AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

Indirect inference with time series observed with error

Eduardo Rossi and Paolo Santucci de Magistris

CREATES Research Paper 2014-57

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Indirect inference with time series observed with error*

Eduardo Rossi[†]
University of Pavia
Italy

Paolo Santucci de Magistris
CREATES, Aarhus University
Denmark

January 12, 2016

Abstract

We analyze the properties of the indirect inference estimator when the observed series are contaminated by measurement error. We show that the indirect inference estimates are asymptotically biased when the nuisance parameters of the measurement error distribution are neglected in the indirect estimation. We propose to solve this inconsistency by jointly estimating the nuisance and the structural parameters. The range of applicability of this methodology is supported by theoretical results based on several examples for both discrete and continuous-time models. Indirect inference is used to estimate the parameters of stochastic volatility models with auxiliary specifications based on realized volatility measures. Monte Carlo simulations show the bias reduction of the indirect estimates obtained when the microstructure noise is explicitly modeled. Finally, an empirical application illustrates the relevance of a realistic specification of the microstructure noise distribution to match the features of the observed log-returns at high frequencies.

Keywords: Indirect inference, measurement error, stochastic volatility, realized volatility
J.E.L. classification: C13, C15, C22, C58

*We are grateful to Torben Andersen, Roberto Renò, Viktor Todorov, the participants to the 2014 Conference on Indirect Estimation Methods in Finance and Economics (Constance) and SoFiE Conference 2015 (Aarhus) for useful comments on previous versions of this paper. We also thank the Editors and two anonymous referees for very useful suggestions which have improved the earlier version of the paper. Eduardo Rossi acknowledges financial support from the MIUR PRIN project MISURA - Multivariate Statistical Models for Risk Assessment. Paolo Santucci de Magistris acknowledges financial support from CREATES - Center for Research in Econometric Analysis of Time Series, funded by the Danish National Research Foundation (DNRF78).

[†]Dipartimento di scienze economiche e aziendali. Via San Felice 5, 27100 Pavia, Italy. Tel.: +39 0382/986207 Fax: +39 0382/304226. E-mail: erossi@eco.unipv.it.

1 Introduction

A common feature of many economic and financial time series is that they are recorded with errors or frictions. In some cases they are estimated rather than observed exactly. For instance, in macroeconomics the error in the measurement of GDP is a well-known problem, see Aruoba et al. (2013). In finance, the estimates of the relative risk aversion parameter in the C-CAPM are highly dependent on the quality of the measurement of the aggregated consumption, so that proxies for the latter are adopted in empirical studies, see for example Savov (2011). Moreover, asset returns sampled at high frequency are subject to a vast number of frictions, so that the observed transaction log-price is the sum of an unobservable efficient price and a noise component due to the imperfections of the trading process. The analysis of models with errors in variables started historically with the investigation of relations between statistical variables when all or some of them are subject to errors of measurement, e.g. see Anderson (1984). Regressions for time series models with errors in measurement have been discussed by Hannan (1963), Moran (1971), Grether and Maddala (1973) and Robinson (1986). Identifiability problems for such models appear in Anderson and Deistler (1984), Maravall (1979), Nowak (1985) and Solo (1986). In a pure time-series framework, Chanda (1995) studied the identifiability and the estimation of ARMA models which have errors in variables. Chanda (1996) deals with autoregressive models and establishes the asymptotic properties of the OLS estimator based on a set of modified Yule-Walker equations. Staudenmayer and Buonaccorsi (2005) study the estimation of parameters in autoregressive models when measurement errors are uncorrelated but possibly heteroskedastic. Recently, more general hypotheses on measurement error have been considered. For instance, Komunjer and Ng (2014) study the general conditions for identification of single and multiple-equation dynamic models subject to serially correlated measurement error. Song et al. (2015) consider both endogeneity and measurement error in the same variable when the structural model is nonlinear. In a time-series framework, well known econometric methods, such as the Kalman filter or the instrumental variables, can be employed to deal with the measurement error problem and to provide parameter estimates that are consistent also in presence of errors-in-variables. However there are limits to the range of applicability of the standard methodologies. For example, the Kalman filter, in its basic form, can only be adopted in a linear-Gaussian state-space framework. Alternatively, several filtering techniques for non-linear and non-Gaussian cases have been proposed, see for example Durbin and Koopman (2012, part II). However, the practical implementation of these methodologies is often complicated and computationally intensive. Finally, the instrumental variables, such as the lagged observed series, are very weak when the signal is not persistent (see Hansen and Lunde, 2014).

We propose an alternative and general methodology to deal with the errors-in-variables problem that is valid also when the likelihood function (or any other criterion function that might form the basis of estimation) is analytically intractable or too difficult to evaluate. The proposed methodology is based on indirect inference. Indirect inference has been introduced in the econometric literature by Smith (1993), Gouriéroux et al. (1993), Bansal et al. (1995) and Gallant and Tauchen (1996), and is surveyed in Gouriéroux and Monfort (1996) and Jiang and Turnbull (2004) to deal with the problem of estimating the parameters of an economic/financial model, called *structural* model, for which is too complicated to obtain a closed-form expression for the probability distributions associated with the observable variables. The estimation consists of two stages. First, an auxiliary model is estimated on the observed data. Then an analytical or simulated mapping, called *binding function*, of the structural model parameters to the auxiliary statistic is calculated. Indirect inference chooses the parameters of the economic model so that these two estimates of the parameters of the auxiliary model are as close as possible. The indirect inference estimators are typically placed into one of two categories: score-based estimators made popular by Gallant and Tauchen (1996), or distance-based estimators proposed by Smith (1993) and refined by Gouriéroux et al. (1993). The simulated score-based estimators have the computational advantage that the auxiliary parameters are estimated from the observed data only once. On the other hand, the distance-based estimators must re-estimate the auxiliary parameters from simulated data at each step of the optimization algorithm. The indirect inference methods have been successfully employed in the estimation of continuous-time models for asset prices and volatility, that have experienced important developments in the last twenty years, see among others Gallant et al. (1997), Chernov et al. (2003). Since the transition density functions are often unknown and the stochastic volatility (SV) process is not directly observable, the estimation with latent variables can be carried out using simulation methods.

In this article, we first formally study how neglecting the presence of measurement error in the observed time series affects the indirect inference estimates. We shows that the errors-in-variables cause the auxiliary model to produce an inconsistent functional estimator of the theoretical binding function, as the latter depends on the nuisance parameters governing the noise. This discrepancy makes the indirect inference estimates inconsistent and biased in finite samples. Indeed, as noted by Ghysels and Khalaf (2003) and Dridi et al. (2007) indirect inference theory, as originally proposed by Gouriéroux et al. (1993), Gouriéroux and Monfort (1996) and Gallant and Tauchen (1996), does not take nuisance parameters formally into consideration. The theoretical analysis is also supported by examples based on ARMA and continuous time stochastic processes.

Second, we propose a simple solution to the inconsistency of the indirect inference estimator that is to explicitly account for the presence of measurement error, and to treat it as a structural feature. This means that the nuisance parameters characterizing the conditional distribution of the noise must be estimated jointly with the structural parameters. This implies that the simulated trajectories of the structural model must be contaminated by measurement error. It follows that the simulated binding function will explicitly depend on the noise parameters, thus leading to a consistent matching of the auxiliary estimates. The main advantage of this approach is that it is fully built within the indirect inference framework, so that the asymptotic consistency and normality of this estimator can be easily proved under correct specification of the contaminating processes. Close to the approach taken here is the simulated minimum distance by Gospodinov et al. (2015) proposed for the case of ARDL models with measurement errors when external instruments may not be available or are weak.

Third, the issue of identification is discussed in detail. Indeed, while the indirect inference framework provides a general setup to tackle the problem of measurement errors, choosing an auxiliary model able to identify both the structural and the noise parameters may be non trivial. A crucial assumption in the indirect inference framework is that the binding function must be locally injective to guarantee identification. We prove that when the structural model is an ARMA(r, l) with $r > l$ contaminated by i.i.d. noise with non-zero variance, the Jacobian of the binding function generated by autoregressive auxiliary models with $m > r + l$ lags has full-column rank. Since many economic and financial models have an ARMA representation, this result gives a necessary condition for the identifiability of the structural models in presence of measurement error.

Fourth, the proposed method is employed in the estimation of continuous-time SV models based on ex-post measurement of daily volatility like realized volatility (RV). When the volatility is generated by the Heston (1993) model, then the binding function relative to the HAR-RV model of Corsi (2009) can be written in terms of the SV parameters so that the identification condition is formally verified. The analysis is further extended to the case of leverage, drift and price jumps, modeled as a compound Poisson process with independent Gaussian innovations. We show that all parameters of the Heston model with drift and jumps, including those governing the jumps and the microstructure noise, are identifiable by a multivariate auxiliary model for daily returns, RV and signed jump variations with the leverage parameter is identified by the contemporaneous correlation between daily returns and RV. A set of Monte Carlo simulations confirm that accounting for the microstructure noise produces unbiased SV parameter estimates, when employing intraday returns at very high frequencies (e.g. 30 seconds). On the contrary, neglecting the microstructure noise leads to severe biases especially in the estimates of the long-run mean of volatility. Alternatively, sampling returns at low intradaily frequencies attenuates the impact of the market microstructure noise on the volatility estimates. In terms of efficiency of the estimates, sampling at the highest frequency has the advantage that no data is discarded. On the other hand, additional nuisance parameters need to be estimated.

Finally, an empirical study based on the RV series of JP Morgan corroborates the evidence emerged in the theoretical study and highlights the advantages and limits of the proposed methodology. In particular, neglecting the microstructure noise makes the estimates of the SV parameters highly dependent on the choice of the sampling frequency adopted in the construction of RV. Indeed, it is not possible to reconcile the estimates of the auxiliary parameters obtained with RV based on log-returns sampled at 5 seconds and 5 minutes, unless the microstructure noise is explicitly modeled. Moreover, using the data at very high frequency allows to verify if the assumed structural model and the generation of the microstructure noise are coherent descriptions of the properties of the observed log-returns. Sampling at low frequencies does not provide sufficient information on the price generation mechanism in a realistic scenario and makes the identification of price jumps very difficult. On the contrary, sampling the returns at 5-seconds allows to evaluate the impact of price decimalization on the indirect inference estimates. This also explains why in previous studies, based on daily returns only, the jump parameters were weakly identified, and hence were constrained, see e.g. Chernov et al. (2003). It turns out that, when the signed jump variation is adopted to disentangle price jumps from volatility dynamics and microstructure noise is in the form of a bid-ask spread, the fit of the model is dramatically improved.

The paper is organized as follows. Section 2 illustrates the effect of the presence of measurement error on the indirect inference estimator in the pure time series case. We formally prove the inconsistency of the indirect inference estimates in this case and we illustrate it by means of several examples. In Section 3 we prove the consistency of the proposed estimator that accounts for the noise and we illustrate its reliability in finite samples by means of a number of example both in discrete and in continuous time. In Section 4 we illustrate, both theoretically and by means of Monte Carlo simulations, the peculiar problems that arise in the indirect estimation of continuous-time SV model by means of RV, and Section 5 reports the evidence based on real data. Finally, Section 6 concludes.

2 The effect of measurement error on the indirect inference estimator

Following the framework and notation of Gouriéroux et al. (1993), we first present the properties of the indirect inference method in presence of measurement error. The parameters of interest are those in the vector θ which characterizes the data-generating process of the unobserved series y_t. Here we consider the case of a discrete-time process y_t which is contaminated by an error term. To simplify the notation and the exposition of the results, we consider only the dependence on past values of y_t, i.e. the pure time series case.

Assumption 1 *The process $\{y_t\}$ is a strictly stationary and ergodic process with transition density $p(y_t|\underline{y_{t-1}};\theta)$, where $\underline{y_{t-1}} = (y_{t-1}, \dots, y_{t-l})$, that is difficult or impossible to evaluate analytically. The vector containing the true structural parameters is $\theta_0 \in \Theta \subseteq \mathbb{R}^p$.*

Assumption 2 *A sample of T observations $\{x_t\}_{t=1}^T$ is observed as*

$$x_t = g(y_t, u_t) \quad t = 1, 2, \dots, T \tag{1}$$

Assumption 3 *The term u_t is the measurement error which is supposed to be covariance stationary with a known conditional distribution, i.e. $f(u_t|u_{t-1}, u_{t-2}, \dots; \psi_0)$, where $\psi_0 \in \Psi \subseteq \mathbb{R}^h$.*

Assumption 1 is rather standard in this framework, as indirect inference requires the process y_t to be stationary with constant moments to be asymptotically valid. Assumption 2 is quite general, it allows a non-linear mapping between the observed series x_t, the signal y_t and u_t. Generally $g(\cdot)$ is a linear/additive function or can be reduced to be linear, i.e. $x_t = y_t + u_t$. For example, suppose that the observed stock price P_t is equal to a latent efficient price times an error term with positive support, $P_t = P_t^* \cdot \tilde{\epsilon}_t$, then the efficient log-price, p_t^*, is contaminated by an additive measurement error term, i.e. $p_t = p_t^* + \epsilon_t$, where $\epsilon_t = \log(\tilde{\epsilon}_t)$. Assumption 3 characterizes the dynamic features of measurement error, which depends on a number of true nuisance parameters, contained in the vector ψ_0, and it does not exclude correlation between the signal and the noise and autocorrelation in u_t. As it will be clear from the discussion in Section 3, the knowledge of the functional form of the conditional distribution of u_t is crucial when we want to simulate from it.

We now investigate the impact of neglecting the possible presence of measurement error when carrying out indirect inference on θ_0 employing observations of the contaminated process x_t. The indirect inference consists of two steps: the estimation of the auxiliary (or instrumental) model and the calibration. The auxiliary model is defined by a conditional probability density function $f(x_t|\underline{x_{t-1}}; \beta)$ which depends on a q-dimensional parameter vector, $\beta \in \mathcal{B} \subseteq \mathbb{R}^q$. This density has a convenient analytical expression. The number of parameters in the auxiliary model must be at least as large as the number of parameters in the economic model, i.e., $q \geq p$. The auxiliary model is, in general, incorrectly specified, i.e. need not describe accurately the conditional distribution of x_t. The parameters of the auxiliary model can be estimated using the observed data by maximizing the log-likelihood function or any other criterion function, $Q_T(\underline{x_T}; \beta)$, which satisfies some technical assumptions, see Gouriéroux and Monfort (1996, p.85), i.e.

$$\hat{\beta}_T = \arg\max_{\beta} Q_T(x_1, \ldots, x_T; \beta). \tag{2}$$

In a likelihood setting, identification requires the true densities of the data being "smoothly embedded" within the scores of the auxiliary model, see Gallant and Tauchen (1996). The criterion is assumed to tend asymptotically (and uniformly almost certainly) to a non-stochastic limit (see Gouriéroux et al., 1993, Assumption 2)

$$\lim_{T\to\infty} Q_T(x_1, \ldots, x_T; \beta) = Q_\infty(\theta_0, \psi_0, \beta). \tag{3}$$

When the series is measured without noise, this limit depends only on the unknown auxiliary parameter β and on the true parameter of interest $\theta_0 \in \Theta \subseteq \mathbb{R}^p$. However when the series at hand is contaminated by noise this limit depends also on the true nuisance parameter vector, $\psi_0 \in \Psi \subseteq \mathbb{R}^h$, where h is the dimension of ψ_0. For example, when u_t is assumed to be an $i.i.d.N(0, \sigma_u^2)$, then $\psi_0 = \sigma_{u,0}^2$ and $h = 1$. As in Gouriéroux et al. (1993) we assume that this limit criterion is continuous in β and has a unique maximum

$$\beta_0 = \arg\max_{\beta\in\mathcal{B}} Q_\infty(\theta_0, \psi_0, \beta)$$

The binding function, i.e. the link between the auxiliary model parameters and the structural parameters, is given by

$$b(\theta, \psi) = \arg\max_{\beta\in\mathcal{B}} Q_\infty(\theta, \psi, \beta) \tag{4}$$

it follows that

$$\beta_0 = b(\theta_0, \psi_0).$$

$\hat{\beta}_T$ is a consistent estimator of $b(\theta_0, \psi_0)$ which is an unknown function that depends on θ_0 and ψ_0.

In the second step of the procedure, we simulate S trajectories from the DGP of y_t and the estimation of the auxiliary model is carried out on each simulated series. The auxiliary estimator based on the s-th simulated path of the signal's DGP for some θ is

$$\hat{\beta}_T^s(\theta) = \arg\min_{\beta\in\mathcal{B}} Q_T(y_1^{(s)}(\theta), \ldots, y_T^{(s)}(\theta); \beta).$$

When $T \to \infty$, $\hat{\beta}_T^s(\theta)$ converges to the solution of the limit problem

$$\tilde{b}(\theta) = \arg\max_{\beta \in \mathcal{B}} Q_\infty(\theta, \beta)$$

that is

$$\lim_{T\to\infty} \hat{\beta}_T^s(\theta) = \tilde{b}(\theta)$$

therefore $\hat{\beta}_T^s(\cdot)$ is an inconsistent functional estimator of $b(\theta, \psi)$.

Proposition 2.1 *If the auxiliary estimator has the following Edgeworth expansion, given Assumptions 1-3,*

$$\hat{\beta}_T(\theta_0, \psi_0) = b(\theta_0, \psi_0) + \frac{A(v_x; \theta_0, \psi_0)}{\sqrt{T}} + \frac{B(v_x; \theta_0, \psi_0)}{T} + o\left(\frac{1}{T}\right) \tag{5}$$

then the indirect inference estimator can be expressed as

$$\hat{\theta}_{ST}(\psi_0) = \theta_0 + \frac{a^*}{\sqrt{T}} + \frac{b^*}{T} + o\left(\frac{1}{T}\right) \tag{6}$$

where

$$a^* = \sqrt{T}\left[\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}\right]^{-1} [b(\theta_0, \psi_0) - \tilde{b}(\theta_0)] + \left[\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}\right]^{-1} [A(v_x; \theta_0, \psi_0) - \frac{1}{S}\sum_{s=1}^{S} A(v_s; \theta_0)] \tag{7}$$

and

$$\begin{aligned} b^* = & \left[\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}\right]^{-1} \left[B(v_x; \theta_0, \psi_0) - \frac{1}{S}\sum_{s=1}^{S} B(v_s; \theta_0)\right] \\ & - \left[\frac{\partial b}{\partial \theta'}(\theta_0, \psi_0)\right]^{-1} \left\{\frac{1}{S}\sum_{s=1}^{S} \frac{\partial A(v_s; \theta_0)}{\partial \theta'} a^* + \frac{1}{2}\left(I_q \otimes a^*\right)' F(\theta_0) a^*\right\}, \end{aligned} \tag{8}$$

where

$$F(\theta) = [F_1(\theta)', \ldots, F_q(\theta)']'$$

with $F_i(\theta) = \frac{\partial \tilde{b}_i(\theta)}{\partial\theta\partial\theta'}$, *and* $\tilde{b}_i(\theta)$ *the i-th element of* $\tilde{b}(\theta)$.[1]

The consequence of Proposition 2.1 is that the indirect inference estimator of θ, found by minimizing the distance between $\hat{\beta}_T$ and $\frac{1}{S}\sum_{s=1}^{S} \hat{\beta}_T^s(\theta)$ under a metric given by the positive definite matrix Ω, as

$$\hat{\theta}_{ST}(\psi) = \arg\min_\theta \left\| \hat{\beta}_T - \frac{1}{S}\sum_{s=1}^{S} \hat{\beta}_T^s(\theta) \right\|_\Omega^2$$

is inconsistent. The estimator $\hat{\theta}_{ST}(\psi)$ depends on the data via $\hat{\beta}_T$, and thus on the nuisance parameter ψ. Indeed, the limit of $\hat{\beta}_T$ as $T \to \infty$ is $b(\theta_0, \psi_0)$. Instead, when $S \to \infty$, $\frac{1}{S}\sum_{s=1}^{S} \hat{\beta}_T^s(\theta) \xrightarrow{p} E[\hat{\beta}_T^s(\theta)]$ and $E[\hat{\beta}_T^s(\theta)] = \tilde{b}(\theta)$. This discrepancy induces an asymptotic bias in the indirect inference estimator. When $p = q$, it is possible to show that $p\lim \hat{\theta}_{ST} = \theta_0 + \left[\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}\right]^{-1} [b(\theta_0, \psi_0) - \tilde{b}(\theta_0)]$.[2] This makes clear that the term responsible for the distortion and the inconsistency of $\hat{\theta}_{ST}$, i.e.

[1] The proof of Proposition 2.1 is in the document with supplementary material.

[2] Note that the short-hand notation $\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}$ indicates $\left.\frac{\partial \tilde{b}(\theta)}{\partial \theta'}\right|_{\theta=\theta_0}$. Moreover, the document with additional material reports the details on the derivation of this result.

$\left[\frac{\partial\tilde{b}(\theta_0)}{\partial\theta'}\right]^{-1}[b(\theta_0,\psi_0)-\tilde{b}(\theta_0)]$, does not vanish asymptotically since $b(\theta_0,\psi_0)$ depends also on ψ_0. Notably, this result is general and it holds for any misspecification that involves nuisance parameters, ψ. The solution to this inconsistency is therefore based on the idea of treating measurement error as a potential source of misspecification, thus considering it as a structural feature, see Section 3.

In the following subsections we present a few examples where the distortion in the binding function induced by the presence of measurement error can be computed analytically. In the analysis, the framework consists of a linear measurement equation:

$$x_t = y_t + u_t, \quad t = 1, 2, \ldots, T \tag{9}$$

where u_t is supposed to be independent and identically distributed with $\mathrm{Var}[u_t] < \infty$ and $\mathrm{Var}[u_t] = \sigma_u^2$, independent of all leads and lags of y_t.

2.1 ARMA processes

The identifiability and estimation of ARMA processes contaminated by additive i.i.d. noise has been studied in literature, see among others Chanda (1995), Jones (1980), Lee and Shin (1997) and Maravall (1979). More recently, de Luna and Genton (2001) have proposed an indirect estimation method to robustify the estimation of ARMA under outliers contamination. In this section, we show the consequences for the indirect estimation of ARMA models when the signal is contaminated by measurement error. We derive closed-form expressions of the binding function when an autoregressive auxiliary model is adopted and the ARMA signal is contaminated by an i.i.d. measurement error.

Example 2.1 *Consider a zero-mean stationary ARMA(1,1) signal* y_t

$$(1-\alpha L)y_t = (1+\varphi L)\varepsilon_t, \quad \varepsilon_t \sim i.i.d.N(0,\sigma_\varepsilon^2)$$

The polynomials $1-\alpha z$ *and* $1+\varphi z$ *have roots outside the unit circle. Given the result of Granger and Morris (1976),* x_t *is also an ARMA(1,1). Since measurement error is uncorrelated with constant variance, the error leads to an increase in the variance of* x_t*. The parameter vector of the ARMA(1,1) is* $\theta = (\alpha, \varphi, \sigma_\varepsilon^2)'$*. Suppose the auxiliary model is an AR(2) process*

$$(1-\phi_1 L - \phi_2 L^2)x_t = e_t \tag{10}$$

with parameters in $\beta = (\phi_1, \phi_2, \sigma_e^2)'$ *that can be estimated by OLS. In this case,* $p = q$ *and the model is exactly identified. Let* $\phi = (\phi_1, \phi_2)'$*, the binding function is*

$$b(\theta, \sigma_u^2) = p\lim_{T\to\infty}\begin{bmatrix}\hat{\phi}_T \\ \hat{\sigma}_{e,T}^2\end{bmatrix}$$

with

$$p\lim\hat{\phi}_T = \frac{(\alpha+\varphi)(1+\alpha\varphi)}{(1-\alpha^2)\left[\left(\frac{1+\varphi^2+2\alpha\varphi}{1-\alpha^2}+\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right)^2-\left(\frac{(\alpha+\varphi)(1+\alpha\varphi)}{1-\alpha^2}\right)^2\right]}\times\begin{bmatrix}\left(\frac{1+\varphi^2+2\alpha\varphi}{1-\alpha^2}+\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right)-\frac{\alpha(\alpha+\varphi)(1+\alpha\varphi)}{1-\alpha^2} \\ \alpha\left(\frac{1+\varphi^2+2\alpha\varphi}{1-\alpha^2}+\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right)-\frac{(\alpha+\varphi)(1+\alpha\varphi)}{1-\alpha^2}\end{bmatrix} \tag{11}$$

and

$$p\lim\hat{\sigma}_{e,T}^2 = \left(\sigma_\varepsilon^2\frac{1+\varphi^2+2\alpha\varphi}{1-\alpha^2}+\sigma_u^2\right)\left(1-\left(\sigma_\varepsilon^2\frac{(\alpha+\varphi)(1+\alpha\varphi)}{1-\alpha^2}\right)^2+\left(\sigma_\varepsilon^2\frac{\alpha(\alpha+\varphi)(1+\alpha\varphi)}{1-\alpha^2}\right)^2\right). \tag{12}$$

The function $\tilde{b}(\theta)$*, which is the limit of the estimators of the auxiliary parameters based on the simulated data, is simply obtained by setting* $\sigma_u^2 = 0$ *in (11) and (12), while when* $\sigma_u^2 > 0$*, the difference* $b(\theta, \sigma_u^2)$*-*$\tilde{b}(\theta)$ *is non zero as* $T \to \infty$*.*

2.2 Continuous-time models

Let suppose that we are interested in estimating the parameter vector θ_0 which characterizes now the transition density, often unknown, of a continuous-time process $\{z(t)\}$. Denote with y_t the function of discrete observations on $z(t)$

$$y_t = h(z(t); \Delta), \quad t = 1, 2, \ldots, T$$

where $h(\cdot; \Delta)$ is a known function and Δ is a known parameter which represents the discretization step. In a simple case, the series y_t can be the discretized $z(t)$ process. It is well known that when y_t is observed in place of $z(t)$ standard indirect inference procedures corrects for the discretization error, see Gouriéroux et al. (1993) and Broze et al. (1998). However, there are cases where transformations of discrete realizations of the process $z(t)$, employed in the indirect estimation of continuous-time models, are likely to be contaminated by measurement error. In other words, the observed process x_t is the result of the interaction between the latent signal, $z(t)$, the measurement error, u_t, the function $h(\cdot; \Delta)$ and the discretization step, Δ. Hence, x_t is

$$x_t = g\left[h(z(t); \Delta), u_t\right], \quad t = 1, 2, \ldots, T$$

In cases like this, it is often impossible to obtain a closed form expression of the likelihood function for the parameters which characterize the process $z(t)$ and u_t based only on the observation of x_t. The examples presented below illustrate the impact that measurement error has on the indirect inference estimator.

Example 2.2 *An Ornstein-Uhlenbeck process is the solution of the differential equation:*

$$\mathrm{d}z(t) = k(\omega - z(t))\mathrm{d}t + \sigma \mathrm{d}W(t), \quad t > 0 \tag{13}$$

where $k, \omega, \sigma \geq 0$ and $W(t)$ is a standard Brownian motion on $\mathbb{R}$. The initial value of $z(0)$ is a given random variable (possibly, a constant) taken to be independent of $\{W(t)\}_{t\geq 0}$. The data $z(t)$ is recorded discretely at points $(\Delta, 2\Delta, \ldots, n\Delta)$ in the time interval $[t-1, t]$ with $t = 1, \ldots, T$, that is $z_{t-1+i\Delta}$, for $i = 1, \ldots, n = 1/\Delta$. Let assume without loss of generality that $\Delta = 1$ so that $y_t = z_t$ is the discrete realization of $z(t)$ on the unit interval. The observed series, x_t, is given by (9). As in Gouriéroux and Monfort (1996), the auxiliary model is

$$x_t = x_{t-1} + \beta_1(\beta_2 - x_{t-1}) + \beta_3 e_t, \quad e_t \sim i.i.d.N(0, 1) \tag{14}$$

where the set of auxiliary parameters is $\beta = (\beta_1, \beta_2, \beta_3)'$. Let $\theta_0 = (k_0, \omega_0, \sigma_0)'$ be the vector of unknown true model parameters. The asymptotic bias of the indirect inference estimator $\hat{\theta}_{ST}$ can be derived noting that $\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}^{-1}[b(\theta_0, \psi_0) - b(\theta_0)]$ responsible for the inconsistency is equal to[3]

$$\begin{bmatrix} e^{-k_0} & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\sigma}{2}\left[\left(\frac{2k_0}{1-e^{-2k_0}}\right)^{1/2}\left(\frac{e^{-2k_0}}{k_0} - \frac{1-e^{-2k_0}}{2k_0^2}\right)\right] & 0 & \left(\frac{1-e^{-2k_0}}{2k_0}\right)^{1/2} \end{bmatrix}^{-1} \times$$
$$\begin{bmatrix} e^{-k_0} - \frac{e^{-k_0}\sigma_0^2}{\sigma_0^2+2k_0\sigma_{u,0}^2} \\ 0 \\ \left\{\left(\frac{\sigma_0^2}{2k_0} + \sigma_{u,0}^2\right)\left[1 + \left(\frac{e^{-k_0}\sigma_0^2}{\sigma_0^2+2k_0\sigma_{u,0}^2}\right)^2\right] - \left(\frac{e^{-k_0}\sigma_0^2}{\sigma_0^2+2\kappa_0\sigma_{u,0}^2}\right) e^{-k_0}\frac{\sigma_0^2}{k_0}\right\}^{1/2} - [(\frac{\sigma_0^2}{2k_0}(1 - e^{-2k_0}))]^{1/2} \end{bmatrix} \tag{15}$$

[3]Details on the derivation of an expression of $\frac{\partial \tilde{b}(\theta_0)}{\partial \theta'}^{-1}[b(\theta_0, \psi_0) - \tilde{b}(\theta_0)]$ for the Ornstein-Uhlenbeck process are in the document with supplementary material.

that is non null when $\sigma^2_{u,0} > 0$. Therefore, the indirect inference estimates of k and σ are asymptotically biased. Interestingly, the estimator of the long-run mean parameter ω is not affected by measurement error. This is a consequence of the fact that the error term is an additive i.i.d process with zero mean, so that x_t and y_t have the same long-run mean. Moreover, when $\sigma^2_{u,0} = 0$, i.e. measurement error is absent, the vector $[b(\theta_0, \psi_0) - \tilde{b}(\theta_0)]$ in (15) is zero so that $\hat{\theta}_{ST}$ is consistent.

3 A consistent indirect inference estimator

So far, we have shown that neglecting the measurement error generates a bias in the indirect inference estimator, as a result of an inconsistent functional estimation of the binding function. Therefore, a straightforward solution to the inconsistency caused by the presence of measurement error is to consider the nuisance parameters ψ among the structural parameters that need to be estimated. The $(p+h) \times 1$ parameter vector to be estimated is now denoted by $\zeta = (\theta', \psi')'$. The parameter space of ζ is $\mathcal{Z}$. The auxiliary model is characterized by a criterion function $Q_T(\underline{x}_T, \beta)$, where $\beta \in \mathcal{B}$ with $\mathcal{B}$ compact subset of $\mathbb{R}^q$, with $q \geq p+h$. The proposed indirect inference procedure requires that we can simulate trajectories from the structural model contaminated by measurement error. This means that we simulate $y_t^{(s)}$ from the structural model and the measurement error from the assumed conditional density. The contaminated artificial series, i.e. $x_t^{(s)} = y_t^{(s)} + u_t^{(s)}$, are used in place of $y_t^{(s)}$, thus

$$\hat{\beta}^s_T(\zeta) = \arg\min_{\beta} Q_T(x_1^{(s)}(\zeta), \ldots, x_T^{(s)}(\zeta); \beta).$$

The estimated binding function, $\hat{b}(\zeta)$, which now explicitly depends on θ and ψ, is used to match both sets of parameters.

Assumption 4.v (see Appendix A.1) guarantees that $b(\zeta)$ is locally identified, so that the equation $\beta = b(\zeta)$ admits a unique solution in ζ at the true parameter value, ζ_0. The indirect estimator of ζ, i.e. of the structural and nuisance parameters, is obtained as

$$\hat{\zeta}_{ST} = \arg\min_{\zeta} \Xi(\zeta) \tag{16}$$

with $\Xi(\zeta) = \left\| \hat{\beta}_T - \frac{1}{S}\sum_{s=1}^S \hat{\beta}^s_T(\zeta) \right\|^2_{\Omega}$. Indeed, Assumptions 1-4 guarantee that the indirect inference problem at hand is standard, so that the asymptotic distribution of $\hat{\zeta}_{ST}$ is the same as in Proposition 3 in Gouriéroux et al. (1993), see Appendix A.1. In other words, indirect inference provides asymptotically unbiased and normal estimates in presence of measurement error, if the parameters governing the latter are considered among the structural ones and pseudo-data can be simulated from the contaminated structural model. If the auxiliary model is such that it identifies all structural parameters, i.e. Assumption 4.v is satisfied, then standard theory applies.

In the next Section, we analyze the identification condition for ARMA models. We establish a necessary condition for the identification of the ARMA parameters and the variance of the measurement error, i.e. that Assumption 4.v holds.

3.1 Identification of ARMA models

This section briefly discusses the identification condition for the indirect estimation of a stationary ARMA(r, l) signal observed with additive measurement error:

$$\begin{aligned} x_t &= y_t + u_t \quad t = 1, \ldots, T \\ \alpha(L) y_t &= c + \varphi(L)\varepsilon_t \quad \varepsilon_t \sim i.i.d. N(0, \sigma^2_\varepsilon) \end{aligned} \tag{17}$$

We assume that the roots of $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_r L^r$ and $\varphi(L) = 1 + \varphi_1 L + \ldots + \varphi_l L^l$ all lie outside the unit circle, and there are no common roots. Note that

$$\alpha(L)x_t = c + \varphi(L)\varepsilon_t + \alpha(L)u_t \tag{18}$$

is an ARMA(r,max $\{r,l\}$). The parameter vector to be estimated is $\zeta = (c, \alpha_1, \ldots, \alpha_r, \varphi_1, \ldots, \varphi_l, \sigma_\varepsilon^2, \sigma_u^2)'$. The auxiliary model is an AR(m), i.e.

$$\phi(L)x_t = \phi_0 + e_t, \tag{19}$$

where $E(e_t) = 0$, $E(e_t^2) = \sigma_e^2$ and $\phi(L) = (1 - \phi_1 L - \ldots - \phi_m L^m)$. The $(q \times 1)$ vector of auxiliary parameters is $\beta = (\phi_0, \phi_1, \ldots, \phi_m, \sigma_e^2)'$, with $q = (m+2) \geq p + 1 = 3 + r + l$.

Lemma 2.5 in Chanda (1995) makes clear that the identifiability of $\varphi_1, \ldots, \varphi_l$, σ_ε^2 and σ_u^2 is possible if and only if $r > l$. In the following proposition we explicit the identification condition of ζ_0 (Assumption 4.v) when the auxiliary model is an AR(m).

Proposition 3.1 *Let the structural model be the stationary ARMA(r,l) in* (17) *with measurement error* $u_t \sim WN(0, \sigma_u^2)$ *and the auxiliary model be the AR(m) in* (19) *with* $m > r + l$. *The binding function is*

$$b(\zeta) = \begin{bmatrix} Q_{ZZ}^{-1} Q_{ZX} \\ Q_{XX} - Q_{XZ} Q_{ZZ}^{-1} Q_{ZX} \end{bmatrix}$$

where Q_{ZZ}, Q_{XX} *and* Q_{ZX} *contain the mean, the variance and the autocovariances of the process for* x_t *in* (18) *up to lag* m. *Then the Jacobian matrix*

$$\frac{\partial b(\zeta)}{\partial \zeta'} = \begin{bmatrix} -(Q_{ZX}' Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1}) \frac{\partial vec Q_{ZZ}}{\partial \zeta'} + Q_{ZZ}^{-1} \frac{\partial Q_{ZX}}{\partial \zeta'} \\ \frac{\partial Q_{XX}}{\partial \zeta'} - vec(Q_{ZX} Q_{XZ})'(Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1}) \frac{\partial vec Q_{ZZ}}{\partial \zeta'} - 2 Q_{XZ} Q_{ZZ}^{-1} \frac{\partial Q_{ZX}}{\partial \zeta'} \end{bmatrix}$$

has full column rank in $\zeta_0 \in \mathcal{Z}$ *if* $r > l$.

Proof: See Appendix A.2.

The proposition says that the ARMA structural model is not identified when the number of the moving average parameters is larger or equal to the number of autoregressive coefficients, even if $q \geq p + 1$. Indeed, the Assumption 4.v does not hold for any choice of the auxiliary AR(m) model with $m > l + r$. In other words, if the structural model is not identifiable, then there is no possibility also for the indirect inference estimator to guarantee consistency. Therefore, $r > l$ is a necessary condition for the identification of ARMA plus noise processes by AR(m) models.[4] The case of an MA(1) contaminated by measurement error provides an example of the violation of the identification condition.

Example 3.1 *In the case of MA(1) plus noise, the set of structural parameters is* $\zeta = [\varphi, \sigma_\varepsilon^2, \sigma_u^2]'$. *Hence, the number of auxiliary parameters must be at least 3 to satisfy the order condition* $q \geq p + 1$. *If the auxiliary model is a zero-mean AR(2), the* (3×1) *vector of auxiliary parameters is* $\beta = [\phi_1, \phi_2, \sigma_e^2]'$. *The binding function is given by*

$$b(\zeta) = p \lim_{T \to \infty} \begin{bmatrix} \hat{\phi}_{1,T} \\ \hat{\phi}_{2,T} \\ \hat{\sigma}_{e,T}^2 \end{bmatrix} = \begin{bmatrix} \frac{[(1+\varphi^2)\sigma_\varepsilon^2 + \sigma_u^2]\varphi\sigma_\varepsilon^2}{[(1+\varphi^2)\sigma_\varepsilon^2 + \sigma_u^2]^2 - \varphi^2\sigma_\varepsilon^4} \\ \frac{-\varphi^2\sigma_\varepsilon^4}{[(1+\varphi^2)\sigma_\varepsilon^2 + \sigma_u^2]^2 - \varphi^2\sigma_\epsilon^4} \\ (1+\varphi^2)\sigma_\varepsilon^2 + \sigma_u^2 + 3\frac{[(1+\varphi^2)\sigma_\varepsilon^2 + \sigma_u^2]\varphi^2\sigma_\varepsilon^4}{[(1+\varphi^2)\sigma_\varepsilon^2 + \sigma_u^2]^2 - \varphi^2\sigma_\varepsilon^4} \end{bmatrix}$$

[4] We conjecture that this condition is also sufficient. But the proof is fairly involved, since it would consist of showing that full-column rank of the Jacobian implies $r > l$.

and a closed form expression of $\frac{\partial b(\zeta)}{\partial \zeta'}$ *can be obtained.*[5] *However,* $\frac{\partial b(\zeta)}{\partial \zeta'}$ *has rank equal to 2 for any* $\zeta_0 \in \mathcal{Z}$*, while the number of structural parameters is 3. This is a consequence of the fact that the ARMA condition* $r > l$ *is not satisfied for the MA(1) case, meaning that the structural model cannot be identified by any AR(m) with* $m > 1$*.*

Another example is the identification of the Ornstein-Uhlenbeck process plus noise, presented in Example 2.2, where the measurement error variance is considered as a structural parameter. Note that the discretized Ornstein-Uhlenbeck process is an AR(1), therefore it satisfies the condition of Proposition 3.1.

Example 3.2 *The* (4×1) *vector of structural parameters is* $\zeta = (k, \omega, \sigma, \sigma_u^2)'$*. The auxiliary model in equation* (14)*, which contains only 3 parameters, must be extended to satisfy the order condition,* $q \geq p + 1$*. The additional auxiliary parameter is* $\sigma_x^2 = \mathrm{Var}(x_t)$*, so that the vector of auxiliary parameters becomes* $\beta = (\beta_1, \beta_2, \beta_3, \sigma_x^2)'$*. The binding function results to be*

$$b(\zeta) = \begin{bmatrix} 1 - \frac{e^{-k}\sigma^2}{\sigma^2 + 2k\sigma_u^2} \\ \omega \\ \left[\left(\frac{\sigma^2}{2k} + \sigma_u^2\right)\left[1 + \left(\frac{e^{-k}\sigma^2}{\sigma^2 + 2k\sigma_u^2}\right)^2\right] - \left(\frac{e^{-k}\sigma^2}{\sigma^2 + 2k\sigma_u^2}\right) e^{-k}\frac{\sigma^2}{k}\right]^{1/2} \\ \frac{\sigma^2}{2k} + \sigma_u^2 \end{bmatrix} \tag{20}$$

and the Jacobian matrix $\frac{\partial b(\zeta_0)}{\partial \zeta'}$ *has full rank for any* $\zeta_0 \in \mathcal{Z}$*. Hence, the auxiliary model identifies all the parameters in* ζ*, including the variance of the measurement error.*

An alternative approach to the indirect estimation is the method of Chanda (1995). However, the recursive and nonlinear form of the conditions that define the estimator and the necessity of relying on preliminary estimates of the autoregressive coefficients makes it computationally involved. Komunjer and Ng (2014) discuss the identification of VARX models, when the dependent variables and the covariates can be contaminated by serially correlated error. They obtain necessary order and sufficient rank conditions for local identification of the structural parameters, which exploit the triangular structure of the problem. However, their results only apply to the case of dynamic process without a moving average component, thus ruling out ARMA signals. The identification condition in Proposition 3.1 is similar to the one put forward by Komunjer and Ng (2014) in Proposition 3. Gospodinov et al. (2015) consider the estimation of an ARDL(p,q) model by means of the simulated minimum distance estimator, which is an indirect inference estimator. In their setup the only variable to be measured with error is the covariate. Analogously, to the approach taken here, they consider the joint estimation of the parameters of interest and the nuisance ones. They obtain the identification condition for the ARDL(0,0) and ARDL(1,0) when $p = q$, by establishing the invertibility of the binding function. Therefore, the results of Gospodinov et al. (2015) can be seen as a further example of the applicability of indirect inference to solve errors-in-variables problems.

4 Estimation of SV models with RV

The recent results in the theory of *RV* based on high frequency data open the door to the estimation of continuous-time *SV* models by GMM and indirect inference. Under unrealistic assumptions, e.g. the absence of microstructure noise (MN), the *RV* is an asymptotically unbiased and efficient estimator of *IV*. However, the presence of MN can dramatically affect the consistency of the SV parameter estimates. The MN is generated by structural features of financial markets, like trading

[5]Details on the derivation of these results are shown in the document with supplementary material. Due to space constraints the expression of $\partial b(\zeta)/\partial \zeta'$ is not reported. It is available upon request from the authors.

rules, the bid-ask spread and the discreteness of price changes. Neglecting the MN in calculating RV leads to biased and inconsistent estimates of integrated volatility (IV) as a true measure of daily volatility. Indeed, when the sample interval shrinks to zero, the MN obscures the IV signal. In a GMM framework, Bollerlsev and Zhou (2002) propose a simplified approach to deal with this problem which does not disentangle the discretization error from the MN. Differently, Corsi and Renò (2012), in an indirect inference framework, sample log-returns at low frequencies and do not explicitly model the MN. We call this approach *neutralization*.

The problem of the estimation of the parameters of SV models by RV under the presence of MN has been tackled also in Corradi and Distaso (2006), Todorov (2009) and, more recently, by Creel and Kristensen (2015). Corradi and Distaso (2006) propose a SMM approach and derive a set of sufficient conditions for the asymptotic negligibility of the measurement error, when the moments of the unobservable IV are replaced by the moments of the RV. Todorov (2009) uses a corrected estimator robust to MN and price jumps, but he also derives the moment conditions in closed form so that there is no need for simulation. However, the use of corrected realized estimators, like those proposed in Zhang et al. (2005), Barndorff-Nielsen et al. (2008), Hansen et al. (2008) and Andersen et al. (2012), may not be the best solution in the indirect inference framework proposed in this paper. Indeed, these realized estimators depend in a non trivial way on the nuisance parameters of the MN distribution, and their consistency is derived under some (strong) assumptions about the MN. In this case, the binding function would still depend on the MN nuisance parameters, making impossible to be matched by the simulated trajectories and hence potentially causing inconsistent estimates of the SV parameters. Alternatively, Creel and Kristensen (2015) explicitly deal with the problem of accounting for the intradaily sampling error when realized measures are used in the estimation of continuous-time SV models. They propose a limited information method, based on the Approximate Bayesian Computation, which is an alternative to indirect inference.

Instead, consistently with the general approach outlined in this paper, we suggest to estimate the parameters of the MN distribution jointly with the SV structural parameters. In other words, instead of deriving conditions for the asymptotic negligibility of the measurement error, we show that indirect inference, coupled with a contamination scheme of the simulated trajectories, avoids the need to *neutralize* a-priori the impact of the MN on the volatility estimates. Hence, the auxiliary model can be based on potentially distorted but efficient and simple estimators of IV, like RV, BPV or signed jump variation as in Barndorff-Nielsen et al. (2010) and Patton and Sheppard (2015). This represents a very general approach to the treatment of the problem of measurement error in the estimation of the SV framework, as it is potentially valid for any SV model and contamination scheme. Since the parameters of the MN distribution must be estimated jointly with the structural ones, special attention has to be devoted to the identification issue. In the following section, we closely look at the identification of the Heston SV model parameters with MN.

4.1 Estimation of the Heston model with microstructure noise

The Heston (1993) model is a well known continuous-time stochastic process used to describe the evolution of the volatility of an underlying asset and widely used in option pricing. Assume that $\sigma^2(t)$ follows a square root process as in Heston (1993), then

$$\mathrm{d}p^*(t) = \sigma(t)\mathrm{d}W_1(t) \quad (21)$$

$$\mathrm{d}\sigma^2(t) = \kappa(\omega - \sigma^2(t))\mathrm{d}t + \varsigma\sigma(t)\mathrm{d}W_2(t) \quad (22)$$

where $\kappa > 0$ governs the speed of mean reversion, $\varsigma > 0$ is the volatility of volatility parameter, while $\omega > 0$ is the long run mean of $\sigma^2(t)$, where the latter is the instantaneous volatility and it is independent of the process $W_1(t)$. The assumption that $\mathrm{Corr}[\mathrm{d}W_1(t), \mathrm{d}W_2(t)] = 0$, i.e. absence of leverage, is relaxed in Section 4.2. The condition $2\kappa\omega \geq \varsigma^2$ guarantees that the volatility process is stationary and it can never reach zero. Following the general representation of Meddahi (2003), we assume that the instantaneous volatility can be written as an autoregressive variance process:

$$\sigma^2(t) = a_0 + a_1 P_1(f(t)) \tag{23}$$

where f_t is a state-variable process. The function $P_1(\cdot)$ is defined so that it has the following properties:

$$E[P_1(f_t)] = 0, \quad \mathrm{Var}[P_1(f(t))] = 1, \quad E[P_1(f(t+\Delta))|P_1(f(\tau)), \tau \leq t] = e^{-\lambda_1 \Delta} P_1(f(t)).$$

Following Meddahi (2003), model (22) for $\sigma^2(t)$ can be rewritten as in (23) with $P_1(f(t)) = \frac{\sqrt{2\kappa}}{\sqrt{\omega\varsigma^2}}(\omega - f(t))$, $a_0 = \omega$, $a_1 = -\varsigma\sqrt{\frac{\omega}{2\kappa}}$, $\lambda_1 = \kappa$ and $\sigma^2(t) = f(t)$. Now, we focus on the properties of the ex-post estimates of IV, defined as $IV_t = \int_{t-1}^{t} \sigma^2(u)\mathrm{d}u$, which cumulates the instantaneous volatility over periods of unit length. A non-parametric estimator of IV_t is $RV_t(\Delta) = \sum_{i=1}^{n} r^2_{t-1+i\Delta}$, where $n = 1/\Delta$, and $r_{t-1+i\Delta}$ are the intradaily returns over the intervals $[t-1+(i-1)\Delta; t-1+i\Delta]$, for $i = 1, \ldots, n$. When the MN is present and contaminates the high-frequency returns the observed intradaily price is observed with error, i.e.

$$p_{t,i}(\Delta) = p^*_{t,i}(\Delta) + \epsilon_{t,i}(\Delta) \quad \text{for} \quad t = 1, ..., T \quad \text{and} \quad i = 1, ..., n \tag{24}$$

where $p^*_{t,i}(\Delta)$ is the i-th latent efficient log-price on day t. The term $\epsilon_{t,i}(\Delta)$ is the noise around the true price, with mean 0 and finite fourth moment and it is assumed $i.i.d.$ and independent of the efficient price. Over periods of length Δ, the log-return $r_{t,i}(\Delta) \equiv r_{t-1+i\Delta}$ is given by

$$r_{t,i}(\Delta) = \left(p^*_{t,i}(\Delta) - p^*_{t,i-1}(\Delta)\right) + (\epsilon_{t,i}(\Delta) - \epsilon_{t,i-1}(\Delta)) = r^*_{t,i}(\Delta) + \nu_{t,i}(\Delta) \tag{25}$$

with $\sigma^2_\nu = \mathrm{Var}\left[\nu_{t,i}(\Delta)\right] < \infty$.

When there is no drift in prices, the RV_t is observed with a measurement error, that is due both to the discretization error and to the MN:

$$RV_t(\Delta) = IV_t + u_t(\Delta), \tag{26}$$

where

$$u_t(\Delta) \stackrel{\mathcal{L}}{=} \eta_t(\Delta) + \sum_{i=1}^{n} \nu^2_{t,i}(\Delta) + 2\sum_{i=1}^{n} \sigma_{t,i,\Delta} z_{t,i} \nu_{t,i,\Delta}, \tag{27}$$

where $\eta_t(\Delta) = \sum_{i=1}^{n} \eta_{t-1+i\Delta}$ is the discretization error. Meddahi (2002) proves that $\eta_t(\Delta)$ has a nonzero mean, when the drift in prices is non-zero, and is heteroskedastic. The correlation between IV and $\eta_t(\Delta)$ is zero when there is no leverage effect (Barndorff-Nielsen and Shephard, 2002b and Meddahi, 2002). Assuming that the drift is null and there is no leverage effect, Barndorff-Nielsen and Shephard (2002a) show that, for finite $\Delta > 0$, the discretization error for the interval $[t-1+(i-1)\Delta, t-1+i\Delta]$ can be written as

$$\eta_{t-1+i\Delta}(\Delta) \stackrel{\mathcal{L}}{=} \sigma^2_{t,i}(\Delta)\left(z^2_{t,i} - 1\right) \tag{28}$$

where $z_{t,i}$ is $i.i.d.N(0,1)$ and it is independent of $\sigma^2_{t,i}(\Delta) = \int_{t-1+(i-1)\Delta}^{t-1+i\Delta} \sigma^2(s)\mathrm{d}s$, $\sigma^2_{t,i}(\Delta)$ is the integrated variance over the i-th subinterval of length Δ. Meddahi (2003) proves that when the instantaneous volatility is a square-root process, like in the Heston (1993) model, then both IV and RV have an ARMA(1,1) representation.[6]

[6]As noted by Meddahi (2003), already Bollerlsev and Zhou (2002) explicitly recognized that IV and RV are ARMA(p,p) processes, $p = 1, 2$, when the spot variance depends on p square-root processes.

For a given $\Delta > 0$, the mean and the variance of RV are equal to

$$E[RV_t(\Delta)] = E[IV_t] + E[u_t(\Delta)] \tag{29}$$

$$\mathrm{Var}[RV_t(\Delta)] = \mathrm{Var}[IV_t] + \mathrm{Var}[u_t(\Delta)] \tag{30}$$

with $E[IV_t] = \omega$, $E[u_t(\Delta)] = \Delta^{-1}\sigma_\nu^2$, and $\mathrm{Var}[u_t(\Delta)] = 2\Delta^{-1}E\left[\left(\sigma_{t,i}^2(\Delta)\right)^2\right] + 4\omega\sigma_\nu^2 + \Delta^{-1}(\kappa_\nu - \sigma_\nu^4)$, $\kappa_\nu = E[\nu_{t,i}(\Delta)^4]$, see Rossi and Santucci de Magistris (2014). A closed form expression of the term $2\Delta^{-1}E\left[\left(\sigma_{t,i}^2(\Delta)\right)^2\right]$ as a function of the structural parameters is derived in Meddahi (2002, 2003). It follows that the variance of $RV_t(\Delta)$ is

$$\gamma(0) = \mathrm{Var}[RV_t(\Delta)] = 2\frac{a_1^2}{\kappa^2}[\exp(-\kappa) + \kappa - 1] + \mathrm{Var}[u_t(\Delta)]$$

and the autocovariances of $RV_t(\Delta)$ are

$$\begin{aligned}\gamma(j) &= \mathrm{Cov}[RV_t(\Delta), RV_{t-j}(\Delta)] = \mathrm{Cov}[IV_t, IV_{t-j}] \\ &= a_1^2 \frac{[1 - \exp(-\kappa)]^2 \exp(-\kappa(j-1))}{\kappa^2} \quad j > 0.\end{aligned} \tag{31}$$

We are interested in estimating the parameters of the Heston model in (21)-(22) along with those of conditional distribution of the MN which contaminates the log-returns at high frequency. Estimates of the structural parameter vector, $\zeta = (\kappa, \omega, \varsigma, \sigma_\nu^2)'$, can be obtained by indirect inference using an auxiliary model based on RV. A well known example of a simple reduced-form model for RV is the HAR-RV model of Corsi (2009), which is

$$x_t = \phi_1 + \phi_2 x_{t-1} + \phi_3 x_{t-1}^w + \phi_4 x_{t-1}^m + e_t, \tag{32}$$

where $x_t = RV_t(\Delta)$, $x_t^w = \frac{1}{5}\sum_{t=0}^{4} x_{t-j}$ and $x_t^m = \frac{1}{22}\sum_{t=0}^{21} x_{t-j}$, and $\phi = [\phi_1, \phi_2, \phi_3, \phi_4]'$.[7] The (5×1) vector of auxiliary parameters is $\beta = (\phi', \sigma_e^2)'$.

Proposition 4.1 *(Identification of Heston model with MN) Let the structural model be the Heston model in* (21) *and* (22)*, with the* $RV(\Delta)$ *and the measurement error as in* (26) *and* (27)*, respectively with* $\Delta > 0$*. The auxiliary model is the HAR-RV in* (32)*, which is an AR(22) with restrictions contained in the* (23×4) *matrix, R. The binding function results to be*

$$b(\zeta) = \begin{bmatrix} [R'Q_{ZZ}R]^{-1} R'Q_{ZX} \\ Q_{XX} - [Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'Q_{ZX}] \end{bmatrix} \tag{33}$$

where Q_{ZZ}*,* Q_{XX}*,* Q_{XZ} *are the moment matrices of* RV *and are function of* $\zeta = (\kappa, \omega, \varsigma, \sigma_\nu^2)'$*. The Jacobian matrix* $\frac{\partial b(\zeta)}{\partial \zeta'}$ *has full column rank for any* $\zeta_0 \in \mathcal{Z}$*.*

Proof: See Appendix A.3.

Proposition 4.1 proves that the Heston parameters can be identified when the auxiliary model is the HAR-RV since the rank of the binding function is full. At a first sight, this result seems to contradict the evidence reported in Section 3.1 about the non-identifiability of an ARMA(r, l) plus noise when $l \leq r$. On the contrary, despite RV is the sum of an ARMA(1,1) and a noise term, identification is guaranteed in this case. Indeed, as noted by Barndorff-Nielsen and Shephard (2002b) and Meddahi (2003), the moving-average root of the IV signal is in turn the result of the

[7] In order to derive the binding function in closed form, the dependent variable of the HAR-RV is the RV in levels. This is slightly different from the setup in Corsi (2009) and Corsi and Renò (2012) where the dependent variable is $\log RV$.

state-space representation of an AR(1) signal plus noise, so that the MA parameter is restricted and dependent on the autoregressive root. The consequence is that the AR and MA parameters are no more *functionally independent*, which means that the parameter space dimension is reduced by one. This explains why the ARMA(1,1) plus noise representation of RV is identifiable, in terms of the underlying Heston parameters, by an autoregressive auxiliary model.

4.1.1 Binding function when $\Delta \to 0$

Let's consider now the case $\Delta \to 0$. As noted by Meddahi (2003), in absence of MN, the moving average roots of the RV converge to those of IV when $\Delta \to 0$ and hence the ARMA representations of IV and RV coincide. However, when the MN is present, then the mean and the variance of the measurement error diverge as $\Delta \to 0$, thus completely obscuring the volatility signal. As noted by Bandi and Russell (2006), while the efficient return is of order $O_p(\sqrt{\Delta})$, the MN is of order $O_p(1)$ over any period of time. This means, that, when $\Delta \to 0$, then the MN dominates over the true return process, and longer period returns are less contaminated by noise than shorter period returns. Taking the limits of $E[RV_t]$, $\text{Var}[RV_t]$ and $\text{Cov}[RV_t, RV_{t-j}]$ for $\Delta \to 0$, we get

$$\lim_{\Delta\to 0} E[RV_t(\Delta)] = +\infty, \quad \lim_{\Delta\to 0} \text{Var}[RV_t(\Delta)] = +\infty \quad \lim_{\Delta\to 0} \text{Cov}[RV_t, RV_{t-j}] = \text{Cov}[IV_t, IV_{t-j}]$$

Therefore, the limits in $T \to \infty$ of the elements in Q_{XX}, Q_{ZZ}, Q_{ZX} diverge as $\Delta \to 0$, so the limit of the binding function in (33) is unbounded. This means that it is not possible to identify the parameters of the the Heston SV model using RV in presence of MN when $\Delta \to 0$. As stated in Proposition 4.1, indirect inference with measurement error works without the hypothesis of $\Delta \to 0$, so in this setup, there is no need for infill asymptotics but only T has to diverge for any $\Delta > 0$. This also represents the most realistic scenario in practice. Indeed, the sampling frequency of high-frequency data has a lower bound, so that the binding function is always associated to a finite limit in empirical applications. Clearly, choosing a small Δ would convey more information about MN, in line with Zhang et al. (2005) and Ait-Sahalia et al. (2005), while the volatility signal dominates as Δ increases. For this reason, in empirical applications, when the SV model presents additional features, such as leverage, drift and price jumps, it might be necessary to adopt an auxiliary model based on realized measures obtained from returns computed at different sampling frequencies, as discussed below.

4.2 Leverage and microstructure noise

Now, we assume that in the Heston SV model, i.e. in the system (21)-(22), $\text{Corr}[\mathrm{d}W_1(t), \mathrm{d}W_2(t)] = \rho \mathrm{d}t$, namely there is leverage. In this case,

$$\text{Var}[RV_t(\Delta)] = \text{Var}[IV_t] + \text{Var}[u_t(\Delta)] + 2\,\text{Cov}[IV_t, u_t(\Delta)] \tag{34}$$

Under leverage, see Meddahi (2002, Proposition 4.2)

$$\text{Var}[\eta_{t-1+i\Delta}] = 4\Big(\frac{a_0^2\Delta^2}{2} + \frac{a_1^2}{\kappa^2}(\exp(-\kappa\Delta) - 1 + \kappa\Delta)\Big) + 8\rho^2 a_1 e_{1,1}\frac{e_{1,0}}{\kappa}\left[\frac{\Delta}{\kappa} + \frac{\exp(-\kappa\Delta)(\kappa\Delta)}{\kappa^2}\right] \tag{35}$$

where $e_{1,0} = \sqrt{2\kappa\omega}, e_{1,1} = \varsigma$. The additional term which appears in (34) is the covariance between IV_t and $u_t(\Delta)$ which is equal to

$$\text{Cov}(IV_t, u_t(\Delta)) = 2\Delta^{-1}\rho a_1 e_{1,1}\frac{e_{1,0}}{\kappa}\left[\frac{\Delta}{\kappa} + \frac{\exp(-\kappa\Delta)(\kappa\Delta)}{\kappa^2}\right]. \tag{36}$$

As noted by Meddahi (2002), the correlation between the noise and the integrated volatility tends to zero very quickly as one increases the frequency of intra-daily observations. Therefore, similarly to Bandi and Renò (2012, p.110), we suggest to identify the contemporaneous leverage parameter by exploiting the contemporaneous covariance between the daily returns and $RV_t(\Delta)$, that is given by

$$\mathrm{Cov}(r_t, RV_t(\Delta)) = \rho\frac{a_1 e_{1,0}}{\kappa^2}[\exp(-\kappa) - 1 + \kappa] + \frac{\rho}{\Delta}\frac{a_1 e_{1,0}}{\kappa^2}[\exp(-\kappa\Delta) - 1 + \kappa\Delta],$$

since $\mathrm{Cov}(r_t, \sum_{i=1}^n \nu_{t,i}^2(\Delta)) = 0$ as daily observed returns can be assumed unaffected by MN, without loss of generality, while for a small Δ, $[\exp(-\kappa\Delta) - 1 + \kappa\Delta] \sim \kappa^2\Delta^2/2$ so that the second term doesn't depend on ρ when $\Delta \to 0$. Since the parametric expression of $\mathrm{Cov}(r_t, RV_t(\Delta))$ doesn't involve any MN parameter, the leverage parameter ρ can be isolated from the noise and separately identified by the contemporaneous covariance between daily returns and RV, while the other SV parameters and the variance of MN can be identified by a HAR type auxiliary model as in (32), as shown in Proposition 4.1.

4.3 Drift and jump prices with microstructure noise

We consider an extension of the Heston SV model that allows for non-zero drift and jumps in prices. We assume that the efficient price process $p^*(t)$ follows an Itô semimartingale, see Bates (1996), namely

$$\mathrm{d}p^*(t) = m(t)\mathrm{d}(t) + \sigma(t)\mathrm{d}W_1(t) + \tau(t)\mathrm{d}N(t), \tag{37}$$

which implies that the efficient log-return over an interval of length Δ is

$$r_{t,i}^*(\Delta) = \mu_{t,i} + v_{t,i} + J_{t,i},$$

where $\mu_{t,i} = \int_{t-1+(i-1)\Delta}^{t-1+i\Delta} m(u)\mathrm{d}(u)$, $v_{t,i} = \int_{t-1+(i-1)\Delta}^{t-1+i\Delta} \sigma(u)\mathrm{d}W_1(u)$ and $J_{t,i} = \sum_{j=1}^{N_{t,i}} \tau_j$. $N_{t,i} = N(t-1+i\Delta) - N(t-1+(i-1)\Delta)$ denotes the number of jumps in the i-th subinterval of day t. The jump size $\tau(t)$ is assumed time invariant and for the j-th jump arrival is distributed as $N(\mu_\tau, \sigma_\tau^2)$, while the jump arrival process, $N(t)$ is Poisson distributed, uncorrelated with $W_1(t)$ and $W_2(t)$, with λ, which is the average jump intensity on the unit interval $[t-1, t]$. It follows that $N_{t,i}$ is Poisson with intensity $\lambda\Delta$. We maintain the assumption of no leverage. If the efficient return is measured with MN, then

$$r_{t,i}(\Delta) = \mu_{t,i} + v_{t,i} + J_{t,i} + \nu_{t,i}, \tag{38}$$

therefore,

$$r_{t,i}^2(\Delta) = \int_{t-1+(i-1)\Delta}^{t-1+i\Delta} \sigma^2(u)\mathrm{d}u + \xi_{t,i}(\Delta)$$

where $\xi_{t,i}(\Delta) = J_{t,i}^2 + \mu_{t,i}^2 + \left[v_{t,i}^2 - \int_{t-1+(i-1)\Delta}^{t-1+i\Delta} \sigma^2(u)\mathrm{d}u\right] + \nu_{t,i}^2 + 2\mu_{t,i}v_{t,i} + 2\mu_{t,i}J_{t,i} + 2\mu_{t,i}\nu_{t,i} + 2v_{t,i}J_{t,i} + 2v_{t,i}\nu_{t,i} + 2J_{t,i}\nu_{t,i}$. Assuming that the drift is constant, $m(t) = \mu\ \forall t$, and that $\nu_{t,i} \sim iidN(0, \sigma_\nu^2)$, the expected value and variance of the daily return is

$$\begin{aligned} E[r_t] &= \mu + \lambda\mu_\tau \\ \mathrm{Var}[r_t] &= \omega + \sigma_\nu^2 + \lambda(\sigma_\tau^2 + \mu_\tau^2). \end{aligned}$$

The mean and variance of RV_t are

$$\begin{aligned} E[RV_t(\Delta)] &= \omega + \Delta^{-1}\sigma_\nu^2 + \Delta^{-1}E[J_{t,i}^2] + \Delta\mu^2 + 2\Delta\mu\lambda\mu_\tau \\ \mathrm{Var}[RV_t(\Delta)] &= \mathrm{Var}[IV_t] + 2\Delta^{-1}E\left[\left(\sigma_{t,i}^2(\Delta)\right)^2\right] + 2\Delta^{-1}\sigma_\nu^4 + \Delta^{-1}E[J_{t,i}^4] + 4\omega\sigma_\nu^2 + 4\Delta^{-1}E[J_{t,i}^2]\sigma_\nu^2 \\ &\quad +4\Delta\mu^2\sigma_\nu^2 + 4E[J_{t,i}^3]\mu + 8\Delta\mu\lambda\mu_\tau\sigma_\nu^2 + 4\omega E[J_{t,i}^2] + 4\Delta^2\mu^2\omega + 6\Delta\mu E[J_{t,i}^2] \\ &\quad +8\Delta^2\mu\lambda\mu_\tau\omega - 2\Delta\mu(\mu^2 + 2\lambda\mu_\tau)E[J_{t,i}^2] + 4\Delta^3\mu^3\lambda\mu_\tau - 4\Delta^3\mu^2\lambda^2\mu_\tau^2, \end{aligned}$$

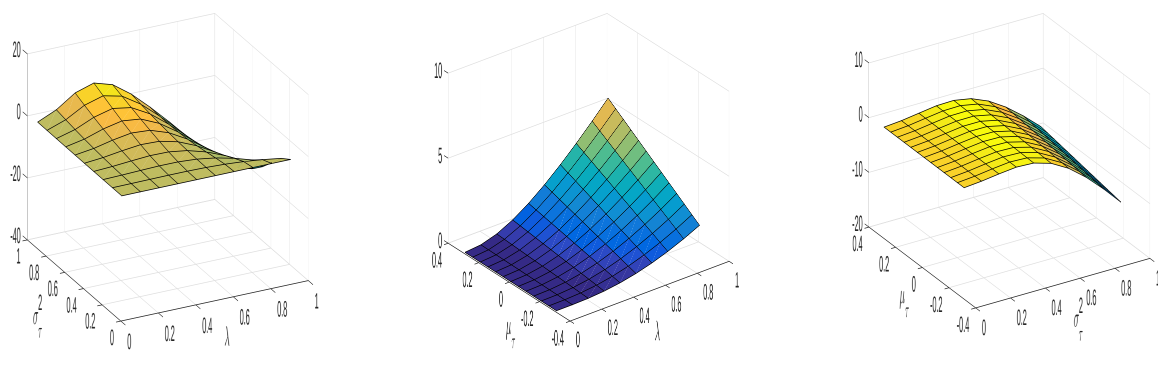
where the moments of the squared jump component are computed as in Lemma A.1. The variance of the realized signed jump variation of Barndorff-Nielsen et al. (2010), defined as $SJ_t(\Delta) = [RS_t^+(\Delta) - RS_t^-(\Delta)]$, is

$$\begin{aligned}\text{Var}[SJ_t(\Delta)] &= 3\Delta^{-1}E\Big[\Big(\sigma_{t,i}^2(\Delta)\Big)^2\Big] + 3\Delta^{-1}\sigma_\nu^4 + 6\omega\sigma_\nu^2 + 6\Delta^{-1}E[J_{t,i}^2]\sigma_\nu^2 + \Delta^{-1}E[J_{t,i}^4] \\ &\quad +6\Delta\mu^2\sigma_\nu^2 + 4\mu E[J_{t,i}^3] + 12\Delta\mu\lambda\mu_\tau\sigma_\nu^2 + 6\omega E[J_{t,i}^2] \\ &\quad +6\Delta^2\mu^2\omega + 6\Delta\mu E[J_{t,i}^2] + 12\Delta^2\mu\lambda\mu_\tau\omega + 4\Delta^3\mu^3\lambda\mu_\tau + \Delta^3\mu^4.\end{aligned}$$

Using these results, we can study the identification of the Heston model with jump prices and non-zero drift. We evaluate if indirect inference, based on an auxiliary specification built on daily returns and realized measures of volatility, is able to identify all the structural parameters. The (8×1) vector of structural parameters with i.i.d. Gaussian MN is $\zeta = (\kappa, \omega, \varsigma, \sigma_\nu^2, \mu, \lambda, \mu_\tau, \sigma_\tau^2)'$. We consider the following multivariate auxiliary model based on daily returns, RV_t and SJ_t

$$\begin{aligned} r_t &= \alpha + e_{r,t}, \\ RV_t(\Delta_1) &= \phi_1 + \phi_2 RV_{t-1}(\Delta_1) + \phi_3 RV_{t-2}(\Delta_1) + e_{RV,t}, \\ SJ_t(\Delta_1) &= e_{SJ_1,t}, \\ SJ_t(\Delta_2) &= e_{SJ_2,t}, \end{aligned}$$

where Δ_1 and Δ_2 are two distinct sampling frequencies. The number of auxiliary parameters is $q = 8$, i.e. we are in the exactly identified case. An evaluation of the rank for the Jacobian of the binding function for any ζ_0 in Ψ is unfeasible in closed form as it would require an analysis in $\mathbb{R}^8$. Therefore, we first compute the Jacobian for all the structural parameters but then we fix $\kappa_0, \omega_0, \varsigma_0, \sigma_{\nu,0}^2$ to the values adopted in the Monte Carlo study in Section 4.4. Moreover, the parameter μ is set equal to 0.2, which leads to a 5% drift on annual basis. We then compute the determinant of the 8×8 Jacobian matrix of the binding function only for varying λ_0, $\mu_{\tau,0}$ and $\sigma_{\tau,0}^2$. Since we have a triplet of parameters, we fix one of the three at a time to obtain a value of the determinant of the Jacobian for each combination of the other two parameters. The results are displayed in Figure 1. From Panel a) it emerges that, for a given $\mu_{\tau,0} = -0.1$, both λ_0 and



(a) $\mu_{\tau,0} = -0.1$ (b) $\sigma_{\tau,0}^2 = 0.5$ (c) $\lambda_0 = 0.3$

Figure 1: Determinant of the Jacobian matrix of the binding function of model (39) for different values of $\lambda \in (0,1)$, $\sigma_{\tau,0}^2 \in (0,1)$ and $\mu_\tau \in (-0.3, 0.3)$. Panel a) plots the determinant when $\mu_{\tau,0} = -0.1$, Panel b) plots the determinant when $\sigma_{\tau,0}^2 = 0.5$ and Panel c) plots the determinant when $\lambda_0 = 0.3$.

$\sigma_{\tau,0}^2$ need to be different from zero to guarantee identification. Not surprisingly, as both λ_0 and $\sigma_{\tau,0}^2$ increase, the determinant moves away from zero. Indeed, identification is easier if the jump term has a greater impact on the return total variability. When instead $\sigma^2_{\tau,0} = 0.5$ in Panel b), the determinant is zero for any $\mu_{\tau,0} \in (-0.3, 0.3)$ when $\lambda_0 = 0$. On the contrary, the determinant is larger than zero when $\mu_t = 0$ and it increases with λ_0. Similarly in Panel c) with $\lambda_0 = 0.3$, the determinant is zero for any $\mu_{\tau,0} \in (-0.3, 0.3)$ when $\sigma^2_{\tau,0} = 0$, while it increases (in absolute value) as $\sigma^2_{\tau,0}$ increases. In other words, variability in the jumps is required for identification purposes, that is both λ_0 and $\sigma^2_{\tau,0}$ need to be larger than zero to guarantee a non-singular Jacobian matrix.

4.4 Monte Carlo simulations

In the Monte Carlo experiments reported below we study the indirect inference estimation of the Heston SV model in (21) and (22) using the series of non-overlapping RV's, i.e. $\{RV_t(\Delta)\}_{t=1}^T$ under the presence of MN. Table 1 reports the Monte Carlo summary statistics of the indirect inference estimates of the Heston model parameters with and without accounting for the presence of MN, using different sampling frequencies, 30 seconds, 1, 5 and 30 minutes to construct the daily RV. The true parameter set is calibrated to values close to those found in empirical works, see Bollerlsev and Zhou (2002) and Garcia et al. (2011), and are relative to percentage returns. The long-run mean parameter, ω, is set equal to 0.5 and 0.8, which corresponds to an annualized volatility of 11.2% and 14.2% respectively. The speed of mean reversion, κ, is equal to either 0.1 or 0.05, while the vol-of-vol parameter ς is either 0.1 or 0.2. For each Monte Carlo replication, a simulated trajectory for the intradaily returns is generated from Euler discretization of model (22) for $T = 1500$ days, with intradaily step of 30 seconds, which corresponds to $n = 780$. The intraday return is contaminated with an additive MN, $\nu_{i,t}$, that is assumed to be i.i.d. Gaussian with mean 0 and variance $\sigma^2_\nu = 0.0005$.[8] This choice for σ^2_ν is in line with the numbers reported in Ait-Sahalia et al. (2005, p.364), and corresponds to a percentage standard deviation of 0.02%. According to Ait-Sahalia et al. (2005), the optimal sampling frequency, under the assumption of constant volatility, should be between 1 and 5 minutes. The RV series is constructed from high-frequency returns with different sub-sampling and, similarly to Corsi and Renò (2012), the HAR-RV model on $\log(RV_t)$ is used as auxiliary model such that the set of auxiliary parameters in the indirect inference estimation is $\beta = \left[\phi_1, \phi_2, \phi_3, \phi_4, \sigma^2_\epsilon\right]'$. The approximation of the binding function is based on $S = 100$ simulated trajectories of equation (22) with the same Euler discretization as the sampling frequency used to compute the RV series on the real data. When the possible presence of MN is neglected in the indirect inference estimation, the set of structural parameters is $\theta = [\kappa, \omega, \varsigma]'$. On the contrary, if the MN is considered in the indirect inference estimation, the set of structural parameters, ζ, also includes σ^2_ν. In this case, the simulated trajectories of the log-returns used in the second stage of the indirect estimation are contaminated with a Gaussian MN with variance σ^2_ν. In both cases, the HAR-RV model is used as auxiliary model on each simulated series, $RV_t^{(s)}$, for $s = 1, \ldots, S$, and the set of parameters β is estimated by OLS.

The results for the indirect inference estimator are presented in the left panel of Table 1 and are based on $M = 1000$ Monte Carlo simulations. In summary, the Monte Carlo results reflect the properties of RV under MN, see Ait-Sahalia et al. (2005). If we don't want to take into account the MN, the best choice, in terms of bias, is to compute the RV using a low sampling frequency, say 5 or 30 minutes. Indeed, sampling at 30 minutes neutralizes the effect of the MN on the estimates of the RV, such that the indirect inference estimates are only slightly affected. For example, the bias of $\hat{\omega}^{30min}$ and $\hat{\varsigma}^{30min}$ are negligible in all cases. On the other hand, sampling at 30 seconds or at 1 minute, but neglecting the presence of MN, induces large biases in the indirect inference estimates of ω, which as expected is over-estimated, and in ς. Surprisingly, the estimates of κ are unbiased for any choice of Δ also when the MN is neglected in the estimation. In terms of efficiency, the estimates corresponding to a 5-minutes sampling scheme without correction are

[8] As in Bandi and Russell (2006), we contaminate the intardaily log-returns instead of the intradaily log-prices. This leads to equivalent results as $\sigma^2_\nu = 2\sigma^2_\epsilon$, where σ^2_ϵ is the variance of $\epsilon_{t,i}$ in $p_{t,i} = p^*_{t,i} + \epsilon_{t,i}$.

generally associated with a lower RMSE than those obtained sampling at 30 minutes. Indeed, the effect of the MN when sampling returns every 5 minutes is rather limited, and the efficiency of RV when n is relatively large leads to more precise estimates of the Heston parameters than those obtained when the sampling frequency is 30 minutes. In other words, the squared bias component of RMSE when sampling at 5 minutes is smaller than the variance component of RMSE when sampling at 30 minutes. Moreover, when sampling at 30 seconds, the squared bias term dominates over the parameter variance in the RMSE, especially for ω and ς.

Turning our attention to the results when the variance of MN is included among the structural parameters, it emerges that sampling at the highest possible frequency provides unbiased estimates of all structural parameters. Instead, the estimates of σ_ν^2 are biased when sampling at low frequencies, since the log-returns do not contain enough information about the MN. Therefore, if one is interested in estimating the MN variance together with the SV parameters, the best option is to sample at the highest frequency available. Notably, in Table 1, the RMSE of the estimates of ω and ς obtained sampling at 30 seconds intervals is significantly smaller than the RMSE obtained by neutralizing MN, i.e. sampling at 5 or 30 minutes. This means that the discretization error, that is a function of the Heston parameters via the term $2\Delta^{-1}E\left[\left(\sigma_{t,i}^2(\Delta)\right)^2\right]$ in the variance of $u_t(\Delta)$, affects more significantly the SV parameter estimates when the latent volatility process is not very persistent and with a smaller long-run mean, as in Table 1. In these cases, the price paid by sampling log-returns at low frequencies is relatively higher than in the setup characterized by a persistent instantaneous volatility. Indeed, when the volatility signal is more persistent, with larger long-run mean and vol-of-vol, as in the bottom panel of Table 1, the values of the RMSE at 30 seconds are roughly the same as those obtained neutralizing the MN and sampling at 5 minutes. This means that the impact of the discretization error is less relevant in this case, and that sampling at low frequency does not induce severe efficiency losses.

The right panel of Table 1 also reports the estimation results of the Heston parameters based on a state-space representation of the process that characterizes $RV_t(\Delta)$, which, as discussed above, has a restricted ARMA(1,1) plus noise representation. Indeed, following Barndorff-Nielsen and Shephard (2002a), the parameters of the Heston model can be estimated by QML exploiting the Kalman filter routine applied to an ARMA(1,1) plus noise in state-space form. The parameters of the ARMA(1,1) are expressed as a function of the Heston's model parameters, such that the log-likelihood function is maximized with respect to the latter. As noted by Barndorff-Nielsen and Shephard (2002a), in absence of MN, the variance of the measurement error, $u_t(\Delta)$, is a functional of the Heston's model parameters and the discretization step only. Therefore, when the impact of MN is negligible (e.g. when sampling at 30 minutes), the Kalman filter is able to identify all the Heston parameters thus providing unbiased estimates. Not surprisingly, the QML estimates $\hat{\theta}_{30min}$ are more efficient than those obtained under indirect inference at the same frequency, although $u_t(\Delta)$ and the innovation of the state variable are non Gaussian. On the contrary, when including the variance of MN among the structural parameters, the estimates $\hat{\zeta}_{30min}$ obtained with the Kalman filter are highly biased and present very large RMSE, especially for ω and ς. Interestingly, the problem remains when increasing the sampling frequency, as the biases and RMSE are still very large for $\hat{\zeta}_{30sec}$. This is a clear indication that the estimation method based on the Kalman filter is unable to assign to the discretization error and MN the correct proportion of variability in the measurement error, $u_t(\Delta)$, thus leading to identification problems that instead do not affect the indirect inference estimator.

	Indirect Inference				Kalman Filter			
	$\kappa = 0.1$	$\omega = 0.5$	$\varsigma = 0.1$	$\sigma_\nu^2 = 0.0005$	$\kappa = 0.1$	$\omega = 0.5$	$\varsigma = 0.1$	$\sigma_\nu^2 = 0.0005$
$\hat{\theta}_{30sec}$	0.1016 (0.0212)	0.8889 (0.3901)	0.0740 (0.0207)	–	0.1037 (0.0166)	0.8874 (0.3886)	0.0751 (0.0251)	–
$\hat{\zeta}_{30sec}$	0.1014 (0.0197)	0.4994 (0.0274)	0.1013 (0.0097)	0.0005 (0.0000)	0.1040 (0.0157)	0.0550 (0.4493)	0.6306 (0.7155)	0.0011 (0.0011)
$\hat{\theta}_{1min}$	0.1014 (0.0212)	0.6925 (0.2009)	0.0838 (0.0189)	–	0.1037 (0.0158)	0.6932 (0.1944)	0.0849 (0.0155)	–
$\hat{\zeta}_{1min}$	0.1015 (0.0202)	0.4972(0.0552)	0.1014 (0.0105)	0.0005 (0.0001)	0.1037 (0.0158)	0.0505 (0.4553)	0.6039 (0.6468)	0.0016 (0.0017)
$\hat{\theta}_{5min}$	0.1010 (0.0227)	0.5346 (0.0817)	0.0956 (0.0138)	–	0.1043 (0.0181)	0.5373 (0.0430)	0.0966 (0.0069)	—
$\hat{\zeta}_{5min}$	0.1010 (0.0228)	0.4978 (0.0721)	0.1018 (0.0149)	0.0006 (0.0002)	0.1043 (0.0181)	0.0537 (0.4513)	0.8143 (0.9733)	0.0062 (0.0063)
$\hat{\theta}_{30min}$	0.0965 (0.0264)	0.4911 (0.0680)	0.0958 (0.0158)	–	0.1075 (0.0265)	0.5043 (0.0224)	0.1008 (0.0124)	—
$\hat{\zeta}_{30min}$	0.0966 (0.0256)	0.4718 (0.1014)	0.1022 (0.0307)	0.0016 (0.0046)	0.1075 (0.0265)	0.0797 (0.4319)	1.2032 (1.6417)	0.0327 (0.0336)
	$\kappa = 0.05$	$\omega = 0.8$	$\varsigma = 0.2$	$\sigma_\nu^2 = 0.0005$	$\kappa = 0.05$	$\omega = 0.8$	$\varsigma = 0.2$	$\sigma_\nu^2 = 0.0005$
$\hat{\theta}_{30sec}$	0.0493 (0.0153)	1.2008 (0.4149)	0.1514 (0.0542)	–	0.0545 (0.0132)	1.1827 (0.3981)	0.1637 (0.0372)	–
$\hat{\zeta}_{30sec}$	0.0500 (0.0161)	0.8217 (0.1123)	0.2102 (0.0356)	0.0005 (0.0001)	0.0544 (0.0131)	0.1090 (0.7061)	1.6184 (1.9916)	0.0014 (0.0014)
$\hat{\theta}_{1min}$	0.0482 (0.0144)	1.0027 (0.2338)	0.1657 (0.0424)	–	0.0545 (0.0132)	0.9858 (0.2156)	0.1795 (0.0230)	–
$\hat{\zeta}_{1min}$	0.1015 (0.0154)	0.4972(0.1166)	0.1014 (0.0318)	0.0005 (0.0001)	0.0545 (0.0132)	0.1050 (0.7101)	1.8146 (2.2139)	0.0023 (0.0023)
$\hat{\theta}_{5min}$	0.0506 (0.0131)	0.8448 (0.1239)	0.1892 (0.0267)	–	0.0551 (0.0141)	0.8290 (0.1082)	0.1962 (0.0143)	—
$\hat{\zeta}_{5min}$	0.0511 (0.0140)	0.8023 (0.1261)	0.2064 (0.0294)	0.0006 (0.0005)	0.0551 (0.0141)	0.1118 (0.7044)	2.2207 (2.8610)	0.0092 (0.0094)
$\hat{\theta}_{30min}$	0.0514 (0.0163)	0.8086 (0.1223)	0.1969 (0.0312)	–	0.0564 (0.0161)	0.7953 (0.1029)	0.2018 (0.0217)	—
$\hat{\zeta}_{30min}$	0.0517 (0.0165)	0.7974 (0.1322)	0.2063 (0.0351)	0.0014 (0.0032)	0.0564 (0.0166)	0.1389 (0.6845)	2.4572 (3.3837)	0.0505 (0.0523)

Table 1: Heston SV model. Mean and RMSE (in parenthesis) of estimated parameters with and without correction for the MN. The true parameter set are $\kappa = (0.10, 0.05)$, $\omega = (0.5, 0.8)$, $\varsigma = (0.1, 0.2)$ and $\sigma_\nu^2 = 0.0005$. The simulation are carried out with Euler discretization scheme with step size corresponding to 30 seconds. Three aggregation levels are considered to construct RV, 30 seconds and 1, 5 and 30 minutes. The number of simulated days is $T = 1500$ with $n = 780$ intradaily observations generated in each day. The number of Monte Carlo simulations is M=1000. The table reports the estimates obtained with indirect inference with the HAR model as auxiliary model and with QML coupled with the Kalman filter.

5 Empirical application

We provide some empirical evidence of the importance of accounting for the MN, when performing indirect inference estimates of SV based on RV series. We estimate the parameters of the two-factor Heston model (TFSV henceforth) on the RV series of JPMorgan (JPM) from July 2, 2003 to June 29, 2007 using intradaily returns sampled at 5-seconds frequency, i.e. $n = 4680$. The choice of the sample period is motivated by the evidence of parameter instability for the TFSV model during the sub-prime crisis, between June-2007 until June 2009, as shown in Grassi and Santucci de Magistris (2015). Instead, in the period 2003-2007, the RV is not subject to major breaks and we expect the parameters to be rather stable through time. The RV series is computed with returns sampled at two frequencies, 5-seconds and 5-minutes, RV^{5s} and RV^{5m} respectively. The dynamics of the two series are reported in Figure 2. The impact of the MN emerges clearly from the graph

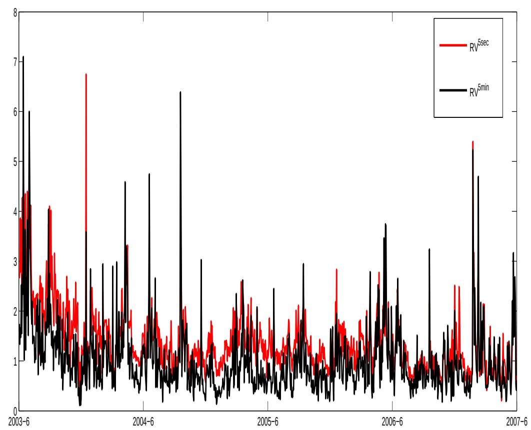


Figure 2: RV of JPM based on 5 seconds (red) and 5 minutes (black) sampling.

since the long run mean of RV^{5s} is shifted upward compared to that of RV^{5m}. Moreover, RV^{5m} is more noisy than RV^{5s}, meaning that the discretization error, denoted by $\eta_t(\Delta)$ in equation (27) seems to have a higher impact on the variance of the measurement error while the MN mainly impacts on the mean of RV. This evidence is also confirmed by the sample statistics reported in Table 2. The moments of the daily *de-volatized* returns, $\tilde{r}_t = r_t/\sqrt{RV_t^{5m}}$, are rather close to those of the standard Gaussian distribution. Notably, the autocorrelation function of RV^{5s} is much higher than that of RV^{5m}, as a consequence of the smaller impact of the discretization error on the variance of the measurement error.

	Mean	SD	SK	KU	AR(1)	AR(20)
$\tilde{r}_t$	0.0516	0.9019	0.0572	2.8371	-0.0074	0.0030
RV_t^{5s}	1.4073	0.6656	2.1826	11.454	0.6423	0.2861
RV_t^{5m}	1.0266	0.7381	2.9010	16.415	0.5204	0.0900
BPV_t^{5s}	0.6408	0.3464	2.7453	19.273	0.5558	0.0949
BPV_t^{5m}	0.9691	0.7239	3.0619	18.300	0.5148	0.0381
SJ_t^{5s}	-0.0102	0.1151	-1.2384	22.312	0.0632	0.0284
SJ_t^{5m}	0.0025	0.3473	-1.7984	31.585	0.0223	0.0481

Table 2: Sample statistics of $\tilde{r}_t = r_t/\sqrt{RV_t^{5m}}$, and realized measures of JPM.

The BPV computed from returns sampled at 5-seconds is clearly downward biased, as its mean and variance are much lower than those of RV^{5m}. Conversely, the moments of RV^{5m} are very close to those of BPV^{5m}. The reason for the bias in BPV^{5s} is the *decimalization* effect, which induces discontinuities in the trajectories of returns sampled at very high-frequencies, so that the product $|r_{t,i-1}| \times |r_{t,i}|$ is often equal to 0. For what concerns the signed jump variation, both SJ_t^{5s} and SJ_t^{5m} are centered around 0 and display little autocorrelation, while SJ_t^{5m} has more variability than SJ_t^{5s}.

5.1 A TFSV model with drift, jumps and leverage

We estimate the following TFSV Heston model, with drift, leverage and price jumps:

$$\begin{aligned}
\mathrm{d}p^*(t) &= \mu \mathrm{d}t + \sigma_1(t)\mathrm{d}W_1(t) + \sigma_2(t)\mathrm{d}W_2(t) + \tau(t)\mathrm{d}N(t) && (39)\\
\mathrm{d}\sigma_1^2(t) &= \kappa_1(\omega - \sigma_1^2(t))\mathrm{d}(t) + \varsigma_1\sigma_1(t)\mathrm{d}W_3(t) && (40)\\
\mathrm{d}\sigma_2^2(t) &= \kappa_2(\omega - \sigma_2^2(t))\mathrm{d}(t) + \varsigma_2\sigma_2(t)\mathrm{d}W_4(t) && (41)\\
\mathrm{Corr}(\mathrm{d}W_1(t), \mathrm{d}W_3(t)) &= \rho_1\mathrm{d}t, \quad \mathrm{Corr}(\mathrm{d}W_2(t), \mathrm{d}W_4(t)) = \rho_2\mathrm{d}t && (42)
\end{aligned}$$

where the parameters κ_1 and κ_2 govern the speed of mean reversion, while ς_1 and ς_2 determine the volatility of the volatility innovations. The parameter ω is the long-run mean of each volatility component and, as in Corsi and Renò (2012), it is assumed to be the same for both $\sigma_1^2(t)$ and $\sigma_2^2(t)$, in order to guarantee identification. $\{W_1(t) : t \geq 0\}$, $\{W_2(t) : t \geq 0\}$, $\{W_3(t) : t \geq 0\}$,$\{W_4(t) : t \geq 0\}$ are standard Brownian motions and $p^*(t)$ denotes the efficient log-price. The leverage effect depends on the parameters ρ_1 and ρ_2. Similarly to Section 4.3, we assume that $N(t) \sim Poisson(\lambda)$ and $\tau(t)$ is time invariant with $\tau_j \sim N(\mu_\tau, \sigma_\tau^2)$. The MN is modeled either as an i.i.d Gaussian variable that is added to the log-returns as in Bandi and Russell (2006), $r_{t,i} = r_{t,i}^* + \nu_{t,i}$, or as the bid-ask spread. In particular, the bid-ask bounce is generated as

$$p_{t,i} = p_{t,i}^* + \frac{\xi}{2}\mathbb{I}_{t,i} \qquad (43)$$

where ξ is the spread, and the order-driven indicator variables $\mathbb{I}_{t,i}$ are independently across t and i and identically distributed with $Pr\{\mathbb{I}_{t,i} = 1\} = Pr\{\mathbb{I}_{t,i} = -1\} = \frac{1}{2}$. This variable takes value 1 when the transaction is buyer-initiated, and -1 when it is seller-initiated.

The parameters of model (39) are collected in $\theta = [\kappa_1, \kappa_2, \omega, \varsigma_1, \varsigma_2, \mu, \rho_1, \rho_2, \lambda, \mu_\tau, \sigma_\tau^2]'$. Depending on the contamination scheme adopted, the structural parameters are collected in the vectors $\zeta_G = [\theta', \sigma_\nu^2]'$ and $\zeta_{BA} = [\theta', \xi]'$, where both ζ_G and ζ_{BA} are (12×1) vectors. Several restrictions of model (39) are considered, as well as alternative specifications for the auxiliary models. Table 3 reports the parameter estimates. The top panel of Table 3 reports the parameter estimates for the simple TFSV Heston model with no drift, no leverage and no jumps. If we neglect MN, the vector of structural parameter is $\theta = [\kappa_1, \kappa_2, \omega, \varsigma_1, \varsigma_2]'$. The auxiliary model adopted in this case is the univariate HARV model in (32) based on RV computed with returns sampled at 5 minutes (Model I) and 5 seconds (Model II). In this case, there are 5 auxiliary parameters and 5 structural parameters, so that the model is exactly identified. The estimates of the parameters strongly depend on the sampling frequency selected to compute RV. Indeed, the long-run mean, ω, is approximately 30% higher for RV^{5s} than RV^{5m}, reflecting the differences observed in the sample statistics. The speed of mean reversion κ_2 is two times lower when RV^{5s} is used instead of RV^{5m} accommodating the higher persistence of RV^{5s}. We also consider a bivariate HAR-RV model as auxiliary (Model III), where the dependent variables are RV^{5m} and RV^{5s}.[9] If the presence of MN is neglected, the criterion function is minimized at $\Xi = 126.9$, which is statistically significant different from zero.

This means that it is not possible to match the moments arising from the two RV series unless the presence of the MN is explicitly accounted for. Indeed, when the noise is modeled as an i.i.d.

[9] The document with additional material contains details on all auxiliary specifications adopted in this section.

Gaussian random variable with variance σ_ν^2 or with bid-ask spread of size ξ, the adherence of the simulated processes to the observed RV series improves significantly and the ξ only leads to a marginal rejection of the model. Looking at the estimates of the parameters, we notice they are almost identical both under i.i.d. Gaussian distribution and bid-ask spread. As noted by Creel and Kristensen (2015), the two contamination schemes have an identical impact on the RV since, by a central limit theorem argument, they are approximately normally distributed. In particular, the long-run mean ω is around 0.46 and the other TFSV parameters are generally within the estimates obtained when the HAR model on RV^{5m} and RV^{5s} are estimated separately. Moreover, the estimates of σ_ν^2 and ξ are almost identical, indicating that the two perturbation schemes have similar effects in the contamination of the latent efficient process.

The bottom panel of Table 3 reports the estimation results when drift, leverage and jumps are also included in the model. First, we start with the univariate auxiliary specification of Corsi and Renò (2012) (Model IV),

$$\begin{aligned}\log(RV_t) \;&=\; \beta_0 + \beta_1 \log(RV_{t-1}) + \beta_2 \log(RV_{t-1,w}) + \beta_3 \log(RV_{t-1,m}) + \\ &\quad + \gamma_1 r^-_{t-1} + \gamma_2 r^-_{t-1,w} + \gamma_3 r^-_{t-1,m} + \delta_1 J_{t-1} + \delta_2 J_{t-1,w} + \delta_3 J_{t-1,m} + e_t,\end{aligned}$$

where $r^-_{t-1}, r^-_{t-1,w}, r^-_{t-1,m}$ are the past negative log-returns aggregated over daily, weekly and monthly horizons, and $J_t = RV_t - BPV_t$ is a proxy for the squared jumps term. In the first auxiliary specification considered, RV_t and BPV_t are based on log-returns sampled every 5 minutes, while in Model V they are based on log-returns computed on grids of 5 seconds. Both specifications constrain the mean of jumps, μ_τ, to be zero, as in Chernov et al. (2003). The parameter estimates and the fit are dramatically affected by the sampling frequency adopted in the estimation. Interestingly, the estimates of λ, which measures the average number of jumps per day, are very high in both cases, but the jump sizes, as measured by σ_τ^2, are almost null, meaning that the price jump component is statistically insignificant.[10] In particular, sampling too sparsely, i.e. at 5 minutes, does not provide sufficient information to disentangle price jumps from discretization error and volatility, thus resulting in insignificant estimates of the jumps, even if the overall fit is quite good, as $\Xi = 6.899$. When sampling more frequently, the volatility signal is strongly affected by the MN, thus resulting in biased estimates of ω, negligible estimates of the jump term and poor fit, since Ξ is above 40. Moreover, the standard errors are very high for almost all parameters, so that only few of them are found significant. This may signal poor identification of the structural parameters.

In line with the theoretical results obtained in Section 4.3, we therefore consider the following multivariate specification (Model VI),

$$\begin{aligned} r_t &= \phi_0 + e_{r,t}, \\ RV_t^{5m} &= \beta_0 + \beta_1 RV_{t-1}^{5m} + \beta_2 RV_{t-1,w}^{5m} + \beta_3 RV_{t-1,m}^{5m} + e_{RV,t}, \\ SJ_t^{5m} &= e_{SJ_m,t}, \\ SJ_t^{5s} &= e_{SJ_s,t}, \\ \Sigma &= \mathrm{Cov}([e_{r,t}, e_{RV,t}, e_{SJ_m,t}, e_{SJ_s,t}]) \end{aligned}$$

so that the set of auxiliary parameters is $\beta = [\phi_0, \beta_0, \beta_1, \beta_2, \beta_3, \mathrm{vech}(\Sigma)]'$, which is a 16×1 vector. Based on this auxiliary model, the estimates of the structural parameters are generally significant, thus signaling a good overall identification. For what concerns the parameters governing the jump term, we find that λ, μ_τ and σ_τ^2 are highly significant in all cases, while the estimates of the other SV parameters are not much affected by the inclusion of jumps compared to the values obtained in absence of jumps. Interestingly, the estimates of the average jump size, μ_τ, is large and negative in all cases, meaning that price jumps are typically associated to bad news.

[10] The product $\hat{\lambda}\hat{\sigma}_\tau^2 = 0.0167$ has a standard error of 0.0202.

	κ_1	κ_2	ω	ς_1	ς_2	ρ_1	ρ_2	μ	λ	μ_τ	σ^2_τ	ξ	σ^2_ν	$\Xi(\hat{\zeta})$	$q-p-h$
							No Drift, No Leverage, No Jumps:								
I	2.3306^a	0.0559^a	0.4905^a	1.7690^a	0.1934^a	*	*	*	*	*	*	*	*	0.000	0
II	2.9712^a	0.0279^b	0.6850^a	1.4301^a	0.1351^a	*	*	*	*	*	*	*	*	0.000	0
III	2.4659^a	0.0232^a	0.6239^a	1.3117^a	0.1388^a	*	*	*	*	*	*	*	*	126.9^a	6
III	2.6781^a	0.0396^a	0.4698^a	1.7758^a	0.1674^a	*	*	*	*	*	*	0.0001^a	*	10.94^c	5
III	2.6917^a	0.0394^a	0.4699^a	1.7849^a	0.1668^a	*	*	*	*	*	*	*	0.0001^a	11.06^c	5
							Drift, Leverage and Jumps:								
IV	3.7082^b	0.0811^a	0.4280	2.3712	0.2161	0.1350	-0.5422	0.0438	13.891^a	*	0.0012	*	*	6.899^a	1
V	5.3100	0.0042	1.5996	2.2868	0.4518	-0.1119	0.1542	0.0035	40.248^a	*	0.0015	*	*	40.49^a	1
VI	5.0016^a	0.0172^a	0.3509^a	4.0755^a	0.2221^a	0.0171	-0.1237^a	0.0494^a	0.2165^a	-0.4298^a	0.0198^a	*	*	63.55^a	3
VI	2.3771^a	0.0111^a	0.4877^a	2.0350^a	0.2228^a	-0.2616^a	-0.0010	0.0499^a	0.0615^a	-0.3440^a	0.1063^a	0.0005^a	*	25.20^a	2
VI	3.7075^a	0.0146^a	0.4150^a	3.0636^a	0.2556^a	-0.1235^a	-0.0160^b	0.0485^a	0.0849^a	-0.5052^a	0.0446^a	*	0.0001^a	51.17^a	2

Table 3: Indirect inference estimates of the TFSV model with jumps and leverage under MN in (39)-(42). Several restrictions on the full model are considered. The asterisk indicates that the parameter is not estimated. The auxiliary model adopted for each structural specification is indicated in the first column. a, b and c stand for significance at 1%, 5% and 10% respectively. The last columns report the value of the criterion function, $\Xi(\hat{\zeta})$, and the difference $q-p-h$, with represent the degrees of freedom of the χ^2 distribution.

When instead we look at the estimated second moment of the jumps, that is $E[J(t)^2] = \lambda\mu_\tau^2 + \lambda(\sigma_\tau^2 + \mu_\tau^2)$, we note that it is quite low and equal to 0.0143 and 0.0275, when MN is included in the structural model. In turns, this implies that the average jump contribution to the total return variability of the jump term is estimated between 1.47% and 3.31%. Interestingly, when the MN is modeled as a bid-ask bound, the estimated jump variability is close to zero, while the parameter ξ is very high, meaning that the variability that otherwise would be attributed to the jumps is instead due to the bouncing between bid and ask prices. Indeed, in absence of MN, the estimated jump average variability is 0.0529, which is associated to a relative jump contribution to return variability of 7.54%, a result in line with the values reported by Huang and Tauchen (2005). In this case, the jump component is responsible for a significant portion of the total return variation so that the estimates of the long-run mean, ω, are much smaller than those obtained without jumps. This means that part of the gap between the unconditional mean of BPV^{5s} and RV^{5s} is attributed to the jump term, see Table 2. On the other hand, the fit of the model is not optimal in all cases, as the criterion function is still too large compared to the Ξ critical values. Notably, when the MN is modeled as bid/ask spread, the criterion function is minimized at 25.20, while when it is modeled as i.i.d. Gaussian noise, the criterion function is minimized at 51.17. An explanation for this difference in the model fit is due to the inability of the contamination method based on i.i.d. Gaussian noise to provide a realistic setup for the generation of high frequency returns under the presence of jumps. Indeed, the log-prices at very high frequencies are characterized by discreteness, due to the *decimalization* and rounding effects, making difficult to disentangle the volatility signal from MN and jumps. This misspecification is responsible for the fact that the criterion function is minimized far from zero, and its value is only marginally lower for the i.i.d. Gaussian noise than that obtained when noise is completely neglected, while the bid/ask spread seems to be more coherent with a realistic data generating process for the high-frequency returns. This evidence confirms the importance of the correct specification of the contamination term to guarantee a good fit. Unfortunately, generating log-prices under a decimalization scheme is not a viable solution in the indirect inference framework as it induces discreetness in the observed log-price thus masking the impact of changes in the SV parameters on the continuous dynamics and leading to problems of identification. As a consequence, the numerical Jacobian of the binding function, necessary for the calculation of the standard errors, contains many values near zero and it is almost singular. This is very informative and may indicate that alternative structural models, involving for example pure-jump Lévy processes as in Barndorff-Nielsen and Shephard (2001) or trawl processes as in Barndorff-Nielsen et al. (2014), may be better suited to model stock returns at very high-frequencies and could possibly be estimated by indirect inference.

6 Conclusions

This paper studies the inconsistency problem of indirect inference estimator caused by measurement error in the observed series. We show that this inconsistency is originated by a mismatch between the binding function implied by the observed data and that obtained by simulation. We propose a general method to deal with this error-in-variable problem in the indirect inference framework. The solution is to jointly estimate the nuisance parameters of the error term distribution and the structural ones. Hence, the simulated series used to match the auxiliary parameters must be contaminated by the noise. Under standard assumptions, this estimator is consistent and asymptotically normal. We show that ARMA models contaminated by i.i.d. noise can be estimated by indirect inference when the autoregressive order of the signal is larger than the moving average. This results may be helpful when estimating economic and financial models by indirect inference, when the variables have an ARMA plus noise reduced form. One of these cases is represented by the Heston SV model, for which we prove that the HAR-RV auxiliary model guarantees the identification. Monte Carlo simulations show the viability of the proposed method in this framework and highlights the trade-off between bias reduction and efficiency as the sampling frequency changes. We also show that the Heston *SV* model with jumps and leverage can be identified by a multivariate model for daily returns, RV and signed jump variation. The empirical application illustrates the practical usefulness of the proposed methodology and the need for a correct specification of the conditional distribution of the microstructure noise term. A detailed analysis of the robustness/sensitivity of the indirect inference estimation to misspecifications of the measurement error is left to future research.

References

Ait-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, 18(2):351–416.

Andersen, T., Dobrev, D., and Schaumbur, E. (2012). Robust volatility estimation using nearest-neighbor truncation. *Journal of Econometrics*, 169:75–93.

Anderson, B. D. O. and Deistler, M. (1984). Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, 5(1):1–13.

Anderson, T. W. (1984). Estimating linear statistical relationships. *The Annals of Statistics*, 12(1):pp. 1–45.

Aruoba, S. B., Diebold, F. X., Nalewaik, J., Schorfheide, F., and Song, D. (2013). Improving GDP measurement: A measurement-error perspective. Working Paper 18954, NBER.

Bandi, F. and Russell, J. (2006). Separating market microstructure noise from volatility. *Journal of Financial Economics*, 79:655–692.

Bandi, F. M. and Renò, R. (2012). Time-varying leverage effects. *Journal of Econometrics*, 169(1):94 – 113. Recent Advances in Panel Data, Nonlinear and Nonparametric Models: A Festschrift in Honor of Peter C.B. Phillips.

Bansal, R., Gallant, A. R., Hussey, R., and Tauchen, G. (1995). Non-parametric estimation of structural models for high-frequency currency market data. *Journal of Econometrics*, 66(1-2):251–287.

Barndorff-Nielsen, O. and Shephard, N. (2001). Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of Royal Statistical Society. Series B*, 63:167–241.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.

Barndorff-Nielsen, O. E., Kinnebrock, S., and Shephard, N. (2010). Measuring downside risk-realised semivariance. In T. Bollerslev, J. R. and Watson, M., editors, *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle.*

Barndorff-Nielsen, O. E., Lunde, A., Shephard, N., and Veraart, A. E. (2014). Integer-valued trawl processes: A class of stationary infinitely divisible processes. *Scandinavian Journal of Statistics*, 41(3):693–724.

Barndorff-Nielsen, O. E. and Shephard, N. (2002a). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of Royal Statistical Society. Series B*, 64:253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5):457–477.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 9(1):69–107.

Bollerlsev, T. and Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109(109):33–65.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*. Berlin:Springer-Verlag.

Broze, L., Scaillet, O., and Zakoian, J.-M. (1998). Quasi-indirect inference for diffusion processes. *Econometric Theory*, pages 161–186.

Chanda, K. C. (1995). Large sample analysis of autoregressive moving-average models with errors in variables. *Journal of Time Series Analysis*, 16(1):1–15.

Chanda, K. C. (1996). Asymptotic properties of estimators for autoregressive models with errors in variables. *The Annals of Statistics*, 24(1):423–430.

Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225 – 257.

Corradi, V. and Distaso, W. (2006). Semi-parametric comparison of stochastic volatility models using realized measures. *Review of Economic Studies*, 73(3):635–667.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7:174–196.

Corsi, F. and Renò, R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business and Economic Statistics*, 30:368–380.

Creel, M. and Kristensen, D. (2015). ABC of SV: Limited information likelihood inference in stochastic volatility jump-diffusion models. *Journal of Empirical Finance*, 31(C):85–108.

de Luna, X. and Genton, M. G. (2001). Robust simulation-based estimation of ARMA models. *Journal of Computational and Graphical Statistics*, 10(2):370–387.

Dridi, R., Guay, A., and Renault, E. (2007). Indirect inference and calibration of dynamic stochastic general equilibrium models. *Journal of Econometrics*, 136(2):397 – 430.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Gallant, A., Hsieh, D., and Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics*, 81(1):159 – 192.

Gallant, A. R. and Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12:657–681.

Garcia, R., Lewis, M.-A., Pastorello, S., and Renault, E. (2011). Estimation of objective and risk-neutral distributions based on moments of integrated volatility. *Journal of Econometrics*, 160(1):22–32.

Ghysels, E. and Khalaf, L. (2003). Simulation based inference in moving average models. *Annales d'economie et de statistique*, 69:85–99.

Gospodinov, N., Komunjer, I., and Ng, S. (2015). Simulated minimum distance estimation of dynamic models with errors-in-variables. Technical report.

Gouriéroux, C. and Monfort, A. (1996). *Simulation-Based Econometric Methods*. Oxford University Press.

Gouriéroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics*, 8:85–118.

Granger, C. W. J. and Morris, M. J. (1976). Time series modelling and interpretation. *Journal of the Royal Statistical Society*, 146:246–257.

Grassi, S. and Santucci de Magistris, P. (2015). It's all about volatility of volatility: Evidence from a two-factor stochastic volatility model. *Journal of Empirical Finance*, 30(0):62 – 78.

Grether, D. M. and Maddala, G. S. (1973). Errors in variables and serially correlated disturbances in distributed lag models. *Econometrica*, 41(2):255–262.

Hannan, E. J. (1963). Regression for time series with errors of measurement. *Biometrika*, 50(3/4):293–302.

Hansen, P., Large, J., and Lunde, A. (2008). Moving average-based estimators of integrated variance. *Econometric Reviews*, 27(1-3):79–111.

Hansen, P. R. and Lunde, A. (2014). Estimating the persistence and the autocorrelation function of a time series that is measured with error. *Econometric Theory*, 30(2010-8):60–93.

Heston, S. (1993). A closed form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–344.

Huang, X. and Tauchen, G. (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics*, 3(4):456–499.

Jiang, W. and Turnbull, B. (2004). The indirect method: Inference based on intermediate statistics - a synthesis and examples. *Statistical Science*, 19:239–263.

Jones, R. H. (1980). Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, 22(3):389–395.

Komunjer, I. and Ng, S. (2014). Measurement errors in dynamic models. *Econometric Theory*, 30:150–175.

Lee, H. J. and Shin, W. D. (1997). Maximum likelihood estimation for ARMA models in the presence of ARMA errors. *Communications in Statistics - Theory and Methods*, 26(5):1057–1072.

Maravall, A. (1979). *Indentification in Dynamic Shock-Error Models*. Berlin: Spriger-Verlag.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, 17:475–508.

Meddahi, N. (2003). ARMA representation of integrated and realized variances. *Econometrics Journal*, 6:335–356.

Moran, P. A. P. (1971). Estimating structural and functional relationships. *Journal of Multivariate Analysis*, (1):232–255.

Nowak, E. (1985). Global identification of the dynamic shock-error model. *Journal of Econometrics*, 27(2):211 – 219.

Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 3:683–697.

Robinson, P. (1986). On the errors-in-variables problem for time series. *Journal of Multivariate Analysis*, 19(2):240–250.

Rossi, E. and Santucci de Magistris, P. (2014). Estimation of long memory in integrated variance. *Econometric Reviews*, 33:7:785–814.

Savov, A. (2011). Asset pricing with garbage. *The Journal of Finance*, 66(1):177–201.

Smith, J. A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8:63–84.

Solo, V. (1986). Identifiability of time series models with errors in variables. *Journal of Applied Probability - Essays in Time Series and Allied Processes Papers in Honour of E .J. Hannan*, 23:63–71.

Song, S., Schennach, S. M., and White, H. (2015). Identification and estimation of nonseparable models with measurement errors. *Quantitative Economics*, forthcoming.

Staudenmayer, J. and Buonaccorsi, J. P. (2005). Measurement error in linear autoregressive models. *Journal of the American Statistical Association*, 100(471):pp. 841–852.

Todorov, V. (2009). Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *Journal of Econometrics*, 148(2):131 – 148.

Zhang, L., Mikland, P., and Ait-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association*, 100:1394–1411.

A Proofs

A.1 A consistent and asymptotically normal indirect inference estimator with ME

Assumption 4 *Similarly to Gouriéroux et al. (1993), we assume that:*

i. the normalized function $Q_T(\underline{x}_T^{(s)}(\zeta), \beta)$ uniformly converges in (ζ, β) to a deterministic function $Q_\infty(\zeta, \beta)$ when T diverges.

ii. The limit function $Q_\infty(\zeta, \beta)$ has a unique maximum with respect to β. The maximum is $b(\zeta) = \arg\max_{\beta \in \mathcal{B}} Q_\infty(\zeta, \beta)$.

iii. The functions $Q_T(\underline{x}_T^{(s)}(\zeta), \beta)$ and Q_∞ are differentiable with respect to β.

iv. The only solution of the asymptotic first order condition is associated with $\beta_0 = b(\zeta_0)$.

v. $b(\zeta)$ is a one-to-one (locally injective) function and $\frac{\partial b(\zeta_0)}{\partial \zeta'}$ is a full-column rank matrix.

Proposition A.1 *Under Assumptions 1-4 the indirect inference estimator $\hat{\zeta}_{ST}$ is consistent. Moreover, under regularity conditions, for $T \to \infty$ and S fixed, the indirect inference estimator $\hat{\zeta}_{ST}$ is asymptotically normal, with*

$$\sqrt{T}(\hat{\zeta}_{ST} - \zeta_0) \xrightarrow{d} N\Big(0, W(S, \Omega)\Big) \tag{A.1}$$

where $W(S, \Omega)$ is given in Gouriéroux and Monfort (1996, p.70).

Proof: See Gouriéroux et al. (1993).
The proof Proposition A.1 follows directly from the results in Gouriéroux et al. (1993).

A.2 Proof of Proposition 3.1

Let $q = m + 2$ and $p + 1 = 3 + r + l$, with $q \geq p + 1$. Given the AR(m) model in (19), the OLS estimates of $\beta = (\phi_0, \phi_1, \ldots, \phi_m)'$ converges in probability, when T diverges, to $Q_{ZZ}^{-1} Q_{ZX}$ (see Proposition 8.10.1 and Theorem 8.1.1 in Brockwell and Davis, 1991), while

$$p \lim_{T\to\infty} \hat{\sigma}_e^2 = Q_{XX} - Q_{XZ} Q_{ZZ}^{-1} Q_{ZX}$$

where

$$Q_{ZZ} = \begin{bmatrix} 1 & \mu_x & \mu_x & \ldots & \mu_x \\ \mu_x & \gamma_x(0) + \mu_x^2 & \gamma_x(1) + \mu_x^2 & \ldots & \gamma_x(m-1) + \mu_x^2 \\ \mu_x & \gamma_x(1) + \mu_x^2 & \gamma_x(0) + \mu_x^2 & \ldots & \gamma_x(m-2) + \mu_x^2 \\ \vdots & & & & \vdots \\ \mu_x & \gamma_x(m-1) + \mu_x^2 & \gamma_x(m-2) + \mu_x^2 & \ldots & \gamma_x(0) + \mu_x^2 \end{bmatrix}$$

and $Q_{XZ} = [\mu_x, \gamma_x(1) + \mu_x^2, \ldots, \gamma_x(m) + \mu_x^2] = Q'_{ZX}$. Since we assume that $E(u_t) = 0$, then $\mu_x = \mu \equiv E[y_t]$. The variance and the autocovariances of x_t are $Q_{XX} \equiv \gamma_x(0) = \gamma(0) + \sigma_u^2$ and $\gamma_x(k) = \gamma(k)$ when $k \neq 0$, where $\gamma(k) = \text{Cov}[y_t, y_{t-k}]$. Thus the binding function is

$$b(\zeta) = \begin{bmatrix} Q_{ZZ}^{-1} Q_{ZX} \\ Q_{XX} - Q_{XZ} Q_{ZZ}^{-1} Q_{ZX} \end{bmatrix}. \tag{A.2}$$

In order to find the Jacobian matrix of $b(\zeta)$ consider the differential for each component of $b(\zeta)$. Since

$$\begin{aligned} \text{d}(Q_{ZZ}^{-1} Q_{ZX}) &= (\text{d}Q_{ZZ}^{-1}) Q_{ZX} + Q_{ZZ}^{-1} (\text{d}Q_{ZX}) \\ &= -(Q_{ZZ}^{-1} \text{d}Q_{ZZ} Q_{ZZ}^{-1}) Q_{ZX} + Q_{ZZ}^{-1} \text{d}Q_{ZX} \end{aligned}$$

and given that $\text{vec}(ABCD) = (D'C' \otimes A)\text{vec}(B)$ for suitably dimensioned matrices, where the *vec* operator transforms a matrix into a vector by stacking the columns of the matrix one underneath the other,

$$\begin{aligned}\text{dvec}(Q_{ZZ}^{-1}Q_{ZX}) &= -(Q'_{ZX}Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1})\text{dvec}(Q_{ZZ}) + Q_{ZZ}^{-1}\frac{\partial Q_{ZX}}{\partial \zeta'}\text{d}\zeta \\ &= \left[-(Q'_{ZX}Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1})\frac{\partial \text{vec}(Q_{ZZ})}{\partial \zeta'} + Q_{ZZ}^{-1}\frac{\partial Q_{ZX}}{\partial \zeta'}\right]\text{d}\zeta\end{aligned}$$

The differential of the second component of $b(\zeta)$ is

$$\begin{aligned}\text{d}[Q_{XX}] - \text{d}[Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}] &= \text{d}[Q_{XX}] - \text{d}[Q_{XZ}]Q_{ZZ}^{-1}Q_{ZX} - Q_{XZ}\text{d}[Q_{ZZ}^{-1}]Q_{ZX} - Q_{XZ}Q_{ZZ}^{-1}\text{d}[Q_{ZX}] \\ &= \text{d}[Q_{XX}] - \left\{Q_{XZ}\text{d}[Q_{ZZ}^{-1}]Q_{ZX}\right\} - 2\left\{Q_{XZ}Q_{ZZ}^{-1}\text{d}[Q_{ZX}]\right\}\end{aligned}$$

where

$$\text{d}[Q_{XX}] = \frac{\partial Q_{XX}}{\partial \zeta'}\text{d}\zeta,$$

$$\begin{aligned}Q_{XZ}\text{d}[Q_{ZZ}^{-1}]Q_{ZX} &= -Q_{XZ}(Q_{ZZ}^{-1}\text{d}Q_{ZZ}Q_{ZZ}^{-1})Q_{ZX} = -\text{tr}\left(Q_{XZ}(Q_{ZZ}^{-1}\text{d}Q_{ZZ}Q_{ZZ}^{-1})Q_{ZX}\right) \\ &= -\text{vec}(Q_{ZX}Q_{XZ})'(Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1})\frac{\partial \text{vec}Q_{ZZ}}{\partial \zeta'}\text{d}\zeta\end{aligned}$$

and

$$\left\{Q_{XZ}Q_{ZZ}^{-1}\text{d}[Q_{ZX}]\right\} = Q_{XZ}Q_{ZZ}^{-1}\frac{\partial (Q_{ZX})}{\partial \zeta'}\text{d}\zeta$$

Finally,

$$\begin{aligned}&\text{d}[Q_{XX}] - \text{d}[Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}] = \\ &\quad \left[\frac{\partial Q_{XX}}{\partial \zeta'} + \text{vec}(Q_{ZX}Q_{XZ})'(Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1})\frac{\partial \text{vec}Q_{ZZ}}{\partial \zeta'} - 2Q_{XZ}Q_{ZZ}^{-1}\frac{\partial Q_{ZX}}{\partial \zeta'}\right]\text{d}\zeta.\end{aligned}$$

The Jacobian matrix is

$$\frac{\partial b(\zeta)}{\partial \zeta'} = \begin{bmatrix} -(Q'_{ZX}Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1})\frac{\partial \text{vec}Q_{ZZ}}{\partial \zeta'} + Q_{ZZ}^{-1}\frac{\partial Q_{ZX}}{\partial \zeta'} \\ \frac{\partial Q_{XX}}{\partial \zeta'} + \text{vec}(Q_{ZX}Q_{XZ})'(Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1})\frac{\partial \text{vec}Q_{ZZ}}{\partial \zeta'} - 2Q_{XZ}Q_{ZZ}^{-1}\frac{\partial Q_{ZX}}{\partial \zeta'} \end{bmatrix}.$$

it can be written as

$$\begin{aligned}\frac{\partial b(\zeta)}{\partial \zeta'} &= \begin{bmatrix} 0 \\ \frac{\partial Q_{XX}}{\partial \zeta'} \end{bmatrix} + \begin{bmatrix} -(Q'_{ZX}Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1}) & Q_{ZZ}^{-1} \\ \text{vec}(Q_{ZX}Q_{XZ})'(Q_{ZZ}^{-1} \otimes Q_{ZZ}^{-1}) & -2Q_{XZ}Q_{ZZ}^{-1} \end{bmatrix} \begin{bmatrix} \frac{\partial \text{vec}Q_{ZZ}}{\partial \zeta'} \\ \frac{\partial Q_{ZX}}{\partial \zeta'} \end{bmatrix} \\ &= A + BC\end{aligned}$$

The rank of A is $p+1$ no matter what it is the value of l and r. The rank of B is $m+2$ which is larger by assumption than $p+1$. It follows that the rank of $\partial b(\zeta_0)/\partial \zeta'$ depends on the column rank of the $((m+1)^2 + (m+1)) \times (p+1)$ matrix C. The rows of the matrix C contain the partial derivatives of $\mu^2 + \gamma(k), k = 0, 1, \ldots, m$ with respect to c, $\alpha = (\alpha_1, \ldots, \alpha_l)'$, $\varphi = (\varphi_1, \ldots, \varphi_r)'$, σ_ε^2 and σ_u^2, i.e.

$$\begin{bmatrix} \frac{\partial(\mu^2+\gamma_x(k))}{\partial c} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \alpha'} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \varphi'} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_\varepsilon^2} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_u^2} \end{bmatrix}$$

where

$$
\begin{aligned}
\frac{\partial(\mu_x^2+\gamma_x(k))}{\partial c} &= \frac{2c}{(1-\sum_{i=1}^{r}\alpha_i)^2}\\
\frac{\partial(\mu_x^2+\gamma_x(k))}{\partial \alpha'} &= 2\mu\frac{\partial \mu}{\partial \alpha'}+\frac{\partial \gamma(k)}{\partial \alpha'}\\
\frac{\partial(\mu_x^2+\gamma_x(k))}{\partial \varphi'} &= \frac{\partial \gamma(k)}{\partial \varphi'}\\
\frac{\partial(\mu_x^2+\gamma_x(k))}{\partial \sigma_\varepsilon^2} &= \frac{\partial \gamma(k)}{\partial \sigma_\varepsilon^2}\\
\frac{\partial(\mu_x^2+\gamma_x(k))}{\partial \sigma_u^2} &= \frac{\partial \gamma_x(k)}{\partial \sigma_u^2} = \begin{cases}1 & k=0\\ 0 & k\neq 0\end{cases}
\end{aligned}
$$

First, consider the case of ARMA(1,0). The rows of C have the following expression

$$
c_k = \begin{bmatrix}\frac{\partial(\mu^2+\gamma_x(k))}{\partial c} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \alpha_1} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_\varepsilon^2} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_u^2}\end{bmatrix}'.
$$

If C has reduced rank then $Cw=0$ for $w\neq 0$, i.e. there exists a vector $w=[w_1,w_2,w_3w_4]'$ such that $c_k'w=0$ for all rows of C. The partial derivatives in c_k are

$$
\begin{aligned}
\frac{\partial(\mu^2+\gamma_x(k))}{\partial c} &= \frac{2c}{(1-\alpha_1)^2}\\
\frac{\partial(\mu^2+\gamma_x(k))}{\partial \alpha_1} &= 2\mu\frac{\partial \mu}{\partial \alpha_1}+\left(k\alpha_1^{k-1}\gamma(0)+\alpha_1^k\frac{\partial \gamma(0)}{\partial \alpha_1}\right)\\
\frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_\varepsilon^2} &= \alpha_1^k\frac{\partial \gamma(0)}{\partial \sigma_\varepsilon^2}
\end{aligned}
$$

The reduced rank condition implies

$$
w_1\Big(\frac{\partial(\mu^2+\gamma_x(0))}{\partial c}\Big)+w_2\Big(\frac{\partial(\mu^2+\gamma_x(0))}{\partial \alpha_1}\Big)+w_3\Big(\frac{\partial(\mu^2+\gamma_x(0))}{\partial \sigma_\varepsilon^2}\Big)+w_4=0 \text{ for } k=0
$$

and for $k>0$

$$
w_1\Big(\frac{\partial(\mu^2+\gamma_x(k))}{\partial c}\Big)+w_2\Big(\frac{\partial(\mu^2+\gamma_x(k))}{\partial \alpha_1}\Big)+w_3\Big(\frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_\varepsilon^2}\Big)=0.
$$

Equating the two expressions above

$$
w_2\Big[k\alpha_1^{k-1}\gamma(0)+\alpha_1^k\Big(\frac{\partial \gamma(0)}{\partial \alpha_1}-\frac{\partial \gamma_x(0)}{\partial \alpha_1}\Big)\Big]+w_3\frac{\partial \gamma_x(0)}{\partial \sigma_\varepsilon^2}(a^k-1)=w_4
$$

it is easy to see that w_4 varies with k which implies that cannot exist any vector $w\neq 0$ which lies in the null column subspace of C. We conclude that in the case of ARMA(1,0) the column rank of C is full. Now, we show that when $r=l=1$, on the contrary, the rank of C is smaller than $p+1$. In this case it is easy to show that there exists a vector $w\neq 0$ such that $Cw=0$. For ARMA(1,1) the rows of $\partial b(\zeta)/\partial\zeta'$ consist of

$$
\begin{bmatrix}\frac{\partial(\mu^2+\gamma_x(k))}{\partial c} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \alpha_1} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \varphi_1} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_\varepsilon^2} & \frac{\partial(\mu^2+\gamma_x(k))}{\partial \sigma_u^2}\end{bmatrix},
$$

where the variance and the autocovariances are

$$
\gamma_x(0)=\sigma_\varepsilon^2\frac{1+\varphi_1^2+2\alpha_1\varphi_1}{1-\alpha_1^2}+\sigma_u^2
$$

$$\gamma(k) = \sigma_\varepsilon^2 \alpha_1^{k-1} \frac{(\alpha_1 + \varphi_1)(1 + \alpha_1\varphi_1)}{1 - \alpha_1^2} \quad k > 1$$

with

$$\frac{\partial\gamma(k)}{\partial\varphi_1} = \sigma_\varepsilon^2 \alpha_1^{k-1} \frac{(1 + \alpha_1\varphi_1) + \alpha_1(\alpha_1 + \varphi_1)}{1 - \alpha_1^2}$$
$$\frac{\partial\gamma(k)}{\partial\sigma_\varepsilon^2} = \alpha_1^{k-1} \frac{(\alpha_1 + \varphi_1)(1 + \alpha_1\varphi_1)}{1 - \alpha_1^2}.$$

The vector w that is orthogonal to the rows of C is

$$w = (0, 0, 1, -\sigma_\varepsilon^2 \frac{(1 + \alpha_1\varphi_1) + \alpha_1(\alpha_1 + \varphi_1)}{(\alpha_1 + \varphi_1)(1 + \alpha_1\varphi_1)}, w_5)'$$

where

$$w_5 = \sigma_\varepsilon^2 \frac{(1 + \alpha_1\varphi_1) + \alpha_1(\alpha_1 + \varphi_1)}{(\alpha_1 + \varphi_1)(1 + \alpha_1\varphi_1)} \frac{\partial\gamma(0)}{\partial\sigma_\varepsilon^2} - \frac{\partial\gamma(0)}{\partial\varphi_1}.$$

For the case $r < l$ an analogous argument shows that the column rank of C is reduced. Thus when $r \leq l$ the rank of C is smaller than $p + 1$.

A.3 Proof of Proposition 4.1

The HAR-RV model can be written as an AR(22) with linear restrictions on the autoregressive parameters

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_{22} x_{t-22} + e_t \tag{A.3}$$

where $\alpha_0 = \phi_1$, $\alpha_1 = (\phi_2 + \phi_3/5 + \phi_4/22)$, $\{\alpha_2, ..., \alpha_5\} = (\phi_3/5 + \phi_4/22)$ and $\{\alpha_6, ..., \alpha_{22}\} = (\phi_4/22)$. Let R be the 23×4 matrix with the linear restrictions, then a compact expression for α is $\alpha = R\phi$. The restricted AR(22) model in (A.3) can be estimated by OLS imposing the restriction contained in the matrix R. The (23×4) matrix R is

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{5} & \frac{1}{22} \\ 0 & 0 & \frac{1}{5} & \frac{1}{22} \\ 0 & 0 & \frac{1}{5} & \frac{1}{22} \\ 0 & 0 & \frac{1}{5} & \frac{1}{22} \\ 0 & 0 & \frac{1}{5} & \frac{1}{22} \\ 0 & 0 & 0 & \frac{1}{22} \\ 0 & 0 & 0 & \frac{1}{22} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \frac{1}{22} \end{bmatrix}. \tag{A.4}$$

The OLS estimate of ϕ is

$$\hat{\phi}_T = \left[R'\Big(\sum_t z_t z_t'\Big)R\right]^{-1} R' \sum_t (z_t x_t)$$

where $z_t = (1, x_{t-1}, \ldots, x_{t-22})'$. Under standard assumptions, it can be shown that $\frac{1}{T}\sum_t (z_t z_t') \xrightarrow{p} E[z_t z_t'] \equiv Q_{ZZ}$, where

$$Q_{ZZ} = \begin{bmatrix} 1 & \mu & \mu & \ldots & \mu \\ \mu & \gamma(0) + \mu^2 & \gamma(1) + \mu^2 & \ldots & \gamma(21) + \mu^2 \\ \mu & \gamma(1) + \mu^2 & \gamma(0) + \mu^2 & \ldots & \gamma(20) + \mu^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu & \gamma(21) + \mu^2 & \gamma(20) + \mu^2 & \ldots & \gamma(0) + \mu^2 \end{bmatrix}$$

and $\mu = E[x_t]$. The probability limit of $\hat{\alpha}$ is therefore

$$p\lim_{T\to\infty} \hat{\alpha}_T = Q_{ZZ}^{-1} Q_{ZX}$$

$Q_{ZX} = E[z_t x_t] = [\mu, \mu^2 + \gamma(1), \dots, \mu^2 + \gamma(22)]'$ and $Q_{XX} = E[x_t^2] = \gamma(0) + \mu^2$. The limit of $\hat{\phi}_T$

$$\begin{aligned} p\lim_{T\to\infty} \hat{\phi}_T &= \left[R' p\lim_{T\to\infty} \Big(\sum_t z_t z_t'\Big) R\right]^{-1} R' p\lim_{T\to\infty} \sum_t (z_t x_t) \\ &= \left[R' Q_{ZZ} R\right]^{-1} R' Q_{ZX} \end{aligned}$$

The matrix Q_{ZZ} and the vector Q_{ZX} both depend on the structural parameters ζ. The estimator of the variance of e_t is

$$\begin{aligned} \hat{\sigma}^2_{e,T} &= \frac{\sum_t \hat{e}_t^2}{T} \\ p\lim_{T\to\infty} \hat{\sigma}^2_{e,T} &= Q_{XX} - Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' Q_{ZX}. \end{aligned}$$

Then the binding function results to be

$$b(\zeta) = \begin{bmatrix} \left[R' Q_{ZZ} R\right]^{-1} R' Q_{ZX} \\ Q_{XX} - [Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' Q_{ZX}] \end{bmatrix}$$

To calculate the derivative of $b(\zeta)$ with respect to ζ we use the differential of $b(\zeta)$:

$$\mathrm{d}\{\left[R' Q_{ZZ} R\right]^{-1} R' Q_{ZX}\} = \mathrm{d}\{\left[R' Q_{ZZ} R\right]^{-1}\} R' Q_{ZX} + \left[R' Q_{ZZ} R\right]^{-1} R' \mathrm{d}\{Q_{ZX}\} \tag{A.5}$$

The differential of first term in the RHS of (A.5)

$$\begin{aligned} \mathrm{d}\{\left[R' Q_{ZZ} R\right]^{-1}\} R' Q_{ZX} &= -(R' Q_{ZZ} R)^{-1} \mathrm{d}\{R' Q_{ZZ} R\} (R' Q_{ZZ} R)^{-1} R' Q_{ZX} \\ &= -(R' Q_{ZZ} R)^{-1} R' \mathrm{d}\{Q_{ZZ}\} R (R' Q_{ZZ} R)^{-1} R' Q_{ZX}, \end{aligned}$$

taking the vec of both sides

$$\begin{aligned} \mathrm{vec}\Big[\mathrm{d}\{\left[R' Q_{ZZ} R\right]^{-1}\} R' Q_{ZX}\Big] &= \mathrm{vec}\Big[-(R' Q_{ZZ} R)^{-1} R' \mathrm{d}\{Q_{ZZ}\} R (R' Q_{ZZ} R)^{-1} R' Q_{ZX}\Big] \\ &= -\Big[Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' \otimes (R' Q_{ZZ} R)^{-1} R'\Big] \mathrm{dvec}(Q_{ZZ}) \\ &= -\Big[Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' \otimes (R' Q_{ZZ} R)^{-1} R'\Big] \frac{\partial \mathrm{vec} Q_{ZZ}}{\partial \zeta'} \mathrm{dvec}\zeta \end{aligned}$$

Thus

$$\begin{aligned} &\mathrm{d}\{\left[R' Q_{ZZ} R\right]^{-1} R' Q_{ZX}\} = \\ &\qquad \Big\{-\Big[Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' \otimes (R' Q_{ZZ} R)^{-1} R'\Big] \frac{\partial \mathrm{vec} Q_{ZZ}}{\partial \zeta'} + (R' Q_{ZZ} R)^{-1} R' \frac{\partial Q_{ZX}}{\partial \zeta'}\Big\} \mathrm{d}\zeta. \end{aligned}$$

Now, the differential of the last row in $b(\zeta)$

$$\mathrm{d}\{Q_{XX} - [Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' Q_{ZX}]\} = \mathrm{d}\{Q_{XX}\} - \mathrm{d}\{[Q_{XZ} R (R' Q_{ZZ} R)^{-1} R' Q_{ZX}]\}$$

with

$$\mathrm{d}\{Q_{XX}\} = \frac{\partial Q_{XX}}{\partial \zeta'} \mathrm{d}\zeta$$

$$\begin{aligned}\mathrm{d}\{[Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'Q_{ZX}]\} &= 2[Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'\mathrm{d}\{Q_{ZX}\}] + [Q_{XZ}R\mathrm{d}\{(R'Q_{ZZ}R)^{-1}\}R'Q_{ZX}] \\ &= 2[Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'\mathrm{d}\{Q_{ZX}\}] \\ &\quad - Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'\mathrm{d}\{Q_{ZZ}\}R(R'Q_{ZZ}R)^{-1}R'Q_{ZX}\end{aligned}$$

The second element in the previous expression can be rewritten, using the trace operator, as

$$\begin{aligned}&Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'\mathrm{d}\{Q_{ZZ}\}R(R'Q_{ZZ}R)^{-1}R'Q_{ZX} \\ &\quad = \mathrm{tr}(Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'\mathrm{d}\{Q_{ZZ}\}R(R'Q_{ZZ}R)^{-1}R'Q_{ZX}) \\ &\quad = \mathrm{vec}(Q_{ZX}Q_{XZ})'[(R(R'Q_{ZZ}R)^{-1}R') \otimes (R(R'Q_{ZZ}R)^{-1}R')]\mathrm{dvec}(Q_{ZZ}).\end{aligned}$$

The complete expression for the differential of the last row of $b(\zeta)$ is then

$$\begin{aligned}\mathrm{d}\{Q_{XX} - [Q_{XZ}R(R'Q_{ZZ}R)^{-1}R'Q_{ZX}]\} &= \frac{\partial Q_{XX}}{\partial \zeta'}\mathrm{d}\zeta - 2[Q_{XZ}R(R'Q_{ZZ}R)^{-1}R']\frac{\partial Q_{ZX}}{\partial \zeta'}\mathrm{d}\zeta \\ &+ \mathrm{vec}(Q_{ZX}Q_{XZ})'[(R(R'Q_{ZZ}R)^{-1}R') \otimes (R(R'Q_{ZZ}R)^{-1}R')]\frac{\partial \mathrm{vec}Q_{ZZ}}{\partial \zeta'}\mathrm{d}\zeta\end{aligned}$$

The Jacobian matrix of $b(\zeta)$

$$\frac{\partial b(\zeta)}{\partial \zeta'} =$$

$$\begin{bmatrix} -\left[Q_{XZ}R(R'Q_{ZZ}R)^{-1}R' \otimes (R'Q_{ZZ}R)^{-1}R'\right]\frac{\partial \mathrm{vec}Q_{ZZ}}{\partial \zeta'} + (R'Q_{ZZ}R)^{-1}R'\frac{\partial Q_{ZX}}{\partial \zeta'} \\ \frac{\partial Q_{XX}}{\partial \zeta'} + \mathrm{vec}(Q_{ZX}Q_{XZ})'[(R(R'Q_{ZZ}R)^{-1}R') \otimes (R(R'Q_{ZZ}R)^{-1}R')]\frac{\partial \mathrm{vec}Q_{ZZ}}{\partial \zeta'} - 2[Q_{XZ}R(R'Q_{ZZ}R)^{-1}R']\frac{\partial Q_{ZX}}{\partial \zeta'}\end{bmatrix}$$

In order to prove that the Jacobian matrix has full-column rank, i.e. equal to 4, we focus on the matrix (as in the Proof of Proposition 3.1)

$$C = \begin{bmatrix}\frac{\partial \mathrm{vec}Q_{ZZ}}{\partial \zeta'} \\ \frac{\partial Q_{ZX}}{\partial \zeta'}\end{bmatrix}.$$

This matrix contains the partial derivatives of the variance and the autocovariances of the RV_t in the case of the Heston model, see (31). Let denote c_j the row of C which contains the partial derivative of $\gamma(j)$, i.e.

$$c_j = \begin{bmatrix}\frac{\partial \gamma(j)}{\partial \kappa} & 2\mu + \frac{\partial \gamma(j)}{\partial \omega} & \frac{\partial \gamma(j)}{\partial \varsigma} & 2\mu\Delta + \frac{\partial \gamma(j)}{\partial \sigma_\nu^2}\end{bmatrix}'.$$

The matrix C has reduced column rank if there exists a vector $w = [w_1, w_2, w_3, w_4]' \neq 0$ such that

$$c_j' w = 0$$

and this must hold for all rows of C, that is $Cw = 0$. Since

$$\frac{\partial \gamma(j)}{\partial \sigma_\nu^2} = 0 \quad \text{for } j > 0$$

we have that for $j > 0$

$$w_1 \frac{\partial \gamma(j)}{\partial \kappa} + w_2 \left(2\mu + \frac{\partial \gamma(j)}{\partial \omega}\right) + w_3 \left(2\mu\Delta + \frac{\partial \gamma(j)}{\partial \varsigma}\right) = 0 \tag{A.6}$$

and for $j = 0$

$$w_1 \frac{\partial \gamma(0)}{\partial \kappa} + w_2 \left(2\mu + \frac{\partial \gamma(0)}{\partial \omega}\right) + w_3 \frac{\partial \gamma(0)}{\partial \varsigma} + w_4 \left(2\mu\Delta + \frac{\partial \gamma(0)}{\partial \varsigma}\right) = 0. \tag{A.7}$$

Now equating (A.6) and (A.7), we get

$$w_1\Big[\frac{\partial\gamma(j)}{\partial\kappa}-\frac{\partial\gamma(0)}{\partial\kappa}\Big]+w_2\Big[\frac{\partial\gamma(j)}{\partial\omega}-\frac{\partial\gamma(0)}{\partial\omega}\Big]+w_3\Big[\frac{\partial\gamma(j)}{\partial\varsigma}-\frac{\partial\gamma(0)}{\partial\varsigma}\Big]-w_4\frac{\partial\gamma(0)}{\partial\sigma_\nu^2}=0. \tag{A.8}$$

where the expressions in parenthesis are functions of j. This means that the value of w_4 which satisfies (A.8) is

$$w_4=\Big[\frac{\partial\gamma(0)}{\partial\sigma_\nu^2}\Big]^{-1}\Big\{w_1\Big[\frac{\partial\gamma(j)}{\partial\kappa}-\frac{\partial\gamma(0)}{\partial\kappa}\Big]+w_2\Big[\frac{\partial\gamma(j)}{\partial\omega}-\frac{\partial\gamma(0)}{\partial\omega}\Big]+w_3\Big[\frac{\partial\gamma(j)}{\partial\varsigma}-\frac{\partial\gamma(0)}{\partial\varsigma}\Big]\Big\}$$

which is obviously not constant since it depends on j. Thus the only vector w which satisfies (A.8) is the null vector. We can conclude that the matrix C has full column rank.

Lemma A.1 *(Moments of the compound Poisson process) Consider the compound Poisson process* $J(t)=\sum_{j=1}^{N(t)}\tau_j$, *where* $N(t)$ *is a homogeneous Poisson process with parameter* $\lambda>0$ *where the jump size is time invariant with* $\tau_j\sim i.i.d.N(\mu_\tau,\sigma_\tau^2)$. *The first four moments of* $J(t)$ *are*

$$\begin{aligned}
E[J(t)]&=\lambda\mu_\tau t\\
E[J(t)^2]&=\lambda^2\mu_\tau^2t^2+\lambda(\sigma_\tau^2+\mu_\tau^2)t\\
E[J(t)^3]&=(\lambda^3t^3+3\lambda^2t^2+\lambda t)\mu_\tau^3+3(\lambda^2t^2+\lambda t)\sigma_\tau^2\mu_\tau\\
E[J(t)^4]&=(\lambda^4t^4+6\lambda^3t^3+7\lambda^2t^2+\lambda t)\mu_\tau^4+6(\lambda^3t^3+3\lambda^2t^2+\lambda t)\mu_\tau^2\sigma_\tau^2+3(\lambda^2t^2+\lambda t)\sigma_\tau^4
\end{aligned}$$

Proof. Since the Gaussian distribution is closed with respect to the sum, it follows that $J(t)|N(t)=j\sim i.i.d.N(j\mu_\tau,j\sigma_\tau^2)$. Therefore, the s-th moment of $J(t)$ is given by the following formula

$$E[J(t)^s]=\sum_{j=0}^{\infty}\frac{\exp(-\lambda t)(\lambda t)^j}{j!}E\Big[\tau_j^s|N_t=j\Big],\quad s=1,2,... \tag{A.9}$$

where $E[\tau_j^s|N(t)=j]$ is the s-th moment of a Gaussian distribution with mean $j\mu_\tau$ and variance $j\sigma_\tau^2$. Integrating out the dependence on the Poisson process gives the desired result.

When the compound Poisson process is $J_{t,i}$ over the interval Δ, the moments are

$$\begin{aligned}
E[J_{t,i}]&=\Delta\lambda\mu_\tau\\
E[J_{t,i}^2]&=\Delta^2\lambda^2\mu_\tau^2+\Delta\lambda(\sigma_\tau^2+\mu_\tau^2)\\
E[J_{t,i}^3]&=(\Delta^3\lambda^3+3\Delta^2\lambda^2+\Delta\lambda)\mu_\tau^3+3(\Delta^2\lambda^2+\Delta\lambda)\sigma_\tau^2\mu_\tau\\
E[J_{t,i}^4]&=(\Delta^4\lambda^4+6\Delta^3\lambda^3+7\Delta^2\lambda^2+\Delta\lambda)\mu_\tau^4\\
&\quad+6(\Delta^3\lambda^3+3\Delta^2\lambda^2+\Delta\lambda)\mu_\tau^2\sigma_\tau^2+3(\Delta^2\lambda^2+\Delta\lambda)\sigma_\tau^4.
\end{aligned}$$

Research Papers 2013

2014-40: Søren Johansen and Lukasz Gatarek: Optimal hedging with the cointegrated vector autoregressive model

2014-41: Laurent Callot and Johannes Tang Kristensen: Vector Autoregressions with Parsimoniously Time Varying Parameters and an Application to Monetary Policy

2014-42: Laurent A. F. Callot, Anders B. Kock and Marcelo C. Medeiros: Estimation and Forecasting of Large Realized Covariance Matrices and Portfolio Choice

2014-43: Paolo Santucci de Magistris and Federico Carlini: On the identification of fractionally cointegrated VAR models with the F(d) condition

2014-44: Laurent Callot, Niels Haldrup and Malene Kallestrup Lamb: Deterministic and stochastic trends in the Lee-Carter mortality model

2014-45: Nektarios Aslanidis, Charlotte Christiansen, Neophytos Lambertides and Christos S. Savva: Idiosyncratic Volatility Puzzle: Infl¤uence of Macro-Finance Factors

2014-46: Alessandro Giovannelli and Tommaso Proietti: On the Selection of Common Factors for Macroeconomic Forecasting

2014-47: Martin M. Andreasen and Andrew Meldrum: Dynamic term structure models: The best way to enforce the zero lower bound

2014-48: Tim Bollerslev, Sophia Zhengzi Li and Viktor Todorov: Roughing up Beta: Continuous vs. Discontinuous Betas, and the Cross-Section of Expected Stock Returns

2914-49: Tim Bollerslev, Viktor Todorov and Lai Xu: Tail Risk Premia and Return Predictability

2014-50: Kerstin Gärtner and Mark Podolskij: On non-standard limits of Brownian semi-stationary

2014-51: Mark Podolskij : Ambit fields: survey and new challenges

2014-52: Tobias Fissler and Mark Podolskij: Testing the maximal rank of the volatility process for continuous diffusions observed with noise

2014-53: Cristina M. Scherrer: Cross listing: price discovery dynamics and exchange rate effects

2014-54: Claudio Heinrich and Mark Podolskij: On spectral distribution of high dimensional covariation matrices

2014-55: Gustavo Fruet Dias and Fotis Papailias: Forecasting Long Memory Series Subject to Structural Change: A Two-Stage Approach

2014-56: Torben G. Andersen, Nicola Fusari and Viktor Todorov: The Risk Premia Embedded in Index Options

2014-57: Eduardo Rossi and Paolo Santucci de Magistris: Indirect inference with time series observed with error