Y_{t-1}$, $\hat{\varepsilon}_t^*=y_t^*-\hat{\phi}^{\prime}Y_{t-1}^*$, and $\varepsilon_t^*=y_t^*-\phi^{\prime}Y_{t-1}^*$. We show that (i) $n^{-1}\sum_{t=1}^{n}Y_{t-1}^{*}Y_{t-1}^{*\prime}\xrightarrow{P^*}A$ in probability, and (ii) $n^{-1/2}\sum_{t=1}^{n}Y_{t-1}^{*}\hat{\varepsilon}_t^{*}\Rightarrow^{d_{P^*}}N\left(0,B\right)$ in probability. Conditional on the original data, for any $\delta>0$,

$$n^{-1}\sum_{t=1}^{n}Y_{t-1}^{*}Y_{t-1}^{*\prime}-A=\left\{n^{-1}\sum_{t=1}^{n}Y_{t-1}^{*}Y_{t-1}^{*\prime}-n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}\right\}+\left\{n^{-1}\sum_{t=1}^{n}Y_{t-1}Y_{t-1}^{\prime}-A\right\}\equiv A_{1n}^*+A_{2n}.$$

Theorem 3.1 shows $A_{2n}\xrightarrow{P}0$. Next we show $A_{1n}^*\xrightarrow{P^*}A$, in probability. Conditional on the data, by

Chebyshev's inequality, it suffices that $E^*\left(A_{2n}^* A_{2n}^{*\prime}\right) = o_P(1)$. But

$$\begin{aligned}
E^*\left(A_{2n}^* A_{2n}^{*\prime}\right) &= n^{-1} E^*\left(n^{-1}\sum_{t=1}^n\sum_{s=1}^n\left(Y_{t-1}^* Y_{t-1}^{*\prime} - n^{-1}\sum_{t=1}^n Y_{t-1}Y_{t-1}^\prime\right)\left(Y_{s-1}^* Y_{s-1}^{*\prime} - n^{-1}\sum_{t=1}^n Y_{t-1}Y_{t-1}^\prime\right)^\prime\right) \\
&= n^{-1}\left\{n^{-1}\sum_{t=1}^n\left(Y_{t-1}Y_{t-1}^\prime - n^{-1}\sum_{t=1}^n Y_{t-1}Y_{t-1}^\prime\right)\left(Y_{t-1}Y_{t-1}^\prime - n^{-1}\sum_{t=1}^n Y_{t-1}Y_{t-1}^\prime\right)^\prime\right\},
\end{aligned}$$

where the middle matrix is $O_P(1)$ given Assumption A (in particular, given A (vi)), delivering the result. Next we show (ii). We can write

$$\begin{aligned}
n^{-1/2}\sum_{t=1}^n Y_{t-1}^*\hat{\varepsilon}_t^* &= n^{-1/2}\sum_{t=1}^n\left(Y_{t-1}^*\varepsilon_t^* - n^{-1}\sum_{t=1}^n Y_{t-1}\varepsilon_t\right) \\
&+ \left(n^{-1}\sum_{t=1}^n Y_{t-1}Y_{t-1}^\prime - n^{-1}\sum_{t=1}^n Y_{t-1}^* Y_{t-1}^{*\prime}\right)\sqrt{n}\left(\hat{\phi}_n - \phi\right) \equiv B_1 + B_2.
\end{aligned}$$

Since $B_2 \xrightarrow{P^*} 0$ in probability because of the previous argument and $\sqrt{n}\left(\hat{\phi}_n - \phi\right) = O_P(1)$. (ii) follows if we prove that $B_1 \Rightarrow^{d_{P^*}} N(0, B)$ in probability. This follows straightforwardly by an application of Lyapunov's CLT given that $Z_t^* \equiv Y_{t-1}^*\varepsilon_t^* - n^{-1}\sum_{t=1}^n Y_{t-1}\varepsilon_t$ is (conditionally) i.i.d. with mean zero and variance $Var^*\left(Z_t^*\right) = n^{-1}\sum_{t=1}^n Z_t Z_t^\prime$, where $Z_t \equiv Y_{t-1}\varepsilon_t - n^{-1}\sum_{t=1}^n Y_{t-1}\varepsilon_t$, and by Theorem 3.1 $n^{-1}\sum_{t=1}^n Y_{t-1}Y_{t-1}^\prime\varepsilon_t^2 \xrightarrow{P} B$ and $n^{-1}\sum_{t=1}^n Y_{t-1}\varepsilon_t \xrightarrow{P} 0$. ■

Proof of Corollary 3.1. Given the previous results, it suffices to show that $\hat{C}_n^* \xrightarrow{P^*} C$, i.e. (i) $\hat{A}_n^* \xrightarrow{P^*} A$, and (ii) $\hat{B}_n^* \xrightarrow{P^*} B$, in probability, where $B = \tilde{B}$ for the recursive-design WB. We showed (i) in Lemma A.4 for the recursive-design WB, and in Theorems 3.3 and 3.4, for the fixed-design WB and pairwise bootstrap, respectively. Next, we sketch the proof of (ii). For simplicity we take $p = 1$. The proof for general p is similar. For each of the three bootstrap schemes, we can write $\widetilde{\varepsilon}_t^* = \hat{\varepsilon}_t^* - \left(\hat{\phi}_n^* - \hat{\phi}_n\right)y_{t-1}^*$, where $\hat{\varepsilon}_t^* = \hat{\varepsilon}_t\eta_t$ for the recursive-design and fixed-design WB, and $\hat{\varepsilon}_t^* = y_t^* - \hat{\phi}_n y_{t-1}^*$ for the pairwise bootstrap. Thus,

$$\begin{aligned}
\hat{B}_n^* &= \hat{B}_{1n}^* + \hat{B}_{2n}^* + \hat{B}_{3n}^*, \quad \text{with} \\
\hat{B}_{1n}^* &= n^{-1}\sum_{t=1}^n y_{t-1}^{*2}\hat{\varepsilon}_t^{*2},\ \hat{B}_{2n}^* = -2n^{-1}\sum_{t=1}^n y_{t-1}^{*3}\hat{\varepsilon}_t^*\left(\hat{\phi}_n^* - \hat{\phi}_n\right),\ \text{and}\ \hat{B}_{3n}^* = n^{-1}\sum_{t=1}^n y_{t-1}^{*4}\left(\hat{\phi}_n^* - \hat{\phi}_n\right)^2.
\end{aligned}$$

It is enough to show that with probability approaching one, (a) $\hat{B}_{1n}^* \xrightarrow{P^*} B$, (b) $\hat{B}_{2n}^* \xrightarrow{P^*} 0$, and (c) $\hat{B}_{3n}^* \xrightarrow{P^*} 0$. For the fixed-design WB, starting with (a), note that $y_{t-1}^* = y_{t-1}$, and therefore $\hat{B}_{1n}^* - B = n^{-1}\sum_{t=1}^n y_{t-1}^2\hat{\varepsilon}_t^2\left(\eta_t^2 - 1\right) + n^{-1}\sum_{t=1}^n y_{t-1}^2\hat{\varepsilon}_t^2 - B \equiv \chi_{1n} + \chi_{2n}$. Under our assumptions $\chi_{2n} \xrightarrow{P} 0$. Since $\hat{\varepsilon}_t = \varepsilon_t - \left(\hat{\phi}_n - \phi\right)y_{t-1}$, we can write $\chi_{1n} = n^{-1}\sum_{t=1}^n y_{t-1}^2\varepsilon_t^2\left(\eta_t^2 - 1\right) - 2\left(\hat{\phi}_n - \phi\right)n^{-1}\sum_{t=1}^n y_{t-1}^3\varepsilon_t\left(\eta_t^2 - 1\right) + \left(\hat{\phi}_n - \phi\right)^2 n^{-1}\sum_{t=1}^n y_{t-1}^4\left(\eta_t^2 - 1\right)$. We can show that each of these terms is $o_{P^*}(1)$, in probability. For the first term, write $z_t = y_{t-1}^2 \varepsilon_t^2 \left(\eta_t^2 - 1\right)$, and note that z_t is (conditionally) a m.d.s. with respect to $\mathcal{F}_\eta^t = \sigma\left(\eta_t, \ldots, \eta_1\right)$. Thus, by Andrews' (1988) LLN, it follows that $n^{-1}\sum_{t=1}^n z_t \xrightarrow{P^*} 0$, in probability, provided that $E^* \left|z_t\right|^r = O_P(1)$, or $E\left(E^* \left|z_t\right|^r\right) = O(1)$, for some $r > 1$, which holds under our moment conditions (in particular, the existence of $4r$ moments of ε_t suffices). A similar argument applies to the last two terms of χ_{1n}, where we note that $\hat{\phi}_n - \phi \xrightarrow{P} 0$. For (b), and given $\hat{\phi}_n^* - \hat{\phi}_n = o_{P^*}(1)$, it suffices that $n^{-1}\sum_{t=1}^n y_{t-1}^3 \hat{\varepsilon}_t^* = O_{P^*}(1)$, in probability, or that $E^*\left|n^{-1}\sum_{t=1}^n y_{t-1}^3 \hat{\varepsilon}_t^*\right| = O_P(1)$. This condition holds under Assumption A (first apply the triangle inequality, then use the definition of $\hat{\varepsilon}_t$, and finally apply repeatedly the Cauchy-Schwartz inequality to the sums involving products of y_{t-1} and/or ε_t.) For (c), by a reasoning similar to (b), it suffices that $n^{-1}\sum_{t=1}^n y_{t-1}^4 = O_P(1)$, which holds under our moment conditions. For the pairwise bootstrap, we proceed similarly, but rely on the (conditional) independence of $\left(y_t^*, y_{t-1}^*\right)$ to obtain the results. In particular, for (a), following Theorem 3.3 we can define $\hat{\varepsilon}_t^* = \varepsilon_t^* - \left(\hat{\phi}_n - \phi\right) y_{t-1}^*$, with $\varepsilon_t^* = y_t^* - \phi y_{t-1}^*$, which implies $\hat{B}_{1n}^* = \chi_{1n} + \chi_{2n}$. We can show that $\chi_{2n} = o_{P^*}(1)$, whereas $\chi_{1n} = n^{-1}\sum_{t=1}^n z_{1t}^* + \zeta_n$ where $z_{1t}^* = y_{t-1}^{*2}\varepsilon_{t-1}^{*2} - n^{-1}\sum_{t=1}^n y_{t-1}^2 \varepsilon_t^2$ and $\zeta_n = n^{-1}\sum_{t=1}^n y_{t-1}^2 \varepsilon_t^2$. By Theorem 3.1 $\zeta_n \xrightarrow{P} B$. Since z_{1t}^* is a uniformly square-integrable m.d.s. (conditional on the original data) Andrews' LLN implies that the first term of χ_{1n} is $o_{P^*}(1)$, in probability. For the recursive-design WB, for part (a), note that we can write $\hat{B}_{1n}^* = \chi_{1n} + \chi_{2n}$, where $\chi_{1n} = \sum_{j=1}^{n-1} \hat{b}_j^2 \left(n^{-1}\sum_{t=j+1}^n \hat{\varepsilon}_{t-j}^{*2} \hat{\varepsilon}_t^{*2}\right)$, and $\chi_{2n} = n^{-1}\sum_{t=1}^n \sum_{i \neq j, i,j=1}^{t-1} \hat{b}_j \hat{b}_i \hat{\varepsilon}_{t-i}^* \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^{*2}$. Now, using arguments analogous to those used in the proof of Lemmas A.4 and A.5 we can show that $\chi_{1n} \xrightarrow{P^*} \tilde{B}$, and $\chi_{2n} \xrightarrow{P^*} 0$, in probability. Similar arguments apply for (b) and (c).

Proof of Lemma A.1. The proof follows closely that of Lemma A.1 of Kuersteiner (2001). We reproduce his steps under our weaker Assumption A. In particular, we show that for all $\lambda \in \mathbb{R}^m$ such that $\lambda'\lambda = 1$ we have $n^{-1/2}\sum_{t=1}^n \lambda' Y_t \Rightarrow N\left(0, \lambda'\Omega_m\lambda\right)$, where $Y_t = \left(\varepsilon_t\varepsilon_{t-1}, \ldots, \varepsilon_t\varepsilon_{t-m}\right)'$. Noting that $\{Y_t, \mathcal{F}_t\}$ is a vector m.d.s., we check the m.d.s. CLT conditions (cf. Davidson, 1994, Theorem 24.3). Let $Z_t = \lambda' Y_t$. We check: (i) $n^{-1}\sum_{t=1}^n \left[Z_t^2 - E\left(Z_t^2\right)\right] \xrightarrow{P} 0$, where $E\left(Z_t^2\right) = \lambda' E\left(Y_t Y_t'\right)\lambda = \lambda'\Omega_m\lambda$; and (ii) $n^{-1/2}\max_{1\leq t\leq n}\left|Z_t\right| \xrightarrow{P} 0$. To see (i), note that $n^{-1}\sum_{t=1}^n \left[Z_t^2 - E\left(Z_t^2\right)\right] = A_1 + A_2$, with

$$A_1 = n^{-1}\sum_{t=1}^n \left[Z_t^2 - E\left(Z_t^2|\mathcal{F}_{t-1}\right)\right] \quad \text{and } A_2 = n^{-1}\sum_{t=1}^n \left[E\left(Z_t^2|\mathcal{F}_{t-1}\right) - E\left(Z_t^2\right)\right].$$

First consider A_1. Since $\{Z_t, \mathcal{F}_{t-1}\}$ is a m.d.s we have that $Z_t^2 - E\left(Z_t^2|\mathcal{F}_{t-1}\right)$ is an L_1-mixingale with mixingale constants $c_t = E\left|Z_t^2 - E\left(Z_t^2|\mathcal{F}_{t-1}\right)\right|$: $E\left|E\left(Z_t^2 - E\left(Z_t^2|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t-k}\right)\right| \leq c_t\xi_k$, $k = 0, 1, \ldots$, with $\xi_k = 1$ for $k = 0$ and $\xi_k = 0$ otherwise. Thus, we apply Andrews' LLN for L_1-mixingales (Andrews, 1988) to show $A_1 \xrightarrow{P} 0$. It suffices that for some $r > 1$ $E\left|Z_t^2\right|^r \leq K < \infty$ and $n^{-1}\sum_{t=1}^n c_t < \infty$. Now,

$E|Z_t|^{2r} = E\left|\sum_{i=1}^m \lambda_i \varepsilon_t \varepsilon_{t-i}\right|^{2r} \leq \left(\sum_{i=1}^m |\lambda_i| \left\|\varepsilon_t \varepsilon_{t-i}\right\|_{2r}\right)^{2r} < K$ by repeated application of Minkowski and Cauchy-Schwartz, given Assumption A(vi). The second condition on $\{c_t\}$ follows similarly. Next we consider A_2. We have that

$$A_2 = \lambda' n^{-1} \sum_{t=1}^n \left(E\left(Y_t Y_t' | \mathcal{F}_{t-1}\right) - E\left(Y_t Y_t'\right)\right) \lambda = \lambda' \left[n^{-1} \sum_{t=1}^n \varepsilon_{t-i}\varepsilon_{t-j} E\left(\varepsilon_t^2 | \mathcal{F}_{t-1}\right) - \sigma^4 \tau_{i,j}\right]_{i,j=1,\ldots,p} \lambda \xrightarrow{P} 0,$$

given Assumption A(v). This proves (i). To prove (ii), note that by Markov's inequality, for any $\eta > 0$ and for some $r > 1$,

$$P\left(\frac{1}{\sqrt{n}} \max_{1\leq t\leq n} |Z_t| > \eta\right) \leq \sum_{t=1}^n P\left(|Z_t| > n^{1/2}\eta\right) \leq \eta^{-2r} n^{-r} \sum_{t=1}^n E|Z_t|^{2r} \leq K\eta^{-2r} n^{1-r} \to 0. \blacksquare$$

Proof of Lemma A.2. First we consider (i) with $j = 0$, without loss of generality. By definition, $\hat{\varepsilon}_t^* \equiv \hat{\varepsilon}_t \eta_t$, and thus

$$n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^{*2} - \sigma^2 = \left[n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^2 \left(\eta_t^2 - 1\right)\right] + \left[n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^2 - \sigma^2\right] \equiv F_{1n}^* + F_{2n},$$

with the obvious definitions. Under our assumptions $F_{2n} = o_P(1)$. So it suffices to show that $P^*[|F_{1n}^*| > \delta] = o_P(1)$, for any $\delta > 0$, or, by Chebyshev's inequality, that $E^*\left((F_{1n}^*)^2\right) = o_P(1)$. Let $z_t^* \equiv \hat{\varepsilon}_t^2\left(\eta_t^2 - 1\right)$ and note that $E^*(z_t^* z_s^*) = 0$ for $t \neq s$, $E^*\left(z_t^{*2}\right) = \hat{\varepsilon}_t^4 E^*\left(\eta_t^4 - 2\eta_t^2 + 1\right) = \hat{\varepsilon}_t^4\left(E^*\left(\eta_t^4\right) - 1\right)$. Thus,

$$E^*\left[(F_{1n}^*)^2\right] = E^*\left(n^{-2} \sum_{t=1}^n \sum_{s=1}^n z_t^* z_s^*\right) = n^{-1}\left(n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^4 \left(E^*\left(\eta_t^4\right) - 1\right)\right) \leq n^{-1} K \left(n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^4\right) = o_P(1),$$

where the last inequality holds by $E^*\left(\eta_t^4\right) \leq \Delta < \infty$ and $n^{-1}\sum_{t=1}^n \hat{\varepsilon}_t^4 = O_P(1)$, given that $E|\varepsilon_t|^4 < K < \infty$ and that $\hat{\phi}_n \to \phi$ in probability. For (ii), by a similar reasoning, it suffices to note that

$$E^*\left[\left(n^{-1} \sum_{t=j+1}^n \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^*\right)^2\right] = n^{-2} \sum_{t=j+1}^n \hat{\varepsilon}_{t-j}^2 \hat{\varepsilon}_t^2 E^*\left(\eta_t^2 \eta_{t-j}^2\right) = n^{-2} \sum_{t=j+1}^n \hat{\varepsilon}_{t-j}^2 \hat{\varepsilon}_t^2 = o_P(1).$$

For (iii), note that

$$n^{-1} \sum_{t=\max(i,j)+1}^n \hat{\varepsilon}_t^{*2} \hat{\varepsilon}_{t-i}^* \hat{\varepsilon}_{t-j}^* - \sigma^4 \tau_{ij} 1(i=j) = n^{-1} \sum_{t=\max(i,j)+1}^n \hat{\varepsilon}_t^2 \hat{\varepsilon}_{t-i} \hat{\varepsilon}_{t-j} \left(\eta_t^2 \eta_{t-i} \eta_{t-j} - 1(i=j)\right)$$

$$+ n^{-1} \sum_{t=\max(i,j)+1}^n \left(\hat{\varepsilon}_t^2 \hat{\varepsilon}_{t-i} \hat{\varepsilon}_{t-j} - \sigma^4 \tau_{ij}\right) 1(i=j) \equiv G_{1n}^* + G_{2n}.$$

Under our assumptions, for any fixed i, j,

$$n^{-1}\sum_{t=\max(i,j)+1}^{n}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-j}=n^{-1}\sum_{t=\max(i,j)+1}^{n}\varepsilon_t^2\varepsilon_{t-i}\varepsilon_{t-j}+R_n,$$

where the remainder R_n involves products of elements of $\hat{\phi}_n-\phi$, which are $o_P(1)$ under our assumptions, with averages of products of elements of Y_{t-1} and ε_t, up to the fourth order, which are bounded in probability, given that $E\left|\varepsilon_t\right|^4<\Delta<\infty$. Thus, $R_n=o_P(1)$, and since $n^{-1}\sum_{t=\max(i,j)+1}^{n}\varepsilon_t^2\varepsilon_{t-i}\varepsilon_{t-j}\to\sigma^4\tau_{ij}$ (cf. proof of Lemma A.1), it follows that $G_{2n}=o_P(1)$. So, if we let $z_t^{*(i,j)}=\hat{\varepsilon}_t^{*2}\hat{\varepsilon}_{t-i}^{*}\hat{\varepsilon}_{t-j}^{*}-1\,(i=j)$, it suffices that

$$\begin{aligned}
P^*\left(\left|G_{1n}^*\right|>\delta\right) &\leq \frac{1}{\eta^2n^2}E^*\left[\sum_{t=\max(i,j)+1}^{n}\sum_{s=\max(i,j)+1}^{n}E^*\left(z_t^{*(i,j)}z_s^{*(i,j)}\right)\right]\\
&= \frac{1}{\eta^2n^2}\sum_{t=\max(i,j)+1}^{n}\hat{\varepsilon}_t^4\hat{\varepsilon}_{t-i}^2\hat{\varepsilon}_{t-j}^2E^*\left[\left(\eta_t^2\eta_{t-i}\eta_{t-j}-1\,(i=j)\right)^2\right]\\
&\leq \frac{K}{\eta^2n}\left(n^{-1}\sum_{t=\max(i,j)+1}^{n}\hat{\varepsilon}_t^4\hat{\varepsilon}_{t-i}^2\hat{\varepsilon}_{t-j}^2\right),
\end{aligned}$$

where the equality holds because $E^*\left(z_t^{*(i,j)}z_s^{*(i,j)}\right)=0$ for $s\neq t$ by the properties of $\{\eta_t\}$, and the second inequality uses the fact that $E^*\left|\eta_t\right|^4<\Delta<\infty$. Under Assumption A strengthened by A′ (vi′), we can show that $n^{-1}\sum_{t=\max(i,j)+1}^{n}\hat{\varepsilon}_t^4\hat{\varepsilon}_{t-k}^2\hat{\varepsilon}_{t-l}^2=O_P(1)$, which implies that $P^*\left(\left|G_{1n}^*\right|>\delta\right)=o_P(1)$. In fact, given that $\hat{\varepsilon}_t=\varepsilon_t-\left(\hat{\phi}_n-\phi\right)'Y_{t-1}$, it follows that $n^{-1}\sum_{t=\max(i,j)+1}^{n}\hat{\varepsilon}_t^4\hat{\varepsilon}_{t-i}^2\hat{\varepsilon}_{t-j}^2=n^{-1}\sum_{t=\max(i,j)+1}^{n}\varepsilon_t^4\varepsilon_{t-i}^2\varepsilon_{t-j}^2+o_P(1)$. In particular, the remainder contains terms involving products of elements of $\hat{\phi}-\phi$ (which are $o_P(1)$) with terms involving averages of cross products of elements of Y_{t-1} and ε_t, up to the eighth order, which are $O_P(1)$, given $E\left|\varepsilon_t\right|^8\leq\Delta<\infty$. This assumption also ensures that $n^{-1}\sum_{t=\max(i,j)+1}^{n}\varepsilon_t^4\varepsilon_{t-i}^2\varepsilon_{t-j}^2=O_P(1)$, by repeated application of the Markov and Cauchy-Schwartz inequalities. ■

Proof of Lemma A.3. Let $\mathcal{F}_t^*=\sigma\left(\eta_t,\eta_{t-1},\ldots,\eta_1\right)$, and define $Y_t^*=\left(\hat{\varepsilon}_t^*\hat{\varepsilon}_{t-1}^*,\ldots,\hat{\varepsilon}_t^*\hat{\varepsilon}_{t-m}^*\right)'$. Conditional on the original sample, we have $E^*\left(Y_t^*|\mathcal{F}_{t-1}^*\right)=E^*\left(\hat{\varepsilon}_t^*|\mathcal{F}_{t-1}^*\right)\left(\hat{\varepsilon}_{t-1}^*,\ldots,\hat{\varepsilon}_{t-m}^*\right)'=0$ since $E^*\left(\hat{\varepsilon}_t^*|\mathcal{F}_{t-1}^*\right)=E^*\left(\hat{\varepsilon}_t\eta_t|\mathcal{F}_{t-1}^*\right)=\hat{\varepsilon}_tE^*\left(\eta_t|\mathcal{F}_{t-1}^*\right)=0$, where $E^*\left(\eta_t|\mathcal{F}_{t-1}^*\right)=E^*\left(\eta_t\right)=0$, by the independence and mean zero properties of $\{\eta_t\}$. Thus, $\{Y_t^*,\mathcal{F}_t^*\}$ is a vector m.d.s. We now apply Theorem A.1 to $Z_t^*=\lambda'Y_t^*$ for arbitrary $\lambda\in\mathbb{R}^m$, $\lambda'\lambda=1$. First, note that $\bar{\sigma}_n^{*2}\equiv Var^*\left(n^{-1/2}\sum_{t=m+1}^{n}Z_t^*\right)=\lambda'n^{-1}\sum_{t=m+1}^{n}E^*\left(Y_t^*Y_t^{*\prime}\right)\lambda\equiv\lambda'\Omega_{n,m}^*\lambda$, where by direct evaluation and using the independence and zero properties of $\{\eta_t\}$,

$$\Omega^*_{n,m} = diag\left(n^{-1}\sum_{t=m+1}^{n}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-1}^2,\ldots,n^{-1}\sum_{t=m+1}^{n}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-m}^2\right).$$

Under our assumptions, we can show $n^{-1}\sum_{t=m+1}^{n}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-i}^2 \xrightarrow{P} \sigma^4\tau_{ii}$, $i = 1,\ldots,m$, which implies $\Omega^*_{n,m} \xrightarrow{P} \tilde{\Omega}_m \equiv \sigma^4 diag\left(\tau_{11},\ldots,\tau_{mm}\right)$. Thus, to verify the first condition of the CLT it suffices that

$$\lambda'\left[n^{-1}\sum_{t=m+1}^{n}Y_t^*Y_t^{*\prime} - \tilde{\Omega}_m\right]\lambda \xrightarrow{P^*} 0, \text{ in probability.}$$

A typical element (k,l) of the middle matrix is given by

$$V^*_{n,kl} \equiv n^{-1}\sum_{t=m+1}^{n}\hat{\varepsilon}_t^{*2}\hat{\varepsilon}^*_{t-k}\hat{\varepsilon}^*_{t-l} - \sigma^4\tau_{k,l}1\left(k=l\right),$$

where by Lemma A.2 (iii), under Assumption A strengthened by A$'$ (vi$'$), we have that $V^*_{n,kl} \xrightarrow{P^*} 0$ in probability. Lastly, condition 2. holds if for some $r > 1$, $n^{-r}\sum_{t=m+1}^{n}E^*\left|\lambda' Y_t^*\right|^{2r} = o_P(1)$. We take $r = 2$. By the c$_r$-inequality, we have

$$\begin{aligned} n^{-r}\sum_{t=m+1}^{n}E^*\left|\lambda' Y_t^*\right|^{2r} &= n^{-r}\sum_{t=m+1}^{n}E^*\left|\sum_{i=1}^{m}\lambda_i\hat{\varepsilon}_t^*\hat{\varepsilon}^*_{t-i}\right|^{2r} \leq m^{2r-1}\sum_{i=1}^{m}\left|\lambda_i\right|^{2r}n^{-r}\sum_{t=m+1}^{n}E^*\left|\hat{\varepsilon}_t^*\hat{\varepsilon}^*_{t-i}\right|^{2r} \\ &\leq n^{-(r-1)}m^{2r-1}\sum_{i=1}^{m}\left|\lambda_i\right|^{2r}n^{-1}\sum_{t=m+1}^{n}\left|\hat{\varepsilon}_t\hat{\varepsilon}_{t-i}\right|^{2r}E^*\left|\eta_t\right|^{2r}E^*\left|\eta_{t-i}\right|^{2r} = o_P(1), \end{aligned}$$

given in particular that $n^{-1}\sum_{t=m+1}^{n}\left|\hat{\varepsilon}_t\hat{\varepsilon}_{t-i}\right|^{2r} = O_P(1)$. ■

Proof of Lemma A.4. We can write $y_t^* = \sum_{j=0}^{t-1}\hat{\psi}_j\hat{\varepsilon}^*_{t-j}$, $t = 1,\ldots,n$, where $\left\{\hat{\psi}_j\right\}$ are defined by $\hat{\psi}_j = \sum_{i=1}^{\min(j,p)}\hat{\phi}_i\hat{\psi}_{j-1}$, with $\hat{\psi}_0 = 1$ and $\hat{\psi}_j = 0$ for $j < 0$. It follows that $Y^*_{t-1} = \sum_{j=1}^{t-1}\hat{b}_j\hat{\varepsilon}^*_{t-j}$, for $t = 2,\ldots,n$, where $\hat{b}_j = \left(\hat{\psi}_{j-1},\ldots,\hat{\psi}_{j-p}\right)'$. Note $Y_1^* = 0$, given the zero initial conditions. Hence,

$$\begin{aligned} n^{-1}\sum_{t=1}^{n}Y^*_{t-1}Y^{*\prime}_{t-1} &= T^*_{1n} + T^*_{2n}, \text{ with } T^*_{1n} = \sum_{j=1}^{n-1}\hat{b}_j\hat{b}_j'\left(n^{-1}\sum_{t=j+1}^{n}\hat{\varepsilon}^{*2}_{t-j}\right), \text{ and} \\ T^*_{2n} &= \sum_{k=1}^{n-2}\sum_{j=1}^{n-k-1}\left(\hat{b}_j\hat{b}_{j+k}' + \hat{b}_{j+k}\hat{b}_j'\right)\left(n^{-1}\sum_{t=1+k}^{n-j}\hat{\varepsilon}^*_{t-j}\hat{\varepsilon}^*_t\right). \end{aligned}$$

Next, we show: (a) $T^*_{1n} \xrightarrow{P^*} A \equiv \sigma^2\sum_{j=1}^{\infty}b_jb_j'$, and (b) $T^*_{2n} \xrightarrow{P^*} 0$, in probability. To prove (a), consider for fixed $m \in \mathbb{N}$,

$$T^*_{1n} = T^{*m}_{1n} + R^{*m}_{1n}, \text{ with } T^{*m}_{1n} = \sum_{j=1}^{m-1}\hat{b}_j\hat{b}_j'\left(n^{-1}\sum_{t=j+1}^{n}\hat{\varepsilon}^{*2}_{t-j}\right), \text{ and } R^{*m}_{1n} = \sum_{j=m}^{n-1}\hat{b}_j\hat{b}_j'\left(n^{-1}\sum_{t=j+1}^{n}\hat{\varepsilon}^{*2}_{t-j}\right).$$

By Lemma A.2.(i), for each $j = 1, \ldots, m$, m fixed, $n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^{*2} \xrightarrow{P^*} \sigma^2$, in probability; also, under Assumption A, $\hat{\psi}_j \xrightarrow{P} \psi_j$, implying $\hat{b}_j \xrightarrow{P} b_j$. Thus, by Slutsky's theorem, $T_{1n}^{*m} \xrightarrow{P^*} \sum_{j=1}^{m-1} b_j b_j' \sigma^2 \equiv A_m$, in probability. Since $\{\psi_j\}$ are absolutely summable, it follows that $A_m \to A$ as $m \to \infty$, by dominated convergence. Thus, $T_{1n}^{*m} \xrightarrow{P^*} A$, in probability. Choose $\lambda \in \mathbb{R}^p$ arbitrarily such that $\lambda'\lambda = 1$. By BD's Proposition 6.3.9, it now suffices to show that, for any $\delta > 0$, $\lim_{m\to\infty} \limsup_{n\to\infty} P^* \left(\left| \lambda' R_{1n}^{*m} \lambda \right| > \delta \right) = 0$, in probability, or $\lim_{m\to\infty} \limsup_{n\to\infty} E^* \left(\left| \lambda' R_{1n}^{*m} \lambda \right| \right) = 0$, in probability, by Markov's inequality. Using the triangle inequality and the properties of $\{\eta_t\}$, we get

$$E^* \left(\left| \lambda' R_{1n}^{*m} \lambda \right| \right) \leq \sum_{j=m}^{n-1} \left| \lambda' \hat{b}_j \hat{b}_j' \lambda \right| E^* \left(n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^{*2} \right) = \sum_{j=m}^{n-1} \left| \lambda' \hat{b}_j \hat{b}_j' \lambda \right| n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^{2}.$$

Given that $\hat{\varepsilon}_t = \varepsilon_t - \left(\hat{\phi}_n - \phi \right)' Y_{t-1}$, and that $\hat{\phi}_n - \phi \xrightarrow{P} 0$, we can show $n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^{2} = O_P(1)$. Thus, conditional on the sample, and for all n sufficiently large,

$$E^* \left(\left| \lambda' R_{1n}^{*m} \lambda \right| \right) \leq K \sum_{j=m}^{n-1} \left| \lambda' \hat{b}_j \hat{b}_j' \lambda \right| \leq K \sum_{k=1}^{p} \sum_{l=1}^{p} \left| \lambda_k \lambda_l \right| \sum_{j=m}^{\infty} \left| \hat{\psi}_{j-k} \hat{\psi}_{j-l} \right|$$

Under our assumptions, $\sum_{j=1}^{p} \left| \hat{\phi}_j - \phi_j \right| = o_P(1)$, so there exists n_1 such that $\sup_{n \geq n1} \sum_{j=1}^{\infty} \left| \hat{\psi}_j \right| < \infty$ in probability (cf. Bühlmann, 1995, Lemma 2.2.). This implies $\sup_{n \geq n_1} \sum_{j=m}^{\infty} \left| \hat{\psi}_{j-k} \hat{\psi}_{j-l} \right| = o_P(1)$ as $m \to \infty$, which completes the proof that $T_{1n}^{*} \xrightarrow{P^*} A$, in probability. Finally, to show (b), consider first for fixed $m \in \mathbb{N}$, $T_{2n}^{*m} = \sum_{k=1}^{m-2} \sum_{j=1}^{m-k-1} \hat{b}_j \hat{b}_{j+k}' \left(n^{-1} \sum_{t=1+k}^{n-j} \hat{\varepsilon}_{t-j}^{*} \hat{\varepsilon}_{t}^{*} \right)$. For fixed j and k, it follows by Lemma A.2.(ii) that $n^{-1} \sum_{t=1+k}^{n-j} \hat{\varepsilon}_{t-j}^{*} \hat{\varepsilon}_{t}^{*} \xrightarrow{P^*} 0$, in probability. Since $\hat{b}_j \hat{b}_{j+k}' \xrightarrow{P} b_j b_{j+k}$, we have that $T_{2n}^{*m} \xrightarrow{P^*} 0$, in probability. To complete the proof of (b) we need to show that each of the following

$$\begin{aligned} R_{2,1n}^{*m} &= \sum_{k=m-1}^{n-1} \sum_{j=1}^{n-k-1} \left(\hat{b}_j \hat{b}_{j+k}' + \hat{b}_{j+k} \hat{b}_j' \right) \left(n^{-1} \sum_{t=1+k}^{n-j} \hat{\varepsilon}_{t-j}^{*} \hat{\varepsilon}_{t}^{*} \right), \text{ and} \\ R_{2,2n}^{*m} &= \sum_{k=1}^{m-2} \sum_{j=m-k}^{n-k-1} \left(\hat{b}_j \hat{b}_{j+k}' + \hat{b}_{j+k} \hat{b}_j' \right) \left(n^{-1} \sum_{t=1+k}^{n-j} \hat{\varepsilon}_{t-j}^{*} \hat{\varepsilon}_{t}^{*} \right), \end{aligned}$$

satisfies the condition $\lim_{m\to\infty} \limsup_{n\to\infty} P^* \left(\left| \lambda' R_{2,in}^{*m} \lambda \right| > \delta \right) = 0$ in probability, for $i = 1, 2$, where λ and δ are as above. This can be verified analogously to above, using in particular the fact that $\sum_{k=1}^{\infty} k \left| \psi_k \right| < \infty$. ■

Proof of Lemma A.5. As in the proof of Lemma A.4, we have $Y_{t-1}^* = \sum_{j=1}^{t-1} \hat{b}_j \hat{\varepsilon}_{t-j}^*$, where $\hat{b}_j = \left(\hat{\psi}_{j-1}, \ldots, \hat{\psi}_{j-p} \right)'$, with $\hat{\psi}_0 = 1$ and $\hat{\psi}_j = 0$ for $j < 0$. Let $Z_t^* = Y_{t-1}^* \hat{\varepsilon}_t^* = \sum_{j=1}^{t-1} \hat{b}_j \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^*$, for

$t = 2, \ldots, n$, and note that $Z_1^* = 0$. Thus,

$$n^{-1/2} \sum_{t=1}^{n} Z_t^* = n^{-1/2} \sum_{t=2}^{n} \sum_{j=1}^{t-1} \hat{b}_j \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^* = \sum_{j=1}^{n-1} \hat{b}_j n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^* \equiv \mathcal{X}_n^*.$$

For fixed $m \in \mathbb{N}$, let $\mathcal{X}_{n,m}^* \equiv \sum_{j=1}^{m-1} \hat{b}_j n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^*$. Next we show: (a) for m fixed, $\mathcal{X}_{n,m}^* \Rightarrow^{d_{P^*}} N\left(0, \tilde{B}_m\right)$, as $n \to \infty$, where $\tilde{B}_m = \sum_{j=1}^{m} b_j b_j' \sigma^4 \tau_{jj}$; (b) $\tilde{B}_m \to \tilde{B}$ as $m \to \infty$, and (c) $\lim_{m\to\infty} \limsup_{n\to\infty} P^*\left(\left|\mathcal{X}_n^* - \mathcal{X}_{n,m}^*\right| > \kappa\right) = 0$ for any $\kappa > 0$. For (a), write

$$\mathcal{X}_{n,m}^* = \sum_{j=1}^{m-1} b_j n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^* + \sum_{j=1}^{m-1} \left(\hat{b}_j - b_j\right) n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^* \equiv Q_{1n}^* + Q_{2n}^*.$$

By Lemma A.3, under Assumption A strengthened by A(vi′), we have that for fixed j $n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^*$ $\Rightarrow^{d_{P^*}} N\left(0, \sigma^4 \tau_{jj}\right)$, where the elements of the $m \times 1$ vector composed of these sums are asymptotically uncorrelated. Hence, $Q_{1n}^* \Rightarrow^{d_{P^*}} N\left(0, \tilde{B}_m\right)$, in probability, where $\tilde{B}_m = \sum_{j=1}^{m} b_j b_j' \sigma^4 \tau_{jj}$. Next, note $Q_{2n}^* \xrightarrow{P^*} 0$ in probability, since $\hat{b}_j - b_j \xrightarrow{P} 0$ and $n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^* = O_{P^*}(1)$ for each $j = 1, \ldots, m$. The asymptotic equivalence lemma now implies (a). (b) follows by dominated convergence given the summability of $\left\{\psi_j\right\}$ and the uniform boundedness of $\sigma^4 \tau_{jj}$. To prove (c), note it suffices that $\lim_{m\to\infty} \limsup_{n\to\infty} E^*\left(\left|\mathcal{X}_n^* - \mathcal{X}_{n,m}^*\right|^2\right) = 0$, by Chebyshev's inequality. Equivalently, we consider for any $\lambda \in \mathbb{R}^p$, such that $\lambda'\lambda = 1$,

$$E^*\left(\left|\lambda'\left(\mathcal{X}_n^* - \mathcal{X}_{n,m}^*\right)\right|^2\right) = E^*\left(\sum_{j=m}^{n-1} \sum_{i=m}^{n-1} \lambda' \hat{b}_j \hat{b}_i' \lambda Z_{nj}^* Z_{ni}^*\right),$$

where $Z_{nj}^* \equiv n^{-1/2} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^* \hat{\varepsilon}_t^*$. Since $E^*\left(Z_{nj}^* Z_{ni}^*\right) = 0$ for $i \neq j$ and $E^*\left(Z_{nj}^{*2}\right) = n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^2 \hat{\varepsilon}_t^2$, it follows that

$$E^*\left(\left|\lambda'\left(\mathcal{X}_n^* - \mathcal{X}_{n,m}^*\right)\right|^2\right) = \sum_{j=m}^{n-1} \lambda' \hat{b}_j \hat{b}_j' \lambda \left(n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^2 \hat{\varepsilon}_t^2\right).$$

Using the definition of $\hat{\varepsilon}_t$, i.e. $\hat{\varepsilon}_t = \varepsilon_t \eta_t - \left(\hat{\phi}_n - \phi\right)' Y_{t-1}$, and the fact that $\hat{\phi}_n - \phi \xrightarrow{P} 0$, we can show $n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^2 \hat{\varepsilon}_t^2 = n^{-1} \sum_{t=j+1}^{n} \varepsilon_{t-j}^2 \varepsilon_t^2 + o_P(1)$. This implies $n^{-1} \sum_{t=j+1}^{n} \hat{\varepsilon}_{t-j}^2 \hat{\varepsilon}_t^2 = O_P(1)$, given that $n^{-1} \sum_{t=j+1}^{n} \varepsilon_{t-j}^2 \varepsilon_t^2 \xrightarrow{P} \sigma^4 \tau_{jj}$, and $\sigma^4 \tau_{jj}$ are uniformly bounded by assumption. The proof of (c) now follows exactly the argument used in Lemma A.4 when dealing with R_{1n}^{*m}. ■

References

[1] Andrews, D.W.K. (1993): "Exactly Median-Unbiased Estimation of First-Order Autoregressive/Unit Root Models", *Econometrica*, 61, 139-166.

[2] Andrews, D.W.K. and H.Y. Chen (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models", *Journal of Business and Economic Statistics*, 12, 187-204.

[3] Bekaert, G., and R.J. Hodrick (2001): "Expectations Hypothesis Tests", *Journal of Finance*, 56, 1357-1394.

[4] Bekaert, G., R.J. Hodrick and D.A. Marshall (1997): "On Biases in Tests of the Expectations Hypothesis in the Term Structure of Interest Rates", *Journal of Financial Economics*, 44, 309-348.

[5] Berkowitz, J, I Birgean and L. Kilian (2000): "On the Finite-Sample Accuracy of Nonparametric Resampling Algorithms for Economic Time Series", in T.B. Fomby and C. Hill (eds.): *Advances in Econometrics: Applying Kernel and Nonparametric Estimation to Economic Topics*, vol. 14, JAI Press, Connecticut, 77-107.

[6] Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 37, 307-327.

[7] Bollerslev, T. (1990): "Modeling the Coherence on Short-Run Nominal Exchange Rates: A Multivariate GARCH Model", *Review of Economics and Statistics*, 72, 498-505.

[8] Bollerslev, T., R.F. Engle and J. M. Wooldridge (1988): "A Capital-Asset Pricing Model with Time-Varying Covariances", *Journal of Political Economy*, 96, 116-131.

[9] Bose, A. (1988): "Edgeworth Correction by Bootstrap in Autoregression", *Annals of Statistics*, 16, 1709-1722.

[10] Braun, P.A., D.B. Nelson and A.M. Sunier (1995): "Good News, Bad News, Volatility and Betas", *Journal of Finance*, 50, 1575-1603.

[11] Chesher, A., and I. Jewitt (1987): "The Bias of a Heteroskedasticity Consistent Covariance Matrix Estimator", *Econometrica*, 55, 1217-1222.

[12] Davidson, J. (1994): *Stochastic Limit Theory*, Oxford University Press, New York.

[13] Davidson, R. and E. Flachaire (2000): "The Wild Bootstrap, Tamed at Last", mimeo, GREQAM.

[14] Deo, S.R. (2000): "Spectral Tests of the Martingale Hypothesis under Conditional Heteroskedasticity", *Journal of Econometrics*, 99, 291-315.

[15] Doan, T.A. (1990): "RATS User's Manual", VAR Econometrics, Evanston, IL.

[16] Durrett, R. (1996): *Probability: Theory and Examples*, Duxury Press, California.

[17] Eichenbaum, M. and C.I. Evans (1995): "Some Empirical Evidence of the Effects of Shocks to Monetary Policy on Exchange Rates", *Quarterly Journal of Economics*, 110, 975-1010.

[18] Eicker, F. (1963): "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions", *Annals of Mathematical Statistics*, 34, 447-456.

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

[20] Engle, R.F. (1990): "Discussion: Stock Market Volatility and the Crash of '87", *Review of Financial Studies*, 3, 103-106.

[21] Engle, R.F., and V.K. Ng (1993): "Measuring and Testing the Impact of News on Volatility", *Journal of Finance*, 48, 1749-1778.

[22] Fama, E., and K. French (1988): "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, 22, 3-25.

[23] Freedman, D.A. (1981): "Bootstrapping Regression Models", *Annals of Statistics*, 9, 1218-1228.

[24] Glosten, L.R., R. Jaganathan, and D.E. Runkle (1993): "On the Relation between the Expected Value and the Volatility of Nominal Excess Returns on Stocks", *Journal of Finance*, 48, 1779-1801.

[25] Godfrey, L.G. and C.D. Orme (2001): "Significance Levels of Heteroskedasticity-Robust Tests for Specification and Misspecification: Some Results on the Use of the Wild Bootstraps, mimeo, University of York.

[26] Goetzmann, W.N., and P. Jorion (1993): "Testing the Predictive Power of Dividend Yields", *Journal of Finance*, 48, 663-679.

[27] Goetzmann, W.N. and P. Jorion (1995): "A Longer Look at Dividend Yields", *Journal of Business*, 68, 483-508.

[28] Hansen, B.E. (1999): "The Grid Bootstrap and the Autoregressive Model", *Review of Economics and Statistics*, 81, 594-607.

[29] Härdle, W., J. Horowitz, and J.-P. Kreiss (2001): "Bootstrap Methods for Time Series", mimeo, Institute for Statistics and Econometrics, Humboldt-Universität zu Berlin.

[30] He, C. and T. Teräsvirta (1999): "Properties of moments of a family of GARCH processes", *Journal of Econometrics*, 92, 173-192.

[31] Hodrick, R.J. (1992): "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement", *Review of Financial Studies*, 5, 357-386.

[32] Inoue, A. and L. Kilian (2002): "Bootstrapping Autoregressions with Possible Unit Roots", *Econometrica,* 70, 377-391.

[33] Kilian, L. (1999): "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?", *Journal of Applied Econometrics*, 14, 491-510.

[34] Kreiss, J.P. (1997): "Asymptotic Properties of Residual Bootstrap for Autoregressions", mimeo, Institute for Mathematical Stochastics, Technical University of Braunschweig, Germany.

[35] Kuersteiner, G.M. (2001): "Optimal Instrumental Variables Estimation for ARMA Models", *Journal of Econometrics*, 104, 359-405.

[36] Lamoureux, C.G., and W.D. Lastrapes (1990): "Persistence in Variance, Structural Change, and the GARCH Model", *Journal of Business and Economics Statistics*, 8, 225-234.

[37] Ledoit, O., P. Santa-Clara and M. Wolf (2001): "Multivariate GARCH Modeling with an Application to International Stock Markets," Working Paper No. 578, Department of Economics, Universitat Pompeu Fabra.

[38] Liu, R.Y. (1988): "Bootstrap Procedure under Some Non-i.i.d. Models", *Annals of Statistics*, 16, 1696-1708.

[39] Lütkepohl, H. (1990): "Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models", *Review of Economics and Statistics*; 72, 116-125.

[40] MacKinnon, J.G., and H. White (1985): "Some Heteroskedasticity Consistent Covariance Matrix Estimators with Improved Finite-Sample Properties", *Journal of Econometrics*, 29, 305-325.

[41] Mammen, E. (1993): "Bootstrap and Wild Bootstrap for High Dimensional Linear Models", *Annals of Statistics*, 21, 255-285.

[42] Mark, N.C. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability", *American Economic Review*, 85, 201-218.

[43] Meese, R. and K. Rogoff (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?", *Journal of International Economics*, 14, 3-24.

[44] Milhøj, A. (1985): "The moment structure of ARCH processes", *Scandinavian Journal of Statistics*, 12, 281-292.

[45] Nicholls, D.F., and A.R. Pagan. (1983): "Heteroskedasticity in Models with Lagged Dependent Variables", *Econometrica*, 51, 1233-1242.

[46] Shephard, N. (1996): "Statistical Aspects of ARCH and Stochastic Volatility" in Cox, D.R., et al (eds), *Time Series Models in Econometrics, Finance and Other Fields*, Chapman & Hall, London.

[47] Sims, C.A. and T. Zha (1999): "Error Bands for Impulse Responses", *Econometrica*, 67, 1113-1155.

[48] White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, 817-838.

[49] White, H. (1999): *Asymptotic Theory for Econometricians*, Second Edition, Academic Press, London.

[50] Wright, J.H. (2001): "Exact Confidence Intervals for Impulse Responses in a Gaussian Vector Autoregression", Board of Governors, International Finance Discussion Papers 682.

[51] Wright, J.H. (2000): "Confidence Intervals for Univariate Impulse Responses with a Near Unit Root", *Journal of Business and Economic Statistics*, 18, 368-373.

[52] Wu, C.F.J. (1986): "Jackknife Bootstrap and other Resampling Methods in Regression Analysis", *Annals of Statistics*, 14, 1261-1295.

[53] Wolf, M. (2000): "Stock Returns and Dividend Yields Revisited: A New Way to Look at an Old Problem", *Journal of Business and Economic Statistics*, 18, 18-30.

The following papers have been published since 2001:

January	2001	Unemployment, Factor Substitution, and Capital Formation	Leo Kaas Leopold von Thadden
January	2001	Should the Individual Voting Records of Central Banks be Published?	Hans Gersbach Volker Hahn
January	2001	Voting Transparency and Conflicting Interests in Central Bank Councils	Hans Gersbach Volker Hahn
January	2001	Optimal Degrees of Transparency in Monetary Policymaking	Henrik Jensen
January	2001	Are Contemporary Central Banks Transparent about Economic Models and Objectives and What Difference Does it Make?	Alex Cukierman
February	2001	What can we learn about monetary policy transparency from financial market data?	Andrew Clare Roger Courtenay
March	2001	Budgetary Policy and Unemployment Dynamics	Leo Kaas Leopold von Thadden
March	2001	Investment Behaviour of German Equity Fund Managers – An Exploratory Analysis of Survey Data	Torsten Arnswald
April	2001	The information content of survey data on expected price developments for monetary policy	Christina Gerberding
May	2001	Exchange rate pass-through and real exchange rate in EU candidate countries	Zsolt Darvas

July	2001	Interbank lending and monetary policy Transmission: evidence for Germany	Michael Ehrmann Andreas Worms
September	2001	Precommitment, Transparency and Montetary Policy	Petra Geraats
September	2001	Ein disaggregierter Ansatz zur Berechnung konjunkturbereinigter Budgetsalden für Deutschland: Methoden und Ergebnisse *	Matthias Mohr
September	2001	Long-Run Links Among Money, Prices, and Output: World-Wide Evidence	Helmut Herwartz Hans-Eggert Reimers
November	2001	Currency Portfolios and Currency Exchange in a Search Economy	Ben Craig Christopher J. Waller
December	2001	The Financial System in the Czech Republic, Hungary and Poland after a Decade of Transition	Thomas Reininger Franz Schardax Martin Summer
December	2001	Monetary policy effects on bank loans in Germany: A panel-econometric analysis	Andreas Worms
December	2001	Financial systems and the role of banks in monetary policy transmission in the euro area	M. Ehrmann, L. Gambacorta J. Martinez-Pages P. Sevestre, A. Worms
December	2001	Monetary Transmission in Germany: New Perspectives on Financial Constraints and Investment Spending	Ulf von Kalckreuth
December	2001	Firm Investment and Monetary Transmission in the Euro Area	J.-B. Chatelain, A. Generale, I. Hernando, U. von Kalckreuth P. Vermeulen

* Available in German only.

January	2002	Rent indices for housing in West Germany 1985 to 1998	Johannes Hoffmann Claudia Kurz
January	2002	Short-Term Capital, Economic Transformation, and EU Accession	Claudia M. Buch Lusine Lusinyan
January	2002	Fiscal Foundation of Convergence to European Union in Pre-Accession Transition Countries	László Halpern Judit Neményi
January	2002	Testing for Competition Among German Banks	Hannah S. Hempell
January	2002	The stable long-run CAPM and the cross-section of expected returns	Jeong-Ryeol Kim
February	2002	Pitfalls in the European Enlargement Process – Financial Instability and Real Divergence	Helmut Wagner
February	2002	The Empirical Performance of Option Based Densities of Foreign Exchange	Ben R. Craig Joachim G. Keller
February	2002	Evaluating Density Forecasts with an Application to Stock Market Returns	Gabriela de Raaij Burkhard Raunig
February	2002	Estimating Bilateral Exposures in the German Interbank Market: Is there a Danger of Contagion?	Christian Upper Andreas Worms
February	2002	The long-term sustainability of public finance in Germany – an analysis based on generational accounting	Bernhard Manzke
March	2002	The pass-through from market interest rates to bank lending rates in Germany	Mark A. Weth

April	2002	Dependencies between European stock markets when price changes are unusually large	Sebastian T. Schich
May	2002	Analysing Divisia Aggregates for the Euro Area	Hans-Eggert Reimers
May	2002	Price rigidity, the mark-up and the dynamics of the current account	Giovanni Lombardo
June	2002	An Examination of the Relationship Between Firm Size, Growth, and Liquidity in the Neuer Markt	Julie Ann Elston
June	2002	Monetary Transmission in the New Economy: Accelerated Depreciation, Transmission Channels and the Speed of Adjustment	Ulf von Kalckreuth Jürgen Schröder
June	2002	Central Bank Intervention and Exchange Rate Expectations – Evidence from the Daily DM/US-Dollar Exchange Rate	Stefan Reitz
June	2002	Monetary indicators and policy rules in the P-star model	Karl-Heinz Tödter
July	2002	Real currency appreciation in accession countries: Balassa-Samuelson and investment demand	Christoph Fischer
August	2002	The Eurosystem's Standing Facilities in a General Equilibrium Model of the European Interbank Market	Jens Tapking
August	2002	Imperfect Competition, Monetary Policy and Welfare in a Currency Area	Giovanni Lombardo

August 2002	Monetary and fiscal policy rules in a model with capital accumulation and potentially non-superneutral money	Leopold von Thadden
September 2002	Dynamic Q-investment functions for Germany using panel balance sheet data and a new algorithm for the capital stock at replacement values	Andreas Behr Egon Bellgardt
October 2002	Tail Wags Dog? Time-Varying Information Shares in the Bund Market	Christian Upper Thomas Werner
October 2002	Time Variation in the Tail Behaviour of Bund Futures Returns	Thomas Werner Christian Upper
November 2002	Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form	Sílvia Gonçalves Lutz Kilian

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt
GERMANY