

Simulation-Based Exact Tests With Unidentified Nuisance Parameters Under the Null Hypothesis: the Case of Jumps Tests in Models With Conditional Heteroskedasticity.

Lynda Khalaf,¹ Jean-Daniel Saphores,²
and Jean-François Bilodeau.³

September 27, 2000

¹Assistant Professor, GREEN and Economics Department, Université Laval, Ste Foy, Québec G1K 7P4, Canada. Tel: (418) 656 2131 - 2409. Fax: (418) 656 7412. E-mail for correspondence: lkha@ecn.ulaval.ca

²Assistant Professor. Formerly: GREEN and Economics Department, Université Laval, Ste Foy, Québec G1K 7P4, Canada. Now: School of Social Ecology, University of California, Irvine 92626. E-mail for correspondence: saphores@uci.edu

³Formerly Graduate Student, GREEN and Economics Department, Université Laval, Ste Foy, Québec, G1K 7P4. Now: Division de l'analyse et de l'élaboration de la politique de la demande, Ressources Naturelles Canada, 580, rue Booth, Ottawa, Ontario, Canada K1A 0E4.

Abstract

We use the Monte-Carlo (MC) test technique to find valid p-values when testing for jumps in discretized jump-diffusion models. While the distribution of the LR statistic for this test is typically non-standard, we show that the MC p-value is finite sample exact if no other (identified) nuisance parameter is present under the null. Otherwise, we derive nuisance-parameter free bounds and obtain exact bounds p-values. We illustrate our approach with an analysis of weekly spot prices of copper, nickel, gold, and crude oil. We find significant jumps in all of these time series. These results have implications for the short term valuation and management of these resources.

Keywords: Monte-Carlo test; bounds test; exact test; jump process; conditional heteroscedasticity.

1 Introduction

Testing problems where nuisance parameters are unidentified under the null hypothesis are pervasive in econometrics. Prominent examples include tests for structural change (Andrews and Ploberger 1994) and ARCH-in-mean tests (Bera and Ra 1995). As it is now well known, when nuisance parameters are present only under the alternative hypothesis, the tests limiting null distributions are not generally chi-square. Indeed, they can take a much more complex form, *e.g.* Andrews' (1993) sup- χ^2 distribution and Hansen's (1996) χ^2 processes. More importantly, as emphasized by Hansen (1996), in several situations, the relevant limiting distributions are nuisance parameter dependent which precludes the construction of specialized critical points tables.

One early approach to dealing with this problem is the asymptotic bounds procedures proposed in Davies (1977, 1987). Hansen (1996) and Andrews (1999) have recently proposed simulation-based procedures to approximate asymptotic p -values, which is valid in settings more general than Davies'. To date, however, no finite sample exact procedure is available for such non-regular test problems.

In continuous-time finance, an important class of tests involving nuisance parameters unidentified under the null looks at the presence of jumps in the sample paths of variables such as interest rates, stock and commodity prices,

or exchange rates. Since Merton (1976) proposed to model stock prices with Poisson jumps superimposed on a geometric Brownian motion, it is well known that the presence of jumps in the sample paths of the underlying variables may have serious impacts on option prices and on hedging strategies based on dynamic portfolio adjustments. Merton's initial model has since been extended to include for example mean-reversion and conditional heteroskedasticity, which is common in high frequency data (e.g. see Bollerslev, Chou, and Kroner 1992 or Amin and Ng 1993). Examples of jump/diffusion models include Jarrow and Rosenfeld (1984), Ball and Torous (1985), Ahn and Thompson (1988), Akgiray and Booth (1988), Jorion (1988), Brorsen and Yang (1994) and Bates (1996a and 1996b). Although recognized as a potential difficulty (e.g., see Brorsen and Yang 1994) un-identification problems have not received the attention they deserve in empirical applications of jump-tests. In fact, most of the references just cited inappropriately use χ^2 critical points.

In a recent paper, Drost, Nijman, and Werker (1998) (thereafter DNW) propose a kurtosis-based test for jumps in a semiparametric weak-GARCH(1,1) setting where the unidentification problem at hand does not come up. They argue that jump-GARCH processes may be identified through an unbounded kurtosis parameter; conformably, they define the no-jump null in terms of

(over-identifying) constraints on the kurtosis and GARCH parameters.¹ These constraints are then tested via a standard Wald-type criterion, using Gaussian GARCH Quasi-Maximum-likelihood (QMLE) parameter and variance-covariance estimators, and the delta-method; the associated QMLE-based asymptotic null distribution is thus standard χ^2 . DNW report favorable simulation evidence but warn that their experiments may not be sufficient to establish the test's overall finite sample validity. Although the DNW χ^2 test is very attractive, some serious problems associated with the delta method (which underlies the test) have recently been pointed out (e.g., see Dufour 1997). These problems are caused by identification difficulties and are not restricted to small samples. Identification difficulties may be encountered here at relevant parameter boundaries, as pointed out in the sequel. Moreover, it is not clear that the DNW approach can be readily generalized to higher order GARCH models.

In this paper, we propose instead a simulation-based Likelihood Ratio

¹To define of a weak-GARCH(1,1) process (Drost and Nijman 1993, thereafter DN), consider a continuous-time process $\{Y_t : t \geq 0\}$. The implied discrete-time process at frequency $h > 0$ is denoted by $y_{(h)t} = Y_t - Y_{t-h}$ ($\frac{t}{h}$ integer), and μ_t is its mean. If $y_{(h)t} - \mu_t$ is a symmetric process, for which the best linear predictor of its variance, $\sigma_{(h)t}^2$, in terms of a constant and lagged values of $y_{(h)t}$ is $\sigma_{(h)t}^2 = \psi_h + \alpha_h y_{(h)t-h}^2 + \beta_h \sigma_{(h)t-h}^2$, (where $\frac{t}{h}$ is an integer, and ψ_h , α_h , and β_h are parameters depending only on h), then $\{y_{(h)t}, \frac{t}{h} \text{ integer}\}$ is weak-GARCH(1,1). If the kurtosis of $y_{(h)t}$ remains bounded as $h \downarrow 0$, DN call $\{Y_t; t \geq 0\}$ a GARCH(1,1)-diffusion because the Kolmogorov criterium implies that its sample path be continuous; otherwise, DN call $\{Y_t; t \geq 0\}$ a jump-GARCH(1,1)-diffusion since its sample path contains discontinuities (jumps). For their test, DNW rely on a result by Drost and Werker (1996), who show that if $\{Y_t; t \geq 0\}$ is a GARCH(1,1)-diffusion, then the kurtosis of $y_{(h)t}$ for any $h > 0$ can be explicitly written in terms of μ_h , ψ_h , α_h , and β_h ; it is not the case if $\{Y_t; t \geq 0\}$ is a jump-GARCH(1,1)-diffusion.

(LR) jump test strategy. We obtain either finite sample exact p-values or finite sample exact bounds p-values (using the terminology of Dufour, 1997) for a large class of parametric GARCH(p,q) models. This class includes strictly stationary, augmented-GARCH(p,q) models (Duan 1997), to which jump processes have been added.² Duan (1997) shows that, as the data sampling frequency increases, strictly stationary augmented-GARCH models converge to a class of diffusions that contains the bivariate diffusion models of Hull and White (1987), Scott (1987), Wiggins (1987), Stein and Stein (1991), and Heston (1993).

To circumvent the unidentified nuisance parameter problem, we follow Dufour (1997) and use pivotal or *boundedly* pivotal test statistics, *i.e.* statistics whose null distributions are either nuisance-parameters-free or which can be bounded by nuisance-parameter-free distributions. We first show that the LR test statistics we rely on satisfy the boundedly pivotal property by deriving explicitly exact pivotal bounds on their null distributions. Then, to obtain exact tests, we apply the Monte Carlo (MC) test procedure (Dufour 1995), which easily gives exact p -values when the null distribution of the underlying test statistic is nuisance-parameter-free. Hence, while in many cases unidentified nuisance parameters cause severe complications, our MC

²The class of augmented-GARCH processes includes, for example, the LGARCH, the MGARCH, the TS-GARCH, the EGARCH, the GJR-GARCH, the NGARCH, the VGARCH, or the TGARCH. See Duan (1997) for references. Augmented-GARCH processes are formally defined in Section 2.

test based procedure exploits the fact that the null DGP does not depend on these parameters. The main arguments underlying this key result are presented below. Furthermore, as will become clear from our presentation, the method is not limited by the tractability of the null distribution.

An important related difficulty is the case of *restricted* alternative hypotheses. As is well known (e.g., see Andrews 1999, 2000), problems similar to unidentification occur when the null hypothesis sets values on the *boundary* of the parameter space. As shown by Andrews (2000), even bootstraps fail in such cases. This question is also relevant in the jump tests case considered here. It turns out, however, that the method we adopt to deal with unidentified nuisance parameters does not have the problems of standard asymptotic tests, in the presence of restricted alternatives. In view of Andrews' results, our strategy may be seen as a useful alternative to the standard bootstrap in the context of jump tests.

Finally, note that whereas the results in Hansen (1996) are not directly applicable to the test problem considered here, our procedure may be applied in Hansen's framework provided distributional assumptions are imposed so that simulated samples conformable with the null can be drawn. The procedures proposed in Andrews (1999) are - in principle - applicable here. These involve simulating the supremum of quadratic forms based on: (i) a random term whose construction requires solving a *restricted* minimization

problem (over a convex cone), and (ii) first and second order derivatives of the likelihood function. While Hansen and Andrews basically approximate the limiting distributions' tail probabilities, the approach we propose simply requires simulated values of LR-based statistics and gives finite sample exact p -values.³

Our proposed methodology is then applied to investigate the existence of jumps (discontinuities) in the sample paths of four commodities, based on weekly spot prices: crude oil (West Texas Intermediate or WTI), copper, nickel, and gold. We model the logarithm of these spot prices using an Ornstein-Uhlenbeck mean reverting process (MRM) with normally distributed Bernoulli jumps, with and without a GARCH(1,1) error structure. While financial models are often based on the geometric Brownian motion, there are strong arguments (e.g., see Schwartz 1997) in favor of mean reversion in commodity prices. We find statistically significant jumps in all four of these commodities.

This paper is organized as follows. Our testing methodology is presented in Section 2. Section 3 describes our empirical application and discusses our results. Our conclusions are summarized in Section 4. The expressions of our likelihood functions can be found in an Appendix.

³For other applications of MC tests in econometrics, see Dufour and Kiviet (1998), Dufour and Khalaf (1997, 2000), or Dufour, Farhat, Gardiol and Khalaf (1998).

2 Methodology

2.1 Testing for jumps: some difficulties.

To set focus, let us first consider a stationary AR(1) model with GARCH (1,1) errors and Poisson jumps. If we denote the observed series by p_t , $t = 1, \dots, T$, this model may be written as

$$(2.1) \quad p_t = a_0 + a_1 p_{t-1} + \sqrt{h_t} z_t + \sum_{i=1}^{n_t} \ln Y_{ti},$$

$$(2.2) \quad h_{t+1} = \alpha_0 + h_t(\alpha_1 z_t^2 + \alpha_2),$$

where $|a_1| < 1$ for stationarity, $z_t \stackrel{iid}{\sim} N(0, 1)$, n_t is the number of jumps which occur between t and $t - 1$, and Y_{ti} ($i = 1, \dots, n_t$) is the size of the i^{th} jump in the time interval $(t - 1; t)$. We also assume that the arrival of jumps follows a Poisson process with arrival rate λ , and that the Y_{ti} 's are (independently) lognormally distributed with mean θ and variance δ^2 .⁴ When $a_1 = 1$, we obtain a Random Walk with drift, jumps, and GARCH innovations for $x_t \equiv p_t - p_{t-1}$; also setting α_2 to zero gives the Jump/ARCH process considered by Jorion (1988); and imposing $\alpha_1 = \alpha_2 = 0$ yields a discretized version of Merton's Jump/Geometric-Brownian-Motion model. If on the other hand we constrain a_1 to be between 0 and 1, we get a discretized Ornstein-Uhlenbeck

⁴ n_t is an integer random variable. If $n_t = 0$, the sum of $\ln Y_{ti}$ is zero. If we constrain n_t to be 0 or 1, we have a Bernoulli process.

model with jumps and GARCH(1,1) innovations. Although the issues we raise next and the methodology we propose apply to a large class of models, we focus for now on (2.1)-(2.2) for concreteness.

The parameters of the mixture model (2.1)-(2.2) may be estimated by numerical maximization of the log-likelihood function. To test the null hypothesis

$$(2.3) \quad H_0 : \lambda = 0 \quad (\text{no jump}) ,$$

the LR statistic takes the form

$$(2.4) \quad LR = 2[L_{JAG} - L_{AG}],$$

where L_{AG} and L_{JAG} are respectively the maximum of the log-likelihood function (MLF) under the null and the alternative hypothesis. As recalled in the introduction, the standard regularity conditions ensuring that the LR statistic is asymptotically χ^2 distributed under the null hypothesis are not verified. One reason is that there are two nuisance parameters θ and δ , which are not identified under H_0 (i.e., when we set $\lambda = 0$, the likelihood function no longer depends on these two parameters). Another reason is that the value of λ tested under H_0 is on the boundary of the parameter space. As a consequence, the asymptotic distribution of the LR statistic under H_0 is non-standard and quite complex. Its χ^2 approximation is no longer valid.

2.2 Finite sample exact tests: a few definitions.

The main contribution of this paper is to show how the Monte Carlo test method (Dufour 1995) may be easily applied to solve the non-regular problem described above with finite sample exact test results. Since the issues of exact size control, exact p-values, and nuisance parameters motivate our approach, let us first recall the definition of an *exact* test (or p-value) as it applies here. For theory and further discussion, see Lehmann (1986, Chapter 3).

Definition 2.1 *Consider a test problem pertaining to a parametric model, i.e. the case where the data generating process [DGP] is determined up to a finite number of unknown real parameters $\omega \in \Omega$. Let Ω_0 refer to the subspace of Ω compatible with the null hypothesis H_0 under test. Without loss of generality, consider a test statistic with critical region $S \geq c$. To obtain an α -level test, c must be chosen so that*

$$(2.5) \quad \sup_{\omega \in \Omega_0} P_{\omega}(S \geq c) \leq \alpha .$$

This test will have size α if and only if

$$(2.6) \quad \sup_{\omega \in \Omega_0} P_{\omega}(S \geq c) = \alpha .$$

Consequently, in carrying out an exact test, two difficulties must be dealt with. First, we need to find the analytic distribution of S , where S is typically (as with the LR jump test under study) a complicated function of the observations. In the next section, we show how the MC test method can be used to tackle this first difficulty. Second, we also need to maximize the rejection probabilities over the relevant nuisance parameter space. This second problem may be pervasive in general but it is trivially solved if S is *pivotal*, i.e. if its (finite sample) null distribution is nuisance parameter-free. It is clear, however, that the LR statistic we consider here is nuisance parameter dependent. It is thus essential to recall the fundamental property associated with size or level constraints: to obtain a useful test which satisfies (2.6) or (2.5), the statistic S must be *boundedly pivotal*, i.e. its null distribution must admit a nuisance parameter-free bound. Indeed, if the maximum of the rejection probabilities (over the nuisance parameter space compatible with H_0) is unbounded, then it is impossible to find a useful cut-off point which controls level or size. Below, we show that the LR jump test is boundedly pivotal regardless of the aforementioned identification difficulties. For a formal treatment of *boundedly pivotal* statistics, see Dufour (1997). For completion, and since this concept is essential in our study, we also provide the following operational definition:

Definition 2.2 *Consider the test based on S underlying (2.6)-(2.5) and sup-*

pose that it is possible to find another pivotal statistic \bar{S} such that

$$(2.7) \quad \forall \omega \in \Omega_0, \forall c, P_\omega(S \geq c) \leq P(\bar{S} \geq c)$$

under the null. Then S is said to be *boundedly pivotal*.

Inequality (2.7) implies that if a constant c is chosen so that $P(\bar{S} \geq c) = \alpha$ then $\forall \omega \in \Omega_0, P_\omega(S \geq c) \leq \alpha$. This means that \bar{S} 's critical points are level-correct (i.e. satisfy the level constraint (2.5)) if used in conjunction with the test based on S . Of course the key here is the pivotal characteristic of \bar{S} .

In practice, to obtain an exact test, whereas size control - if possible - is desirable, level control *is required*. For in many cases of practical interest including the problem under study, it is very difficult to achieve size control. This is mainly due to nuisance parameters. Indeed, we have already pointed out the pervasive role of nuisance parameters θ and δ which are unidentified under the null. From (2.5), it is clear that a_1, α_1 and α_2 also need to be dealt with in order to obtain an exact test.⁵ The identifiability of these parameters under the null does not rule out the dependence problem. In this regard, it is important to remember that the nuisance parameter space associated with a_1, α_1 and α_2 defines important properties of the model under test here, including stationarity in the mean and/or the conditional variance.

⁵Location-Scale invariance, which is straightforward to see here, takes care of the intercept parameters. We deal further with this point below.

The approach we follow in this paper is to obtain a level-correct p-value (in line with (2.5)) which is *invariant* to these characteristic. There is thus no special difficulty with values of a_1 and/or $(\alpha_1 + \alpha_2)$ equal or close to one (i.e. unit roots or near-unit roots in the mean and/or IGARCH or near-IGARCH error processes).⁶ Furthermore, the well known identification difficulties with α_2 , which arises if $\alpha_1 = 0$, are also not present with our test method since we derive p-values that do not depend on these parameters. In the same vein, no conditions on the existence of the errors' fourth moments are required. These conditions on the parameters (which, one must recall, count as nuisance parameters here) are often included as regularity assumptions when a standard test approach is pursued. In particular, DNW's test depends crucially on the assumption of finite fourth moment (this is also a major assumption which underlies aggregation in continuous time GARCH). Since invariance (under the null) to nuisance parameters whose parameter space typically includes near-unidentification regions is a key issue in our study, we refer the reader to Dufour (1997) where the consequences of this problem in finite samples are discussed in details.⁷

⁶An issue that will not be addressed in this paper is the unit-root and/or (G)ARCH pre-test effect. We should mention, however, that the methodology we propose to test for jumps may be applied to test $a_1 = 1$ or $\alpha_1 = \alpha_2 = 0$ within the defined framework (i.e. imposing or ignoring jumps).

⁷Dufour (1997) shows that "assuming such conditions away" through a priori regularity conditions may lead to spurious test results.

2.3 Exact jump tests

Let us now prove an equality that serves to establish the boundedly pivotal property of the LR no-jump test statistic (2.4).

Proposition 2.1 *In the context of the jump test described by (2.1)-(2.2)-(2.3), consider the LR no-jump criterion (2.4). Let L_{RW} denote the MLF imposing (2.3) and $a_0 = 0$, $a_1 = 1$, $\alpha_1 = \alpha_2 = 0$ (we assume a no-GARCH no-JUMP Random Walk). Now define*

$$(2.1) \quad LR_B = 2[L_{JAG} - L_{RW}],$$

the LR statistic for testing the Random Walk null model against the mixed model. Then

$$(2.2) \quad \forall c, P(LR \geq c) \leq P(LR_B \geq c).$$

Proof. By construction, we have that $L_{JAG} \geq L_{AG} \geq L_{RW}$. This implies that

$$(2.3) \quad [L_{JAG} - L_{AG}] \leq [L_{JAG} - L_{RW}],$$

so that $LR \leq LR_B$. Inequality (2.2) follows immediately.

In conjunction with Proposition 2.1, to prove the boundedly pivotal characteristic of LR , we have yet to show how a nuisance-parameter-free cut-off (or alternatively a p-value) may be obtained through the bounding statistic

LR_B . To obtain the desired p-value, we propose to apply the technique of Monte Carlo tests.

Monte Carlo tests, which were originally suggested by Dwass (1957) and Barnard (1963), have recently been generalized to the nuisance parameters-dependent case by Dufour (1995). These tests are based on a fundamental distributional result concerning the ranks of a set of exchangeable random variables. We state this result without proof in the following lemma and refer the reader to Dufour (1995) for a formal discussion.

Lemma 2.2 *Let Z_0, Z_1, \dots, Z_N be exchangeable real random variables, and let R_j be the rank of Z_j in the series $\{Z_0, \dots, Z_N\}$ assuming a non-decreasing ordering. Then*

$$P [R_j / (N + 1) \geq x] = \begin{cases} 1 & , \text{ if } x \leq 0 \\ (1 + I [(N + 1)(1 - x)]) / (N + 1) & , \text{ if } x = 0 \end{cases}$$

where $I(x)$ is the largest integer less than or equal to x .

Lemma 2.2 is the key result to obtain exact LR-based Monte Carlo tests.

Proposition 2.3 *Let Z_0 denote the value of a continuous test statistic computed from the observed data. Suppose that it is possible to obtain N i.i.d. random variables with the same distribution as Z_0 under the relevant null*

hypothesis. Denote these realizations by Z_j , $j = 1, \dots, N$. Now calculate

$$(2.4) \quad \hat{p}_N(Z_0) = \frac{N\hat{G}_N(Z_0) + 1}{N + 1},$$

$$(2.5) \quad \hat{G}_N(Z_0) = \frac{1}{N} \sum_{i=1}^N I_{[0, \infty]}(Z_i - Z_0),$$

$$I_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

Then, for all $0 < \alpha < 1$ where N is such that $\alpha(N + 1)$ is an integer:

$$(2.6) \quad P_{(H_0)}[\hat{p}_N(Z_0) \leq \alpha] = \alpha.$$

Proposition 2.3 obtains immediately from Lemma 2.2 in the case of continuous statistics, if we realize that $N\hat{G}_N(Z_0)$ is the number of simulated criteria $\geq Z_0$ so that $N - N\hat{G}_N(Z_0) + 1$ gives the rank of Z_0 in the series Z_0, Z_1, \dots, Z_N . The formula for $\hat{p}_N(Z_0)$ gives the *empirical probability* to observe a value as extreme or more extreme than Z_0 under the null. Consequently, $\hat{p}_N(Z_0)$ may be viewed as a randomized MC p -value, so that the MC test's critical region of size α may be defined as:

$$\hat{p}_N(Z_0) \leq \alpha.$$

The fact that the latter critical region has size α exactly (with respect to the definitions introduced above) obtains from (2.6).⁸

In the case of jump tests, if a_1 , α_1 , and α_2 are known (e.g. in Merton's model), to apply Proposition 2.3, we first construct Z_1, \dots, Z_N as N realizations of the no-jump LR, obtained by Monte Carlo methods; this involves obtaining simulated samples from the no-jump data generating process and computing the corresponding test statistics. Then (2.4) yields an exact (size correct) p-value. The key here is that the MC p -value (calculated as just described) does not depend in any way on the "problematic" parameters θ and δ^2 .⁹ Furthermore, the boundary restriction does not intervene here, since the only elements of proof concern the conditions underlying Lemma 2.2. In this case, these conditions rest on the assumption that a_1 , α_1 , and α_2 are known, so that no unknown parameter needs to be dealt with in order to generate Z_1, \dots, Z_N .

If the model considered does not set the values of a_1 , α_1 , and α_2 , then they need to be estimated from the data and the Monte Carlo test described above cannot be carried out. It is important to note that if the above routine is implemented with estimates of a_1 , α_1 , and α_2 , as would be the case with

⁸The notation $P_{(H_0)}$ implies reference to the distribution in question, for all DGPs compatible with the null hypothesis. Refer to Dufour (1995) for more details.

⁹The model, and test problem, are invariant to transformations of the form $p_t^* = c^*p_t + d^*$, where $c^* > 0$ and $d^* \in \Re$. It follows that LR statistics are location-scale invariant; see Dagenais and Dufour (1991). Consequently, obtaining simulated samples from the null DGP poses no estimation problems here.

a standard parametric bootstrap, then (2.6) is not preserved. To circumvent this problem, we propose to exploit Lemma 2.2 in conjunction with a pivotal bounding statistic. Our purpose is again to devise a p-value which is θ and δ^2 free, and which does not depend on a_1 , α_1 , and α_2 . From a theoretical perspective, this ensures complete pivotality (i.e. invariance under the null to all intervening parameters, whether identified or not), hence exactness in the above defined sense. From an empirical perspective, this renders our test independent of the stationarity of the DGP, which guards against pre-test effects.

Proposition 2.4 *Let Z_0 denote the value of a continuous test statistic computed from the observed data. Suppose that it is possible to obtain N i.i.d. random variables from a distribution which bounds the distribution of Z_0 , in the sense of (2.7), under the null hypothesis of interest. Denote these realizations by Z_j , $j = 1, \dots, N$. Let*

$$(2.7) \quad \tilde{p}_N(Z_0) = \frac{N\hat{G}_N(Z_0) + 1}{N + 1}$$

where $\hat{G}_N(\cdot)$ is defined by (2.5). Then for all $0 < \alpha < 1$ where N is chosen so that $\alpha(N + 1)$ is an integer:

$$(2.8) \quad P_{(H_0)} [\tilde{p}_N(Z_0) \leq \alpha] \leq \alpha.$$

Proposition 2.4 derives from Definition 2.2 and Lemma 2.2 in the case of continuous statistics. Indeed, using Definition 2.2, the N random variables Z_j , $j = 1, \dots, N$ may be seen as realizations of the bounding statistic underlying (2.7). The formula for $\tilde{p}_N(Z_0)$ gives a *bounds Monte Carlo p-value*. Consequently, a bounds MC test of level α may be defined as:

$$\tilde{p}_N(Z_0) \leq \alpha.$$

The fact that the latter critical region has level α exactly (with respect to the definitions introduced above) obtains from (2.8).

For jump tests, we apply Proposition 2.4 along with Proposition 2.1. Specifically, Z_1, \dots, Z_N may be chosen as N realizations of the no-jump bounding statistic LR_B as defined by (2.1); these may be generated by drawing simulated samples from a no-jump DGP with $a_1 = 1$, $\alpha_1 = \alpha_2 = 0$. To see this, recall that the constrained MLF underlying the LR bounding statistic satisfies $a_1 = 1$, $\alpha_1 = \alpha_2 = 0$. Then (2.7) yields a nuisance-parameter free, level correct, exact p-value for the general LR test. The key here is again that no unknown parameter needs to be dealt with to draw realizations from the bounding statistic LR_B . The dependence on all parameters which intervene in the null distribution of LR (including θ and δ^2) is dealt with through the bound. This takes care of the boundary restriction since the level control

conditions implicit in Proposition 2.4 depend only on the availability of i.i.d. realizations from LR_B .

It is important at this stage to contrast our procedure with a standard parametric bootstrap. As pointed out above, for the jump/GARCH (random walk or stationary AR) case, a parametric bootstrap may be obtained as follows:

1. Generate replications of LR drawing from the no-jump AR/GARCH process with estimated parameters. Constrained maximum likelihood estimates are a natural choice here. Let η denote the set of intervening AR/GARCH nuisance parameters and let $\hat{\eta}$ refer to their relevant estimates. Furthermore, let Z_0 denote the observed LR and let Z_j^c , $j = 1, \dots, N$ refer to the simulated LR , where the superscript c is used here to emphasize conditioning on $\hat{\eta}$.
2. Calculate the empirical p-value $\hat{p}_N(Z_0|\hat{\eta})$, where

$$(2.9) \quad \hat{p}_N(Z_0|\eta) = \frac{N\hat{G}_N(Z_0|\eta) + 1}{N + 1},$$

$$(2.10) \quad \hat{G}_N(Z_0|\eta) = \frac{1}{N} \sum_{i=1}^N I_{[0, \infty]}(Z_i^c - Z_0).$$

3. Reject the null hypothesis if

$$(2.11) \quad \hat{p}_N(Z_0|\hat{\eta}) \leq \alpha.$$

From (2.9), the relation between (2.11) and the p-values underlying Propositions 2.3 and 2.4 becomes clear. One fundamental difference must be born in mind, however. Whereas for a standard parametric bootstrap the series Z_j^c , $j = 1, \dots, N$ underlying (2.11) is generated for a given $\hat{\eta}$, since the simulated statistic is not a pivot, for the MC method, the simulated Z_j , $j = 1, \dots, N$, which appear in Propositions 2.3 and 2.4 are draws from pivotal statistics and do not require estimates of nuisance parameters. This fact has important implications for the properties of tests based on (2.11), in finite sample. Three considerations arise in this regard.

First, since pivotality (the building block of Propositions 2.3 and 2.4), is not preserved with (2.9), the associated bootstrap test is not necessarily level correct in the sense of (2.5): the level property $P_{(H_0)}[\hat{p}_N(Z_0|\hat{\eta}) \leq \alpha] \leq \alpha$ may thus not hold.

Second, the conditions which guarantee the asymptotic validity of 2.11 may also not hold here. Given various regularity conditions, asymptotic validity means that:

$$(2.12) \quad \lim_{T \rightarrow \infty} \{P[\hat{p}_N(Z_0|\hat{\eta}) \leq \alpha] - P[\hat{p}_N(Z_0|\eta) \leq \alpha]\} = 0 ,$$

where $\hat{p}_N(Z_0|\eta)$ is the empirical p-value one would obtain if the nuisance parameters were known. The bootstrap may thus not yield a consistent es-

estimate of the statistic's cut-off point.¹⁰ Dufour (1995) proposes consistency conditions which require $T \rightarrow \infty$ (for a given N) and refers to results from the bootstrap literature on typical conditions which require both $T \rightarrow \infty$ and $N \rightarrow \infty$. In particular, one important condition requires that the statistic's limiting null distribution depends continuously on the intervening parameters; asymptotically pivotal statistics provide a trivial case where this holds. Here, the statistics in question are not asymptotically pivotal due to the presence of the parameters unidentified under the null. Problems may also stem from the AR/GARCH parameters which admit almost unidentified regions (e.g. at the relevant boundaries), thus causing possible discontinuities in the null limiting distribution of the LR.

Third, as demonstrated in Andrews (2000), the same conditions which cause the failure of standard asymptotics when the null hypothesis restricts a parameter to a boundary value (here, the null hypothesis is $\lambda = 0$) may distort the size of tests based on $\hat{p}_N(Z_0|\hat{\eta})$.

For all these reasons, the reliability of the standard bootstrap p-value in finite samples is questionable, whereas our proposed pivot-based procedure is level correct for finite N and T .

Let us emphasize, however, that the above warnings on the standard bootstrap pertain only to rejections. Indeed, Dufour and Khalaf (1997, 2000)

¹⁰The literature on bootstrap refinements is enormous; the reader may refer to Dufour (1995) for useful references.

show formally that non-rejections are *exactly* conclusive in the sense that if $\hat{p}_N(Z_0|\hat{\eta}) > \alpha$, then, *a fortiori*,

$$\left\{ \sup_{\eta} \hat{p}_N(Z_0|\eta) \right\} > \alpha.$$

Now, from Definition 2.1, we can see that a decision based on the largest p-value is level correct, for finite N and T .

This result may be used to circumvent often raised power problems for bounds p-values, since bounds may turn out to be too conservative. Specifically, we recommend using our bounds p-value in conjunction with a parametric bootstrap p-value, as follows:

- If the bounds p-value $\leq \alpha$, then conclude the test is significant.
- If the bootstrap p-value $> \alpha$, then conclude the test is not significant.
- If the bounds p-value exceeds α whereas the bootstrap p-value is less than α , the researcher may adopt a conservative decision based on the former, or a liberal decision, based on the latter.

Such a decision rule, which goes back to the well known Durbin-Watson test, is a sound way to circumvent the pervasive nuisance parameter problem associated with jump tests in discretized jump-diffusion models.

2.4 Generalization

Our methodology can easily be generalized to a wide class of mixed models, written as a combination of a discretized continuous process and a jump process, provided they verify the following assumptions:

1. both the discretized continuous and jump processes are sufficiently specified to permit the formulation of *simulatable* likelihood functions (i.e. it is possible to draw simulated samples from them). Approximate discretizations are valid, provided a *simulated maximum likelihood* (SML) approach is applicable. In what follows, the term *discretized continuous* process implies possibly an approximate discretization. We denote the maxima of the quasi-likelihood functions, with and without jumps, by L_{Mixed} and L_{Smooth} respectively; the term quasi-likelihood is more appropriate since we may have approximate discretizations.
2. the discretized continuous process is nested, possibly imposing boundary constraints, within the mixed model;
3. the discretized continuous process (no jump) is *simulatable*;
4. the mixed model admits a restricted version which is a *simulatable* location-scale model; and finally,
5. the mixed model is *additively separable*, in the sense that the likeli-

hood imposing no-jump constraints does not depend on the nuisance parameters associated with the jump component.

Assumption 1 is fundamental because we adopt quasi maximum-likelihood-based tests, where the quasi-approach results from possibly approximate discretizations. Assumption 2 justifies the use of the quasi-likelihood ratio criterion. Boundary restrictions are typical in the context of jump tests and are formally dealt with here. Assumption 3 is necessary because our proposed methodology requires to draw samples from the relevant no-jump DGP. Assumption 4 relates to our proposed bound and ensures that Proposition 2.1 holds.¹¹ Assumption 5 prevents the presence of nuisance parameters under the null hypothesis of no jumps. It is usually satisfied when the jump component of a mixture model is a Poisson or a Bernoulli process.

In the context of continuous time modelling, where jump-diffusion models have become popular, it is essential to insure that the no-jump component of the mixed process analyzed does converges towards a diffusion as the frequency of observations goes to infinity. Indeed, Drost and Nijman (1993) find that some weak-GARCH models do not converge to diffusions when the frequency of observations goes to infinity.¹² On the other hand, Duan (1997), generalizing Nelson's results (1990), shows that there is a very wide class of

¹¹For discretized jump-diffusion models with GARCH errors, the trendless geometric Brownian motion is a restricted simulatable location-scale model.

¹²See the introduction for a definition of weak GARCH models.

discrete time GARCH models that converge to diffusions as the frequency of observations increases. This is the case for augmented GARCH(p,q) processes that are stationary when mean-adjusted.¹³ The class of augmented GARCH(p,q) models contains many parametric GARCH models such as the LGARCH (Bollerslev 1986), the EGARCH (Nelson 1991), the NGARCH and the VGARCH (Engle and Ng, 1993).

Also allowing for simulatable likelihood functions thus makes our method usable with a very wide class of models. For jump test problems where Assumptions 1 through 5 are satisfied, we can then generalize Proposition

¹³If $\{p_t\}$ is an augmented GARCH(p,q) process, it can be described by the following equations:

$$p_t = \mu_t + h_t \sqrt{\varepsilon_t}, \quad h_t = \begin{cases} |\lambda \phi_t - \lambda + 1|^{\frac{1}{\lambda}}, & \text{if } \lambda \neq 0, \\ \exp(\phi_t - 1), & \text{if } \lambda = 0. \end{cases}$$

$$\phi_t = \sum_{i=1}^p \alpha_1^{(i)} \phi_{t-i} + \sum_{i=1}^q [\alpha_2^{(i)} |\varepsilon_t - c|^\delta + \alpha_3^{(i)} \max(0, c - \varepsilon_t)^\delta] \phi_{t-i}$$

$$+ \sum_{i=1}^q [\alpha_4^{(i)} f(|\varepsilon_t - c|; \delta) + \alpha_5^{(i)} f(\max(0, c - \varepsilon_t); \delta)] + \alpha_0$$

In the above:

- μ_t , the conditional mean, is just required to be F_{t-1} measurable (in practice, it is often a function of h_t); the corresponding mean adjusted process is thus $p_t - \mu_t$;
- ε_t is i.i.d. distributed with zero mean and unit variance and is measurable with respect to F_t ; to derive his convergence results, Duan (1997) makes the additional assumption that ε_t is a standard normal variate;
- F_t is a filtration;
- $f(z; \delta)$ is the well known Box-Cox transformation:

$$f(z; \delta) = \begin{cases} (z^\delta - 1)/\delta, & \text{for } z \geq 0, \delta > 0 \\ Ln(z), & \text{for } z > 0, \delta = 0 \end{cases} ,$$

- finally, $\lambda, \alpha_0, \alpha_1^{(i)}$ ($i = 1, \dots, p$), as well as $\alpha_2^{(j)}, \alpha_3^{(j)}, \alpha_4^{(j)}$, and $\alpha_5^{(j)}$ ($j = 1, \dots, q$) are model parameters.

2.1 as follows:

Proposition 2.5 *Consider a jump test problem where Assumptions 1-5 are satisfied, and let GLR denote the Quasi-LR no-jump criterion:*

$$(2.13) \quad GLR = 2 [L_{Mixed} - L_{Smooth}].$$

Let L_{LS} denote the quasi-MLF associated with the nested location-scale model underlying Assumption 4 and define:

$$(2.14) \quad GLR_B = 2 [L_{Mixed} - L_{LS}].$$

Then, $\forall c$:

$$(2.15) \quad P(GLR \geq c) \leq P(GLR_B \geq c).$$

It follows that the test methodology based on Proposition 2.4 is applicable, using realizations from the null distribution of GLR_B .

To justify the extension of our approach to approximate discretizations is straightforward. First, observe that inequality (2.15) holds simply because a constrained maximum is always less than or equal to its unconstrained counterpart. Whether the objective function is a likelihood or an approximate likelihood does not matter provided the nesting conditions implicit in Assumptions 1-5 are satisfied. Second, the exact size control property im-

plied by inequality (2.8) in Proposition 2.4 does not depend on the fact that statistics are of the LR-type. All that is needed are N realizations from a distribution which bounds - in the sense of (2.7) - the distribution of the statistic under the null hypothesis of interest. This conditions obtains from proposition 2.5. Exactness is thus not lost in a SML framework.¹⁴ For further reference on MC tests in approximate models, see Dufour and Valery (2000).

We do not claim that our proposed bound is the best available. The existence of a bound is, however, an important theoretical result, for as emphasized by Dufour (1997), if a test admits no bound, then it is impossible to correct its size. Thus, if a jump test problem belongs to the above class, our results ensure first that the LR statistic is a valid test criterion and second that it makes sense to attempt to correct its size.

3 Empirical illustration

To illustrate our proposed methodology, we consider weekly observations of spot prices for three commercial commodities (crude oil (West Texas Intermediate, WTI), copper, and nickel), and one precious metal (gold). Daily

¹⁴It is important to distinguish the term *exact test* from the usual reference to *exact discretizations*. The term *exact test* was formally defined above and relates to size control for a finite sample size. Our point here is that exact tests are possible even with approximate likelihood discretizations.

WTI spot prices (provided by Natural Resources Canada), cover the period extending from 01/02/86 to 05/13/99. Daily prices of copper and nickel, obtained from the London Metal Exchange, extend from 01/03/89 to 06/30/99. Daily closing prices of gold, from the New York Metals Exchange, go from 01/03/89 to 10/15/99. Weekly series are constructed from daily data by taking the Wednesday price to avoid beginning or end of the week effects. In the rare instances where the Wednesday price is missing, we use the Tuesday price instead. The four weekly time series analyzed are shown on Figures 1 to 4.

While the geometric Brownian motion (GBM) is often used in finance as a basic model for stock prices or exchange rates, there are theoretical reasons for modeling the spot price of commodities with a mean-reverting process instead. First, the GBM implies that the volatility of spot prices increases without bounds as the time horizon increases. Second, as recalled for example in Schwartz (1997), when the price of a commodity is relatively high, supply tends to increase as higher cost producers enter the market, which ends up lowering prices. Conversely, when prices are relatively low, supply over time tends to decrease as higher cost producers exit the market, which puts an upward pressure on prices. As illustrated by the case of crude oil, the existence of a partial cartel can also lead spot prices to revert to an agreed upon target price.

A casual look at commodity spot prices time series also seems to show that, from time to time, there are changes of unusual magnitude following the arrival of information unanticipated by markets. For crude oil, such changes could be triggered by the announcement of lower output targets by OPEC, or the beginning of a war on the territory of a major oil producer. For gold, it could be the decision by central banks to sell part of their stock on the world market. These changes, whose magnitude is unusually large, could be modelled by a jump process. In addition, as suggested by Vlaar and Palm (1993) in the context of exchange rates, after a jump has taken place, spot price volatility may be high. Volatility may then return to lower values when a new equilibrium is reached. This feature suggests including GARCH errors in our model.

Table 1 presents summary statistics for the logarithm of the spot prices of the series considered. The departures from the normal density are apparent from the high values of the Jarques-Bera statistics, which are all significant at 5 %. Three series are skewed (Copper, Nickel, and Gold), while WTI is not significantly skewed (at 5 %). The excess kurtosis coefficient is significant at 5 % except for Gold.

To model these data, we thus fit a discretized Ornstein-Uhlenbeck (thereafter MRM) process with normally distributed Bernoulli jumps, with and without GARCH(1,1) effects, to the logarithm of the spot price of each of

our commodity spot price time series. Assuming a Bernoulli instead of a Poisson process for jump simplifies the estimation of the parameters, and it is easier to interpret.

If P_t follows a discretized lognormal Ornstein-Uhlenbeck with Bernoulli jumps and GARCH(1,1) errors, then $x_t = Ln(P_t)$ can be written:

$$(3.16) \quad \begin{aligned} x_t &= \mu(1 - e^{-\kappa}) + e^{-\kappa}x_{t-1} + \sqrt{h_t}z_t + \delta_{1n_t} \ln Y_t, \\ h_{t+1} &= \alpha_0 + h_t(\alpha_1 z_t^2 + \alpha_2), \end{aligned}$$

where $z_t \stackrel{iid}{\sim} N(0, 1)$; n_t is either zero or one if a jump occurs between t and $t - 1$; and δ_{1n_t} is 0 if n_t differs from 1, and 1 otherwise. We also assume that the probability of arrival of a jump in a unit time interval is λ , and that the jump size Y_t is (independently) lognormally distributed with mean θ and variance δ^2 . For a Bernoulli jump process, we get (3.16) from (2.1)-(2.2) by setting $a_0 = \mu(1 - e^{-\kappa})$ and $a_1 = e^{-\kappa}$, where $\kappa > 0$. To get a discretized MRM model with jumps and no ARCH effects, we simply set $\alpha_1 = \alpha_2 = 0$. In this case, we designate $\alpha_0 > 0$ by σ^2 . Drost and Werker (Section 5, 1996) or Duan (1997) show that the diffusion limit of the no-jump part of (3.16) is the stochastic volatility model:

$$dp_t = \varsigma(v - p_t)dt + \sigma_t dW_{1,t}$$

$$d\sigma_t^2 = \beta(\alpha - \sigma_t^2)dt + \gamma^2\sigma_t^2dW_{2,t}.$$

Our models are estimated by maximum likelihood using the software package Gauss.¹⁵ Expressions of the likelihood functions are presented in the Appendix. Results are presented in Table 2. We report estimates and standard errors, the LR test, the bootstrap and bounds MC p-values.¹⁶

From Table 2, we first observe that there is ample evidence of statistically significant jumps in each of the time series investigated, with and without GARCH effects. Indeed, both the Monte-Carlo and Monte-Carlo bound p-values are 0.01 for 100 replications for all time series and all models considered. We see that the jump frequency, given by λ , is usually not affected by the inclusion of a GARCH error term. For copper and nickel, we find $\lambda \approx 0.2$ (one jump every 5 weeks), while for gold λ is approximately 0.05 (one jump every 20 weeks). The frequency of arrival of jumps is highest for WTI, where $\lambda \approx 0.37$, which represents a jump at least every third week.

We also notice that the jumpless MRM-GARCH model for gold seem non-

¹⁵As remarked by Ball and Torous (1985), the likelihood functions of jump-diffusion models usually have a local maximum at $\lambda = 0$ (the no jump case). To guard against numerical maximization problems, we use, for observed and simulated bootstrap maximum likelihood functions: (i) the procedure OPTMUM in GAUSS with several starting points of likely parameter values for each iteration (up to 42 for the GARCH model), and (ii) further iterations from a Simulated Annealing algorithm (a global optimizer) to validate convergence. Our GAUSS programs are available upon request.

¹⁶Although we report asymptotic standard errors as is usual in this literature, we warn against their use for a t -tests. As argued in Section 2, asymptotic SE based t -tests may be seriously flawed in the presence of the type of identification problems we are dealing with here. Following Dufour (1997), we would rather use LR tests for hypotheses of interest.

stationary since $\alpha_1 > 1$. In addition, although ARCH effects seem present in the GARCH-MRM models (with and without jumps), the GARCH parameter does not appear to differ from zero.¹⁷ This result is compatible with the findings of Schwartz (1997).

Finally, if we compare our MRM results with those Schwartz (1997) obtained for futures prices, we find much smaller values for the coefficient of mean reversion: κ varies between 0.006 and 0.013 for copper, it is approximately 0.016 for nickel, 0.008 for gold, and 0.04 for WTI. By contrast, Schwartz estimates values of κ of 0.37 for copper and 0.30 for oil for the MRM. This can be explained by the higher volatility of spot prices compared to futures prices, but Schwartz's estimates of κ may also be biased since his approach assumes that spot prices do not contain jumps.

4 Conclusions

When nuisance parameters are unidentified under the null, conventional asymptotics fail even for large samples. Such problems frequently occur in jump/diffusion models. In this paper we propose an approach relying on (exact) boundedly pivotal statistics, which combines exact bounds and MC test procedures based on the LR no-jump test statistic. On theoretical grounds,

¹⁷Formal GARCH and stationarity tests are beyond the scope of this paper. We emphasize that our jump tests are robust to both characteristics.

we establish *explicitly* the criterion's boundedly pivotal characteristic and then we apply simulation-based methods to the LR and bounding statistics to obtain size-correct p -values. Although the problem we consider is highly non-regular, the solution we propose is computationally attractive and it is finite sample exact. It can be generalized to other non identification problems in the presence of nuisance parameters.

We illustrate our methodology with jump tests applied to weekly spot price series for crude oil, copper, nickel, and gold. We fit to these data a mean-reverting process with normally distributed, Bernoulli jumps, with and without a GARCH(1,1) error structure. We find statistically significant jumps for the four time series considered. These findings have implications for pricing derivative instruments, adopting hedging strategies, and calculating the value of portfolios involving these non renewable resources.

Acknowledgments

This work was supported by *Natural Resources Canada* and the *Fonds FCAR* (Government of Québec). The authors also thank Jean-Marie Dufour, John Galbraith, Victoria Zindy–Walsh, participants at the 1999 congress of the *Société Canadienne de Science Économique*, the 1999 and 2000 congresses of the *Canadian Economic Association*, and the 2000 congresses of the *Society for Computational Economics* and the *Econometric Society*. As usual, remaining errors are our responsibility.

References

- [1] Andrews, D. W. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* 61, 821-856.
- [2] Andrews, D. W. K. (1999): "Testing When a Parameter is on the Boundary of the Maintained Hypothesis," Working Paper, Cowles Foundation for Research In Economics, Yale University.
- [3] Andrews, D. W. K. (2000): "Inconsistency of the Bootstrap When a Parameter is on the Boundary of the Parameter Space," *Econometrica* 68, 399-405.
- [4] Andrews, D. W. K. and Ploberger, W. (1994). "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative", *Econometrica* 62, 1383-1414.
- [5] Ahn, C. M. and Thompson, H. E. (1988): "Jump-Diffusion Processes and the Term Structure of Interest Rates," *Journal of Finance* 43, 155-174.
- [6] Akgiray V. and Booth, G. G. (1988): "Mixed Diffusion-Jump Process Modelling of Exchange Rate Movements," *The Review of Economics and Statistics* 70, 631-637.

- [7] Amin, K. I., and Ng, V. K. (1993): "Option Valuation with Systematic Stochastic Volatility," *Journal of Finance* 48, 881-910.
- [8] Ball, C. and Torous, W. (1985): "On Jumps in Common Stock Prices and their Impact on Call Option Pricing," *Journal of Finance* 40, 155-173.
- [9] Barnard, G. A. (1963): "Comment on 'The Spectral Analysis of Point Processes' by M. S. Bartlett," *Journal of the Royal Statistical Society, Series B* 25, 294.
- [10] Bates, D. S. (1996a): "Dollar Jump Fears, 1984-1992: Distributional Abnormalities Implicit in Currency Futures Options," *Journal of International Money and Finance* 15, 65-93.
- [11] Bates, D. S. (1996b): "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *The Review of Financial Studies* 9, 69-107.
- [12] Bera, A. K. and Ra, S. (1995): "A Test for the Presence of Conditional Heteroskedasticity within ARCH-M Framework," *Econometric Reviews* 14, 473-485.
- [13] Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 31, 307-327.

- [14] Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992): "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52, 5-59.
- [15] Brorsen, B. W. and Yang, S. R. (1994): "Nonlinear Dynamics and the Distribution of Daily Stock Index Returns," *The Journal of Financial Research* 17, 187-203.
- [16] Campbell, J. W., Lo A. W. and MacKinlay A. C. (1997): *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.
- [17] Dagenais, M. G. and Dufour, J.-M. (1991): "Invariance, Nonlinear Models, and Asymptotic Tests," *Econometrica* 59, 1601-1615.
- [18] Davies, R. B. (1977): "Hypothesis Testing When a Parameter is Present Only Under the Alternative," *Biometrika* 64, 247-254.
- [19] Davies, R. B. (1987): "Hypothesis Testing When a Parameter is Present Only Under the Alternatives," *Biometrika* 74, 33-43.
- [20] Drost, F. C., and Nijman, T. E. (1993): "Temporal Aggregation of GARCH Processes," *Econometrica* 61, 909-927.
- [21] Drost, F. C., Nijman, T. E. and Werker, B. J. M. (1998): "Estimation and Testing in Models Containing Both Jump and Conditional

- Heteroskedasticity,” *Journal of Business and Economic Statistics* 16, 237-243.
- [22] Drost, F. C. and Werker, B. J. M. (1996): “Closing the GARCH Gap: Continuous Time GARCH Modeling,” *Journal of Econometrics* 74, 31-57.
- [23] Duan, J.-C. (1997): “Augmented GARCH(p,q) Process and its Diffusion Limit,” *Journal of Econometrics* 79, 97-127.
- [24] Dufour, J.-M. (1995): “Monte Carlo Tests in the Presence of Nuisance Parameters with Econometric Applications,” Working Paper, Centre de recherche et développement en économie (C.R.D.E.) and Département de sciences économiques, Université de Montréal.
- [25] Dufour, J.-M. (1997): “Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models,” *Econometrica* 65, 1365-1387.
- [26] Dufour, J.-M., Farhat, A., Gardiol, L., and Khalaf, L. (1998): “Simulation-Based Finite Sample Normality Tests in Linear Regressions,” *The Econometrics Journal* 1, 154-173.
- [27] Dufour, J.-M., and Khalaf, L. (1997): “Simulation-Based Finite and Large Sample Inference Methods in Multivariate Regressions and Seem-

- ingly Unrelated Regressions,” *Cahier* 309, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.
- [28] Dufour, J.-M., and Khalaf, L. (2000): “Monte Carlo Test Methods in Econometrics,” in Baltagi B. (Ed.) *Companion to Theoretical Econometrics*, Blackwell; forthcoming.
- [29] Dufour, J.-M., and Kiviet, J. F. (1998): “Exact Inference Methods for First-Order Autoregressive Distributed Lag Models,” *Econometrica* 66, 79–104.
- [30] Dufour, J.-M., and Valery P. (1995): “Monte Carlo Tests Applied to Models Estimated by Indirect Inference”, Working Paper, Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal.
- [31] Dwass, M. (1957): ”Modified Randomization Tests for Nonparametric Hypotheses,” *Annals of Mathematical Statistics* 28, 181-187.
- [32] Engle, R. F., and Ng, V. K. (1993): ”Measuring and Testing the Impact of News on Volatility,” *Journal of Finance* 48, 1749-1778.
- [33] Hansen, B. E. (1996): “Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis,” *Econometrica* 64, 413-430.

- [34] Heston, S. L. (1993): "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency options," *Review of Financial Studies* 6, 327-343.
- [35] Hull, J, and White, A. (1987): "The Pricing of Options and Assets with Stochastic Volatilities," *Journal of Finance* 42, 281-300.
- [36] Jarrow, R. A. and Rosenfeld, E. R. (1984): "Jump Risks and the Intertemporal Capital Asset Pricing Model," *The Journal of Business* 57, 337-351.
- [37] Jorion, P. (1988): "On Jump Processes in the Foreign Exchange and Stock Markets," *The Review of Financial Studies* 1, 427-445.
- [38] Karlin, S. and H. M. Taylor (1981): *A Second Course in Stochastic Processes*, Academic Press, San Diego.
- [39] Lehmann, E. L. (1986): *Testing Statistical Hypotheses*, 2nd ed. Wiley, New York.
- [40] Merton, R. C. (1976): "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics* 3, 125-144.
- [41] Nelson, D. B. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics* 45, 7-38.

- [42] Nelson, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica* 59, 347-370.
- [43] Schwartz, E. S. (1997): "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *Journal of Finance* 52, 923-973.
- [44] Scott, L. (1987): "Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application," *Journal of Financial and Quantitative Analysis* 22, 419-438.
- [45] Stein, E.M., and Stein, J.C. (1991): "Stock Price Distributions with Stochastic Volatility: An Analytic Approach," *Review of Financial Studies* 4, 727-752.
- [46] Vlaar, P. J. G., and F. C. Palm (1993): "The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroskedasticity, and Jumps," *Journal of Business and Economic Statistics* 11 (3), 351-360.
- [47] Wiggins, J. B. (1987): "Option Values Under Stochastic Volatility: Theory and Empirical Estimates," *Journal of Financial Economics* 19, 351-372.

Table 1: Summary Statistics of the Logarithm of Weekly Spot Prices

	Copper	Nickel	Gold	WTI
Mean	7.721	8.871	5.900	2.922
Standard deviation	0.212	0.307	0.138	0.193
Skewness	-0.453	0.491	-0.453	0.143
Excess kurtosis	-0.516	0.850	0.044	1.399
Jarques-Bera	24.818	38.548	19.279	59.21

The data consists of weekly Wednesday spot prices. There are 548 observations for Copper and Nickel, 563 observations for Gold, and 697 observations for WTI. Following Campbell, Lo and MacKinlay (1997, pages 18-20), the distributions of skewness and excess kurtosis under normality may be approximated as $N(0,6/T)$ and $N(0,24/T)$, respectively. The approximate distribution for Jarques-Bera is $\chi^2(2)$. We see that the skewness, excess kurtosis, and Jarques-Bera statistics reported in Table 1 are all significant at conventional significance levels, except WTI's skewness and Gold's excess kurtosis.

Table 2: Parameter Estimates and Bernoulli Jump Test Results

	Copper		Nickel	
	No Jumps	Jumps	No Jumps	Jumps
MRM				
μ	7.600 (0.125)	7.579 (0.152)	8.734 (0.107)	8.643 (0.175)
κ	0.013 (0.007)	0.011 (0.006)	0.017 (0.006)	0.012 (0.005)
σ	0.032 (0.001)	0.023 (0.002)	0.039 (0.001)	0.029 (0.002)
λ		0.214 (0.098)		0.188 (0.091)
θ		-2E-4 (6.7E-3)		0.002 (0.009)
δ		0.048 (0.008)		0.059 (0.010)
MLE	1110.03	1134.87	1003.82	1025.18
LR	49.67 [0.01, 0.01]		42.72 [0.01, 0.01]	
MRM GARCH				
μ	7.499 (0.298)	7.410 (0.386)	8.685 (0.123)	8.544 (0.207)
κ	0.006 (0.006)	0.006 (0.006)	0.015 (0.006)	0.012 (0.006)
α_0	1.8E-4 (6.5E-5)	7.9E-5 (4.3E-5)	3.2E-4 (1.3E-4)	6.1E-4 (1.9E-4)
α_1	0.119 (0.039)	0.113 (0.045)	0.109 (0.037)	0.075 (0.044)
α_2	0.679 (0.089)	0.699 (0.085)	0.654 (0.108)	0.074 (0.161)
λ		0.199 (0.124)		0.239 (0.120)
θ		0.003 (0.008)		0.007 (0.008)
δ		0.039 (0.008)		0.053 (0.009)
MLE	1136.61	1150.97	1019.75	1032.55
LR	28.71 [0.01, 0.01]		25.59 [0.01, 0.01]	

Standard errors are in parenthesis. We report $[\hat{p}_N(LR|MLE), \tilde{p}_N(LR)]$, i.e. the bootstrap and bounds p-values defined by (2.9) and (2.7) respectively.

Table 2 (Continued): Parameter Estimates and Bernoulli Jump Test Results

	Gold		WTI	
	No Jumps	Jumps	No Jumps	Jumps
MRM				
μ	5.854 (0.090)	5.687 (0.231)	2.909 (0.049)	2.974 (0.042)
κ	0.009 (0.006)	0.005 (0.004)	0.043 (0.011)	0.048 (0.011)
σ	0.019 (0.001)	0.013 (0.001)	0.055 (0.002)	0.030 (0.003)
λ		0.072 (0.022)		0.348 (0.070)
θ		0.008 (0.009)		-0.009 (0.007)
δ		0.053 (0.009)		0.077 (0.006)
MLE	1431.00	1545.58	1043.93	1104.41
LR	229.17 [0.01, 0.01]		120.97 [0.01, 0.01]	
MRM GARCH				
μ	5.548 (0.220)	5.725 (0.177)	2.941 (0.038)	3.040 (0.057)
κ	0.007 (0.004)	0.005 (0.004)	0.042 (0.012)	0.046 (0.012)
α_0	1.8E-4 (1.9E-5)	1.4E-4 (1.5E-5)	1.3E-4 (3.3E-5)	0.0E0 (3.5E-5)
α_1	1.001 (0.177)	0.064 (0.031)	0.137 (0.027)	0.139 (0.030)
α_2	0.000 (- -)	0.000 (- -)	0.807 (0.032)	0.807 (0.035)
λ		0.065 (0.022)		0.374 (0.158)
θ		0.006 (0.010)		-0.012 (0.006)
δ		0.054 (0.009)		0.040 (0.006)
MLE	1460.16	1547.48	1155.39	1168.98
LR	174.63 [0.01, 0.01]		27.19 [0.01, 0.01]	

Standard errors are in parenthesis. We report $[\hat{p}_N(LR|MLE), \tilde{p}_N(LR)]$, i.e. the bootstrap and bounds p-values defined by (2.9) and (2.7) respectively. For Gold, we estimate MRM-ARCH models since α_2 appears to be zero.

Appendix: Likelihood Functions

This appendix provides the expressions of the log-likelihood functions used in this article. $\phi(z)$ designates the density of the standard normal distribution.

As mentioned in the introduction, based on the possibility of entry and exit of producers and on the availability of substitutes, it is likely that the price of non-renewable resources follow a mean reverting process. To obtain a simple and tractable mean reverting process for a price P , we suppose that $X = \ln(P)$ follows the Ornstein-Uhlenbeck process (Karlin and Taylor, 1981): $dX = \kappa(\mu - X)dt + \sigma dw$, where dw is an increment of a standardized Wiener Process. From Karlin and Taylor (1981), we know that, conditional on $X = x_0$ at time 0, X_T is normally distributed with mean $\mu + (x_0 - \mu)e^{-\kappa T}$ and variance $\sigma^2 \frac{(1-e^{-2\kappa T})}{2\kappa}$. If at time t ($0 < t < T$) we add a normally distributed jump (with mean θ and variance σ^2) to X_t , it is easy to show that X_T (conditional on the arrival of a $N(\theta, \sigma^2)$ jump at t) is now normally distributed with mean $\mu + (x_0 - \mu)e^{-\kappa T} + \theta e^{-\kappa(T-t)}$ and variance $\sigma^2 \frac{(1-e^{-2\kappa T})}{2\kappa} + \delta^2 e^{-2\kappa(T-t)}$. To derive $f_{X_T}(\cdot)$, the unconditional (with respect to the time of arrival of a jump) distribution of X_T for a single jump occurring between 0 and T , we assume that the density of the arrival of jumps is the rescaled exponential distribution $\frac{\lambda e^{-\lambda t}}{1-e^{-\lambda T}}$. We obtain:

$$f_{X_T}(x) = \int_0^T \frac{\phi\left(\frac{x_t - \mu(1-e^{-\kappa T}) - x_{t-1}e^{-\kappa T} - \theta e^{-\kappa(T-t)}}{\sqrt{\sigma^2 \frac{1-e^{-2\kappa T}}{2\kappa} + \delta^2 e^{-2\kappa(T-t)}}}\right)}{\sqrt{\sigma^2 \frac{1-e^{-2\kappa T}}{2\kappa} + \delta^2 e^{-2\kappa(T-t)}}} \frac{\lambda e^{-\lambda t}}{1-e^{-\lambda T}} dt$$

Although the integral above cannot be calculated analytically, it is easy to estimate numerically. A simple Gauss-Legendre scheme with four points over the interval $[0, 1]$ (i.e. $T=1$, see below) was found to be sufficiently accurate for the parameter values encountered in this study. Thus, if we assume that jumps arrive according to a Bernoulli process with parameter λ , the log-likelihood function for an Ornstein-Uhlenbeck process with Bernoulli jumps is given by:

$$l_B(\vec{\Theta}_B; \vec{x}) = \sum_{t=1}^T \ln \left[(1-\lambda) \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1}}{h_2}\right)}{h_2} + \lambda \int_0^1 \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1} - h_3 e^{\kappa t}}{\sqrt{h_2^2 + h_4 e^{2\kappa t}}}\right)}{\sqrt{h_2^2 + h_4 e^{2\kappa t}}} \frac{\lambda e^{-\lambda t}}{1-e^{-\lambda}} dt \right]$$

$$\vec{\Theta}_B = (\mu, \kappa, \sigma^2, \lambda, \theta, \delta^2), \quad \kappa \geq 0, \quad \lambda \geq 0,$$

$$h_0 = \mu(1 - e^{-\kappa}), \quad h_1 = e^{-\kappa}, \quad h_2 = \sigma \left(\sqrt{\frac{1 - e^{-2\kappa}}{2\kappa}} \right), \quad h_3 = \theta e^{-\kappa}, \quad h_4 = \delta^2 e^{-2\kappa}.$$

Adding a GARCH(1,1) error structure leads to the log-likelihood function for

an Ornstein-Uhlenbeck with Bernoulli jumps and GARCH(1,1) errors:

$$l_{BG}(\overrightarrow{\Theta_{BG}}; \vec{x}) = \sum_{t=1}^T \ln \left[(1-\lambda) \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1}}{\sqrt{h_t}}\right)}{\sqrt{h_t}} + \lambda \int_0^1 \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1} - h_3 e^{\kappa t}}{\sqrt{h_t + h_4 e^{2\kappa t}}}\right)}{\sqrt{h_t + h_4 e^{2\kappa t}}} \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda}} dt \right]$$

$$\begin{aligned} \overrightarrow{\Theta_{BG}} &= (\mu, \kappa, \alpha_0, \alpha_1, \alpha_2, \lambda, \theta, \delta^2), \quad \kappa \geq 0, \quad \alpha_0 > 0, \quad \alpha_1 \geq 0, \quad \alpha_2 \geq 0, \quad \lambda \geq 0, \\ h_0 &= \mu(1 - e^{-\kappa}), \quad h_1 = e^{-\kappa}, \quad h_3 = \theta e^{-\kappa}, \quad h_4 = \delta^2 e^{-2\kappa}, \\ h_t &= \alpha_0 + \alpha_1(x_{t-1} - h_0 - h_1 x_{t-2})^2 + \alpha_2 h_{t-1}, \quad t = 3, \dots, T. \end{aligned}$$

Since the Monte-Carlo method requires evaluating these likelihood functions a great many times, it is useful to assume that, if there is a jump in the interval (0, 1) it takes place at time 1, in order to reduce computing time substantially. This assumption eliminates the integral in the expression of $f_{X_T}(x)$, which simplifies the expressions of l_B and l_{BG} :

$$l_B^S(\overrightarrow{\Theta_B}; \vec{x}) = \sum_{t=1}^T \ln \left[(1 - \lambda) \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1}}{h_2}\right)}{h_2} + \lambda \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1} - \theta}{\sqrt{h_2^2 + \delta^2}}\right)}{\sqrt{h_2^2 + \delta^2}} \right]$$

$$l_{BG}^S(\overrightarrow{\Theta_{BG}}; \vec{x}) = \sum_{t=1}^T \ln \left[(1 - \lambda) \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1}}{\sqrt{h_t}}\right)}{\sqrt{h_t}} + \lambda \frac{\phi\left(\frac{x_t - h_0 - h_1 x_{t-1} - \theta}{\sqrt{h_t + \delta^2}}\right)}{\sqrt{h_t + \delta^2}} \right]$$

For our empirical illustration, we find that maximizing l_B^S and l_{BG}^S gives almost exactly the same results as maximizing l_B and l_{BG} .

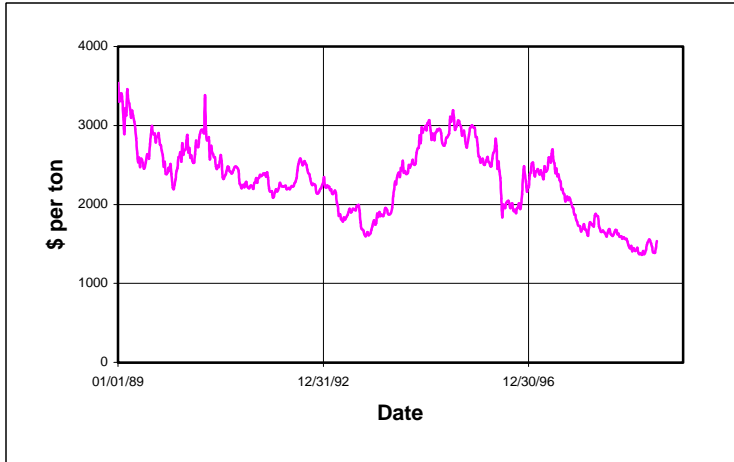


Fig. 1: Weekly Spot Price for Copper

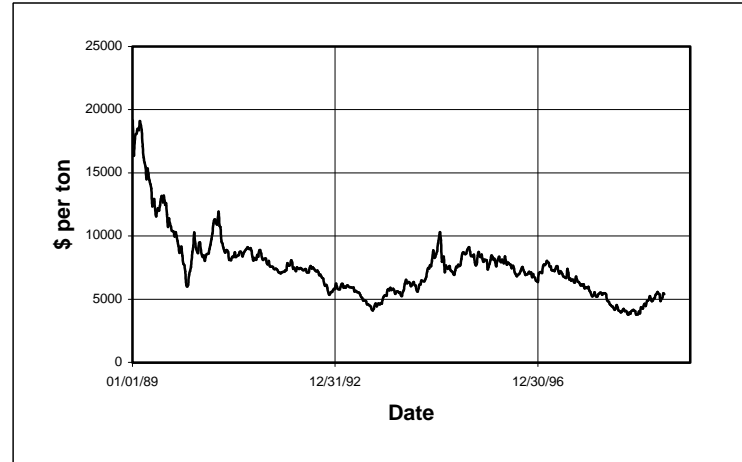


Fig. 2: Weekly Spot Price for Nickel

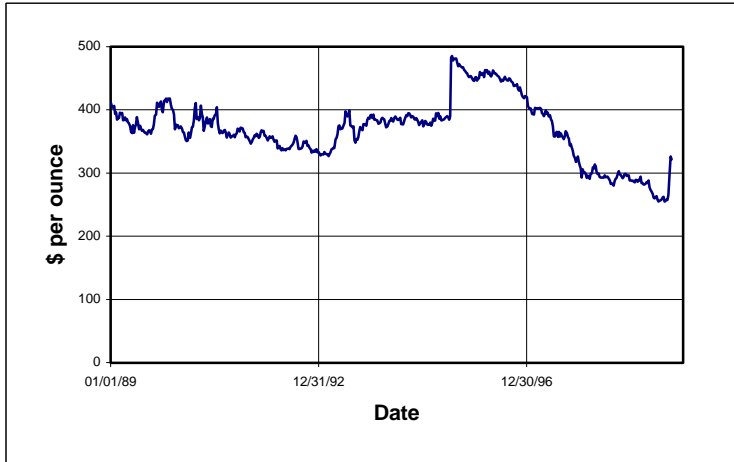


Fig. 3: Weekly Spot Price for Gold

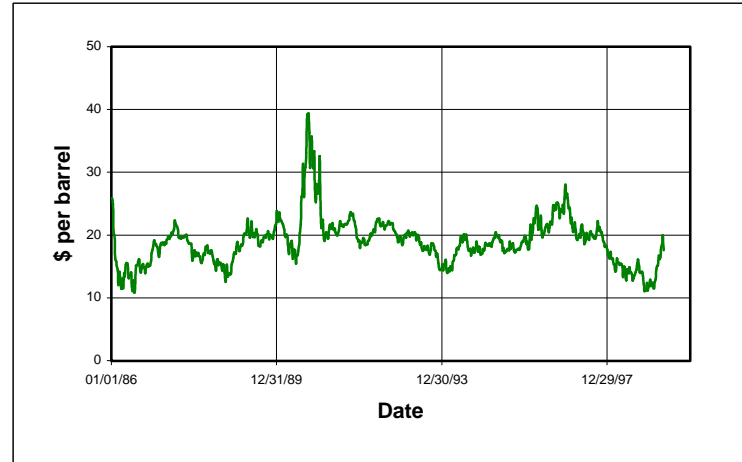


Fig. 4: Weekly Spot Price for WTI