Identifying Volatility Risk Price Through Leverage Effect

Xu Cheng*, Eric Renault†, Paul Sangrey‡

This Version: January 29, 2023

Abstract

In option pricing models with stochastic volatility, changes in volatility affect risk premia through two channels: (1) the investor’s willingness to bear high volatility in order to get high expected returns, as measured by the underlying asset return risk price, and (2) the investor’s direct aversion to changes in future volatility, as measured by the volatility risk price. Disentangling these channels is difficult and poses a subtle identification problem. With a general exponentially affine Stochastic Discount Factor (SDF) model, we focus on the identification of two risk aversion parameters $\zeta_1$ and $\zeta_2$ that drive respectively the price of volatility risk and risk aversion to volatility in asset returns. Our main theoretical result shows that the presence of leverage effect (instantaneous causality between asset return and its volatility process) reverses two common beliefs regarding the identification of the price $\zeta_1$ of volatility risk. First, observations of the variance premium (difference between the risk neutral and historical expectations of the underlying asset return realized variance) may not allow for identification of the price $\zeta_1$ of volatility risk. Due to leverage, the variance premium actually depends on both risk aversion parameters $\zeta_1$ and $\zeta_2$. Second, observations of the price of the underlying asset (without any option price data) may actually allow for identification of the price of volatility risk, because this risk is latent in the asset return due to leverage. We quantify this identification challenge by adopting the discrete-time exponentially affine model of Han, Khrapov, and Renault (2020), an extension of the idea of “realizing smiles” as proposed by Corsi, Fusari, and La Vecchia (2013) for option pricing with data on realized volatility. In particular, we develop a minimum distance criterion that links the asset return risk price, the volatility risk price, and the leverage effect to well-behaved reduced-form parameters that govern the return and volatility’s joint distribution. The link functions are almost flat if the leverage effect is close to zero, raising an issue of (nearly) weak identification of the volatility risk price. We extend the conditional quasi-likelihood ratio test that Andrews and Mikusheva (2016) develop in a nonlinear GMM setting to a minimum distance framework. The resulting conditional quasi-likelihood ratio test is uniformly valid. We invert this test to derive robust confidence sets that provide correct coverage for the risk prices regardless of the leverage effect’s magnitude.

Keywords: weak identification, robust inference, stochastic volatility, leverage effect, market return, risk premium, volatility risk premium, risk price, confidence set, asymptotic size.

JEL Classification: C12, C14, C38, C58, G12.

*University of Pennsylvania, Department of Economics, 133 South 36th Street, Philadelphia, PA 19104; Email: xucheng@econ.upenn.edu
†University of Warwick, Department of Economics, Coventry, West Midlands, CV4 7AL, UK; Email: eric.renault@warwick.ac.uk
‡Noom Inc, 450 W 33rd St, New York, NY, 10001; Email: paul@sangrey.io
1 Introduction

When asking the question “Interpretable Asset Markets”, Bansal, Khatchatrian, and Yaron (2005)’s main conclusion is that “financial markets dislike economic uncertainty”. They document that “asset valuations drop as economic uncertainty rises”. This means that, to understand the risk premium of an asset price, one must consider not only the volatility of this asset price but also shocks to fundamentals that affect agents’ views about economic uncertainty. This explains the popularity of the long-run risk model of Bansal and Yaron (2004), which introduces a second risk factor in the pricing kernel beyond the future short-term return of the asset of interest.

As argued by Bollerslev, Tauchen, and Zhou (2009), the importance of economic uncertainty leads to “a two-factor structure for the endogenously determined equity risk premium in which the factors are directly related to the underlying volatility dynamics of consumption growth”. According to these authors, the so-called “variance risk premium”, defined as “the difference between the risk-neutralized expected return variation and the realized return variation”, “effectively isolates the factor associated with the volatility of consumption growth volatility”. A positive risk variance premium is then interpreted as “an insurance premium for an asset whose payoff is high when return variation is large (…), when investors perceive that the danger of big shocks to the state of the economy is high, the hedging premium increases, resulting in large variance premium” (Drechsler and Yaron, 2010).

In this paper, we follow the idea that volatility of volatility is a convenient way to capture investors’ aversion to economic uncertainty; however, we question the belief that the variance risk premium is an unambiguous way to econometrically identify investors’ specific risk aversion to volatility of volatility. More precisely, we start with a general exponentially affine pricing kernel with two risk factors: the daily asset return and its ex-post volatility. The latter is measured by the realized return variation constructed from high frequency intraday data. Obviously, there are alternative ways to characterize the two-dimensional vector space where the pricing kernel lies; however, as suggested by the empirical study of Bollerslev, Tauchen, and Zhou (2009), volatility of volatility is informative to predict stock market returns over the quarterly horizon. The methodology of our paper is general. It could be easily adapted if the realized return variation is replaced by a different risk factor.

The starting point of this paper is the remark that, to interpret asset prices based on two risk factors, a key concept is the instantaneous causality between these two factors in the sense of Pierce and Haugh (1977). When the two factors are future asset return and its ex-post volatility, as in this paper, this instantaneous causality relationship is usually dubbed “leverage effect” in the sense of Black (1976). The main general message of this paper
is to stress that the presence of leverage effect reverses both sides of the common belief that identification of risk aversion to volatility of volatility must be based on observations from derivative markets. We recall that the variance risk premium is measured as the difference between the risk-neutralized expected return variation and the realized return variation. Bollerslev, Tauchen, and Zhou (2009) refer to Britten-Jones and Neuberger (2000) to measure the market’s risk neutral expectation of the total returns variation between time \( t \) and \( t + 1 \) conditional on time \( t \) information from a portfolio of European calls, which relies on an ever-increasing number of calls with strikes spanning from zero to infinity. They note that this model-free measure provides a natural empirical analog to the risk-neutral expectation of future return variations. As such, the variance risk premium is naturally measured by the difference between this risk-neutral expectation and the ex-post realized return variation.

The main theoretical message of this paper regarding the impact of leverage effect on the accuracy of the aforementioned common belief is twofold. On the negative side, we argue that it is not true that the variance risk premium unambiguously identifies the risk aversion to volatility of volatility. Due to leverage effect, the volatility factor on the time interval \([t, t+1]\) is correlated with the return itself, conditional on time \( t \) information. Therefore, the variance risk premium combines the effect of two risk aversion parameters: our parameter \( \zeta_1 \) for aversion to the risk of volatility variations and parameter \( \zeta_2 \) for aversion to the volatility level in return. On the positive side, we show that, precisely because the volatility factor on the time interval \([t, t+1]\) is correlated, conditional on time \( t \) information, with the return itself, observations of return alone (without any observation from derivative markets) could be sufficient to identify both risk aversion parameters. The asset pricing equation in the presence of leverage effect depends on both risk aversion parameters, which may allow econometric identification of these parameters from the joint distribution of return and its volatility.

The econometric focus is to devise a valid inferential methodology to confirm the feasibility of the positive side of the theoretical message above. This is challenging for the following reason: On the one hand, although we expect a negative leverage effect, this parameter is difficult to quantify empirically and its estimate usually is small (Aït-Sahalia, Fan, and Li, 2013). On the other hand, we know that in the absence of leverage effect, risk aversion for volatility of volatility is not identified from time series of the asset only. Therefore, when the true leverage effect is small, the return data only provide a limited amount of information about the aversion to volatility risk, compared to the finite-sample noise in the data. This low signal-to-noise ratio, as modelled by (nearly) weak identification, may invalidate standard inference based on the Generalized Method of Moments (GMM)

We provide an identification-robust confidence set for the structural parameters that measure the return risk price, the volatility risk price, and the leverage effect. The robust confidence set provides correct asymptotic coverage uniformly over a large set of models allowing for any magnitude of the leverage effect. This uniform validity is crucial for the confidence set to have good finite-sample coverage (Mikusheva, 2007; Andrews and Guggenberger, 2010). In contrast, standard confidence sets based on the GMM estimator and its asymptotic normality do not have uniform validity in the presence of a small leverage effect. This issue affects all the structural parameters because they are estimated simultaneously.

We achieve robust inference in two steps. First, we establish a minimum distance criterion using link functions between the structural parameters and a set of reduced-form parameters that determine the joint distribution of the return and volatility. The structural model implies that the link functions are zero when evaluated at the true values of the structural parameters and the reduced-form parameters. Identification and estimation of these reduced form parameters are standard and are not affected by the presence of a small leverage effect. However, the link functions are almost flat in one of the structural parameters when the leverage effect is small, resulting in weak identification. Second, given this minimum distance criterion, we invert the conditional quasi-likelihood ratio (QLR) test of Andrews and Mikusheva (2016) to construct a robust confidence set. The key feature of this test is that it treats the flat link functions as an infinite-dimensional nuisance parameter. The critical value is constructed by conditioning on a sufficient statistic for this nuisance parameter, and it is known to yield a valid test regardless of the nuisance parameter’s value. Andrews and Mikusheva (2016) develop this test in a GMM framework. We show it works in a minimum distance context such as the one considered here and provide conditions for its asymptotic validity. For practitioners, we provide a detailed algorithm for the construction of this simulation-based robust confidence set. It is worth noting that this strategy of identification robust inference in a minimum-distance context may be applied in other fields, as in Magnusson and Mavroeidis (2010) and Magnusson (2010).

Our empirical results relate to the extant literature on the econometric analysis of the effect of uncertainty in volatility on risk premia. While our theoretical result about identification is model-free (up to the maintained assumption of an exponential affine pricing kernel), link functions for the minimum distance estimation must be built from a parametric model of the joint distribution of return and volatility. Following Renault (1997) and Garcia and Renault (1998), we preclude any Granger causality relationship from the asset return
to its volatility, to maintain a natural homogeneity relationship for option pricing formulas. Our bivariate discrete time model is inspired by the continuous model of Heston (1993) and belongs to the general class of Compound AutoRegressive Models (CAR) of Darolles, Gouriéroux, and Jasiak (2006). Following Gouriéroux and Jasiak (2006), we consider the simplest version (the so-called ARG(1) model), where the transition dynamics of the volatility process are driven by a Gamma distribution. Given the past and current volatility, the log-excess return follows a Gaussian distribution. Such a conditionally Gaussian model of asset returns paves the way for skewness and fat tails not only in the unconditional distribution of return but also in the conditional distribution given past observations that define the investors’ information set. Such a stochastic volatility model with conditional skewness has also been studied by Feunou and Tédongap (2012) (see also Feunou, Fontaine, Taamouti, and Tédongap, 2013), but with independent models for historical and risk neutral distributions, without connecting them by a convenient Stochastic Discount Factor (SDF).

Our empirical model is an extension of that in Corsi, Fusari, and La Vecchia (2013). The proposed extension is motivated by the need to properly accommodate leverage effect as well as the need to illustrate our identification statement. First, we parameterize the impact of leverage effect in order to capture what is suggested by usual continuous time modelling in the option pricing literature with stochastic volatility. On the one hand, as in continuous time models, the knowledge of the contemporaneous volatility innovation would reduce the conditional variance of the return innovation because the two innovations are correlated, i.e., the instantaneous causality effect. On the other hand, this knowledge would also introduce an additional term in the return drift because the knowledge of future volatility has an impact on expected return. As stressed by Bollerslev, Litvinova, and Tauchen (2006), this effect of leverage on expected return is hard to disentangle in discrete time from volatility feedback due to risk compensation. Bollerslev, Litvinova, and Tauchen (2006) stress that only continuous time observations would allow us to clearly disentangle the two effects by detecting the direction of causality. This identification issue is carefully taken into account in the definition of our three structural parameters: the two risk aversion parameters and the leverage effect parameter that is identified in the risk neutral world.

Our empirical model also extends that in Corsi, Fusari, and La Vecchia (2013) in another respect to illustrate our identification result. By introducing additional state variables in the conditional mean and variance of the Gaussian log-return, we manage to obtain identification of the two risk aversion parameters from observation of the equity return only (jointly with its realized variance), in view of the no-arbitrage condition equilibrium. In contrast, Corsi, Fusari, and La Vecchia (2013) acknowledge that in this model, “the SDF is compatible with the no arbitrage conditions, provided that suitable parameter restrictions are fulfilled”. For
them, the parameter of risk aversion for volatility of volatility ‘must be calibrated’.

Our model can be seen as a particular case of the model put forward by Han, Khrapov, and Renault (2020), where the two risk aversion parameters are estimated by resorting to additional option price data. In this paper, we simplify the model of the realized variance process by assuming that it is a Markov process. It would be more realistic to see it as an ARMA(1,1) process, as in Han, Khrapov, and Renault (2020). Considering this more general model would not invalidate the identification result and the minimum distance approach that we propose, although the link functions would be much more complicated. For the sake of clarity, we choose to keep the simplest AR(1) model for realized variance, which provides closed form formulas for the different steps of our identification robust inference procedure.

The rest of the paper is organized as follows. Section 2 discusses the theoretical identification in a model-free framework where the conditional probability distributions of interest, both in the historical and the risk-neutral world, are characterized by their conditional Laplace transforms. These conditional Laplace transforms describe the joint distribution of some asset return and volatility factor that are connected by an exponential affine SDF. We see that, as in continuous time with the Girsanov theorem, the change of measure provided by the exponentially affine SDF is leverage effect preserving: there is instantaneous causality between the two variables in the risk-neutral world if and only if this causality exists in the historical world. We prove our identification claims in this general setting.

In section 3, we specify a general CAR model for the joint historical distribution of return and volatility, with conditional normality for log-return. The exponential affine structure is well-suited to characterize identification of the risk aversion parameters through state variables. Such an identification result is not possible in the case considered by Corsi, Fusari, and La Vecchia (2013). An ARG(1) specification for the volatility factor provides a fully parametric model for the purpose of empirical analysis. Section 4 sets the focus on risk neutral distribution that is characterized by suitably risk-neutralized CAR and ARG(1) distributions. Moreover, the risk-neutral framework allows us to characterize without ambiguity the amplitude of leverage effect.

Section 5 provides identification of the three structural parameters of interest through convenient link functions connecting them to reduced form parameters. Section 6 states the asymptotic distribution of the GMM estimators of reduced form parameters. A detailed algorithm to construct confidence sets for structural parameters that are identification robust is given in Section 7. Section 8 shows that the method works well in simulation while Section 9 applies the methods to data on the S&P 500 and provides estimates of the risk prices. Section 10 concludes. Proofs are given in the appendix.
2 General Framework

2.1 Exponentially Affine SDF and Conditional Laplace Transforms

The focus of interest of this paper is the econometric identification of the compensation required by investors to bear the volatility risk that we dub the "price of volatility risk". This price is a structural parameter that shows up in a Stochastic Discount Factor (SDF). More precisely, we consider the bivariate process of an asset return $r_t$ and its volatility factor $\sigma_t^2$ and define its dynamics through the conditional Laplace transform. The return process $\{r_t\}$ stands for a log-return in excess of a risk free rate $r_{f,t}$ while the volatility factor process $\{\sigma_t^2\}$ appears in the definition of the conditional variance of the asset return in a way specified below. We assume throughout that the representative investor's information set $I(t)$ at time $t$ contains at least the past and present observations of the joint process $W_t = (r_t, \sigma_t^2)$ such that $\{W_\tau, \tau \leq t\} \subset I(t)$. We then have the historical joint dynamics as defined by the conditional Laplace transform

$$L_t(u, v) = E[\exp\left(-u\sigma_{t+1}^2 - vr_{t+1}\right) | I(t)] \forall u \in \mathbb{R}, v \in \mathbb{R},$$

and the risk neutral joint dynamics as defined by the conditional Laplace transform

$$L^*_t(u, v) = E^*[\exp\left(-u\sigma_{t+1}^2 - vr_{t+1}\right) | I(t)] \forall u \in \mathbb{R}, v \in \mathbb{R}. \quad (2.2)$$

The bridge between the two conditional distributions is given by the SDF $M_{t+1}(\zeta)$ through the identity

$$\exp(-r_{f,t})E^*[\exp\left(-u\sigma_{t+1}^2 - vr_{t+1}\right) | I(t)] = E[M_{t+1}(\zeta) \exp\left(-u\sigma_{t+1}^2 - vr_{t+1}\right) | I(t)], \quad (2.3)$$

where $r_{f,t}$ stands for the continuously compounded risk free interest rate between $t$ and $t+1$, and by definition

$$\exp(-r_{f,t}) = E[M_{t+1}(\zeta) | I(t)]. \quad (2.4)$$

We maintain the assumption of an exponentially affine SDF

$$M_{t+1}(\zeta) = \exp(-r_{f,t})M_{0,t}(\zeta) \exp\left[-\zeta_1\sigma_{t+1}^2 - \zeta_2 r_{t+1}\right] \quad (2.5)$$

where we have

$$[M_{0,t}(\zeta)]^{-1} = E[\exp\left(-\zeta_1\sigma_{t+1}^2 - \zeta_2 r_{t+1}\right) | I(t)] = L_t(\zeta_1, \zeta_2) \quad (2.6)$$
by (2.4) and (2.5). The vector \( \zeta = (\zeta_1, \zeta_2)' \) encapsulates the two prices of risk: price \( \zeta_1 \) of risk on \( \sigma^2_{t+1} \) (volatility risk) and price \( \zeta_2 \) of risk on \( r_{t+1} \) (return risk).

The exponentially affine SDF is consistent with most popular asset pricing models, such as the intertemporal expected utility model and the recursive model with Epstein-Zin-Weil preferences. Furthermore, it is especially convenient for pricing with conditional Laplace transforms because the time \( t \) price of the payoff \( \exp(-u\sigma^2_{t+1} - vr_{t+1}) \) at time \( t+1 \) is simply

\[
E[M_{t+1}(\zeta) \exp (-u\sigma^2_{t+1} - vr_{t+1}) | I(t)] = \exp(-r_{f,t}) \frac{\mathcal{L}_t(u + \zeta_1, v + \zeta_2)}{\mathcal{L}_t(\zeta_1, \zeta_2)}. \tag{2.7}
\]

by (2.5) and (2.6). In other words, the bridge between the historical and risk-neutral conditional distributions, given in (2.3), can be written in terms of the conditional Laplace transforms

\[
\mathcal{L}^*_t(u, v) = \frac{\mathcal{L}_t(u + \zeta_1, v + \zeta_2)}{\mathcal{L}_t(\zeta_1, \zeta_2)}. \tag{2.8}
\]

Throughout we maintain the natural assumption that the risk neutral distribution, as characterized by the conditional Laplace transform \( \mathcal{L}^*_t(u, v) \), does not depend on the preference parameters \( \zeta_1 \) and \( \zeta_2 \).

### 2.2 Leverage Effect Preserving Change of Measure

In continuous time models, the leverage effect is defined as the instantaneous correlation between the asset return and its volatility process. By virtue of the Girsanov theorem, the leverage effect is unchanged when moving from the historical measure to the risk neutral measure. Interestingly, the fact that the pricing kernel is exponentially affine is sufficient (and necessary to a large extent) for a similar property in discrete time, regarding the presence of leverage effect.

**Proposition 1.** Given \( I(t) \), \( r_{t+1} \) and \( \sigma^2_{t+1} \) are conditionally independent for the historical distribution if and only if they are conditionally independent for the risk neutral distribution.

When the conditional independence condition in Proposition 1 is violated, we say that there exists leverage effect. The proof of proposition 1 is obvious since independence is characterized by the factorization of the Laplace transform function.

Combining independence and (2.8), we have

\[
\mathcal{L}_{\sigma,t}(u + \zeta_1)\mathcal{L}_{r,t}(v + \zeta_2) = \mathcal{L}^*_{\sigma,t}(u)\mathcal{L}^*_{r,t}(v)\mathcal{L}_{\sigma,t}(\zeta_1)\mathcal{L}_{r,t}(\zeta_2), \quad \forall \ u, v \in \mathbb{R}, \tag{2.9}
\]
where $L_{\sigma,t}(\cdot)$ (resp., $L_{r,t}(\cdot)$) stands for the historical conditional Laplace transforms of $\sigma_{t+1}^2$ (resp., $r_{t+1}$) given $I(t)$, and $L_{\sigma,t}^*(u)$ and $L_{r,t}^*(v)$ are their risk-neutral counterparts. By considering this formula for $u = 0$ or $v = 0$, we obtain

$$L_{\sigma,t}(u + \zeta_1) = L_{\sigma,t}^*(u)L_{\sigma,t}(\zeta_1), \ \forall \ u \in \mathbb{R}, \ (2.10)$$

$$L_{r,t}(v + \zeta_2) = L_{r,t}^*(v)L_{r,t}(\zeta_2), \ \forall \ v \in \mathbb{R},$$

under independence. Since the risk neutral distribution does not depend on the preference parameters $\zeta_1$ and $\zeta_2$, we have the following conclusion in Proposition 2 below.

**Proposition 2.** When there is no leverage effect,

(i) the historical conditional distribution of $\sigma_{t+1}^2$ given $I(t)$ depends on the pricing kernel only through the parameter $\zeta_1$;

(ii) the historical conditional distribution of $r_{t+1}$ given $I(t)$ depends on the pricing kernel only through the parameter $\zeta_2$.

Proposition 2 explains why it is common knowledge that one can identify $\zeta_1$, price of the volatility risk, by comparing the risk neutral and the historical distribution of the volatility factor. A common practice is to draw this comparison through the variance premium.

Proposition 2 also explains why it is common knowledge that one cannot identify $\zeta_1$ when only observing the return on the underlying asset. In the absence of leverage effect, the historical distribution of the return only depends on $\zeta_2$, price of the return risk. However, we emphasize hereafter that this common knowledge becomes misleading in the presence of leverage effect.

### 2.3 Identification of Risk Prices

We first show that, in contrast to a common belief, observations of the risk neutral expectation of future variance, through the VIX, may not provide identification of $\zeta_1$, price of the volatility risk. Traditional approaches to identification of risk prices go through the risk neutral conditional expectations of the asset return and its volatility. From (2.8), we can characterize the two risk neutral distributions by

$$L_{\sigma,t}^*(u) = \frac{L_t(u + \zeta_1, \zeta_2)}{L_t(\zeta_1, \zeta_2)}, \quad L_{r,t}^*(v) = \frac{L_t(\zeta_1, v + \zeta_2)}{L_t(\zeta_1, \zeta_2)}. \quad (2.11)$$

The risk neutral expectation of the volatility factor is given by

$$E^*[\sigma_{t+1}^2 | I(t)] = \frac{\partial L_{\sigma,t}^*(0)}{\partial u} = \frac{1}{L_t(\zeta_1, \zeta_2)} \frac{\partial L_t(\zeta_1, \zeta_2)}{\partial u}. \quad (2.12)$$
In the presence of leverage effect, i.e., \( \mathcal{L}_t(\zeta_1, \zeta_2) \neq \mathcal{L}_{\sigma,t}(\zeta_1)\mathcal{L}_{r,t}(\zeta_2) \), \( \mathcal{L}_{r,t}(\zeta_2) \) cannot be factorized out for simplification between the numerator and the denominator of (2.12). Thus, the risk neutral expectation of future variance should depend on both \( \zeta_1 \) and \( \zeta_2 \) for a given historical Laplace transform. It is not clear how these two parameters could be disentangled.

Next, we show that, again in contrast to a common belief, observations of the price of the asset may actually provide identification of \( \zeta_1 \), price of the volatility risk. By definition, the price of the underlying asset satisfies that the risk neutral expectation of the excess return \( \exp(r_{t+1}) \) equals to unity, i.e.,

\[
\mathcal{L}^*_r(-1) = E^*[\exp(r_{t+1})|I(t)] = 1. \tag{2.13}
\]

Combining this equality with the Laplace transform in (2.8) with \( u = 0 \) and \( v = -1 \), the pricing equation of the underlying asset satisfies

\[
\mathcal{L}_t(\zeta_1, \zeta_2 - 1) - \mathcal{L}_t(\zeta_1, \zeta_2) = 0. \tag{2.14}
\]

In the presence of leverage effect, i.e., \( \mathcal{L}_t(\zeta_1, \zeta_2) \neq \mathcal{L}_{\sigma,t}(\zeta_1)\mathcal{L}_{r,t}(\zeta_2) \), \( \mathcal{L}_{\sigma,t}(\zeta_1) \) cannot be factorized out for simplification of (2.14). Thus, the difference in (2.14) should depend on both \( \zeta_1 \) and \( \zeta_2 \) for a given historical Laplace transform.

We set the focus of this paper on a model where the bivariate process \( W_t = (r_t, \sigma^2_t) \) is Markov of order one. In this case, identification of the coefficients of the two state variables in equation (2.14) in general provide two independent equations that allow for identification of both risk price parameters \( \zeta_1 \) and \( \zeta_2 \).

To sum up, the presence of leverage effect reverse both sides of common beliefs. Not only do observations of the variance premium not provide identification of the price of volatility risk \( \zeta_1 \), but also, on the contrary, observations of the return of the asset of interest do provide identification of both risk prices \( \zeta_1 \) and \( \zeta_2 \).

### 3 Distributional Assumptions for Identification of Risk Prices

#### 3.1 Conditional Normality Given Current Volatility Factor

Ané and Geman (2000) (see also Clark (1973)) for a seminal contribution) argue that for log-returns, conditional normality can be recovered when conditioning on a measure of the market activity. We follow Corsi, Fusari, and La Vecchia (2013) and Han, Khrapov, and...
Renault (2020) to adopt such a measure of market activity, the integrated variance $IV_{t+1}$ over the time interval $[t, t+1]$. In other words, we consider $\sigma_{t+1}^2 = IV_{t+1}$ and assume that for the historical distribution, given the extended information set

$$I^\sigma(t) = I(t) \lor \{\sigma_{t+1}^2\}, \quad (3.1)$$

the log-excess return $r_{t+1}$ follows a Gaussian distribution characterized by the conditional Laplace transform

$$L_{r,t}^{(\sigma)}(v) = \exp \left\{ -v \mu_t(\sigma_{t+1}^2) + \frac{v^2}{2} \omega_t(\sigma_{t+1}^2) \right\}, \quad \text{where} \quad \mu_t(\sigma_{t+1}^2) = E[r_{t+1} \mid I^\sigma(t)], \quad \omega_t(\sigma_{t+1}^2) = Var[r_{t+1} \mid I^\sigma(t)]. \quad (3.2)$$

Therefore, by the law of iterated expectations, the joint conditional Laplace transform is given by

$$L_t(u, v) = \exp \left\{ -u \sigma_{t+1}^2 - v \mu_t(\sigma_{t+1}^2) + \frac{v^2}{2} \omega_t(\sigma_{t+1}^2) \right\} \left| I^\sigma(t) \right|. \quad (3.3)$$

It is worth pointing out that such a model paves the way for skewness and fat tails not only in the unconditional distribution of returns but also in the conditional distribution given the past observations that define the investor’s information set $I(t)$. The mixing variable $\sigma_{t+1}^2$, which we dub the volatility factor, maintains some randomness in the conditional mean and variance of the Gaussian distribution.

### 3.1.1 Conditional Variance

Since our focus of interest is the leverage effect, we extend the model of Consi, Fusari, and La Vecchia (2013) in two respects. First, Consi, Fusari, and La Vecchia (2013) assume that

$$\omega_t(\sigma_{t+1}^2) = \sigma_{t+1}^2. \quad (3.4)$$

Instead, we assume

$$\omega_t(\sigma_{t+1}^2) = \xi \sigma_{t+1}^2 + X_t, \quad (3.5)$$

where $\xi \in [0, 1]$ is some leverage parameter and $X_t$ stands for some state variable belonging to the past information set $I(t)$. The rationale is as follows. Up to additional state variables included in $X_t$, the conditional variance is equal to $\sigma_{t+1}^2$ only in the absence of leverage effect. A strong leverage effect reduces the rescaling coefficient $\xi$ because it entails a strong correlation (in absolute value) between the return and the contemporaneous volatility that yields a reduction of the forecast error of the return.
The role of the state variable $X_t$ is to enable us to enforce the identification condition

$$Var[r_{t+1} | I(t)] = E[\sigma^2_{t+1} | I(t)].$$

(3.6)

Although this identification condition in (3.6), with $\sigma^2_{t+1} = IV_{t+1}$ being the integrated variance, is conformable to common continuous time stochastic volatility models, it is obviously at odds with the condition $Var[r_{t+1} | I^a(t)] = \sigma^2_{t+1}$ put forward by Corsi, Fusari, and La Vecchia (2013) under (3.4). Their framework implies instead

$$E \{ Var[r_{t+1} | I^a(t)] | I(t) \} = E[\sigma^2_{t+1} | I(t)].$$

(3.7)

By replacing (3.4) with the new assumption in (3.5) with a state variable $X_t$, we obtain

$$Var[r_{t+1} | I(t)] = E \{ Var[r_{t+1} | I^a(t)] | I(t) \} + Var \{ E[r_{t+1} | I^a(t)] | I(t) \}$$

$$= E[\omega_t (\sigma^2_{t+1}) | I(t)] + Var[\mu_t (\sigma^2_{t+1}) | I(t)]$$

$$= \xi E[\sigma^2_{t+1} | I(t)] + X_t + Var[\mu_t (\sigma^2_{t+1}) | I(t)],$$

meaning that the identification condition in (3.6) is tantamount to

$$X_t = (1 - \xi) E[\sigma^2_{t+1} | I(t)] - Var[\mu_t (\sigma^2_{t+1}) | I(t)].$$

(3.9)

3.1.2 Conditional Mean

While Corsi, Fusari, and La Vecchia (2013) assume that

$$\mu_t (\sigma^2_{t+1}) = \left( \zeta_2 - \frac{1}{2} \right) \sigma^2_{t+1},$$

(3.10)

we write more generally

$$\mu_t (\sigma^2_{t+1}) = \psi \sigma^2_{t+1} + Z_t, \quad \psi = \left( \zeta_2 - \frac{1}{2} \right) \xi + LEV$$

(3.11)

for some state variable $Z_t$. The rationale of this specification is as follows: (i) The risk premium included in the asset return is proportional to the conditional variance $\omega_t (\sigma^2_{t+1})$. (ii) The proportionality coefficient is the price $\zeta_2$ for the risk on return, up to the Jensen effect – the convexity adjustment between $r_{t+1}$ and its exponential given by the coefficient
-1/2. With our variance specification \( \omega_t \sigma^2_{t+1} = \xi \sigma^2_{t+1} + X_t \), it would lead to

\[
\mu_t \sigma^2_{t+1} = \left( \zeta_2 - \frac{1}{2} \right) \xi \sigma^2_{t+1} + \left( \zeta_2 - \frac{1}{2} \right) X_t.
\] (3.12)

Note that this specification is consistent with Corsi, Fusari, and La Vecchia (2013), who set the focus in the particular case \( \xi = 1, X_t = 0 \). Besides introducing an unconstrained state variable \( Z_t \) (not necessarily proportional to \( X_t \)), our main addition in (3.11) is the term \( \text{LEV} \).

As pointed out by Bollerslev, Litvinova, and Tauchen (2006), in discrete time, it is hardly possible to disentangle the volatility feedback effect (a consequence of risk compensation through parameter \( \zeta_2 \)) and the impact of leverage effect on expected return that we propose to characterize by an additional parameter denoted by \( \text{LEV} \). Note that \( \text{LEV} \) must be a characteristic of the risk-neutral distribution that captures leverage effect, independent of the risk aversion parameters (see section 4 for details).

### 3.2 Identification Through State Variables

Introducing the state variables \( X_t \) and \( Z_t \), as given in (3.5) and (3.11), is necessary for the identification of the two risk prices \( \zeta_1 \) and \( \zeta_2 \). In order to see this, consider the alternative specification without these state variables. In this case, \( \omega_t \sigma^2_{t+1} = \xi \sigma^2_{t+1} \) and \( \mu_t \sigma^2_{t+1} = \psi \sigma^2_{t+1} \). Plugging them into the Laplace transform \( L_t(u, v) \) in (3.3) gives

\[
L_t(u, v) = E \left[ \exp \left\{ -\sigma^2_{t+1} \left( u + v \psi - \frac{v^2}{2} \xi \right) \right\} \right] I(t).
\] (3.13)

Together with the identification condition in (2.14), i.e., \( L_t(\zeta_1, \zeta_2) = L_t(\zeta_1, \zeta_2 - 1) \), the absence of state variables \( X_t \) and \( Z_t \) leads to the pricing equation

\[
E \left[ \exp \left\{ -\sigma^2_{t+1} \left( \zeta_1 + \zeta_2 \psi - \frac{\zeta_2^2}{2} \xi \right) \right\} \right] I(t),
\]

\[
= E \left[ \exp \left\{ -\sigma^2_{t+1} \left( \zeta_1 + (\zeta_2 - 1) \psi - \frac{(\zeta_2 - 1)^2}{2} \xi \right) \right\} \right] I(t).
\] (3.14)

This pricing equation is fulfilled as long as \( \psi = \zeta_2 - 1/2 \) and \( \xi = 1 \), leading Corsi et al. (2013) to conclude that “the SDF is compatible with the no-arbitrage conditions, provided that suitable parameter restrictions are satisfied”. The parameter \( \zeta_1 \) remains a free parameter that “must be calibrated”. While this statement seems to confirm the common view that price of the volatility risk cannot be identified from underlying asset returns, it is striking that it can be fundamentally changed by the introduction of state variables.
By introducing the state variables, i.e., $\omega_t \left( \sigma^2_{t+1} \right) = \xi \sigma^2_{t+1} + X_t$ and $\mu_t \left( \sigma^2_{t+1} \right) = \psi \sigma^2_{t+1} + Z_t$, the Laplace transform becomes

$$L_t(u, v) = E \left[ \exp \left\{ -\sigma^2_{t+1} \left( u + v \psi - \frac{v^2}{2} \xi \right) \right\} \left| I(t) \right] \exp \left\{ -v Z_t + \frac{v^2}{2} X_t \right\}. \quad (3.15)$$

Together with the identification condition $L_t(\zeta_1, \zeta_2) = L_t(\zeta_1, \zeta_2 - 1)$, this specification with state variables yields the pricing equation

$$E \left[ \exp \left\{ -\sigma^2_{t+1} \left[ \zeta_1 + \zeta_2 \psi - \frac{\zeta_2^2}{2} \xi \right] \right\} \left| I(t) \right] \exp \left\{ -\zeta_2 Z_t + \frac{\zeta_2^2}{2} X_t \right\} \right. \quad (3.16)$$

That is

$$E \left[ \exp \left\{ -\sigma^2_{t+1} \left[ \zeta_1 + \zeta_2 \psi - \frac{\zeta_2^2}{2} \xi \right] \right\} \left| I(t) \right] \exp \left\{ -(\zeta_2 - 1) Z_t + \frac{(\zeta_2 - 1)^2}{2} X_t \right\}. \quad (3.17)$$

Obviously, if the two state variables $X_t$ and $Z_t$ are not proportional, the previous solution $\psi = \zeta_2 - 1/2$ (when $\xi = 1$) does not work anymore. In this case, the solution to (3.17) could allow us to identify both $\zeta_1$ and $\zeta_2$.

### 3.3 What State Variables X and Z?

As explained above (see section 3.1.1.), our choice of the state variable $X_t$ is implied by the main identification condition, i.e.,

$$Var(r_{t+1} | I(t)) = E[\sigma^2_{t+1} | I(t)] \iff X_t = (1 - \xi) E[\sigma^2_{t+1} | I(t)] - \psi^2 Var[\sigma^2_{t+1} | I(t)] \quad (3.18)$$

following (3.9) and (3.11). Hence, the choice of the state variable $X_t$ depends on a dynamic model for the first two conditional moments of the volatility factor $\sigma^2_{t+1}$. We choose an exponential affine model defined by the conditional Laplace transform

$$E[\exp \left\{ -u \sigma^2_{t+1} \right\} | I(t)] = \exp \left\{ -a(u) \sigma^2_t - b(u) \right\}, \quad (3.19)$$
which leads to
\[
E[\sigma_{t+1}^2 | I(t)] = a'(0)\sigma_t^2 + b'(0), \quad (3.20)
\]
\[
Var[\sigma_{t+1}^2 | I(t)] = -a''(0)\sigma_t^2 - b''(0).
\]

**Remark 3.1.** Corsi, Fusari, and La Vecchia (2013) choose a more general specification with \( J = 4 \) state variables such that
\[
E[\sigma_{t+1}^2 | I(t)] = \sum_{j=1}^{J} a^{(1)}_j IV_{j,t} + b^{(1)},
\]
\[
Var[\sigma_{t+1}^2 | I(t)] = -\sum_{j=1}^{J} a^{(2)}_j IV_{j,t} - b^{(2)}. \quad (3.21)
\]

For Corsi, Fusari, and La Vecchia (2013), \( IV_{1,t} = \sigma_t^2 = IV_t \)— integrated variance in the previous period, \( IV_{2,t} = IV_t^{(\omega)} \) — medium term volatility factor, \( IV_{3,t} = IV_t^{(m)} \) — long term volatility factor, and \( IV_{4,t} = 1_{[r_t < 0]} IV_t \) — leverage effect factor. Introducing the asset return \( r_t \) in the conditional moments of the volatility factor (through the indicator function \( 1_{[r_t < 0]} \) in the factor \( IV_{4,t} \)) is at odds with the noncausality argument put forward in Renault (1997) and Garcia and Renault (1998). To maintain the natural homogeneity property for option pricing formulas, we preclude any Granger causality relationship from the return to the volatility.

**Remark 3.2.** Even though the general setting of remark 3.1. with multiple state variables would not cause any specific difficulty for our analysis, we set the focus on the simplest framework possible to illustrate our identification issue, that is
\[
J = 1, IV_{1,t} = \sigma_t^2 = IV_t, \quad (3.22)
\]
leading to the model in (3.20). A weakness of this model is to restrict \( \sigma_t^2 = IV_t \) as an heteroskedastic AR(1) process. Both theoretical arguments (temporal aggregation of spot volatility) and empirical goodness of fit would rather suggest an ARMA(1,1) dynamics. While the multiplicity of state variables \( (J > 1) \), as in Corsi, Fusari, and La Vecchia (2013) is a way to preclude univariate AR(1) dynamics of the volatility factor, a more parsimonious alternative, as proposed by Han, Khrapov, and Renault (2020), is to introduce an AR(1) volatility factor \( \tilde{\sigma}_t^2 \) such that
\[
\sigma_{t+1}^2 = IV_{t+1} = \tilde{\sigma}_{t+1}^2 + f\tilde{\sigma}_t^2 + g, \quad (3.23)
\]
where $f$ and $g$ are two additional free parameters. Han, Khropov, and Renault (2020) document that this ARMA(1, 1) model for the integrated variance has a better empirical performance than the AR(1) model. Again, even though this extension is not difficult to handle, we choose to set the focus on the simplest AR(1) model.

Our choice (3.19) of an exponential affine model for the volatility factor $\sigma_t^2$, jointly with the assumption of conditional normality of $r_{t+1}$ given the current volatility factor $\sigma_{t+1}^2$, defines a bivariate exponential affine model for $(\sigma_{t+1}^2, r_{t+1})$. This model is dubbed the Compound AutoRegressive Model (CAR) by Darolles, Gouriéroux, and Jasiak (2006):

1. The conditional distribution of $\sigma_{t+1}^2$ given $I(t)$ is characterized by the conditional Laplace transform (3.19);
2. The conditional distribution of $r_{t+1}$ given $I(t)$ and $\sigma_{t+1}^2$ is characterized by the conditional Laplace transform

$$E[\exp \{-vr_{t+1}\} | I^\sigma(t)] = \exp \{-\alpha(v)\sigma_{t+1}^2 - \beta(v)\sigma_t^2 - \gamma(v)\} \quad (3.24)$$

for some functions $\alpha(v), \beta(v), \gamma(v)$, where the maintained assumption of conditional Gaussianity of $r_{t+1}$ given $I^\sigma(t)$ is tantamount to the fact that the three functions $\alpha(u), \beta(u)$ and $\gamma(u)$ are all quadratic.

The conditional moment formulas in (3.20) and the identification condition in (3.18) implies the following choice for the state variable $X_t$:

$$X_t = (1 - \xi) \{ a'(0)\sigma_t^2 + b'(0) \} + \psi^2 \{ a''(0)\sigma_t^2 + b''(0) \} \quad (3.25)$$

Furthermore, the specification in (3.24) implies

$$\mu_t (\sigma_{t+1}^2) = E[r_{t+1} | I^\sigma(t)] = \alpha'(0)\sigma_{t+1}^2 + \beta'(0)\sigma_t^2 + \gamma'(0), \quad (3.26)$$
$$\omega_t (\sigma_{t+1}^2) = Var[r_{t+1} | I^\sigma(t)] = -\alpha''(0)\sigma_{t+1}^2 - \beta''(0)\sigma_t^2 - \gamma''(0). \quad (3.27)$$

Because $\alpha(v), \beta(v)$ and $\gamma(v)$ are all quadratic, and the state variables $X_t$ and $Z_t$ have been introduced such that

$$\mu_t (\sigma_{t+1}^2) = \psi\sigma_{t+1}^2 + Z_t, \quad (3.27)$$
$$\omega_t (\sigma_{t+1}^2) = \xi\sigma_{t+1}^2 + X_t,$$
we combine (3.25) – (3.27) and obtain

\[ \alpha(v) = \psi v - \frac{\xi}{2} v^2, \]  

\[ \beta(v) = \beta v - \frac{1}{2} \left[ (1 - \xi) a'(0) + \psi^2 a''(0) \right] v^2, \]  

\[ \gamma(v) = \gamma v - \frac{1}{2} \left[ (1 - \xi) b'(0) + \psi^2 b''(0) \right] v^2, \]

and

\[ Z_t = \beta \sigma_t^2 + \gamma. \]

**Remark 3.3.** We end up with state variables \( X_t \) and \( Z_t \) that are both affine functions of the volatility factor \( \sigma_t^2 \). While \( X_t \) depends in particular on the volatility dynamics through the functions \( a(\cdot) \) and \( b(\cdot) \), \( Z_t \) depends on the coefficients \( \beta \) and \( \gamma \) of the conditional Laplace transform of the returns given \( I^\sigma(t) \). By construction, when the coefficients \( \beta \) and \( \gamma \) are free parameters, the variable \( \zeta_2 Z_t - (\zeta_2 - \frac{1}{2}) X_t \) is not identically zero in general, making sensible the identification strategy for prices of risk, as previously devised from (3.17).

**Remark 3.4.** The joint distribution of \( (r_{t+1}, \sigma_{t+1}^2) \) is then characterized by two vector of reduced form parameters: (i) a vector \( \omega_1 \) that characterizes the functions \( a(\cdot) \) and \( b(\cdot) \) for defining the dynamics of the volatility factor in (3.19), and (ii) a vector \( \omega_2 = (\gamma, \beta, \psi, \xi)' \) which jointly with \( \omega_1 \) defines the conditional distribution of the return \( r_{t+1} \) given past and current volatility factor through (3.24) and (3.28).

### 3.4 A Fully Parametric Model

All the discussion of section 4 below is valid without restrictions on the functions \( a(\cdot) \) and \( b(\cdot) \) that define the volatility dynamics. However, for estimation and inference, we specify a discrete time model inspired by the continuous time model of Heston (1993). Following Gouriéroux and Jasiak (2006), we consider the simplest version where the transition dynamics are driven by Gamma distributions as in the affine model of Heston (1993) and its precursor Feller (1951)’s square root process. More precisely, we use the ARG(1) model defined by Gouriéroux and Jasiak (2006) as follows: (i) The conditional distribution of \( \sigma_{t+1}^2 \) given some mixing variable \( U_t \) is Gamma distribution with the shape parameter \( (\delta + U_t) \) and a scale parameter \( c \). (ii) The conditional distribution of \( U_t \) given \( \sigma_t^2 \) is the Poisson distribution with the parameter \( \rho \sigma_t^2/c \). This parametric model is nested in the general
affine model defined in (3.19) with the specification

\[ a(u) = \frac{\rho u}{1 + cu}, \quad b(u) = \delta \log (1 + cu). \tag{3.30} \]

In this parametric model, the reduced form parameters \( \omega_1 = (\rho, c, \delta)' \) are identified by the first two conditional moments:

\[
egin{align*}
E[\sigma_{t+1}^2 | I(t)] &= \rho \sigma_t^2 + \delta c, \tag{3.31} \\
Var[\sigma_{t+1}^2 | I(t)] &= 2\rho c \sigma_t^2 + \delta c^2.
\end{align*}
\]

The interpretation of the three parameters in \( \omega_1 \) is as follows: (i) \( \rho \) is the autocorrelation parameter in the heteroskedastic AR(1) process \( \sigma_t^2 \); (ii) \( c \) is a scale parameter, and (iii) \( \delta \) is the location parameter. Together, they determine the unconditional mean \( E[\sigma_{t+1}^2] = \frac{\delta c}{1 - \rho} \).

4 Risk-Neutral Distributions

As discussed, the discrete time framework makes it hardly possible to separately identify the leverage effect and the volatility feedback due to risk compensation without additional assumptions. We address this issue by making sure that the leverage effect is a risk-neutral characteristic, as such it does not depend on preference parameters.

4.1 Risk-Neutral Distribution of Volatility Factor

The risk-neutral conditional Laplace transform of \( \sigma_{t+1}^2 \) given \( I(t) \) is

\[
\mathcal{L}_{\sigma,t}^*(u) = \frac{\mathcal{L}_t(u + \zeta_1, \zeta_2)}{\mathcal{L}_t(\zeta_1, \zeta_2)},
\tag{4.1}
\]

as given in (2.11). The numerator of (4.1) can be made explicit with the conditional Laplace transforms defined in (3.19) and (3.24):

\[
\begin{align*}
\mathcal{L}_t(u + \zeta_1, \zeta_2) &= E \left[ \exp \left\{ - (u + \zeta_1) \sigma_{t+1}^2 - \zeta_2 r_{t+1} \right\} \right| I(t)] \\
&= E \left[ \exp \left\{ - (u + \zeta_1) \sigma_{t+1}^2 \right\} E \left[ \exp \left\{ - \zeta_2 r_{t+1} \right\} \right| I(t)] \\
&= E \left[ \exp \left\{ - (u + \zeta_1) \sigma_{t+1}^2 \right\} \exp \left\{ - \alpha (\zeta_2) \sigma_{t+1}^2 - \beta (\zeta_2) \sigma_t^2 - \gamma (\zeta_2) \right\} \right| I(t)] \\
&= \exp \left\{ - \beta (\zeta_2) \sigma_t^2 - \gamma (\zeta_2) \right\} \exp \left\{ - a \left[ u + \zeta_1 + \alpha (\zeta_2) \right] \sigma_t^2 - b \left[ u + \zeta_1 + \alpha (\zeta_2) \right] \right\}.
\end{align*}
\]
We get the denominator of (4.1) by setting \( u = 0 \) and obtain
\[
L_{\sigma,t}^*(u) = \exp \left\{ -a [u + \zeta_1 + \alpha (\zeta_2)] \sigma_t^2 + a [\zeta_1 + \alpha (\zeta_2)] \sigma_t^2 \right\}
\exp \left\{ -b [u + \zeta_1 + \alpha (\zeta_2)] + b [\zeta_1 + \alpha (\zeta_2)] \right\}.
\]

Analogous to (3.19), the risk-neutral conditional Laplace transform of the volatility factor can be written as
\[
E^* \left[ \exp \left\{ -u \sigma_{t+1}^2 \right\} \mid I(t) \right] = \exp \left\{ -a^* (u) \sigma_t^2 - b^* (u) \right\},
\]
where
\[
a^* (u) = a [u + \zeta_1 + \alpha (\zeta_2)] - a [\zeta_1 + \alpha (\zeta_2)],
\]
\[
b^* (u) = b [u + \zeta_1 + \alpha (\zeta_2)] - b [\zeta_1 + \alpha (\zeta_2)].
\]

4.2 Risk-Neutral Distribution of Asset Return Conditional on Volatility Factor

To characterize the joint risk-neutral distribution of \((\sigma_{t+1}^2, r_{t+1})\) given \( I(t) \), we also need the conditional risk neutral distribution of \( r_{t+1} \) given \( I^\sigma(t) = I(t) \lor \{ \sigma_{t+1}^2 \} \). To do so we will resort to the following lemma.

**Lemma 1.** Let \( X \) and \( Y \) be two scalar random variables defined on the same probability space. If the joint distribution of \((X, Y)\) is characterized by the Laplace transform \( L(\cdot, \cdot) \), the Laplace transform \( L_{Y|X=x}(v) \) of the conditional distribution of \( Y \) given \( X = x \) is the only function \( g(v|x) \) such that, for all \((u, v)\),
\[
L(u, v) = E \left[ \exp \left\{ -uX \right\} g(v|X) \right].
\]

Application of the law of iterated expectations obviously shows that the condition in Lemma 1 is necessary. It is also sufficient by applying results in Bierens (1982).

We show in the appendix that, for all \((u, v)\), \( L_t^*(u, v) = E^* \left[ \exp \left\{ -u \sigma_{t+1}^2 g_t(v|\sigma_{t+1}^2) \right\} \mid I(t) \right] \), where \( g_t(v|\sigma_{t+1}^2) = \exp \left\{ -\alpha^* (v) \sigma_{t+1}^2 - \beta^* (v) \sigma_t^2 - \gamma^* (v) \right\} \) and
\[
\alpha^* (v) = \alpha (v + \zeta_2) - \alpha (\zeta_2),
\]
\[
\beta^* (v) = \beta (v + \zeta_2) - \beta (\zeta_2),
\]
\[
\gamma^* (v) = \gamma (v + \zeta_2) - \gamma (\zeta_2).
\]
From Lemma 1, we conclude that the risk-neutral conditional Laplace transform of $r_{t+1}$ given $I^*(t)$ is
\[ L^*_{r,t}(u,v) = \exp \{-\alpha^* (v) \sigma_{t+1}^2 - \beta^* (v) \sigma_t^2 - \gamma^*(v) \}. \] (4.7)

### 4.3 Leverage Effect Characterized by Risk-Neutral Distributions

As discussed in (3.11), we expect that the expected return given the current volatility factor $\sigma_{t+1}^2$ features a term $\psi \sigma_{t+1}^2$ (plus the impact of some past and present state variables), where the parameter $\psi$ encapsulates two effects: One is the risk compensation as in Corsi, Fusari, and La Vecchia (2013). The other, denoted by $LEV$, captures the leverage effect, independently of risk aversion parameters. In other words, $LEV$ is a characteristic of the risk-neutral distribution which, conformable to the popular idea of leverage effect (Black (1976)), measures the negative impact of $\sigma_{t+1}^2$ in the risk-neutral return forecast
\[ E^*[\exp \{r_{t+1}\} | I^*(t) ] = \exp \{-\alpha^* (-1) \sigma_{t+1}^2 - \beta^* (-1) \sigma_t^2 - \gamma^*(-1) \}. \] (4.8)

The amount of leverage effect is then defined as
\[ LEV = -\alpha^* (-1) \leq 0. \] (4.9)

We specify the leverage effect characterized by the risk-neutral distributions as a separate parameter, different from the preference parameters.

**Remark 4.1.** We have stressed in section 2.2 that the change of measure preserves the leverage effect. As shown in Proposition 1, $r_{t+1}$ and $\sigma_{t+1}^2$ are conditionally independent given $I(t)$ for the historical distribution if and only if they are conditionally independent for the risk-neutral distribution. In our context, it means that the function $\alpha(\cdot)$ is constant, equal to $\alpha(0) = 0$, if and only if the function $\alpha^*(\cdot)$ is constant, equal to $\alpha^*(0) = 0$. This equivalence is clear from the mapping formula $\alpha^*(v) = \alpha (v + \zeta_2) - \alpha (\zeta_2)$ following (4.6).

**Remark 4.2.** Although the presence of leverage effect is equivalent in the risk-neutral and the historical worlds, it is in the risk-neutral world that the amount of leverage effect $LEV = -\alpha^*(-1) \leq 0$ must be measured. Its historical counterpart $-\alpha (-1)$ underestimates (in absolute value) the leverage effect due to the volatility feedback created by risk aversion. More precisely, following (3.28), we have $\alpha(v) = \psi v - \frac{\xi}{2} v^2$. Therefore,
\[ -\alpha^*(-1) = LEV = \alpha (\zeta_2) - \alpha (\zeta_2 - 1) = \psi + \xi \left(\frac{1}{2} - \zeta_2\right), \] (4.10)
whereas
\[-\alpha(-1) = \psi + \frac{\xi}{2}, \tag{4.11}\]

This implies that
\[-\alpha(-1) = \text{LEV} + \xi\zeta_2. \tag{4.12}\]

The formula in (4.12) is important since it shows that the historical quantity $-\alpha(-1)$ in general underestimates the absolute value of true leverage $\text{LEV} = -\alpha^*( -1)$ because $\xi\zeta_2 \geq 0$.

We recall from (3.5) that $\xi \in [0, 1]$ characterizes the variance reduction (in asset return) provided by the simultaneous knowledge of $\sigma^2_{t+1}$: the stronger the leverage effect, the smaller the parameter $\xi$. In other words, it is precisely when the leverage effect is small that it is hardly identified by $-\alpha(-1)$ since it is hidden by the positive volatility feedback term $\xi\zeta_2$.

This will lead us to consider in section 5 that we have three structural parameters: $\zeta_1$, $\zeta_2$ and $\phi \in [0, 1]$, the latter being a normalization of $\text{LEV}$. Out of these three structural parameters, two of them are possibly weakly identified from asset return time series: the price $\zeta_1$ of volatility risk and the leverage effect parameter $\phi$. By contrast, the variance reduction coefficient $\xi$, caused by the leverage effect, is a reduced form parameter for which identification strength should not be an issue.

5 Identification of Structural Parameters

5.1 Identification of Prices of Risk

The asset pricing equation can be written as
\[E^*\{\exp(\gamma_{t+1})|I(t)\} = 1, \tag{5.1}\]

which, from (4.7), gives
\[
\exp\{\beta^* (-1) \sigma^2_t + \gamma^*(-1)\} = E^*\{\exp(-\alpha^* (-1) \sigma^2_{t+1})|I(t)\}, \tag{5.2}
\]

which is equivalent to the conjunction of two equations
\[a^* [\alpha^* (-1)] = -\beta^* (-1), \tag{5.3}\]
\[b^* [\alpha^* (-1)] = -\gamma^* (-1). \]
In terms of the characteristics of historical distributions, they are equivalent to

\[
\begin{align*}
    a [\zeta_1 + \alpha (\zeta_2 - 1)] - a [\zeta_1 + \alpha (\zeta_2)] &= \beta (\zeta_2) - \beta (\zeta_2 - 1), \\
    b [\zeta_1 + \alpha (\zeta_2 - 1)] - b [\zeta_1 + \alpha (\zeta_2)] &= \gamma (\zeta_2) - \gamma (\zeta_2 - 1)
\end{align*}
\]

following (4.5) and (4.6). Therefore, we end up with two equations with two unknown \(\zeta_1\) and \(\zeta_2\) that should in general allow us to identify both prices of risk if and only if

\[
LEV = \alpha (\zeta_2) - \alpha (\zeta_2 - 1) \neq 0. \tag{5.5}
\]

**Remark 5.1.** Corsi, Fusari, and La Vecchia (2013) assume that the functions \(\beta(\cdot)\) and \(\gamma(\cdot)\) are identically zero. Since the functions \(a(\cdot)\) and \(b(\cdot)\) must be strictly increasing, the only solution to (5.4), with a zero right hand side term, is such that \(\alpha (\zeta_2) = \alpha (\zeta_2 - 1)\). Therefore, there is no possibility of identification of the price \(\zeta_1\) of volatility risk.

**Remark 5.2.** While both prices of risk \(\zeta_1\) and \(\zeta_2\) may be identified by the two equations (5.4), it is worth questioning the identification strength of these equations. For this purpose, let us considering our ARG(1) specification introduced in section 3.4. Following \(a(u) = \frac{\rho u}{1 + cu}\) and \(b(u) = \delta \log [1 + cu]\), we have

\[
\begin{align*}
    a [\zeta_1 + \alpha (\zeta_2 - 1)] - a [\zeta_1 + \alpha (\zeta_2)] &= -\rho \frac{LEV}{[1 + c(\zeta_1 + \alpha (\zeta_2))] [1 + c(\zeta_1 + \alpha (\zeta_2 - 1))]}, \\
    b [\zeta_1 + \alpha (\zeta_2 - 1)] - b [\zeta_1 + \alpha (\zeta_2)] &= \delta \log \left[1 - c \frac{LEV}{1 + c(\zeta_1 + \alpha (\zeta_2))}\right]. \tag{5.6}
\end{align*}
\]

With the notations

\[
\begin{align*}
    \tilde{\beta} &= \beta (\zeta_2) - \beta (\zeta_2 - 1), \\
    \tilde{\gamma} &= \gamma (\zeta_2) - \gamma (\zeta_2 - 1), \tag{5.7}
\end{align*}
\]

(5.4), (5.6) and (5.7) imply that identification of structural parameters are based on

\[
\begin{align*}
    \tilde{\beta} [1 + c(\zeta_1 + \alpha (\zeta_2))] [1 + c(\zeta_1 + \alpha (\zeta_2 - 1))] &= -\rho LEV, \tag{5.8} \\
    \tilde{\gamma} &= \delta \log \left[1 - c \frac{LEV}{1 + c(\zeta_1 + \alpha (\zeta_2))}\right].
\end{align*}
\]

Identification weakness for the structural parameter \(\zeta_1\) is obvious from (5.8) when the leverage coefficient \(LEV\) is close to zero. We see that \(LEV\) being close to zero implies that \(\tilde{\beta}\) and \(\tilde{\gamma}\) are close to zero. As explained in remark 5.1, \(\tilde{\beta} = \tilde{\gamma} = 0\) implies that only \(\zeta_2\) is identified (strongly in general) by \(\alpha (\zeta_2) = \alpha (\zeta_2 - 1)\).
5.2 Links Functions to Identify Leverage Effect

5.2.1 Continuous Time Intuition

As explained above, structural parameters must include not only the two prices of risk \( \zeta_1 \) and \( \zeta_2 \), but also a leverage effect parameter \( \phi \). This parameter \( \phi \) to be defined, should be a proper normalization of the parameter \( LEV \). The difficulty with discrete time modeling is that leverage effect kicks in for two parameters corresponding to the conditional variance and conditional mean, respectively. On the one hand, in the conditional variance \( \omega_t \left( \sigma_{t+1}^2 \right) \), the parameter \( \xi \in [0, 1] \) characterizes the extent of variance reduction brought by the knowledge of the contemporaneous value \( \sigma_{t+1}^2 \) of the volatility factor. On the other hand, in the conditional mean \( \mu_t \left( \sigma_{t+1}^2 \right) \), the parameter \( LEV \leq 0 \) characterizes the drift update (proportional to \( \sigma_{t+1}^2 \)) that is implied by the knowledge of the contemporaneous value \( \sigma_{t+1}^2 \) of the volatility factor.

To understand how these two impacts of leverage effect should be linked together, it is worth making the analogy with the standard continuous time model:

\[
\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t^S, \\
\frac{d\sigma_t}{\sigma_t} = \gamma_t dt + \chi_t dW_t^\sigma, \\
Corr \left[ dW_t^S, dW_t^\sigma \right] = \phi \in [-1, 0].
\]

In this model, the orthogonal decomposition

\[
dW_t^S = \phi dW_t^\sigma + \sqrt{1 - \phi^2} dW_t^\perp
\]

implies that

\[
Var \{dW_t^S | dW_t^\sigma\} = (1 - \phi^2) dt, \\
E \left\{ \frac{dS_t}{S_t} \bigg| dW_t^\sigma \right\} = \mu_t dt + \phi \sigma_t dW_t^\sigma.
\]

5.2.2 First Link Function: Variance Reduction

We compare the continuous time formula

\[
Var \{dW_t^S | dW_t^\sigma\} = (1 - \phi^2) dt
\]

with the discrete time one

\[
\omega_t \left( \sigma_{t+1}^2 \right) = \xi \sigma_{t+1}^2 + X_t.
\]
While the state variable $X_t$ is needed for correcting some bias due to the discrete time modelling (confusion between integrated variance and spot volatility), identification between (5.12) and (5.13) leads us to define the structural parameter $\phi$ through the link function:

$$\xi = 1 - \phi^2. \quad (5.14)$$

It is worth stressing that (5.14) is a relation between the reduced form parameter $\xi$ and the structural parameter $\phi$. While a consistent estimator $\hat{\xi}$ of $\xi$ can easily be deduced from the estimation of the parametric model that characterizes the joint dynamics of $(\sigma_{t+1}^2, r_{t+1})$ (see next section), a consistent estimator $\hat{\phi}$ of the structural parameter $\phi$ can be deduced as

$$\hat{\phi} = -\sqrt{1 - \min(\hat{\xi}, 1)}, \quad (5.15)$$

where the negative solution of (5.14) has been chosen based on our prior knowledge of leverage effect as a negative correlation. Even though $\hat{\phi}$ is not the estimator of $\phi$ we will eventually use (in particular because we want to also refer to the second link function associated to update of drift), it allows us to sketch the double non-standard issue we meet for inference about the structural parameter $\phi$ when its true value is in a tiny neighborhood of zero. First, in the link function (5.14), the derivative of $\xi$ w.r.t. to $\phi$ is arbitrarily close to zero when $\phi$ is close to zero. Hence, similarly to situations described by Dufour (1997), $\phi$ can be arbitrarily poorly identified when the reduced form parameter $\xi$ is arbitrarily close to 1. Second, (5.15) shows that asymptotic normality will be a poor approximation of the actual distribution of $\hat{\phi}$ in finite sample when $\phi$ is close to zero.

### 5.2.3 Second Link Function: Drift Update

We compare the continuous time formula

$$E \left\{ \frac{dS_t}{S_t} \left| dW_t^\sigma \right. \right\} = \mu_t dt + \phi \sigma_t dW_t^\sigma \quad (5.16)$$

with the discrete time one

$$\mu_t (\sigma_{t+1}^2) = \psi \sigma_{t+1}^2 + Z_t, \quad (5.17)$$

$$\psi = \left( \zeta_2 - \frac{1}{2} \right) \xi + LEV.$$
We can overlook at this stage the first term \((\zeta^2 - 1/2)\xi\) in \(\psi\), since, as already explained, this term is a discrete time artifact that fails to disentangle the volatility feedback effect (with Jensen correction) and the leverage effect.

The leverage effect term in (5.17) is \(LEV\sigma_{t+1}^2\), which must be identified with the innovation term \(\phi\sigma_t dW_t^\sigma\) in (5.16). In other words, there should be a proportionality coefficient \(k\) that connects the risk-neutral characteristic \(LEV\) and the historical leverage effect \(\phi\) such that

\[
LEV = k\phi. \tag{5.18}
\]

To figure out this coefficient \(k\), it is worth noting the following. On the one hand, when the volatility feedback is overlooked, we have \(\psi = LEV\). On the other hand, for a properly defined coefficient \(k_t\), we have

\[
\psi = k_t Corr[r_{t+1}, \sigma_{t+1}^2 | I(t)], \tag{5.19}
\]

where \(Corr[r_{t+1}, \sigma_{t+1}^2 | I(t)]\) stands for the instantaneous correlation between return and volatility, a quantity that is typically seen as the right measure of leverage effect in continuous time models. In other words, if the sequence \(k_t\) were time invariant, it would be natural to see the leverage effect parameter \(\phi \in [-1, 0]\) as equal to this correlation. However, \(k_t\) is possibly time-varying, since by definition,

\[
Cov[r_{t+1}, \sigma_{t+1}^2 | I(t)] = Cov\{E[r_{t+1} | I^\sigma(t)], \sigma_{t+1}^2 | I(t)\} = \psi Var[\sigma_{t+1}^2 | I(t)], \tag{5.20}
\]

which implies

\[
Corr^2[r_{t+1}, \sigma_{t+1}^2 | I(t)] = \psi^2 \frac{Var[\sigma_{t+1}^2 | I(t)]}{Var[r_{t+1} | I(t)]} = \psi^2 \frac{Var[\sigma_{t+1}^2 | I(t)]}{E[\sigma_{t+1}^2 | I(t)]}. \tag{5.21}
\]

Hence,

\[
k_t^2 = \frac{E[\sigma_{t+1}^2 | I(t)]}{Var[\sigma_{t+1}^2 | I(t)]} = \frac{E[IV_{t+1} | I(t)]}{Var[IV_{t+1} | I(t)]}. \tag{5.22}
\]

**Remark 5.3.** Han, Khрапов, and Renault (2020) document empirically with observations on S&P500 that it is approximately true that there is a time-invariant proportionality between the two time series \(E[IV_{t+1} | I(t)]\) and \(Var[IV_{t+1} | I(t)]\).

**Remark 5.4.** Obviously, with our simplified \(ARG(1)\) model for \(\sigma_{t+1}^2 = IV_{t+1}\), it is not true that the ratio of \(E[\sigma_{t+1}^2 | I(t)]\) over \(Var[\sigma_{t+1}^2 | I(t)]\) is time invariant. As already mentioned,
Han, Khrapov, and Renault (2020) consider a more general model:

$$\sigma_{t+1}^2 = IV_{t+1} = \tilde{\sigma}_{t+1}^2 + f\tilde{\sigma}_t^2 + g$$  \hspace{1cm} (5.23)

for some ARG(1) process $\tilde{\sigma}_{t}^2$. Then the parameters $f$ and $g$ can be chosen such that $E[\sigma_{t+1}^2 | I(t)]$ over $\text{Var}[\sigma_{t+1}^2 | I(t)]$ is time invariant. However, as already acknowledged we will overlook this possibility of generalization.

In our simplified model, we follow the above logic by assuming

$$LEV = k\phi$$  \hspace{1cm} (5.24)

with

$$k^2 = \frac{E[\sigma_{t+1}^2]}{E\{\text{Var}[\sigma_{t+1}^2 | I(t)]\}},$$  \hspace{1cm} (5.25)

which would coincide with $k_i^2$ in the case of a time invariant sequence.

### 5.2.4 Link Functions for Leverage Parameter $\phi$ in Our Fully Parametric Model

We deduce from former subsections that the two reduced form leverage parameters $\xi$ and $\psi$ are related to the structural parameter $\phi$ by two link functions:

$$\xi = 1 - \phi^2,$$  \hspace{1cm} (5.26)

$$\psi = k\phi + \xi\left(\zeta_2 - \frac{1}{2}\right),$$

with

$$k^2 = \frac{1}{c(1+\rho)},$$  \hspace{1cm} (5.27)

where (5.27) follows by simple calculation from the application of (5.25) to our ARG(1) model.

### 6 Estimation of Reduced Form Parameters

As discussed in Remark 3.4, we have two sets of reduced-form parameters that characterize the joint distribution of $(r_{t+1}, \sigma_{t+1}^2)$: (i) $\omega_1 = (\rho, c, \delta)'$ that characterizes the conditional distribution of $\sigma_{t+1}^2$ given $I(t)$, and (ii) $\omega_2 = (\gamma, \beta, \psi, \xi)'$, which jointly with $\omega_1$, defines the conditional distribution of the return $r_{t+1}$ given $I_0^n$. As such, the reduced-form parameter $\omega = (\omega_1, \omega_2)'$ is identified by the conditional mean and variance of the volatility factor and
the return. Specifically, following (3.32), the reduced-form parameter $\omega$ satisfies

$$
E[u_{1,t+1}(\omega)|I(t)] = 0, \quad \text{with} \quad u_{1,t+1}(\omega) = \sigma^2_{t+1} - [\rho c + \sigma^2_{t+1}], \quad (6.1)
$$

$$
E[u_{2,t+1}(\omega)|I(t)] = 0, \quad \text{with} \quad u_{2,t+1}(\omega) = \sigma^4_{t+1} - [\rho^2 c^2 (1 + \delta) \sigma^2_t + c^2 \delta (1 + \delta)],
$$

and, following (3.27) and (3.28),

$$
E[\varepsilon_{1,t+1}(\omega)|I^\sigma(t)] = 0, \quad \text{with} \quad \varepsilon_{1,t+1}(\omega) = r_{t+1} - (\psi \sigma^2_{t+1} + \beta \sigma^2_t + \gamma), \quad (6.2)
$$

$$
E[\varepsilon_{2,t+1}(\omega)|I^\sigma(t)] = 0, \quad \text{with} \quad \varepsilon_{2,t+1}(\omega) = r^2_{t+1} - (\psi \sigma^2_{t+1} + \beta \sigma^2_t + \gamma)^2 - Var[r^2_{t+1}|I^\sigma(t)],
$$

where

$$
Var[r^2_{t+1}|I^\sigma(t)] = \xi \sigma^2_{t+1} + [(1 - \xi) a'(0) + \psi^2 a''(0)] \sigma^2_t + (1 - \xi) b'(0) + \psi^2 b''(0)
$$

$$
= \xi \sigma^2_{t+1} + [(1 - \xi) \rho - 2 \psi^2 \rho c] \sigma^2_t + (1 - \xi) \delta c - \psi^2 \delta c^2.
$$

These conditional moments are based on the martingale differences sequences $u_{t+1}(\omega) = (u_{1,t+1}(\omega), u_{2,t+1}(\omega))'$ and $\varepsilon_{t+1}(\omega) = (\varepsilon_{1,t+1}(\omega), \varepsilon_{2,t+1}(\omega))'$. Furthermore, they form a set of sequential moments as in Ai and Chen (2012) since $u_{t+1}(\omega) \in I^\sigma_t$. It is obvious from (6.1) and (6.2) that the reduced-form parameter $\omega$ can be easily identified. Without loss of generality, we assume that they are strongly identified such that the standard GMM estimator has an asymptotic normal distribution.

We estimate $\omega$ based on the unconditional moment restrictions

$$
E[h_t(\omega)] = 0, \quad \text{where} \quad h_t(\omega) = [h'_{1,t}(\omega), h'_{2,t}(\omega)]', \quad (6.4)
$$

$$
h_{1,t}(\omega) = [x_{1,t} \otimes u_{1,t+1}(\omega_1), z_{1,t} \otimes u_{2,t+1}(\omega_1)]',
$$

$$
x_{1,t} = (1, \sigma^2_t), \quad z_{1,t} = (1, \sigma^2_t, \sigma^4_t),$$

$$
h_{2,t}(\omega) = [x_{2,t} \otimes e_{1,t+1}(\omega), z_{2,t} \otimes e_{2,t+1}(\omega)]',
$$

$$
x_{2,t} = (1, \sigma^2_t, \sigma^2_{t+1}), \quad z_{2,t} = (x_{2,t}, \sigma^4_t, \sigma^2_{t+1}, \sigma^4_{t+1}).
$$

Let $\widehat{\omega}$ denote the efficient two-step GMM estimator based on the moments in (6.4). Let $P$ denote the distribution of the data $W = \{(r_{t+1}, \sigma^2_{t+1}) : t \geq 1\}$ and $\mathcal{P}$ denote the parameter space of $P$. We make the following high-level assumption on the reduced-form parameter.

**Assumption R.** The following conditions hold uniformly over $P \in \mathcal{P}$. For some positive definite matrix $\Omega$, its estimator $\widehat{\Omega}$, and some fixed constant $0 < C < \infty$.

(i) $T^{1/2}\Omega^{-1/2}(\widehat{\omega} - \omega) \rightarrow_d N(0, I)$.

(ii) $\widehat{\Omega} - \Omega \rightarrow_p 0$.  

27
(iii) \( \lambda_{\text{min}}(\Omega) \geq C^{-1} \) and \( \lambda_{\text{max}}(\Omega) \leq C \).

7 Identification Robust Inference for Structural Parameters

The true values of the structural parameter \( \theta = (\zeta_1, \zeta_2, \phi)' \) and the reduced-form parameter \( \omega \) satisfy the link function \( g(\theta, \omega) = 0 \), where this four-dimensional link function are given by the two identification conditions in (5.4) and the two additional equalities in (5.26). Once the reduced-form parameter \( \omega \) is identified and estimated, following the discussions in Section 6, we rely on this link function for the identification and inference of the structural parameter \( \theta \).

In a standard problem without any identification issues of the structural parameter, we can estimate \( \theta \) by the minimum distance estimator and construct tests and confidence sets for \( \theta \) using an asymptotically normal approximation for \( T^{1/2} (\hat{\theta} - \theta) \). However, this standard method does not work here. This link function only provides weak identification of \( \zeta_1 \), the price of volatility risk, under mild leverage effect, as discussed in Remark 5.2. In this case, \( g(\theta, \hat{\omega}) \) is almost flat in \( \zeta_1 \) and the minimum distance estimator of \( \hat{\zeta}_1 \) may not even be consistent, see Stock and Wright (2000). To make the problem even more complicated, the inconsistency of \( \hat{\zeta}_1 \) has a spillover effect on \( \hat{\zeta}_2 \) and \( \hat{\phi} \), making their distribution non-normal even in large samples, as demonstrated in Andrews and Cheng (2012).

Let \( g_0(\theta) \) denote the link function \( g(\theta, \omega) \) evaluated at the true value of \( \omega \) and \( \hat{g}(\theta) \) denote its counterpart evaluated at the GMM estimator \( \hat{\omega} \). Let \( G(\theta, \omega) \) denote the partial derivative of \( g(\theta, \omega) \) wrt \( \omega \), abbreviated as \( G_0(\theta) \) when evaluated at the true value of \( \omega \) and as \( \hat{G}(\theta) \) when evaluated at \( \hat{\omega} \).

We construct a confidence set for \( \theta \in \Theta := [-M_1, 0] \times [0, M_2] \times [-1 + \epsilon, 0] \) by inverting the test \( H_0 : \theta = \theta_0 \) vs \( H_1 : \theta \neq \theta_0 \), where \( M_1 \) and \( M_2 \) are large positive constants and \( \epsilon \) is a small positive constant. The test statistic is a QLR statistic that takes the form

\[
\text{QLR}(\theta_0) = T \hat{g}(\theta_0)' \hat{\Sigma}(\theta_0, \theta_0)^{-1} \hat{g}(\theta_0) - \min_{\theta \in \Theta} T \hat{g}(\theta)' \hat{\Sigma}(\theta, \theta)^{-1} \hat{g}(\theta),
\]

where \( \hat{\Sigma}(\theta_1, \theta_2) = \hat{G}(\theta_1) \hat{\Omega}^1 \hat{G}(\theta_2)' \) and \( \hat{\Omega} \) is a consistent estimator of \( \Omega \).

Before presenting the robust confidence set based on the QLR statistic, we first introduce some useful quantities and provide a heuristic discussion of the identification problem and its consequences. Define

\[
\eta_T(\theta) = T^{1/2} \left[ \hat{g}(\theta) - g_0(\theta) \right] = G_0(\theta) \Omega^{1/2} \cdot \xi_T + o_p(1),
\]

(7.2)
where $\xi_T \to_d N(0, I)$ following Assumption R. Thus, $\eta_T(\cdot)$ weakly converges to a Gaussian process $\eta(\cdot)$ with covariance function $\Sigma(\theta_1, \theta_2) = G_0(\theta_1)G_0(\theta_2)'$. Following (7.2), we can write $T^{1/2}\hat{g}(\theta) = \eta_T(\theta) + T^{1/2}g_0(\theta)$, where $\eta_T(\theta)$ is the noise from the reduced-form parameter estimation and $T^{1/2}g_0(\theta)$ is the signal from the link function. Under weak identification, $g_0(\theta)$ is almost flat in $\theta$, modeled as the signal $T^{1/2}g_0(\theta)$ being finite even for $\theta \neq \theta_0$ and $T \to \infty$. Thus, the signal and the noise are of the same order of magnitude, yielding an inconsistent minimum distance estimator $\hat{\theta}$.

The identification strength of $\theta_0$ is determined by the function $T^{1/2}g_0(\theta)$. However, this function is unknown and cannot be consistently estimated (due to $T^{1/2}$). Thus, we take the conditional inference procedure as in Andrews and Mikusheva (2016) and view $T^{1/2}g_0(\theta)$ as an infinite dimensional nuisance parameter for the inference of $\theta_0$. The goal is to construct robust confidence set for $\theta_0$ that has correct size asymptotically regardless of this unknown nuisance parameter.

Andrews and Mikusheva (2016) provide the conditional QLR test in a nonlinear GMM problem, where $\hat{g}(\theta)$ is replaced by a sample moment. The same method can be applied to the present nonlinear minimum distance problem. Following Andrews and Mikusheva (2016), we first project $\hat{g}(\theta)$ onto $\hat{g}(\theta_0)$ and construct a residual process

$$\hat{r}(\theta) = \hat{g}(\theta) - \hat{\Sigma}(\theta, \theta_0)\hat{\Sigma}(\theta_0, \theta_0)^{-1}\hat{g}(\theta_0).$$

(7.3)

The limiting distributions of $\hat{r}(\theta)$ and $\hat{g}(\theta_0)$ are Gaussian and independent. Thus, conditional on $\hat{r}(\theta)$, the asymptotic distribution of $\hat{g}(\theta)$ no longer depends on the nuisance parameter, $T^{1/2}g_0(\theta)$, making the procedure robust to any identification strength.

Specifically, we obtain the $1 - \alpha$ conditional quantile of the QLR statistic, denoted by $c_{1-\alpha}(r, \theta_0)$, as follows. For $b = 1, \ldots, B$, we take independent draws $\eta^*_b \sim N(0, \hat{\Sigma}(\theta_0, \theta_0))$ and produce a simulated process,

$$g^*_b(\theta) = \hat{r}(\theta) + \hat{\Sigma}(\theta, \theta_0)\hat{\Sigma}(\theta_0, \theta_0)^{-1}\eta^*_b,$$

(7.4)

and a simulated statistic,

$$QLR^*_b(\theta_0) = T\eta^*_b'\hat{\Sigma}(\theta_0, \theta_0)^{-1}\eta^*_b - \min_{\theta \in \Theta} Tg^*_b(\theta)\hat{\Sigma}(\theta, \theta)^{-1}g^*_b(\theta).$$

(7.5)

Let $b_0 = \lceil(1 - \alpha)B\rceil$, the smallest integer greater than or equal to $(1 - \alpha)B$. Then the critical value $c_{1-\alpha}(r, \theta_0)$ is the $b_0^{th}$ smallest value among $\{QLR^*_b, b = 1, \ldots, B\}$. Finally, we construct a robust confidence set for $\theta$ by collecting the null values that are not rejected,
i.e., the nominal level $1 - \alpha$ confidence set is

$$CS_T = \{\theta_0 : QLR_T(\theta_0) \leq c_{1-\alpha}(r, \theta_0)\}.$$ (7.6)

**Assumption S.** The following conditions hold over $P \in \mathcal{P}$, for any $\theta$ in its parameter space, and any $\omega$ in some fixed neighborhood around its true value, for some fixed $0 < C < \infty$.

1. $g(\theta, \omega)$ is partially differentiable in $\omega$, with partial derivative $G(\theta, \omega)$ that satisfies $\|G(\theta_1, \omega) - G(\theta_2, \omega)\| \leq C\|\theta_1 - \theta_2\|$ and $\|G(\theta, \omega_1) - G(\theta, \omega_2)\| \leq C\|\omega_1 - \omega_2\|$.

2. $C^{-1} \leq \lambda_{\min}(G(\theta, \omega)'G(\theta, \omega)) \leq \lambda_{\max}(G(\theta, \omega)'G(\theta, \omega)) \leq C$.

**Theorem 1.** Suppose Assumption R and Assumption S hold. Then,

$$\liminf_{T \to \infty} \inf_{P \in \mathcal{P}} \Pr(\theta_0 \in CS_T) \geq 1 - \alpha.$$ 

This theorem states that the confidence set constructed by the conditional QLR test has correct asymptotic size. Uniformity is important for this confidence set to cover the true parameter with a probability close to $1 - \alpha$ in finite-samples. This uniform result is established over a parameter space $\mathcal{P}$ that allows for weak identification of the structural parameter $\theta$.

**8 Simulations**

In this section, we investigate the finite-sample performance of the proposed test and show that the asymptotic approximations derived above work well in practice. We also compare it with the standard test that assumes all parameters are strongly identified. The standard test is known to be invalid under weak identification but its degree of distortion is unknown in general. We simulate the data using the parametric model above where the true values of the parameters are given in Table 1 based on the values used by Han, Khrapov, and Renault (2020). To investigate the robustness of the procedure with respect to various identification strengths, we vary both $\phi$ and $T$. Specifically, we consider $\phi \in \{-0.40, -0.10, -0.01\}$ and $T \in \{2,000; 10,000\}$. The number of data points in the empirical section is $\approx 4,200$ for comparison.

1To avoid boundary issues with respect to the estimate of $c$ and $\delta$ in finite-sample, we reparameterize the moment conditions and link functions in terms of $\log(c)$, $\log(c) + \log(\delta)$, and $\logit(\rho)$. This reparameterization forces the scale parameters to be positive and $\rho$ to lie in $(0, 1)$. We find that the resulting estimates for the transformed reduced-form parameters are better approximated by the Gaussian distribution for a given finite sample.
Table 1: Simulation Set-up

<table>
<thead>
<tr>
<th>δ</th>
<th>ρ</th>
<th>c</th>
<th>ζ₁</th>
<th>ζ₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6475</td>
<td>0.50</td>
<td>3.94128 × 10⁻³</td>
<td>-10</td>
<td>1.7680</td>
</tr>
</tbody>
</table>

Parameter Values used by Han, Khrapov, and Renault (2020)

To show the effect of various identification strength, we first vary the true value of φ and plot the distribution of \( \hat{\zeta}_1 \) and \( \hat{\zeta}_2 \) in Figure 1. The reported result is based on 10,000 observations and 2000 simulation repetitions. The black lines in the middle of the figures are the true parameter values. Clearly, the estimators sometimes pile up at the boundaries of the parameter space. As expected, this simulation shows that the Gaussian distribution is not a good approximation for the finite-sample distribution of either of the estimators.

Next, we study the finite-sample size of in the standard QLR test and the proposed
conditional QLR test for a joint test for the three structural parameters. The nominal level of the test is 5%. The critical value of the standard QLR test is the 95% quantile of the $\chi^2$-distribution with 3 degree of freedom. The critical value of the conditional QLR test is obtained by the simulation-based procedure in Section 7, with 1000 simulation repetitions to approximate the quantile of the conditional distribution. The finite-sample size is based on 1000 simulation repetitions.

Table 2: Finite-Sample Size of the Standard and Proposed Tests

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>Standard %</th>
<th>Proposed %</th>
<th>Standard %</th>
<th>Proposed %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T = 2,000$</td>
<td>$T = 10,000$</td>
<td>$T = 2,000$</td>
<td>$T = 10,000$</td>
</tr>
<tr>
<td>-0.01</td>
<td>1.80</td>
<td>5.40</td>
<td>2.35</td>
<td>5.65</td>
</tr>
<tr>
<td>-0.40</td>
<td>3.55</td>
<td>5.10</td>
<td>4.35</td>
<td>4.65</td>
</tr>
</tbody>
</table>

Simulation results show that the standard QLR test under-rejects in finite-sample. This is most severe when the identification is weak, e.g., for $\phi = -0.01$, the rejection rate is 1.8% for $T = 10,000$. The proposed test, however, has finite-sample coverage probabilities close to the nominal level in all cases.

9 Empirical Application

For the empirical application, we use the daily return on the S&P 500 for $r_{t+1}$ and the associated realized volatility computed with high-frequency data for $\sigma^2_{t+1}$. The data is obtained from SPY (SPDR S&P 500 ETF Trust), an exchange-traded fund that mimics the S&P 500. This gives us a market index whose risk is not easily diversifiable and can be used to estimate the prices of risk that investors face in practice. We use the procedure Sangrey (2019) develops to estimate the integrated total volatility, i.e., the instantaneous expectation of the price variance. This measure reduces to the integrated diffusion volatility if prices have continuous paths and it works well in the presence of market microstructure noise.

Since SPY is one of the most liquid assets traded, we can choose the frequency at which we sample the underlying price. To balance market-microstructure noise, computational cost, and efficiency of the resultant estimators, we sample at the 1-second frequency. We annualize the data by multiplying $r_{t+1}$ by 252 and $\sigma^2_{t+1}$ by $252^2$ (i.e., we annualize them). The data starts in 2003 and ends in 2019. Since the asset is only traded during business hours, this leads to 4279 days of data with an average of approximately 24,000 observations per day. We compute $r_{t+1}$ as the daily return from the open to the close of the market, the interval over which we can estimate the volatility. This avoids specifying the relationship
between overnight and intra-day returns. We preprocess the data using the pre-averaging approach as in Podolskij and Vetter (2009) and Aït-Sahalia, Jacod, and Li (2012).

**Figure 2: S&P 500 Volatility and Log-Return**

To see how the data move over time, we plot their time series in Figure 2. We also plot the joint unconditional distribution in Figure 2 to see the static relationship between the two series. The volatility has a long-right tail, a typical gamma-type distribution. The returns have a bell-shaped distribution. They are slightly negatively correlated, as shown by the regression line in the joint plot. This corroborates the work by Bandi and Renò (2012) and Aït-Sahalia, Fan, and Li (2013). We also report a series of summary statistics.

**Table 3: Summary Statistics**

<table>
<thead>
<tr>
<th></th>
<th>$r_{t+1}$</th>
<th>$\sigma_{t+1}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.45</td>
<td>3296.55</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>57.88</td>
<td>3296.55</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>-0.32</td>
<td>12.94</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>12.97</td>
<td>274.67</td>
</tr>
<tr>
<td><strong>Correlation</strong></td>
<td>-0.03</td>
<td></td>
</tr>
</tbody>
</table>

We now report the estimates and confidence intervals for the reduced-form parameters $c$, $\delta$, and $\rho$. The confidence intervals reported here use the Gaussian limiting theory, i.e., the point estimates ±1.96 standard errors.\(^2\) For confidence intervals of the three structural

\(^2\)We first obtain confidence intervals for $\log(c)$ and $\log(c) + \log(\delta)$, and transform them into confidence intervals for $c$ and $\delta$. Similarly, we create the confidence interval for $\rho$ by inverting the interval for $\text{logit}(\rho)$. 

---

33
parameters, we first compute their joint confidence set based on the conditional QLR test and then project it to each of the components. We use 2000 simulations to compute the quantile for the QLR statistic.

Table 5: Structural Parameters

<table>
<thead>
<tr>
<th></th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>(-0.58, -0.17)</td>
</tr>
<tr>
<td>$\zeta_1$</td>
<td>(-31.44, -2.97)</td>
</tr>
<tr>
<td>$\zeta_2$</td>
<td>(0.00, 2.67)</td>
</tr>
</tbody>
</table>

The results in Table 5 have a few notable features. First, and most important, we can reject the hypothesis that the price of volatility risk equals zero with the identification-robust test. The largest value lies within $\zeta_1$’s confidence interval is $-2.97$. The data are unable to reject the combination of large negative values of $\zeta_1$ and zero values of $\zeta_2$. In other words, we cannot reject that the aversion to volatility is capable of explaining the entire equity premium. We can however reject the idea that the aversion to equity risk is capable of explaining the entire equity premium. Our confidence interval for $\phi$ and the upper bound for $\zeta_2$ are consistent with findings in the literature.

10 Conclusion

A common belief is that only option price data may allow for identification of the risk aversion to volatility of volatility. We prove in this paper that, in the presence of leverage effect, observations on the underlying asset return could allow for identification without any option data. Moreover, in contrast with a widespread practice, the presence of leverage effect implies that the variance risk premium does not provide separate identification of the risk aversion to volatility of volatility and the standard risk aversion to volatility level.

While our general identification result is model-free, we provide a simple parametric model for the sake of numerical illustration of our identification robust strategy. This novel inference strategy extends the approach put forward by Andrews and Mikusheva
(2016) from the context of GMM to minimum distance problems. This extension may be useful in other fields of structural econometrics. For the sake of clarity, our empirical illustration is developed in an overly simplified parametric model on purpose. This model could be enriched by more sophisticated dynamics of the volatility factor. An additional useful extension would be the introduction of two volatility factors, one for short term volatility and the other for long term volatility. This extension is natural because the economic uncertainty that we want to capture with volatility of volatility is well understood by the long run risk model (Bansal and Yaron, 2004).

Moreover, following an argument put forward by Bandi and Renò (2016), we suspect that our identification strategy based on leverage effect would be more compelling with a model allowing for jumps both in return and volatility. While accommodating jumps in a discrete time model in general requires the Markov switching regime models, we could resort to the model proposed by Augustyniak, Bauwens, and Dufays (2019).

Of course, although our focus of interest is the possibility of identification without data provided by derivative markets, it does not mean that option price data should be wasted when they are available. They obviously in general allow us to obtain much tighter confidence sets for structural parameters, see e.g., the empirical results of Han, Khrapov, and Renault (2020). Our result must rather be understood as a possibility result. As in other contexts, the possibility of nonparametric identification of a structural model is worth studying even though empirical implementations generally resort to a parametric specification. It could be argued that our results are more robust to misspecification than those obtained with additional option data. For instance, it is well documented that stochastic volatility processes often feature some long memory, see e.g., Comte and Renault (1998). Although long memory is arguably observationally equivalent to structural breaks (or switching regimes) as far as underlying asset returns are concerned, option pricing is much more sensitive to the model choice. Our method avoids this potential model misspecification issue by using asset return data only for inference on the structural parameters.
A.1 Computation of $\mathcal{L}_t^*$ for section 4.2

We know from (2.8) that:

$$\mathcal{L}_t^*(u, v) = \frac{\mathcal{L}_t(u + \zeta_1, v + \zeta_2)}{\mathcal{L}_t(\zeta_1, \zeta_2)}.$$  

The numerator can be written:

$$N_t = E[\exp \left\{ - (u + \zeta_1) \sigma_{t+1}^2 - (v + \zeta_2)r_{t+1} \right\} |I(t)]$$

By (4.6), we get

$$L_t^*(u, v) = \exp \left\{ -\gamma (v + \zeta_2) - \beta (v + \zeta_2) \sigma_t^2 \right\} E[\exp \left\{ -[u + \zeta_1 + \alpha(v + \zeta_2)]\sigma_{t+1}^2 \right\} |I(t)].$$  

Similarly, the denominator can be written:

$$D_t = \exp \left\{ -\gamma (\zeta_2) - \beta (\zeta_2) \sigma_t^2 \right\} E[\exp \left\{ -[\zeta_1 + \alpha(\zeta_2)]\sigma_{t+1}^2 \right\} |I(t)].$$  

By computing the ratio $[N_t/D_t]$ and using (4.6), we get

$$\mathcal{L}_t^*(u, v) = \exp \left\{ -\gamma^* (v) - \beta^* (v) \sigma_t^2 \right\} \frac{B_t}{C_t}.$$  

with

$$B_t = E[\exp \left\{ -[u + \zeta_1 + \alpha(v + \zeta_2)]\sigma_{t+1}^2 \right\} |I(t)],$$

$$C_t = E[\exp \left\{ -[\zeta_1 + \alpha(\zeta_2)]\sigma_{t+1}^2 \right\} |I(t)].$$  

Using the definition (4.6) of $\alpha^*(\cdot)$, we can write

$$B_t = E[\exp \left\{ -[u + \zeta_1 + \alpha^*(v) + \alpha(\zeta_2)]\sigma_{t+1}^2 \right\} |I(t)]$$

By (4.5) we know that

$$a [u + \zeta_1 + \alpha^*(v) + \alpha(\zeta_2)] = a^* [u + \alpha^*(v)] + a [\zeta_1 + \alpha(\zeta_2)],$$

$$b [u + \zeta_1 + \alpha^*(v) + \alpha(\zeta_2)] = b^* [u + \alpha^*(v)] + b [\zeta_1 + \alpha(\zeta_2)].$$
Hence,

\[ B_t = \exp \left\{ -a^* [u + \alpha^*(v)] \sigma_t^2 - a \left[ \zeta_1 + \alpha(\zeta_2) \right] \sigma_t^2 \right\} \exp \left\{ -b^* [u + \alpha^*(v)] - b \left[ \zeta_1 + \alpha(\zeta_2) \right] \right\}. \]  

(A.7)

For \( u = v = 0 \), we get

\[ C_t = \exp \left\{ -a \left[ \zeta_1 + \alpha(\zeta_2) \right] \sigma_t^2 - b \left[ \zeta_1 + \alpha(\zeta_2) \right] \right\}. \]  

(A.8)

By computing the ratio, we end up with

\[ \frac{B_t}{C_t} = \exp \left\{ -a^* [u + \alpha^*(v)] \sigma_t^2 \right\} \exp \left\{ -b^* [u + \alpha^*(v)] \right\}. \]  

(A.9)

Hence, as announced in section 4.2, we have shown that

\[ \mathcal{L}_t^*(u, v) = \exp \left\{ -\gamma^*(v) \right\} \exp \left\{ -a^* [u + \alpha^*(v)] \sigma_t^2 \right\} \exp \left\{ -b^* [u + \alpha^*(v)] \right\} \]  

\[ = E^* \left[ \exp \left\{ -u \sigma_{t+1}^2 g_t(v/\sigma_{t+1}^2) \right\} | I(t) \right] \]  

(A.10)

with

\[ g_t(v/\sigma_{t+1}^2) = \exp \left\{ -a^* (v) \sigma_{t+1}^2 - \beta^*(v) \sigma_{t+1}^2 - \gamma^*(v) \right\}. \]

QED

A.2 Proof of Theorem 1

We obtain this result by applying Theorem 1 of Andrews and Mikusheva (2016). We now verify Assumptions 1–3 in Andrews and Mikusheva (2016). To show weak convergence of \( \eta_T(\cdot) \) to \( \eta(\cdot) \) uniformly over \( \mathcal{P} \), note that by a second-order Taylor expansion,

\[ \eta_T(\lambda) = T^{1/2} \left[ \tilde{g}(\lambda) - g_0(\lambda) \right] = G_0(\lambda) \Omega^{1/2} \xi_T + \delta_T, \]  

where

\[ \xi_T = \Omega^{-1/2} T^{1/2} \left( \tilde{\omega} - \omega_0 \right), \]  

and \( \delta_T = (G(\lambda, \tilde{\omega}) - G(\lambda, \omega_0)) T^{1/2} (\tilde{\omega} - \omega_0) \)  

(A.1)

and \( \tilde{\omega} \) is between \( \tilde{\omega} \) and \( \omega_0 \). Because \( \|G(\lambda, \tilde{\omega}) - G(\lambda, \omega_0)\| \leq C \| \tilde{\omega} - \omega_0 \| \), \( \delta_T = o_p(1) \) uniformly over \( \mathcal{P} \) following Assumption R. To show \( G_0(\lambda) \Omega^{1/2} \xi_T \) weakly converges to \( \eta(\cdot) \), it is sufficient to show (i) the pointwise convergence

\[ \left( \frac{G_0(\lambda_1) \Omega^{1/2} \xi_T}{G_0(\lambda_2) \Omega^{1/2} \xi_T} \right) \rightarrow_d \left( \frac{\eta(\lambda_1)}{\eta(\lambda_2)} \right), \]  

(A.2)

37
which follows from Assumption R, and (ii) the stochastic equicontinuity condition, i.e., for every $\varepsilon > 0$ and $\xi > 0$, there exists a $\delta > 0$ such that

$$
\limsup_{T \to \infty} \Pr \left( \sup_{P \in \mathcal{P}} \sup_{\|\lambda_1 - \lambda_2\| \leq \delta} \left\| G_0(\lambda_1) \Omega^{1/2} \xi_T - G_0(\lambda_2) \Omega^{1/2} \xi_T \right\| > \varepsilon \right) < \xi. \quad (A.3)
$$

For some $C < \infty$, we have $\left\| G_0(\lambda_1) - G_0(\lambda_2) \right\| \leq C \left\| \lambda_1 - \lambda_2 \right\|$ under a uniform bound for the derivative in Assumption S, and we have $\left\| \Omega^{1/2} \right\| \leq C$ under Assumption R. Thus,

$$
\limsup_{T \to \infty} \Pr \left( \sup_{P \in \mathcal{P}} \sup_{\|\lambda_1 - \lambda_2\| \leq \delta} \left\| G_0(\lambda_1) \Omega^{1/2} \xi_T - G_0(\lambda_2) \Omega^{1/2} \xi_T \right\| > \varepsilon \right)
\leq \limsup_{T \to \infty} \Pr \left( C^2 \sup_{P \in \mathcal{P}} \left\| \xi_T \right\| > \frac{\varepsilon}{\delta} \right). \quad (A.4)
$$

Because $\xi_T = O_p(1)$ uniformly over $P \in \mathcal{P}$, there exists $\delta$ such that $\varepsilon/\delta$ is large enough to make the right hand side of the inequality in (A.4) smaller than $\xi$. Assumptions 2 and 3 of Andrews and Mikuevsheva (2016) follow directly from Assumption R and Assumption S.

QED

References


