Autocontours: Dynamic Specification Testing

Gloria GONZÁLEZ-RIVERA

Department of Economics, University of California, Riverside, Riverside, CA 92521 (*gloria.gonzalez@ucr.edu*)

Zeynep SENYUZ

Department of Economics, University of New Hampshire, Durham, NH 03824 (*zeynep.senyuz@unh.edu*)

Emre YOLDAS

Department of Economics, Bentley University, Waltham, MA 02452 (*eyoldas@bentley.edu*)

We propose a new battery of dynamic specification tests for the joint hypothesis of iid-ness and density function based on the fundamental properties of independent random variables with identical distributions. We introduce a device—the autocontour—whose shape is very sensitive to departures from the null in either direction, thus providing superior power. The tests are parametric with asymptotic t and chi-squared limiting distributions and standard convergence rates. They do not require a transformation of the original data or a Kolmogorov style assessment of goodness-of-fit, explicitly account for parameter uncertainty, and have superior finite sample properties. An application to autoregressive conditional duration (ACD) models for trade durations shows that the difficulty with the assumed densities lies on the probability assigned to very small durations. Supplemental materials for this article are available online.

KEY WORDS: Autoregressive conditional duration model; Bootstrap; Parameter uncertainty; Probability contour plot.

1. INTRODUCTION

We propose a new battery of tests for dynamic specification that rely on the fundamental properties of independent random variables with identical distributions. We focus on models in which all the dependence is contained in the first and second moments such that for a process $\{y_t\}$ we have $y_t = \mu_t(\boldsymbol{\theta}_{01}, \Im_{t-1}) + \sigma_t(\boldsymbol{\theta}_{02}, \Im_{t-1})\varepsilon_t$ where $\mu_t(\cdot)$ is the conditional mean and $\sigma_t^2(\cdot)$ is the conditional variance, both functions of an information set $\Im_{t-1}$, $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}'_{01}, \boldsymbol{\theta}'_{02})'$ is a parameter vector, and ε_t is an innovation that is iid. The innovation ε_t is characterized by a parametric pdf, say $f(\varepsilon_t)$. In this context, we formulate a collection of statistics for the joint hypothesis of iid-ness and density functional form of the innovation as specification tests of the full model.

There is an extensive literature on testing a *single* hypothesis, either density functional form or iid behavior. There are numerous parametric and nonparametric tests for density functional form, most of which assume iid observations as a starting point as in Andrews (1997), or independence as in Fernandes and Grammig (2005). On the other hand, there are several diagnostic checks that do not explicitly consider the density functional form. Brock et al. (1996), Hong and Lee (2003), and Chen (2008), among others, test for iid-ness, Escanciano (2008) focused on the specification of conditional mean and variance, and Meitz and Terasvirta (2006) provided Lagrange multiplier (LM) tests of dependence in the context of duration models. Testing the *joint* hypothesis of iid-ness and density functional form arise concurrently with the interest in density forecast evaluation methods based on Rosenblatt's probability integral transform. If the proposed density forecast, say $f_t(\cdot)$, is correct, then the transformed random variables, $u_t = F_t(y_t | \Im_{t-1}; \boldsymbol{\theta}_0)$, are iid uniformly distributed, see Diebold, Gunther, and Tay (1998), Berkowitz (2001), Chen and Fan (2004), who tested the joint hypothesis of iid-ness and uniformity without accounting for parameter uncertainty. Bai (2003) dealt with parameter uncertainty, but his test lacks power against violations of

independence (see Corradi and Swanson 2006). Hong and Li (2005) proposed a nonparametric-kernel-based test statistic that has power against violations of both independence and density functional form. Their procedure is asymptotically distribution free, but it is dependent on the choice of the bandwidth.

In this article we propose a battery of tests for the joint hypothesis of iid-ness and density functional form that are very powerful against violations of both. The proposed tests focus on fundamental properties of independent random variables with identical distributions. Let the process under consideration be $\{\varepsilon_t\}$ with density $f(\cdot)$. The random variables in this process are independent if and only if their multivariate distribution is equal to the product of their marginal distributions, in which case the null hypothesis simply boils down to $f(\varepsilon_{t-k_1}, \varepsilon_{t-k_2}, \ldots, \varepsilon_{t-k_m}) = f(\varepsilon_{t-k_1})f(\varepsilon_{t-k_2})\cdots f(\varepsilon_{t-k_m})$, $\{k_j\}_{j=1}^{m} \in \mathbb{N}$. The specification tests we propose are based on a new concept that we term *autocontour*. Under the null, we horizontally slice the joint density at different levels and project the resulting segments down to the hyperplane $(\varepsilon_{t-k_1}, \varepsilon_{t-k_2}, \ldots, \varepsilon_{t-k_m})$. The projection is the autocontour containing a known percentage of the observations. Based on the sample estimates of these percentages we construct a battery of t-statistics and chi-squared statistics, which have standard asymptotic distributions. The shape of the autocontour will change whenever there is a departure from the null in any direction, and by doing so it will provide superior power to the tests. Our tests can be applied to primitive series and model residuals, in which case we need to address the parameter uncertainty problem. We show that a general bootstrap procedure to obtain asymptotic variance matrices delivers standard asymptotic tests with very good finite sample

performance. In comparison with the existent tests, the autocontour tests have several advantages as they are parametric with standard convergence rates and standard limiting distributions that deliver superior power. They are computationally easy to implement as they are based on counting processes. In addition, they do not require either a transformation of the original data or an assessment of Kolmogorov goodness-of-fit, and explicitly account for parameter uncertainty.

The structure of the article is as follows. In Section 2, we formalize the notion of autocontour and present the testing methodology. In Section 3 we explicitly deal with the parameter uncertainty problem. In Section 4, we provide Monte Carlo evidence on the size and power properties of our tests and offer a comparison with the nonparametric tests of Hong and Li (2005). Section 5 contains an empirical application in the context of autoregressive conditional duration (ACD) models. We conclude in Section 6. All mathematical proofs can be found in Appendix B.

2. THE JOINT TEST OF DENSITY AND INDEPENDENCE

The class of dynamic models that we are interested in are of the following form:

$$y_t = \mu_t(\boldsymbol{\theta}_{01}, \mathfrak{I}_{t-1}) + \sigma_t(\boldsymbol{\theta}_{02}, \mathfrak{I}_{t-1})\varepsilon_t, \qquad t = 1, \ldots, T, \quad (2.1)$$

where $\mathfrak{I}_{t-1}$ denotes the information set available at time $t-1$, $\mu_t(\cdot)$ and $\sigma_t(\cdot)$ are fully parameterized by $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_{01}', \boldsymbol{\theta}_{02}')'$ and measurable with respect to $\mathfrak{I}_{t-1}$, and $\{\varepsilon\}_{t=1}^T$ is a series of iid innovations having a particular density function $f(\cdot)$. Usually, ε_t is assumed to have zero mean and unit variance, but for nonnegative data it will naturally have a nonzero mean. We assume first that ε_t is observable, i.e., $\boldsymbol{\theta}_0$ is known. Later on we will relax this assumption to account for the effects of estimation on the distribution of our test statistics.

2.1 Autocontour

Under correct dynamic specification the null hypothesis in its most general form is stated as, $H_0 : \varepsilon_t$ is iid with density $f(\cdot)$, where H_1 is the negation of H_0. Under this null hypothesis the multivariate density function for an m-dimensional vector $(\varepsilon_{t-k_1}, \varepsilon_{t-k_2}, \ldots, \varepsilon_{t-k_m})$ is written as $f(\varepsilon_{t-k_1}, \varepsilon_{t-k_2}, \ldots, \varepsilon_{t-k_m}) = f(\varepsilon_{t-k_1})f(\varepsilon_{t-k_2})\cdots f(\varepsilon_{t-k_m})$. We define the (α, m) autocontour, ACR_α^m, as the set of points in the hyperplane $(\varepsilon_{t-k_1}, \varepsilon_{t-k_2}, \ldots, \varepsilon_{t-k_m})$ that results from horizontally slicing the multivariate density function at a fixed value, say $\bar{f}_\alpha$, to guarantee that the resulting set contains $\alpha\%$ of observations, that is,

$$ACR_\alpha^m := \left\{ B\big(\varepsilon_{t-k_1}, \ldots, \varepsilon_{t-k_m}\big) \subset \mathbb{R}^m \bigg| \right.$$

$$\int_{l_1}^{u_1} \cdots \int_{l_m}^{u_m} f\big(\varepsilon_{t-k_1}\big) \cdots$$

$$\left. \times f\big(\varepsilon_{t-k_m}\big) \, d\varepsilon_{t-k_1} \cdots d\varepsilon_{t-k_m} \le \alpha \right\}, \quad (2.2)$$

where B is a set in $\mathbb{R}^m$, $u_m = u_m(\varepsilon_{t-k_1}, \ldots, \varepsilon_{t-k_{m-1}})$, $l_m = l_m(\varepsilon_{t-k_1}, \ldots, \varepsilon_{t-k_{m-1}})$, and $\{k_j\}_{j=1}^m \in \mathbb{N}$.

Our testing methodology focuses on the bivariate autocontour because, as we will show in the forthcoming sections, the tests enjoy superior power in the most empirically relevant cases. Monte Carlo simulations illustrate the consistency of the tests against a wide set of alternatives. Tests based on higher-dimensional autocontours can deliver more power, but at the cost of higher-dimensional integration problems in the construction of the autocontour and loss of the graphical representation that is useful in choosing the modeling strategy. Thus, bivariate autocontours allow the combination of formal statistical testing, graphical techniques, and easy implementation, without sacrificing desirable properties of the tests.

The bivariate autocontour is given by

$$ACR_{\alpha,k} := \left\{ B(\varepsilon_t, \varepsilon_{t-k}) \subset \mathbb{R}^2 \bigg| \right.$$

$$\left. \int_{l_0}^{u_0} \int_{l_k(\varepsilon_t)}^{u_k(\varepsilon_t)} f(\varepsilon_t)f(\varepsilon_{t-k}) \, d\varepsilon_t \, d\varepsilon_{t-k} \le \alpha \right\}, \quad (2.3)$$

where B is a set on the plane $\mathbb{R}^2$ and the limits of integration are such that the contour shape of the hypothesized density is preserved. Therefore, we are interested in testing the following null hypothesis:

$$H_0 : f(\varepsilon_t, \varepsilon_{t-k}) = f(\varepsilon_t)f(\varepsilon_{t-k}) \quad \text{for } k = 1, \ldots, K < \infty. \quad (2.4)$$

Note that the null should hold as $k \to \infty$. However, in practice k is limited by the sample size. Thus, we choose K sufficiently large so that we can handle relevant dependence structures while keeping the theoretical framework tractable. In the following, we present the equations representing the boundaries of the autocontours, i.e., when the integral in Equation (2.3) holds with equality, corresponding to the most commonly encountered densities in financial econometrics (see Appendix A for details on construction of the autocontour equations for these densities):

(i) *Standard Normal distribution*: $\varepsilon_t^2 + \varepsilon_{t-k}^2 = a_\alpha$,

(ii) *Student-t distribution*: $1 + (\varepsilon_t^2 + \varepsilon_{t-k}^2)/\nu + (\varepsilon_t^2 \varepsilon_{t-k}^2 / \nu^2) = a_\alpha$,

(iii) *Exponential distribution*: $\varepsilon_t + \varepsilon_{t-k} = a_\alpha$,

(iv) *Weibull distribution*: $(1-\kappa)[\ln(\varepsilon_t) + \ln(\varepsilon_{t-k})] + c^\kappa(\varepsilon_t^\kappa + \varepsilon_{t-k}^\kappa) = a_\alpha$.

In Figure 1, we show the graphical contours corresponding to these cases.

2.2 Test Statistics and Asymptotic Distributions

For a given autocontour $ACR_{\alpha_i,k}$, we define a binary variable as follows

$$I_t^{k,i} = \mathbb{1}\big((\varepsilon_t, \varepsilon_{t-k}) \notin ACR_{\alpha_i,k}\big), \qquad t = k+1, \ldots, T, \quad (2.5)$$

where $\mathbb{1}(\cdot)$ is the usual indicator function, $k = 1, \ldots, K$ and $i = 1, \ldots, C$, i.e., K is the number of lags and C is the number of autocontours. Hence, this Bernoulli random variable

(a) Standard Normal

(b) Student-t ($\nu = 5$)

(c) Exponential ($\beta = 1$)

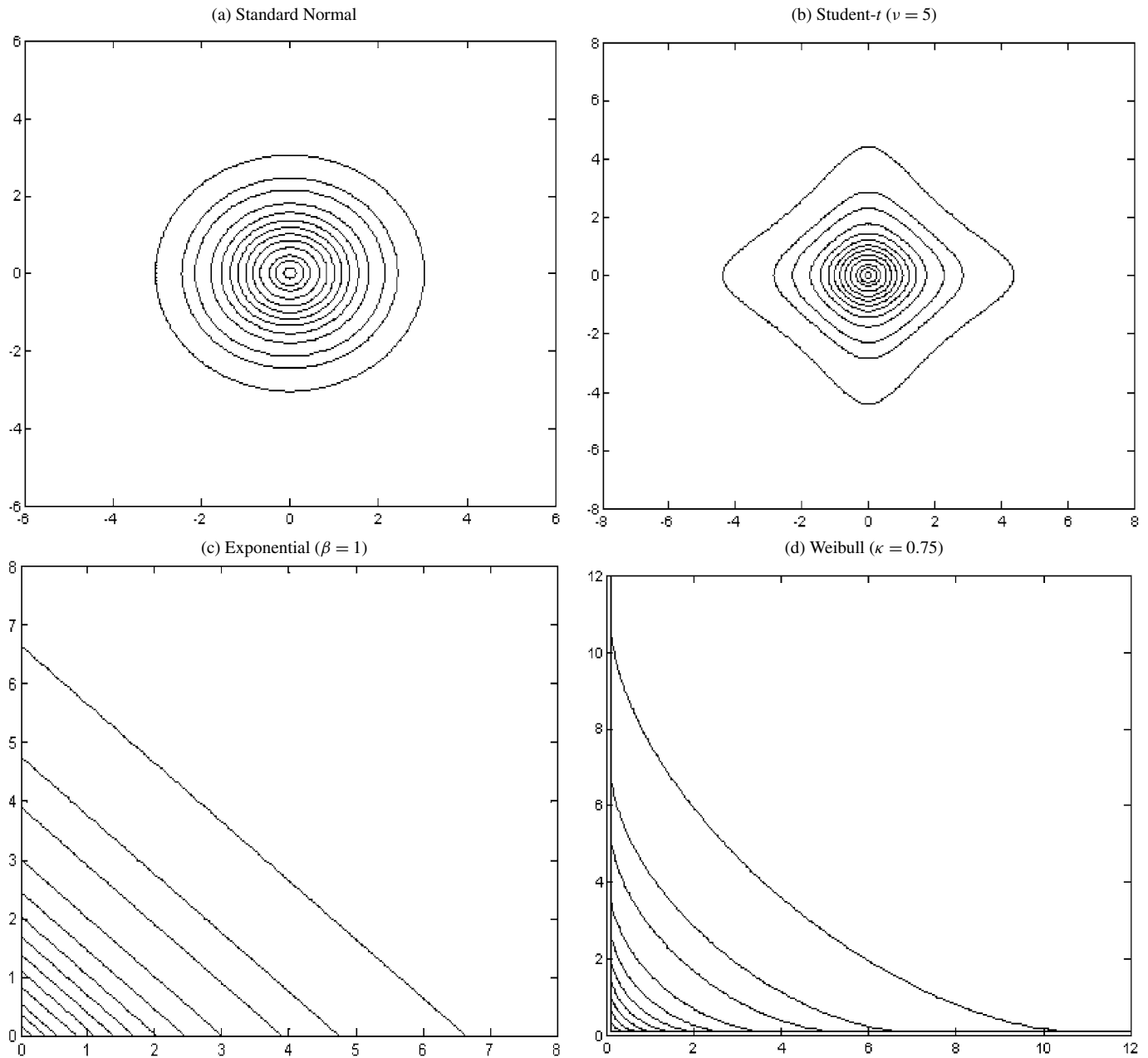(d) Weibull ($\kappa = 0.75$)

Figure 1. Sample autocontours of bivariate distributions under independence. *Notes*: Autocontours are presented for the following coverage levels (%): 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, and 99.

takes on value 1 if an observation falls outside the auto-contour and 0 otherwise. Note that this indicator can also be constructed the other way around, i.e., taking on value 1 if an observation falls inside the autocontour and 0 otherwise, producing the same set of results in a symmetric fashion. Since $ACR_{\alpha_i,k}$ contains $\alpha_i\%$ of observations, we expect to have $(1 - \alpha_i)\%$ outside the autocontour. Let $p_i \equiv 1 - \alpha_i$. Under the null we have $E[I_t^{k,i}] = p_i$ and $\text{Var}(I_t^{k,i}) = p_i(1 - p_i)$. Furthermore, the indicator is a linearly dependent process with a moving average (MA) structure. Its autocovariance function is given by $\gamma_h^i = \mathbb{1}(h = k)[P(I_t^{k,i} = 1, I_{t-h}^{k,i} = 1) - p_i^2]$, i.e., there is dependence only at the kth lag. Our first test statistic is a t-statistic based on this indicator series.

Proposition 1. Let $\hat{p}_i^k = \frac{1}{T-k}\sum_{t=1}^{T-k} I_t^{k,i}$. Under the null hypothesis given in Equation (2.4),

$$t_{k,i} = \frac{\sqrt{T-k}(\hat{p}_i^k - p_i)}{\sigma_{k,i}} \xrightarrow{d} \text{N}(0, 1),$$

where $\sigma_{k,i}^2 = p_i(1 - p_i) + 2\gamma_k^i$.

Note that p_i is given under the null and a consistent estimator of the autocovariance term, γ_k^i, can be easily obtained from data. For a given autocontour i, we can examine the lag structure of $t_{k,i}$ for $k = 1, \ldots, K$ and collect those t-statistics in a graph, which we call *autocontourgram*. See Section 5 for various empirical examples.

In the spirit of Box–Pierce–Ljung statistics, the information contained in the individual $t_{k,i}$ statistics can be pooled either

across K lags or across C contours. The following two test statistics consider the joint distribution of t-ratios associated with different lags or contours.

Proposition 2. For a given autocontour i, consider all lags up to K. Let $q_{k,i} = \sqrt{T-k}(\hat{p}_i^k - p_i)$, $k = 1, \ldots, K$, and stack them in a vector $\mathbf{q}_i = (q_{1,i}, \ldots, q_{K,i})'$. Under the null we have $\mathbf{q}_i \xrightarrow{d} N(0, \boldsymbol{\Omega}_i)$ where any element $\omega_{i,kl}$ in $\boldsymbol{\Omega}_i$ is given by

$$
\omega_{i,kl} = \begin{cases} \mathrm{Cov}(I_t^{k,i}, I_t^{l,i}) + \mathrm{Cov}(I_t^{k,i}, I_{t-k}^{l,i}) \\ \quad + \mathrm{Cov}(I_t^{l,i}, I_{t-l}^{k,i}) + \mathrm{Cov}(I_t^{l,i}, I_{t-l+k}^{k,i}) & \text{for } l > k, \\ p_i(1 - p_i) + 2\gamma_k^i & \text{for } l = k. \end{cases}
$$

It directly follows that

$$
Q_i^K = \mathbf{q}_i' \boldsymbol{\Omega}_i^{-1} \mathbf{q}_i \xrightarrow{d} \chi_K^2.
$$

Proposition 3. For a given lag k, consider multiple contours. Let $z_{i,k} = \sqrt{T-k}(\hat{p}_i^k - p_i)$, $i = 1, \ldots, C$, and stack them in a vector $\mathbf{z}_k = (z_{1,k}, \ldots, z_{C,k})'$. Under the null we have $\mathbf{z}_k \xrightarrow{d} N(0, \boldsymbol{\Xi}_k)$ where any element $\xi_{k,ij}$ in $\boldsymbol{\Xi}_k$ is given by

$$
\xi_{k,ij} = \min(p_i, p_j) - p_i p_j \\ \quad + \mathrm{Cov}(I_t^{k,i}, I_{t-k}^{k,j}) + \mathrm{Cov}(I_{t-k}^{k,i}, I_t^{k,j}), \qquad \forall i, j.
$$

Then, it directly follows that

$$
J_k^C = \mathbf{z}_k' \boldsymbol{\Xi}_k^{-1} \mathbf{z}_k \xrightarrow{d} \chi_C^2.
$$

Since our testing framework imposes both implications of the null hypothesis, independence and correct density function, our statistics have power against violations of both. As an example, consider the case where the null assumes that ε_t is iid normal, but in reality there is neglected dependence in the conditional mean. Under the alternative, we have elliptical autocontours as opposed to circles implied by the null. Now suppose the null assumes iid normal innovations, but ε_t is in fact an independent Student-t process. The actual autocontours will look like Figure 1(b), whereas the null implies circles as in Figure 1(a). The discrepancy between the autocontours in these cases is what makes our tests powerful in detecting departures from the null in either direction. Whenever there is neglected dependence, t and J-statistics will exhibit particular patterns signaling the form of the violation, e.g., significance at early lags, but a fast decay in case of linear dependence. Similar arguments apply to the Q-statistics. When the null is rejected because of an incorrectly assumed density function, the t and J-statistics will be significant, but they will not display any specific dynamic patterns. Finally, when the rejection comes from violations of both dependence and the density function, it will be desirable to combine our omnibus tests with tests that are powerful against violations of the null in a single direction, e.g., Escanciano (2008), Fernandes and Grammig (2005), and Hong (1996).

3. PARAMETER UNCERTAINTY

Even though the tests we propose can be applied to raw data, they will be most useful as a diagnostic tool for model specification. Thus, in practice we will be analyzing residuals, $\hat{\varepsilon}_t(\hat{\boldsymbol{\theta}}_T)$, which depend on parameter estimates, instead of the true error $\varepsilon_t(\boldsymbol{\theta}_0)$. Our tests are subject to the uncertainty created by parameter estimation. The following discussion will be centered around the t-statistics, $t_{k,i}$, considered in Proposition 1, but the same conclusions will apply to the Q_i^K and J_k^C tests given in Propositions 2 and 3, respectively.

To understand how parameter estimation affects the tests, let us take a mean value expansion and apply Slutsky's Theorem, e.g., Randles (1982). This yields

$$
\sqrt{T}(\hat{p}_i^k(\hat{\boldsymbol{\theta}}_T) - p_i) \\ = \sqrt{T}(\hat{p}_i^k(\boldsymbol{\theta}_0) - p_i) \\ \quad + \sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0)' \lim_{T \to \infty} E\left[\left. \frac{\partial \hat{p}_i^k(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \right] + o_p(1). \quad (3.1)
$$

Note that we use T instead of $(T-k)$ since they are asymptotically equivalent. We make the following assumptions to obtain the asymptotic distribution under parameter uncertainty:

A1. $\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1})$ where $\mathbf{A} \equiv E[-\mathbf{H}(\boldsymbol{\theta}_0)]$, $\mathbf{B} \equiv E[\mathbf{S}(\boldsymbol{\theta}_0)\mathbf{S}'(\boldsymbol{\theta}_0)]$, and $\mathbf{H}(\boldsymbol{\theta}_0) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{H}_t(\boldsymbol{\theta}_0)$ and $\mathbf{S}(\boldsymbol{\theta}_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{s}_t(\boldsymbol{\theta}_0)$ are the Hessian matrix and the score vector corresponding to quasi-maximum likelihood (QML) estimation.

A2. Let $\mathbf{D} \equiv \lim_{T \to \infty} E[\frac{\partial \hat{p}_i^k(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}]$. Assume $D_j < \infty$ for $j = 1, \ldots, q$ where q is the dimensionality of the parameter space.

A3. $\mathrm{Cov}(I_t^{k,i}, \mathbf{s}_t(\boldsymbol{\theta}_0)) < \infty$ and $\mathrm{Cov}(I_t^{k,i}, \mathbf{s}_{t-k}(\boldsymbol{\theta}_0)) < \infty$.

A1 is based on standard quasi-maximum likelihood estimate (QMLE) arguments. A2 guarantees that the gradient vector is bounded. A3 is a weak assumption that is required to have a well-defined asymptotic variance. A2 and A3 can be straightforwardly verified for commonly used models.

Proposition 4. Under A1–A3 we have

$$
\sqrt{T}(\hat{p}_i^k(\hat{\boldsymbol{\theta}}_T) - p_i) \xrightarrow{d} N(0, \tau_{k,i}^2),
$$

where $\tau_{k,i}^2 = \sigma_{k,i}^2 + \mathbf{D}'\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{D} + 2E[\sqrt{T}(\hat{p}_i^k(\hat{\boldsymbol{\theta}}_T) - p_i) \times \mathbf{S}'(\boldsymbol{\theta}_0)]\mathbf{A}^{-1}\mathbf{D}$.

Proposition 5. For a Gaussian location-scale model $y_t = \mu_0 + \sigma_0\varepsilon_t$, where $\varepsilon_t \sim$ iid N(0, 1), we have

$$
\mathbf{D} = \left(0, -a_{\alpha_i}f_{z_t}(a_{\alpha_i})/\sigma_0^2\right)',
$$

$$
E[\sqrt{T}(\hat{p}_i^k - p_i)\mathbf{S}'(\boldsymbol{\theta}_0)] = \left(0, (E[I_t^{k,i}\varepsilon_t^2] - p_i)/\sigma_0^2\right),
$$

where f_{z_t} is the pdf of z_t, a chi-squared random variable with 2 degrees of freedom.

This proposition states that the estimation of the parameters in the mean does not affect the asymptotic distribution of the test ($D_1 = 0$) for the case when the variance is known. This result also holds for a model with Student-t innovations.

In general, it will be difficult to obtain an empirical counterpart of the gradient vector $\mathbf{D}$. In addition, for some models the covariance terms given in A3 may be difficult to estimate, e.g., simulation-based methods reviewed in Gouriéroux and Monfort (1996). Therefore, we propose to estimate the asymptotic variance $\tau_{k,i}^2$ using a bootstrap procedure. This is a commonly used approach in the literature to overcome the difficulties associated with asymptotic variance estimation in various contexts (see Efron 1979; Buchinsky 1995; and Ledoit, Santa-Clara, and Wolf 2003 among others). The bootstrap estimator of the variance in Proposition 4 is given by

$$\hat{\tau}_{k,i}^2 = \frac{T}{B-1} \sum_{b=1}^{B} \left(\tilde{p}_i^k(\tilde{\boldsymbol{\theta}}_{T,b}) - \frac{1}{B} \sum_{b=1}^{B} \tilde{p}_i^k(\tilde{\boldsymbol{\theta}}_{T,b}) \right)^2, \qquad (3.2)$$

where B is the total number of bootstrap samples and $\tilde{\boldsymbol{\theta}}_{T,b}$ is the parameter estimate from the bth bootstrap sample. For the chi-squared statistics, the covariance matrix estimators are defined analogously. We prefer a parametric bootstrap since the null hypothesis fully specifies the parametric data generating process (DGP) (see Horowitz 2001). In particular, bootstrap samples are obtained from Equation (2.1) by replacing $\boldsymbol{\theta}_0$ with $\hat{\boldsymbol{\theta}}_T$ and generating ε_t from the specified parametric distribution. Under suitable regularity conditions, this estimator should be consistent as proven within the linear regression context for iid observations by Liu and Singh (1992), and for dependent data by Goncalves and White (2005). Alternatively, one can bootstrap the full distribution of the test statistics. However, our test statistics are not asymptotically pivotal under parameter uncertainty, and this implies that the bootstrap distribution does not necessarily provide a superior approximation to the finite sample distributions of test statistics, see Horowitz (2001). Monte Carlo results (to be presented in the following section) indicate that bootstrapping the asymptotic covariance matrices and using standard asymptotic critical values delivers remarkable results in terms of size and power of the tests.

4. MONTE CARLO SIMULATIONS

4.1 Size of the Tests

For size simulations we consider the following three cases:

1. $y_t = \mu + \sigma \varepsilon_t$, $\varepsilon_t \sim$ iid N(0, 1), $\mu = 1.25$, and $\sigma = 2$.
2. $y_t = \mu + \sigma \varepsilon_t \sqrt{(\nu - 2)/\nu}$, $\varepsilon_t \sim$ iid Student-$t(\nu)$, $\mu = 1.25$, $\sigma = 2$, and $\nu = 5$.
3. $y_t = \sigma \varepsilon_t$, $\varepsilon_t \sim$ iid exp(β), $\sigma = 2$, and $\beta = 1$.

The Gauss 7.0 random number generator is used to generate pseudo random numbers from the three distributions. We apply the tests to the properly standardized residuals, $\hat{\varepsilon}_t \equiv (y_t - \hat{\mu})/\hat{\sigma}$. Because of computational considerations, the number of Monte Carlo replications is 1000. The number of bootstrap replications is 500. We consider 13 autocontours ($C = 13$) with coverage levels (%): 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, and 99, spanning the entire density function. The nominal size level is 5%.

In Table 1 we report the results for $t_{1,i}$, $i = 1, \ldots, 13$. There are no systematic deviations from the nominal size either across autocontours or across distributions with some exceptions for the 1% and 99% autocontours when $T = 250$. This result is not surprising because in small samples there is not enough variation at the extreme contours (1% and 99% coverage).

To check the size robustness of the t-statistics for different lags, we exclusively focus on the 50% autocontour ($i = 7$) and consider $k = 1, \ldots, 5$. We report these results in Table 2. Overall the simulated size values are around 5% indicating good size for all distributions. These results are robust across all the autocontours and also available from the authors upon request. We also report the size results for the Q-statistic (Q_7^K, $K = 2, \ldots, 5$) and the J-statistic (J_1^{13} and J_1^7), respectively, in Table 2. Overall, Q-statistics have acceptable size. For the J-statistic, when we consider all 13 autocontours, there is a tendency for the test to over-reject, but when we focus on the middle autocontours by removing the first three (1%, 5%, and 10%) and the last three (90%, 95%, and 99%), the empirical size improves substantially, approaching the nominal size even for small samples.

Table 1. Size of the t-statistics

T	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$	$t_{1,11}$	$t_{1,12}$	$t_{1,13}$
						(a) Normal							
250	5.3	5.9	7.0	5.8	5.1	5.3	5.6	4.9	4.8	4.8	5.8	4.4	2.4
500	4.0	4.6	4.0	5.1	4.4	5.5	5.1	5.4	5.8	4.6	4.3	4.1	3.9
1000	4.7	5.7	6.8	5.5	4.8	4.9	4.4	5.3	5.7	5.7	4.6	4.5	6.0
2000	4.5	5.2	5.9	4.0	4.8	4.0	4.9	5.6	7.2	6.0	4.5	5.9	4.6
						(b) Student-t							
250	3.6	4.4	4.7	4.3	4.1	4.2	4.3	5.6	4.6	5.3	5.2	4.5	2.0
500	4.2	5.5	5.5	5.3	5.5	5.1	5.2	5.2	5.3	4.9	4.7	4.3	5.7
1000	3.9	3.8	3.9	4.4	3.7	3.6	4.1	3.8	4.3	3.5	3.4	3.6	5.8
2000	5.4	5.0	4.0	4.5	4.4	4.0	5.1	5.1	4.6	4.7	5.4	3.2	4.0
						(c) Exponential							
250	4.9	4.7	4.8	4.4	5.2	5.6	5.0	4.1	5.3	4.4	5.4	4.8	4.2
500	4.7	4.8	5.0	4.8	4.8	4.2	6.2	5.6	6.4	4.6	5.4	4.9	5.6
1000	4.0	5.4	4.8	5.1	5.1	5.6	5.4	5.3	4.7	6.4	4.3	4.1	5.0
2000	3.7	5.5	4.5	4.0	4.4	4.9	4.4	4.7	5.4	5.0	3.9	5.6	5.5

NOTE: Simulated size (%) of t-statistics applied to standardized residuals under three DGPs: (1) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t \sim$ iid N(0, 1); (2) $y_t = 1.25 + 2\varepsilon_t\sqrt{3/5}$, $\varepsilon_t \sim$ iid Student-$t(5)$; and (3) $y_t = 1.25\varepsilon_t$, $\varepsilon_t \sim$ iid exp(1). Number of Monte Carlo replications: 1000; bootstrap replications: 500; nominal size: 5%.

Table 2. Size of the t, Q, and J-statistics

T	$t_{1,7}$	$t_{2,7}$	$t_{3,7}$	$t_{4,7}$	$t_{5,7}$	Q_7^2	Q_7^3	Q_7^4	Q_7^5	J_1^{13}	J_1^7
					(a) Normal						
250	5.0	6.2	4.9	6.4	5.8	5.3	6.7	6.4	6.5	6.6	5.7
500	5.0	5.1	6.6	5.0	4.8	6.0	6.5	6.7	6.6	6.2	6.0
1000	4.7	5.4	5.5	5.1	4.7	5.5	6.6	6.4	6.5	6.6	5.4
2000	5.0	4.6	4.8	5.3	4.6	4.4	5.1	4.6	4.5	5.8	5.4
					(b) Student-t						
250	4.5	4.6	4.1	4.6	4.6	4.4	4.3	4.4	5.1	6.0	4.8
500	5.3	5.0	5.5	4.9	4.3	6.1	6.9	7.1	6.5	7.4	6.2
1000	4.2	4.4	4.1	4.4	4.0	5.0	6.5	6.3	6.0	6.1	5.4
2000	5.0	5.4	4.7	5.5	5.2	5.5	5.6	5.4	6.5	6.9	6.3
					(c) Exponential						
250	5.3	5.3	5.6	4.6	5.0	5.9	5.5	5.3	4.7	7.1	5.5
500	7.0	6.2	5.3	4.4	5.4	7.3	6.0	5.6	5.5	6.0	5.4
1000	5.8	4.5	5.0	4.7	5.6	5.1	5.2	6.0	6.1	6.9	5.7
2000	4.5	4.5	5.5	5.3	6.0	4.5	4.4	5.2	5.2	5.2	5.2

NOTE: Simulated size of t, Q, and J-statistics for three DGPs (see the notes to Table 1 for details).

From a practical point of view, one may not want to use autocontours too small or too large when the sample size is small.

4.2 Power and Consistency of the Tests

To analyze the power properties of our test statistics we consider three alternative data generating processes and apply our tests to the standardized residuals generated by the estimation of the following location-scale models, where $\mu = 1.25$ and $\sigma = 2$ in all cases:

1. $y_t = \mu + \sigma \varepsilon_t$, $\varepsilon_t = \phi \varepsilon_{t-1} + u_t$ where $u_t \sim$ iid N$(0, 1 - \phi^2)$.
2. $y_t = \mu + \sigma \varepsilon_t \sqrt{(v - 2)/v}$, where $\varepsilon_t \sim$ iid Student-$t(v)$.
3. $y_t = \mu + \sigma \varepsilon_t$, $\varepsilon_t = \sqrt{h_t} u_t$ where $u_t \sim$ iid N$(0, 1)$, $h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$.

For all three cases the null hypothesis is $\varepsilon_t \sim$ iid N$(0, 1)$. In case 1 we investigate departures from the independence hypothesis by considering different values of the autoregressive parameter, $\phi \in \{0.5, 0.9\}$. In case 2, we maintain the independence hypothesis and investigate departures from the hypothesized density functional form by generating iid data from the Student-t distribution for two different values of the shape parameter, $v \in \{5, 15\}$. The variance is normalized to unity to control for the scale effects. Finally, in case 3 we analyze departures from both dependence and functional form by generating data from a generalized autoregressive conditional heteroscedasticity (GARCH)(1, 1) model with $(\alpha, \beta) \in \{(0.05, 0.9), (0.15, 0.8)\}$. We set $\omega = 1 - \alpha - \beta$ to normalize the unconditional variance to 1.

In Figure 2, we present the power surfaces for $t_{1,i}$, $i = 1, \ldots, 13$, for samples sizes ranging from 250 to 2000 observations and for the 13 autocontours. In the autoregressive case 1, the power of the t-statistics approaches 1 rather quickly for high values of the autoregressive parameter, though there is a substantial drop in power for autocontours around the 80% coverage level. In the Student-t case 2, the power approaches 1 very quickly for low degrees of freedom. For instance, in the 99% autocontour, the power is as high as in the central autocontours

for all sample sizes due to the leptokurtosis of the Student-t density. As anticipated, rejection rates decrease as v gets larger, for which the null and the alternative become less distinguishable. The power surface exhibits the lowest values at the 90% autocontour, for which the null and the alternative hypotheses are very close to each other. In the GARCH case 3, the data are uncorrelated but dependent through the second moments. There is also excess kurtosis relative to the normal distribution. Persistence is the same across the two alternative parameterizations ($\alpha + \beta = 0.95$), but kurtosis and the level of first order autocorrelation in the second moment is increasing with α. As expected, the power of the test is increasing in α. When $\alpha = 0.05$, the excess kurtosis in the data is only 0.16 and the first order autocorrelation in the second moment is only 0.0725, so the departure from the null is fairly weak.

The substantial drop in power at particular coverage levels observed in Figure 2 should not be interpreted as a lack of consistency of the t-test, but rather as the difficulty of any test to discriminate between a null and an alternative hypothesis that are very close to each other for a fixed sample size. To understand why, let us focus on case 2 since the same reasoning will apply to the others. Let $\eta \equiv 1/v$, so that under normality $H_0 : \eta = 0$, and under Student-t $H_1 : \eta > 0$. The asymptotic power of the t-test, say $\Pi_T(\eta)$, for a fixed contour i and a fixed lag k at the α-significance level is given by:

$$\Pi_T(\eta) = P\big[\sqrt{T}|\hat{p}_i^k - p_i(\eta)|$$

$$\geq \sigma_{k,i}(0)z_{1-\alpha/2} - \sqrt{T}|p_i(\eta) - p_i(0)|\big]$$

$$\to 1 - \Phi\left(\frac{\sigma_{k,i}(0)z_{1-\alpha/2} - \sqrt{T}|p_i(\eta) - p_i(0)|}{\sigma_{k,i}(\eta)}\right)$$

$$+ \Phi\left(\frac{\sigma_{k,i}(0)z_{\alpha/2} - \sqrt{T}|p_i(\eta) - p_i(0)|}{\sigma_{k,i}(\eta)}\right).$$

We can see that the power will increase whenever $p_i(\eta) \neq p_i(0)$, and it will be equal to the size whenever $p_i(\eta) = p_i(0)$. Thus, if the null and the alternative hypothesis are very close to each other we will need a substantial number of observations for $\sqrt{T}|p_i(\eta) - p_i(0)|$ not to vanish. In Table 3 we
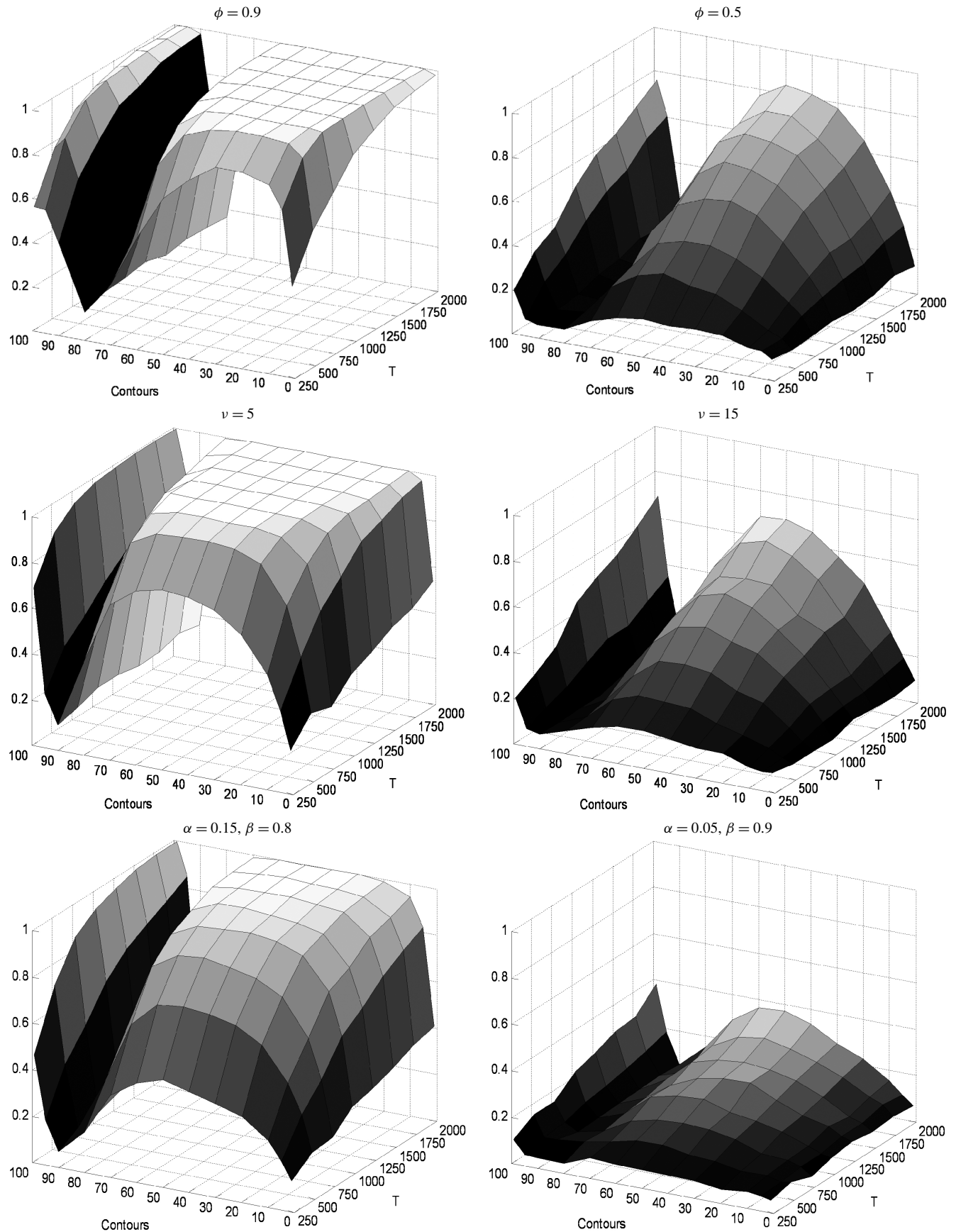
Figure 2. Power of the *t*-statistics. *Notes*: Simulated power of the *t*-statistic for all 13 autocontours and the first lag ($k = 1$) under the following DGPs: (1) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t = \phi\varepsilon_{t-1} + u_t$, $u_t \sim$ iid N$(0, 1 - \phi^2)$; (2) $y_t = 1.25 + 2\varepsilon_t\sqrt{(v-2)/v}$, $\varepsilon_t \sim$ iid Student-$t(v)$; and (3) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t = \sqrt{h_t}u_t$, where $u_t \sim$ iid N$(0, 1)$, $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$. The null hypothesis is $\varepsilon_t \sim$ iid N$(0, 1)$. All test statistics are based on standardized residuals. Number of Monte Carlo replications: 1000; number of bootstrap replications: 500; nominal size: 5%.

Table 3. Difference in violation percentages between Normal and Student-t

i	$p_i(\eta) - p_i(0)$		
	$\nu = 5$	$\nu = 10$	$\nu = 15$
1	−0.0050	−0.0019	−0.0011
2	−0.0236	−0.0089	−0.0055
3	−0.0436	−0.0167	−0.0103
4	−0.0735	−0.0289	−0.0180
5	−0.0911	−0.0368	−0.0230
6	−0.0974	−0.0403	−0.0254
7	−0.0937	−0.0395	−0.0250
8	−0.0809	−0.0344	−0.0218
9	−0.0602	−0.0253	−0.0160
10	−0.0329	−0.0125	−0.0076
11	−0.0017	0.0025	0.0023
12	0.0127	0.0092	0.0066
13	0.0174	0.0095	0.0063

NOTE: $p_i(\eta)$ is the theoretical violation percentage under the alternative hypothesis. $\eta = 1/\nu$, where ν denotes degrees of freedom of the Student-t distribution. $p_i(0)$ is the theoretical violation percentage under the null hypothesis of normality. The true distribution is Student-t with degrees of freedom equal to ν.

show the difference $p_i(\eta) - p_i(0)$ for $\nu = 5, 10, 15$. It is exactly for the 90% autocontour that the difference is the smallest, by as much as one order of magnitude, e.g., for $\nu = 5$, $p_i(\eta) - p_i(0) = -0.0017$. To provide further insight, we replicated the same simulation as in Figure 2 for $\nu = 5$ and for the 90% autocontour ($i = 11$) with sample sizes larger than 2000 observations. We find that a substantial number of observations will be needed to discriminate the null and the alternative hypotheses (more than 125,000 for a rejection rate above 60%), but eventually the power will approach 1 so that the test is consistent.

In Table 4, we report the power of the Q and J-statistics. For Q-statistics we present power results for a different autocontour in each case to offer a comprehensive analysis. In all three cases the rejection rates are high and behave in the right direction. The power is close to 1 for samples larger than 500 when there is high dependence, or a large departure from normality. When the dependence is due to strong autoregressive conditional heteroscedasticity (ARCH) effects, a sample size of 500 provides reasonable power. In case 1 when the dependence is high we observe an increase in rejection rates as K increases. This is expected given the sensitivity of the test to linear dependence. On the other hand, we do not observe a similar

Table 4. Power of the Q and J-statistics

(a) AR(1)

	T	Q_4^2	Q_4^3	Q_4^4	Q_4^5	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\phi = 0.9$	250	83.9	87.5	91.8	93.5	82.8	59.4	43.6	35.2	29.6
	500	99.5	99.7	100.0	100.0	98.0	84.8	69.4	53.2	44.9
	1000	100.0	100.0	100.0	100.0	100.0	99.0	90.6	78.8	65.7
	2000	100.0	100.0	100.0	100.0	100.0	100.0	99.5	94.9	85.5
$\phi = 0.5$	250	11.9	12.6	11.9	11.6	17.1	8.4	8.3	7.0	7.1
	500	19.2	18.2	18.1	17.7	27.8	9.2	7.2	7.7	7.3
	1000	40.0	41.7	42.3	43.9	46.4	11.2	7.5	7.5	6.9
	2000	67.6	72.3	75.7	75.0	75.0	12.8	8.6	4.4	6.3

(b) iid Student-t

	T	Q_8^2	Q_8^3	Q_8^4	Q_8^5	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\nu = 5$	250	73.8	71.0	69.2	67.3	83.4	81.9	81.8	81.1	81.0
	500	94.8	93.5	92.9	92.3	97.3	97.5	96.9	97.3	95.7
	1000	99.8	99.7	99.8	99.8	100.0	99.9	100.0	100.0	100.0
	2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
$\nu = 15$	250	15.8	14.1	12.3	11.3	21.9	22.4	20.2	20.5	19.7
	500	21.5	19.3	17.5	16.1	27.3	28.6	27.7	27.3	28.2
	1000	38.7	36.0	34.8	33.4	47.7	47.4	47.0	45.4	48.3
	2000	67.7	62.6	59.7	57.9	69.3	69.1	69.7	69.0	68.9

(c) GARCH(1, 1)

	T	Q_6^2	Q_6^3	Q_6^4	Q_6^5	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\alpha = 0.15$ $\beta = 0.8$	250	41.3	38.4	36.8	35.6	46.8	45.2	45.9	41.7	40.4
	500	69.5	66.4	64.6	63.0	72.8	70.8	70.9	70.3	68.1
	1000	94.1	92.7	91.5	90.6	94.1	93.4	93.4	92.5	92.2
	2000	99.9	99.8	99.8	99.7	100.0	99.6	99.7	99.4	99.6
$\alpha = 0.05$ $\beta = 0.9$	250	5.2	5.1	5.9	5.7	12.1	10.0	11.3	9.9	10.1
	500	10.5	10.1	9.6	8.9	15.9	13.8	14.7	14.2	14.6
	1000	16.7	14.4	14.2	13.1	20.7	21.2	19.3	19.5	18.7
	2000	29.8	27.5	24.5	22.4	36.1	34.0	33.1	31.1	29.3

NOTE: Simulated power of the Q and J-statistics under three DGPs (see the notes to Figure 2 for details).

Table 5. Size simulations for Hong–Li tests

		$\hat{Q}(k)$			$W(K)$		
T	k	Normal	Student-t	Exponential	Normal	Student-t	Exponential
500	1	3.1	8.4	6.8	3.1	8.4	6.8
	2	4.0	8.1	6.4	4.8	9.8	8.9
	3	3.6	8.8	7.1	5.4	11.8	10.0
	4	4.3	8.5	6.5	6.0	14.0	10.6
1000	1	3.8	9.4	7.3	3.8	9.4	7.3
	2	4.6	9.1	6.2	5.0	12.1	8.7
	3	3.9	9.7	7.5	5.8	14.7	10.1
	4	4.3	8.4	6.0	7.1	15.8	10.8
2000	1	4.9	9.3	7.7	4.9	9.3	7.7
	2	5.4	9.8	7.2	6.6	12.2	9.2
	3	5.0	8.5	6.7	8.0	14.6	11.1
	4	4.3	9.6	6.9	7.7	15.2	12.4

NOTE: Simulated size (%) of the Q and W-statistics of Hong and Li (2005) for three DGPs (see the notes to Table 1 for details). Nominal size is 5%.

pattern in the GARCH(1, 1) case where the dependence comes through higher moments. Regarding the J-statistic, a common characteristic across all three DGPs is that for a given sample size the power is roughly the same as the maximum power of the t-statistics. The pattern on rejection frequencies is the same as that of the individual t-statistics. In case 1, the rejection is stronger when there is high dependence even for small samples. There is a decrease in power as k increases, which is more evident when the autoregressive parameter is small. This is somehow expected because the DGP is only an AR(1) process. In case 2, the test is more powerful for small degrees of freedom, but we should mention that even for $\nu = 15$ the power of the test is still high, around 69% for 2000 observations. In case 3, we have stronger rejections rates when the data exhibit higher levels of kurtosis and stronger autocorrelation in the second moment.

4.3 Comparison of the Autocontour Tests With Hong and Li's Tests

We offer a comparison of our tests with the most recent tests proposed by Hong and Li (2005), who entertained the same joint null hypothesis of iid-ness and density functional form as that of our autocontour tests. There are other parametric and nonparametric tests, like those in Engle and Russell (1998), Fernandes and Grammig (2005), and Meitz and Terasvirta (2006), among others, that focus on a single hypothesis, either dynamic specification or density functional form. Thus, by focusing on a test with the same joint hypothesis we will be able to provide a fair comparison. Moreover, since Hong and Li's test is nonparametric in nature, our comparison will provide an assessment of the advantages of a parametric test versus a kernel-based nonparametric test.

Hong and Li's tests are based on $u_t = F_t(y_t|\mathfrak{I}_{t-1}; \boldsymbol{\theta}_0)$, which must be iid U[0, 1] under correct specification. The test has a null hypothesis of iid U[0, 1] and it compares the estimate of the joint density of $\{u_t, u_{t-j}\}$ with the product of two U[0, 1] densities, which is equal to 1. They propose two tests: (i) for a given displacement k, $\hat{Q}(k) \xrightarrow{d} N(0, 1)$, and (ii) $W(K) = K^{-1/2} \sum_{k=1}^{K} \hat{Q}(k) \xrightarrow{d} N(0, 1)$. Thus, the $\hat{Q}(k)$ test is similar to

our autocontour tests with a fixed displacement k, such as $t_{k,i}$ and J_k^C, and the $W(K)$ is similar to the autocontour test Q_i^K that aggregates over displacements.

We ran the same set of Monte Carlo experiments described previously with the Hong and Li (2003) tests. The results concerning the size of the tests are displayed in Table 5, which are directly comparable with those for our tests reported in Table 2. The autocontour tests, $t_{k,i}$ and Q_i^K, and the HL tests, $\hat{Q}(k)$ and $W(K)$, have a similar size in testing for independence and normality. The main difference arises in testing for Student-t and Exponential distributions. In both instances, the HL tests are heavily oversized. In particular, the $W(K)$ test has a size more than twice the nominal size in the Student-t and Exponential cases. These distortions in size are typical in nonparametric kernel-based tests.

On the power front, the autocontour tests are substantially more powerful in detecting deviations from the density functional form and dependence in higher moments. The HL tests have more power in detecting linear dependence. The simulation results on power differences are displayed in Table 6. The null hypothesis is $\varepsilon_t \sim$ iid N(0, 1). Since the size of the HL tests are reasonable for this null hypothesis we use their asymptotic distribution to assess the power properties. Since we already use the asymptotic distributions of the autocontour tests, both tests are compared on equal grounds. As a side note, an assessment of the power of the HL tests under nonnormal densities will require a size correction through resampling techniques.

In the Gaussian AR(1) case, the HL tests are very powerful even for small autoregressive parameters and small sample sizes. For high levels of dependence both tests are very similar. In the iid Student-t case, the autocontour tests dominate across degrees of freedom and sample sizes. For instance, a Student-t with 15 degrees of freedom is relatively close to a normal density, yet the autocontour tests have power of about 69% versus the HL tests, which have rejection rates around 20%. In the third case, where the DGP is a GARCH(1, 1) process, the autocontour tests are more powerful throughout. As we saw in the previous simulations, the power decreases when the α parameter is small, but even in this case our tests have more than twice the power of the HL tests. Overall, our assessment is that the autocontour tests have considerable advantages over the HL

Table 6. Power comparison of autocontour and Hong–Li tests

(a) AR(1)

T	ϕ	J_1^{13}	$\hat{Q}(1)$	Q_4^4	$W(4)$
500	0.5	27.8	100	18.1	99.9
	0.9	98	100	100	100
1000	0.5	46.4	100	42.3	100
	0.9	100	100	100	100
2000	0.5	75	100	75.7	100
	0.9	100	100	100	100

(b) iid Student-*t*

T	ν	J_1^{13}	$\hat{Q}(1)$	Q_8^2	$W(2)$
500	5	97.3	70.9	94.8	81.2
	15	27.3	8	21.5	11.3
1000	5	100	97.2	99.8	98.8
	15	47.7	11.9	38.7	15.9
2000	5	100	100	100	100
	15	69.3	19	67.7	26.6

(c) GARCH(1, 1)

T	(α, β)	J_1^{13}	$\hat{Q}(1)$	Q_6^2	$W(2)$
500	(0.15, 0.8)	72.8	49.4	69.5	56.3
	(0.05, 0.9)	15.9	4.8	10.5	6.7
1000	(0.15, 0.8)	94.1	81.4	94.1	88.6
	(0.05, 0.9)	20.7	7.3	16.7	10.8
2000	(0.15, 0.8)	100	98.7	99.9	99.7
	(0.05, 0.9)	36.1	13.9	29.5	18.5

NOTE: Simulated power (%) of the autocontour J and Q-statistic versus $\hat{Q}$ and W tests of Hong–Li for three DGPs (see the notes to Figure 2). The null hypothesis is $\varepsilon_t \sim$ iid N(0, 1).

tests as they offer a very good size and high power in almost all instances considered.

5. EMPIRICAL APPLICATION

Our application is concerned with dynamic specification and density functional form in autoregressive conditional duration (ACD) models for trade duration data. Bauwens et al. (2004) and Fernandes and Grammig (2005), among others, called our attention to a void in the literature on duration models, which despite being extremely popular in recent years, have not received much attention in what concerns the testing for model specification. Thus, Bauwens et al. (2004) proposed the evaluation of the density forecast in the manner of Diebold, Gunther, and Tay (1998) as a test on the distributional specification of several ACD models. Fernandes and Grammig proposed a nonparametric procedure to test the specification of the density in ACD models assuming that the conditional mean process is correctly specified. We aim to contribute to this literature by applying our autocontour tests to ACD models. Our contribution brings different aspects to the testing question. Bauwens et al. did not take into account parameter uncertainty and Fernandes and Grammig's approach is nonparametric while our approach takes parameter uncertainty into account; it is parametric and it does not presume that the conditional mean duration is correctly specified. In this sense, our tests may point pitfalls of the model in both directions (mean and density) versus just one.

We consider an ACD model for trade durations (time intervals between transactions) of the Airgas common stock from March 1 to December 31, 2001 with a total of 32,366 observations. Let $t_1, t_2, \ldots, t_i, \ldots$ denote a sequence of transaction times, then durations are defined as $x_i = t_t - t_{i-1}$. Engle and Russell (1998) introduced the ACD model for durations, which is given by,

$$x_i = \psi_i \varepsilon_i,$$
$$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{t-j}, \tag{5.1}$$

where $\psi_i = E[x_i | x_{i-1}, x_{i-2}, \ldots, x_1]$, ε_i is iid with density $f(\cdot)$, and $E[\varepsilon_i] = 1$. We focus on the model proposed by Engle and Russell (2009), which has $p = 3$ and $q = 2$. We estimate the model with quasi-maximum likelihood under exponential density. Through a battery of standard tests, Engle and Russell (2009) pointed out that this specification is successful in capturing the time dependence in durations. With the autocontour tests we aim to analyze the advantages and disadvantages of the three most popular distributional assumptions on ε_i: (i) Exponential, (ii) Weibull, and (iii) Burr. For Weibull and Burr distributions, we need to conduct a second step to obtain the estimates of the shape parameters. The estimates are $\hat{\kappa} = 0.75$ under Weibull and $\hat{\gamma} = 0.05$ and $\hat{\kappa} = 0.77$ under Burr. The Burr density reduces to Weibull when $\gamma = 0$, thus the parameter estimates suggest that the distributions are empirically close to each other.

We restrict our discussion to the t and J-statistics to save space. We summarize the results in Figures 3 and 4. The exponential distribution is overwhelmingly rejected for all coverage levels. Figure 3(a) shows that rejection is the strongest in the lower and the tail contours. For small coverage levels (1%, 5%, and 10%), the exponential density leaves too few observations outside the autocontours (negative t-statistics), thus we need a different density that squeezes the contours inward. On the contrary, for large coverage levels (90%, 95%, and 99%) the exponential leaves too many observations outside the autocontour (positive t-statistics). This means that exponential distribution is not putting enough probability mass on large durations. The Weibull density corrects these distortions to some extent. In Figure 3(b) we see that the Weibull distribution does a good job of assigning higher probabilities to large durations so that the tests at the 90%, 95%, and 99% autocontours are not statistically significant at conventional confidence levels. However, Weibull overcorrects the probability mass of the very small durations as the 1% autocontour t-test changes signs from negative [Figure 3(a)] to positive [Figure 3(b)], but it cannot place enough probability on just small durations, those within the 5% and 10% autoncontours (the t-statistics are still negative). Since $\hat{\kappa} < 1$ the Weibull autocontours have asymptotes, implying that extremely small durations can have substantial probability mass, in fact more than empirically justifiable. This finding on Weibull is in agreement with that of Bauwens et al. (2004) where a different dataset was analyzed. However, we reject the exponential distribution because it does not put enough mass on the very small durations, whereas in Bauwens et al. the exponential is rejected for the opposite reason. Furthermore, they argued that the exponential assumption is adequate for the
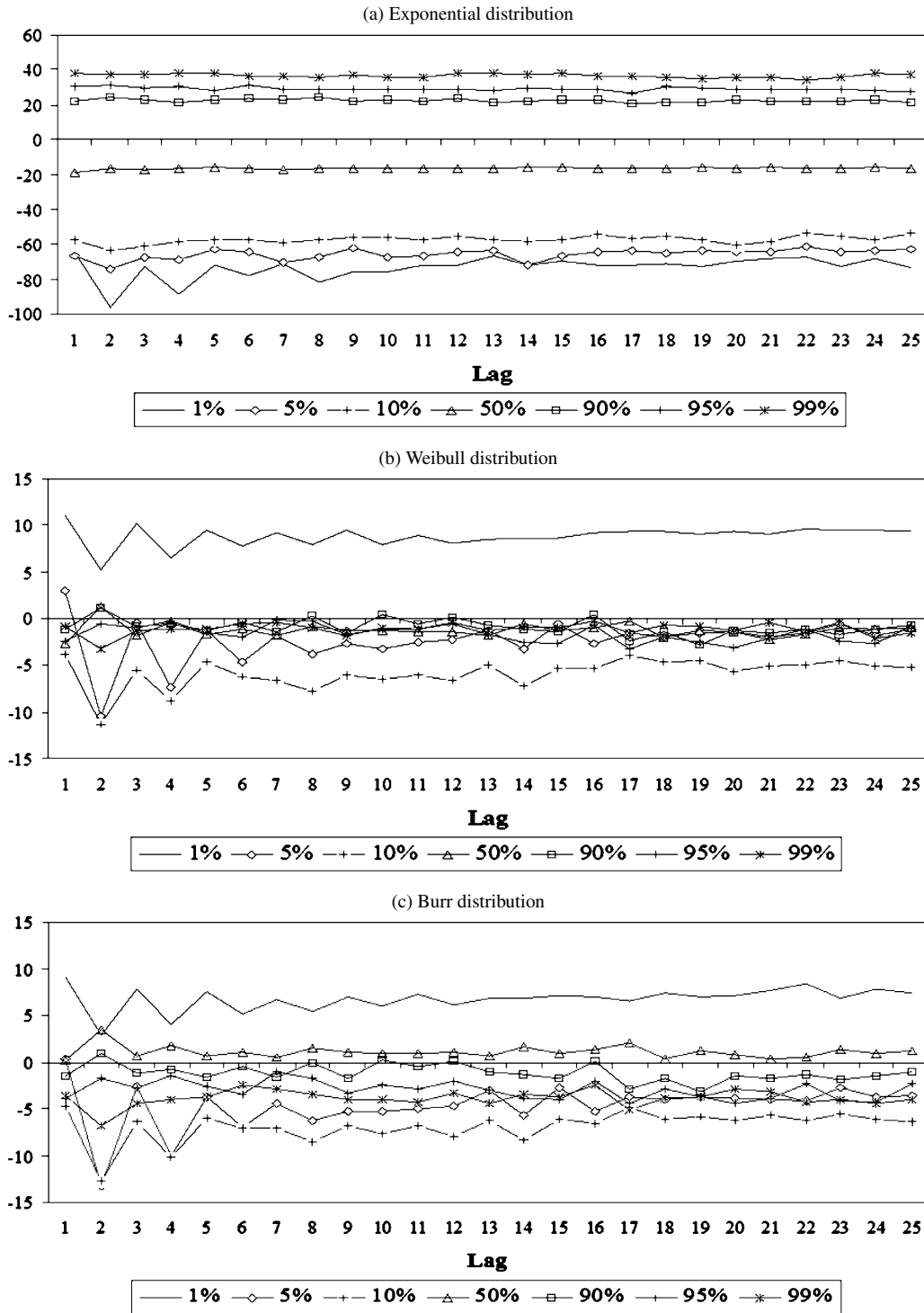
(a) Exponential distribution

(b) Weibull distribution

(c) Burr distribution
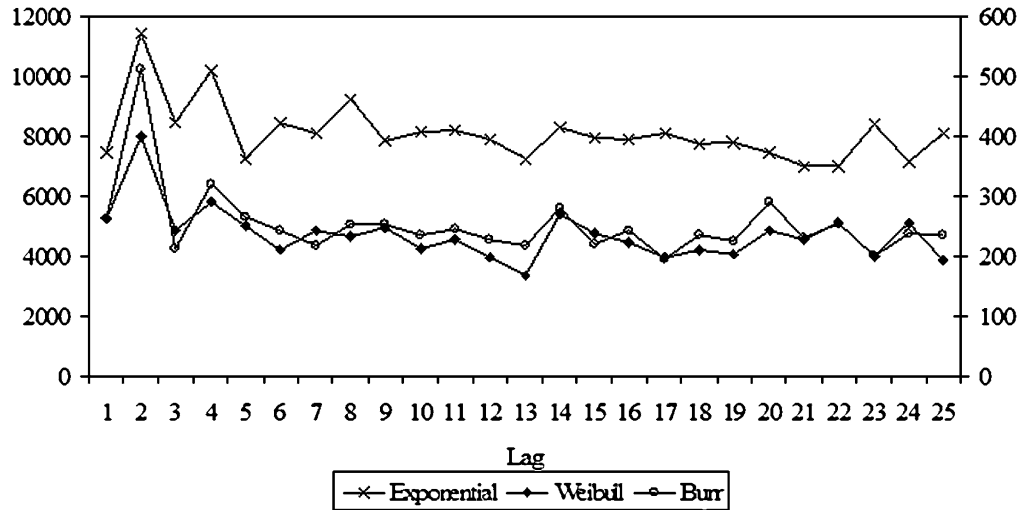

Figure 3. *t*-Statistics for ACD(3, 2) model of Airgas transaction durations. *Notes*: *t*-statistic for selected autocontours for the residuals of the exponential, Weibull, and Burr distributions ACD(3, 2) models fitted to Airgas trade durations. Statistics outside the $[-1.96, 1.96]$ range are significant at 5% level.

tail while we strongly reject this in our analysis. The performance of the Burr distribution is similar to that of Weibull for small and very small autocontours. However, unlike Weibull, Burr overcorrects the tails and the tests keep on rejecting Burr with respect to both 95% and 99% autocontours [Figure 3(c)].

The *J*-statistics, which aggregate information from all autocontours, are plotted in Figure 4. It is clear that all three distri-

butions are rejected by the data even though Weibull and Burr provide a clear improvement over Exponential. This pessimistic view is also shared by Bauwens et al. (2004). However, these results give a direction to search for new densities. The Weibull is a good starting point because it attaches the right probability mass to the tails (very large durations) and to durations within the middle contours, but it needs some corrections for the very

Figure 4. *J*-statistics for ACD(3, 2) model of Airgas transaction durations. *Notes*: *J*-statistic based on all 13 autocontours for the residuals of the Exponential (left-scale), Weibull, and Burr (right-scale) ACD(3, 2) models fitted to Airgas trade durations. 5% critical value is 22.3.

small and small durations. These adjustments can be made non-parametrically so that the proposed density enjoys a parametric and a nonparametric component.

It is illustrative to look at a scatterplot of the data together with the autocontours under various distributional assumptions. In Figure 5 we show a subset of the data (first 7500 observations). In Figure 5(a), we plot the raw durations divided by the sample mean. This is the case where dynamics are ignored and an iid exponential assumption is imposed. In Figures 5(b) through (d) we plot the ACD(3, 2) residuals under Exponential, Weibull, and Burr distributions. In these three panels we can observe the success of the ACD model in capturing the duration dynamics and the failure of the exponential distribution to capture mainly the probability of very small and large durations. Figure 5(c) shows that Weibull provides a great improvement over Exponential. By transforming the boundaries of the autocontours from straight lines to curves bended toward the origin, Weibull assigns relatively more probability mass to both small and large durations. However, because of the autocontour asymptotes, it assigns unnecessarily large probabilities to very small durations and leaves out too many observations outside the 1% autocontour. Figure 5(d) clearly illustrates the failure of Burr to improve over Weibull, mainly when large durations are concerned.

6. CONCLUSION

The methodological advances in two fronts of time series analysis—nonlinear models with nonnormal density functions, and density forecasting—emphasized the need for developing dynamic specification tests for the joint hypothesis of iid-ness and density functional form. In this article we propose a new battery of tests that rely on the fundamental properties of independent random variables with identical distributions and we introduce a graphical device, the *autocontour*. We believe that our approach brings considerable advantages over existing methods. Our tests are very powerful against violations of both hypotheses, iid-ness, and density function; they have parametric

convergence rates and standard limiting distributions; and they take into account the uncertainty introduced by parameter estimation. A comparison with Hong and Li's tests, which are non-parametric, shows that the autocontour tests are more powerful to detect nonlinear dependence and deviations from the hypothesized density function. We applied our tests to ACD models and conclude that none of the conditional densities assumed in the literature provides a good fit to model trade durations. Contrary to the standard findings in returns data, where the problem is the leptokurtosis generated by very large returns, in duration data the problem lies on the probability assigned to very small durations, which is substantially more (Weibull/Burr densities) or substantially less (Exponential density) than empirically is justifiable.

While we introduced our methodology within the context of pair-wise independence, it can also be extended to higher dimensions. On going further than the bivariate case we will be losing the graphical representation of the autocontour, which is helpful for understanding the modeling problem; however, once the analytical functional form of the autocontour is obtained, the indicator variable is easy to construct and the proposed *t*-tests and chi-squared tests will follow naturally.

APPENDIX A: CONSTRUCTION OF THE AUTOCONTOURS

In case of standard normal distribution the joint density of interest is given by

$$f(\varepsilon_t, \varepsilon_{t-k}) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(\varepsilon_t^2 + \varepsilon_{t-k}^2)\right). \qquad (A.1)$$

For a fixed value of this density, say $\overline{f}_\alpha$, we have $\varepsilon_t^2 + \varepsilon_{t-k}^2 = a_\alpha$ where $a_\alpha = -2\ln(2\pi\overline{f}_\alpha)$. Thus, autocontours are circles with radius $\sqrt{a_\alpha}$, which can be computed by numerical integration using the following equation,

$$\int_{-\sqrt{a_\alpha}}^{\sqrt{a_\alpha}} \int_{-g(\varepsilon_t)}^{g(\varepsilon_t)} f(\varepsilon_t, \varepsilon_{t-k})\, d\varepsilon_t\, d\varepsilon_{t-k} = \alpha, \qquad (A.2)$$
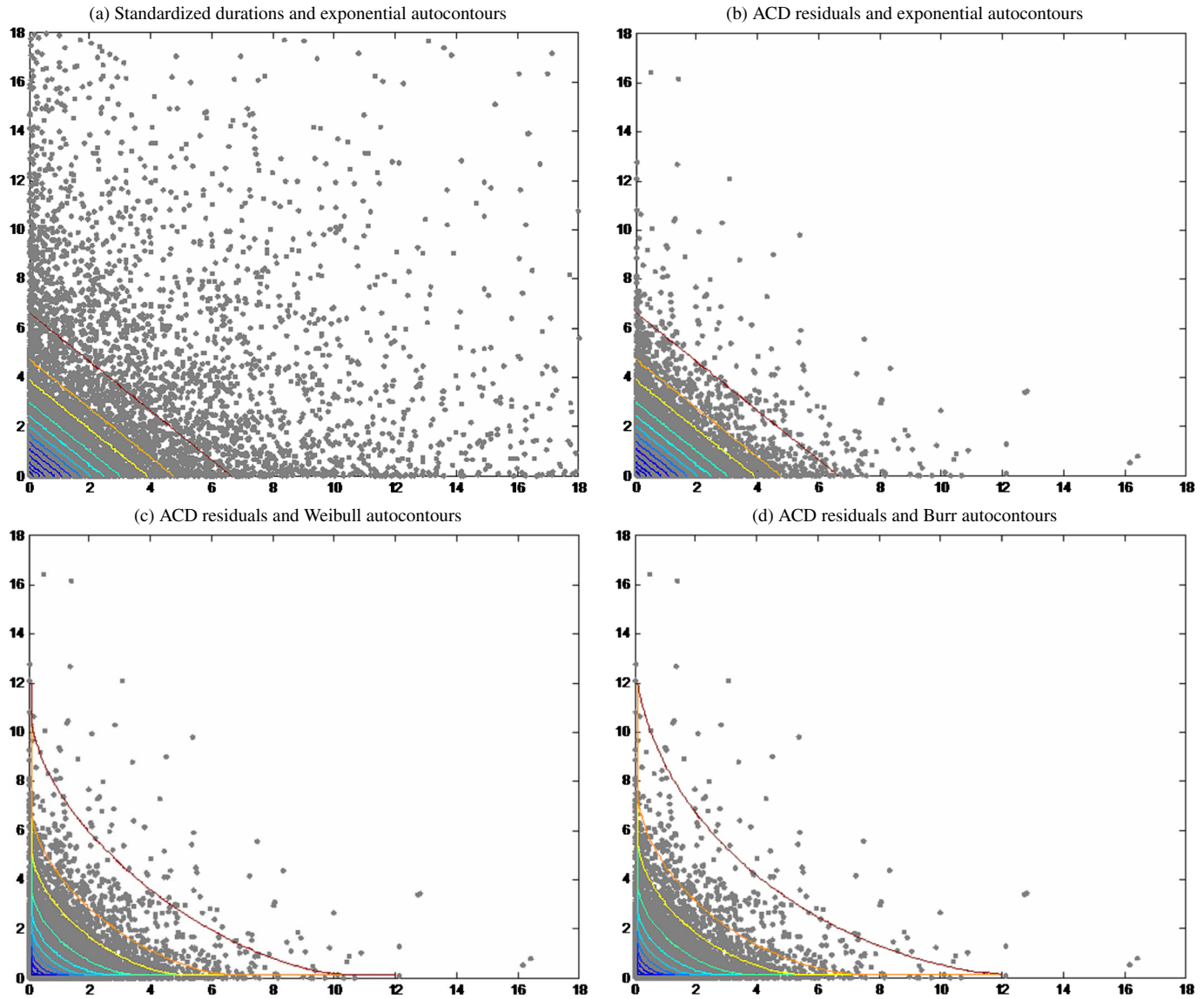
Figure 5. Data and autocontours under different distributions. *Notes*: Panel (a) presents the results for durations standardized by the sample mean. Panels (b), (c), and (d) are based on ACD(3, 2) residuals. ε_t is on the vertical axis and ε_{t-1} is on the horizontal axis. A color version of this figure is available in the electronic version of this article.

where $g(\varepsilon_t) = \sqrt{a_\alpha - \varepsilon_t^2}$. Alternatively, one can get the a values based on the cdf of a chi-squared random variable with 2 degrees of freedom due to the normality assumption, e.g., for 90% autocontour $a_{0.9} = 4.61$.

In case of Student-t and exponential distributions the value of a_α can be computed by numerical integration along the lines illustrated previously. For Weibull and Burr distributions, the autocontours cannot be easily derived by integration, so we need to resort to simulation methods. Further details on both techniques can be found in the Supplemental Appendix.

APPENDIX B: MATHEMATICAL PROOFS

Please note that some of the following proofs are presented in a compact form to save space. The interested reader is referred to the Supplemental Appendix for a complete exposition.

Proof of Proposition 1

The result directly follows from the central limit theorem for covariance stationary processes presented in Anderson (1971).

Proof of Proposition 2

Without loss of generality, let us consider the joint distribution of $q_{k,i}$ and $q_{l,i}$. Let $x \equiv \lambda_1 q_{k,i} + \lambda_2 q_{l,i}$ where λ_1 and λ_2 are arbitrary constants and assume $k < l$. Then

$$x = \lambda_1 \sqrt{T-k}(\hat{p}_i^k - p_i) + \lambda_2 \sqrt{T-l}(\hat{p}_i^l - p_i).$$

Since all results are asymptotic we drop the first $(l-k)$ observations on $I_t^{k,i}$ to use the same scaling factor. Then we have

$$x = \frac{1}{\sqrt{T-l}} \sum_{t=1}^{T-l} [\lambda_1(I_t^{k,i} - p_i) + \lambda_2(I_t^{l,i} - p_i)].$$

Define $e_t = \lambda_1(I_t^{k,i} - p_i) + \lambda_2(I_t^{l,i} - p_i)$, then $x = \frac{1}{\sqrt{T-l}}\sum_{t=1}^{T-l} e_t$. The first two moments of e_t and its autocovariance function are given by

$$E[e_t] = 0,$$

$$\text{Var}(e_t) = (\lambda_1^2 + \lambda_2^2)p_i(1 - p_i)$$
$$+ 2\lambda_1\lambda_2\text{Cov}(I_t^{l,i}, I_t^{k,i}) < \infty,$$

$$\text{Cov}(e_t, e_{t-h}) = \begin{cases} \lambda_1^2\gamma_k^i + \lambda_1\lambda_2\text{Cov}(I_t^{k,i}, I_{t-k}^{l,i}), & \text{if } h = k \\ \lambda_1^2\gamma_l^i + \lambda_1\lambda_2\text{Cov}(I_t^{l,i}, I_{t-l}^{k,i}), & \text{if } h = l \\ \lambda_1\lambda_2\text{Cov}(I_t^{l,i}, I_{t-l+k}^{k,i}), & \text{if } h = l - k \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, $\text{Cov}(e_t, e_{t-h}) < \infty$ for $\forall h$. From the central limit theorem (CLT) for covariance stationary processes, $x \xrightarrow{d} N(0, \sigma_x^2)$ where $\sigma_x^2 = \text{Var}(e_t) + 2\text{Cov}(e_t, e_{t-k}) + 2\text{Cov}(e_t, e_{t-l}) + 2\text{Cov}(e_t, e_{t-l+k})$. Thus, we show that any linear combination of $q_{k,i}$ and $q_{l,i}$ is asymptotically normal, establishing their joint asymptotic normality.

Now let us focus on the elements of the asymptotic covariance matrix. Given that $E[q_{k,i}] = E[q_{l,i}] = 0$, we have $\text{Cov}(q_{k,i}, q_{l,i}) = E[q_{k,i}q_{l,i}]$ where

$$E[q_{k,i}q_{l,i}] = E\left[\frac{1}{\sqrt{T-l}}\sum_{t=1}^{T-l}(I_t^{k,i} - p_i)\frac{1}{\sqrt{T-l}}\sum_{t=1}^{T-l}(I_t^{l,i} - p_i)\right]$$
$$= \text{Cov}(I_t^{k,i}, I_t^{l,i}) + \text{Cov}(I_t^{k,i}, I_{t-k}^{l,i})$$
$$+ \text{Cov}(I_t^{l,i}, I_{t-l}^{k,i}) + \text{Cov}(I_t^{l,i}, I_{t-l+k}^{k,i}) + o(1).$$

Proof of Proposition 3

Consider the joint distribution of $z_{i,k}$ and $z_{j,k}$. Let $x \equiv \lambda_1 z_{i,k} + \lambda_2 z_{j,k}$ and define $e_t = \lambda_1(I_t^{k,i} - p_i) + \lambda_2(I_t^{k,j} - p_j)$. This yields $x = \frac{1}{\sqrt{T-k}}\sum_{t=1}^{T-k} e_t$. The moments of e_t can be obtained as in proof of Proposition 2 and x can be shown to be asymptotically normal. Similar reasoning applies to the elements of the covariance matrix.

Proof of Proposition 4

From assumption A1 in the main text we have,

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = A^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} s_t(\theta_0) + o_p(1).$$

Defining $x \equiv \lambda_1\sqrt{T}(\hat{p}_i^k - p_i) + \lambda_2\sqrt{T}(\hat{\theta}_T - \theta_0)'D$ and suppressing the argument of the score we obtain,

$$x = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\lambda_1(I_t^{k,i} - p_i) + \lambda_2 s_t'A^{-1}D] + o_p(1).$$

Now, let $e_t = \lambda_1(I_t^{k,i} - p_i) + \lambda_2 s_t'A^{-1}D$, then $x = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t + o_p(1)$. The moments of e_t are given by

$$E[e_t] = 0,$$

$$\text{Var}(e_t) = \lambda_1^2 p_i(1 - p_i) + \lambda_2^2 D'A^{-1}BA^{-1}D$$
$$+ 2\lambda_1\lambda_2\text{Cov}(I_t^{k,i}, s_t')A^{-1}D,$$

$$\text{Cov}(e_t, e_{t-h}) = (h = k)[\lambda_1^2\gamma_k^i + \lambda_1\lambda_2\text{Cov}(I_t^{k,i}, s_{t-k}')A^{-1}D].$$

Assumptions A2 and A3 guarantee finiteness of the variance and covariance terms. Applying the CLT, we have $x \xrightarrow{d} N(0, \sigma_x^2)$ where $\sigma_x^2 = \text{Var}(e_t) + 2\text{Cov}(e_t, e_{t-k})$. The variables on the right-hand side of Equation (3.1) are both asymptotically normal due to Proposition 1 and assumption A1. We just showed asymptotic normality of their linear combinations, which completes the proof.

Proof of Proposition 5

Let $\theta_0 = (\mu_0, \sigma_0^2)'$. In this case the indicator series is constructed as follows $I_t^{k,i}(z_t(\theta_0)) = (z_t(\theta_0) - a_{\alpha_i} > 0)$, where $z_t(\theta_0) = \varepsilon_t^2(\theta_0) + \varepsilon_{t-k}^2(\theta_0)$. From the mean value expansion in the text we need the following:

$$V_{12} \equiv \lim_{T\to\infty} E[\sqrt{T}(\hat{p}_i^k(\theta_0) - p_i)S'(\theta_0)], \quad (B.1)$$

$$D \equiv \lim_{T\to\infty} E\left[\frac{\partial\hat{p}_i^k(\theta)}{\partial\theta}\bigg|_{\theta=\theta_0}\right]. \quad (B.2)$$

From the properties of the QML estimator

$$V_{12} = \text{Cov}(I_t^{k,i}(z_t(\theta_0)), s_t'(\theta_0))$$
$$+ \text{Cov}(I_t^{k,i}(z_t(\theta_0)), s_{t-k}'(\theta_0)) + o(1).$$

For the model under consideration, we have $s_t' = (\varepsilon_t/\sigma_0, (\varepsilon_t^2 - 1)/2\sigma_0^2)$. Hence, we need to obtain (i) $E[I_t^{k,i}\varepsilon_t]$, (ii) $E[I_t^{k,i}\varepsilon_{t-k}]$, (iii) $E[I_t^{k,i}\varepsilon_t^2]$, and (iv) $E[I_t^{k,i}\varepsilon_{t-k}^2]$ to get the covariance term. It can be shown that (i) and (ii) are both zero and (iii) and (iv) are given by $1 - \int_{-\sqrt{a_{\alpha_i}}}^{\sqrt{a_{\alpha_i}}}\int_{-g(\varepsilon_t)}^{g(\varepsilon_t)}\varepsilon_t^2 f(\varepsilon_{t-k})f(\varepsilon_t)\,d\varepsilon_{t-k}\,d\varepsilon_t$, where $g(\varepsilon_{t-k}) = \sqrt{a_{\alpha_i} - \varepsilon_{t-k}^2}$. Therefore, we conclude that $V_{12} = (0, (E[I_t^{k,i}\varepsilon_t^2] - p_i)/\sigma_0^2)$.

Now let us concentrate on D. Let $D = (D_1, D_2)'$ then we have

$$D_1 = -\frac{2}{\sigma_0}\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E[\delta(z_t - a_{\alpha_i})(\varepsilon_t + \varepsilon_{t-k})]\right), \quad (B.3)$$

$$D_2 = -\frac{1}{\sigma_0^2}\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E[\delta(z_t - a_{\alpha_i})z_t]\right), \quad (B.4)$$

where $\delta(x)$ is the Dirac delta function (see Arfken 2005 for further details on Dirac delta function and Phillips 1991 for a similar application). To simplify D_1 and D_2, we use the following properties of the Dirac delta function:

$$\delta(x - a) = 0 \quad \text{for } x \neq a, \quad (B.5)$$

$$\int_{-\infty}^{\infty} h(x)\delta(x - a)\,dx = h(a), \quad (B.6)$$

$$\delta(x^2 - a^2) = \frac{1}{2a}[\delta(x - a) + \delta(x + a)], \quad (B.7)$$

where a is a finite constant and h is any continuous function. Using Equations (B.5) through (B.7), we can show that $D_1 = 0$. Applying Equation (B.6) directly to Equation (B.4) yields $E[\delta(z_t - a_{\alpha_i})z_t] = a_{\alpha_i}f_{z_t}(a_{\alpha_i})$, which completes the proof.

SUPPLEMENTAL MATERIALS

Supplemental Appendix: This document contains an extended literature survey, additional Monte Carlo simulation results, and a more detailed exposition of some of the mathematical proofs. (Autocontour_Supp_Apdx.pdf)

ACKNOWLEDGMENTS

REFERENCES

Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, New York: Wiley. [198]

Andrews, D. W. K. (1997), "A Conditional Kolmogorov Test," *Econometrica*, 65, 1097–1128. [186]

Arfken, G. B. (2005), *Mathematical Methods for Physicists*, Amsterdam: Elsevier. [199]

Bai, J. (2003), "Testing Parametric Conditional Distributions of Dynamic Models," *Review of Economics and Statistics*, 85 (3), 532–549. [186]

Bauwens, L., Giot, P., Grammig, J., and Veredas, D. (2004), "A Comparison of Financial Duration Models via Density Forecasts," *International Journal of Forecasting*, 20, 589–609. [195,196]

Berkowitz, J. (2001), "Testing Density Forecasts With Applications to Risk Management," *Journal of Business & Economic Statistics*, 19 (4), 465–474. [186]

Brock, W., Dechert, D., Scheinkman, J., and LeBaron, B. (1996), "A Test for Independence Based on the Correlation Dimension," *Econometric Reviews*, 15, 197–235. [186]

Buchinsky, M. (1995), "Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study," *Journal of Econometrics*, 68, 303–338. [190]

Chen, X., and Fan, Y. Q. (2004), "Evaluating Density Forecasts via the Copula Approach," *Finance Research Letters*, 1, 74–84. [186]

Chen, Y. T. (2008), "A Unified Approach to Standardized Residuals-Based Correlation Tests for GARCH Type Models," *Journal of Applied Econometrics*, 111–133. [186]

Corradi, V., and Swanson, N. (2006), "Predictive Density Evaluation," in *Handbook of Economic Forecasting*, eds. G. Elliot, C. Granger, and A. Timmerman, Amsterdam: Elsevier Science. [186]

Diebold, F. X., Gunther, T., and Tay, A. (1998), "Evaluating Density Forecasts With Applications to Financial Risk Management," *International Economic Review*, 39, 863–388. [186,195]

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, 7 (1), 1–26. [190]

Engle, R. F., and Russell, J. R. (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66 (5), 1127–1162. [194,195]

—— (2009), "Analysis of High Financial Frequency Data," in *Handbook of Financial Econometrics*, eds. Y. Aït-Sahalia and L. Hansen, Amsterdam: North-Holland, pp. 383–426. [195]

Escanciano, J. C. (2008), "Joint and Marginal Specification Tests for Conditional Mean and Variance Models," *Journal of Econometrics*, 143, 74–87. [186,189]

Fernandes, M., and Grammig, J. (2005), "Nonparametric Specification Tests for Conditional Duration Models," *Journal of Econometrics*, 127, 35–68. [186, 189,194,195]

Goncalves, S., and White, H. (2005), "Bootstrap Standard Error Estimates for Linear Regression," *Journal of the American Statistical Association*, 100 (471), 970–978. [190]

Gouriéroux, C., and Monfort, A. (1996), *Simulation Based Econometric Methods (CORE Lectures)*, New York: Oxford University Press. [190]

Hong, Y. (1996), "Consistent Testing for Serial Correlation of Unknown Form," *Econometrica*, 64, 837–864. [189]

Hong, Y., and Lee, T.-H. (2003), "Diagnostic Checking for the Adequacy of Nonlinear Time Series Models," *Econometric Theory*, 19, 1065–1121. [186, 194]

Hong, Y., and Li, H. (2005), "Nonparametric Specification Testing for Continuous-Time Models With Applications to Term Structure of Interest Rates," *Review of Financial Studies*, 18, 37–84. [186,187,194]

Horowitz, J. L. (2001), "The Bootstrap in Econometrics," in *Handbook of Econometrics*, eds. J. Heckman and E. Leamer, Amsterdam: Elsevier Science. [190]

Ledoit, O. P., Santa-Clara, P., and Wolf, M. (2003), "Flexible Multivariate GARCH Modeling With an Application to International Stock Markets," *The Review of Economics and Statistics*, 85 (3), 735–747. [190]

Liu, R. Y., and Singh, K. (1992), "Moving Blocks Jackknifeand Bootstrap Capture Weak Dependence," in *Exploring the Limits of the Bootstrap*, eds. R. LePage and L. Billiard, New York: Wiley, pp. 225–248. [190]

Meitz, M., and Terasvirta, T. (2006), "Evaluating Models of Autoregressive Conditional Duration," *Journal of Business & Economic Statistics*, 24 (1), 104–124. [186,194]

Phillips, P. C. B. (1991), "A Shortcut to LAD Estimator Asymptotics," *Econometric Theory*, 7, 451–463. [199]

Randles, R. H. (1982), "On the Asymptotic Normality of Statistics With Estimated Parameters," *The Annals of Statistics*, 10 (2), 462–474. [189]