

Identification Robust Inference for Risk Prices in Structural Stochastic Volatility Models

Xu Cheng*

Eric Renault[†]

Paul Sangrey[‡]

November 13, 2018

Abstract

In structural stochastic volatility asset pricing models, 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 market 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 that invalidates standard inference. We adopt the discrete-time exponentially affine model of Han, Khrapov, and Renault (2018), which links the identification of volatility risk price to the leverage effect. In particular, we develop a minimum distance criterion that links the market risk price, the volatility risk price, and the leverage effect to the well-behaved reduced-form parameters governing the return and volatility's joint distribution. The link functions are almost flat if the leverage effect is close to zero, making estimating the volatility risk price difficult. We adapt the conditional quasi-likelihood ratio test Andrews and Mikusheva (2016) develop in a nonlinear GMM framework 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 prices regardless of the leverage effect's magnitude.

JEL Codes: C12, C14, C38, C58, G12

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

*University of Pennsylvania, The Perelman Center for Political Science and Economics, 133 South 36th Street, Philadelphia, PA 19104, xucheng@upenn.edu

[†]Brown University, Department of Economics – Box B, 64 Waterman Street, Providence, RI 02912, eric_renault@brown.edu

[‡]University of Pennsylvania, The Perelman Center for Political Science and Economics, 133 South 36th Street, Philadelphia, PA 19104, paul@sangrey.io

Contents

1	Introduction	2
2	Literature Review	3
3	The Model	4
3.2	The Stochastic Discount Factor	5
3.3	Parameterizing the Physical Measure Dynamics	5
3.3.1	Return Dynamics	6
3.3.2	Deriving the Conditional Mean	7
3.4	Identification of the Risk Prices	9
4	The Estimation Criterion in Population	10
4.1	Identified Set	10
5	Estimating the Reduced-Form Parameters	11
5.1	Robust Inference for Risk Price	13
6	Simulations	15
7	Data and Empirical Results	16
7.1	Data	16
7.2	Estimates	18
8	Conclusion	19
	References	21
	Appendices	24
	Appendix A Model Characterization	24
	Appendix B Identification Proofs	26

1 Introduction

The core of modern finance studies how investors optimally trade off risk and return. Standard economic theory predicts investors demand a higher return as compensation for bearing more risk. Hence, we should expect a positive relationship between the mean and volatility of returns. In Sharpe (1964) and Lintner’s (1965) capital asset pricing model (CAPM) the market return’s risk premia is proportional to its variance. However, both empirically and in models with stochastic volatility, the relationship is often nonlinear, (Bansal et al. 2014; Dew-Becker et al. 2017). This nonlinearity is particularly obvious in option pricing models, such as Bates (2008) and Christoffersen, Heston, and Jacobs (2013), where investors care not just about how an asset’s returns co-move with the market’s return but also care how they co-move with the market’s volatility.

In these structural stochastic volatility models, changes in volatility affect risk premia through two channels: (1) the investor’s willingness to tolerate high volatility in order to get high expected returns as measured by the market 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 that invalidates standard inference. We adopt the discrete-time exponentially affine model of Han, Khrapov, and Renault (2018), which links the identification of volatility risk price to the leverage effect. This effect measures the correlation between return and volatility innovations. Although theoretically less than zero, this effect is difficult to quantify empirically, (Bandi and Renò 2012; Aït-Sahalia, Fan, and Li 2013). This low signal-to-noise ratio (weak identification) invalidates standard inference and makes classic asymptotic approximation perform poorly.

We provide identification-robust inference by developing a minimum distance criterion that links the market risk price, the volatility risk price, and the leverage effect to a series of reduced-form parameters that determine the return and volatility’s joint distribution. The link functions are almost flat if the leverage effect is close to zero, making estimating the volatility risk price difficult. We adapt the conditional quasi-likelihood ratio test that Andrews and Mikusheva (2016) develop in a nonlinear general method of moments (GMM) framework 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 prices regardless of the leverage effect’s magnitude.

Before we do that, consider measuring the relationship between market volatility and expected returns when this relationship is linear, and so the market risk price is the only risk price. Even here, unlike the consensus in the theoretical literature, the empirical literature has found pinning down this relationship quite difficult. Not only has its magnitude proven difficult to determine, but various estimates even differ in sign, (Brandt and Kang 2004; Lettau and Ludvigson 2010). This empirical literature focuses on point estimates for the market risk price. However, if individual investors are ambiguity averse as in Hansen and Sargent (2001) and Ju and Miao (2012), they care not just about how the representative investor prices volatility but also about their uncertainty regarding this estimate. This is the first paper to provide identification-robust inference, and hence

valid confidence intervals, for these risk prices.

To further understand the inherent complexity that afflicts this estimation procedure, note that innovations to the volatility process affect the return in three ways, which we must disentangle. First, we must disentangle the volatility feedback effect or leverage — the contemporaneous correlation between the volatility and return innovations — from the risk premium — the relationship between volatility and expected returns. This second relationship is predictable, not contemporaneous. Second, and just as important, we must separate the market risk price and the volatility risk price. In general, and we show this below, the market risk price is strongly identified even in the presence of the volatility risk price. Because the market risk price measures a static tradeoff, there is a simple linear relationship that we can use between the market risk price and the expected return. However, volatility risk introduces nonlinear nuisance terms into this regression. They do not matter asymptotically, but they do in finite samples.

Since this paper’s contributions are in econometric methodology and empirical results, we take a model from the literature that has the various components, instead of developing our own. In particular, we take the model from Han, Khrapov, and Renault (2018) and use it to estimate the relevant parameters. This model has a few nice features. First, it has both market and volatility prices and a leverage effect. As such, it is the natural discrete-time analog of the Heston (1993) option pricing model. It has an exponentially affine stochastic discount factor and shares with Heston (1993) the advantage of having a structure-preserving change of measure between the physical and risk-neutral models. Doing our analysis in discrete-time has two benefits. First, we can more directly compare our results to risk-premia estimates outside of the option pricing literature. Second, the high-frequency jumps should not dramatically affect our results. If we were to use a diffusion process in continuous time, we would severely counterfactually constrain the price’s higher-order moments and likely bias our inference.

As far as estimation is concerned, we derive a series of conditional means and variances. We then take these means and variances and plug them into a minimum-distance criterion. The data we use are $\{r_{t+1}, \sigma_{t+1}^2\}$, where r_{t+1} denotes the daily excess market return on the market and σ_{t+1}^2 denotes the associated realized volatility

2 Literature Review

Our paper lies at the intersection of three strands of the literature. The first is the empirical analysis of the effect of volatility on risk premia. As Lettau and Ludvigson (2010, 620) mention, the evidence here is inconclusive. Papers such as Bollerslev, Engle, and Wooldridge (1988), Harvey (1989), Ghysels, Santa-Clara, and Valkanov (2005), Bali and Peng (2006), and Ludvigson and Ng (2007) find a positive relationship, while Campbell (1987), Breen, Glosten, and Jagannathan (1989), Pagan and Hong (1991), Whitelaw (1994), and Brandt and Kang (2004) find a negative relationship. We contribute this literature by providing valid confidence intervals for the price of market risk, and explicitly accounting for the price of volatility risk. The previous specifications did not account

for this, and so although their estimates were consistent, they suffered from a finite-sample bias.

The second literature argues that we need two factors to explain risk premia dynamics: a market risk factor and a variance risk factor. Most of this literature has been in the options-pricing literature, (Christoffersen, Heston, and Jacobs 2013; Feunou et al. 2014; Dew-Becker et al. 2017). However, there is a sizeable strand of the literature documenting that the difference between the VIX and realized volatility, i.e., the variance risk premium, (Bollerslev, Law, and Tauchen 2008; Drechsler and Yaron 2011), commands a substantial positive premium. Our paper adds to this literature by being the first to provide valid confidence sets for both prices of both market and volatility risk.

The third literature is the weak identification literature, such as Andrews and Mikusheva (2016) and Andrews and Cheng (2012), by adapting the conditional quasi-likelihood test that Andrews and Mikusheva (2016) develop in a generalized method of moments (GMM) framework to a minimum distance one. We also show how to conduct valid inference when the parameter that controls the strength of identification lies on the boundary.

3 The Model

We need a model that relates the prices to the dynamics in the data to estimate the prices. We specify this model using a stochastic discount factor (SDF) and the physical measure. Since we do not have options data, the SDF, which is also called a pricing kernel, is unobserved. Hence, our pricing model must allow us to identify the parameters governing the SDF if we know the parameters that govern the physical measure.

Let \mathcal{F}_t be the representative investor's information set at time t and P_t be the price on the asset in question. Denote the associated return r_{t+1} and volatility σ_{t+1}^2 . Then the SDF — $M_{t,t+1}$ — prices all of the assets. In other words, [definition 1](#) holds. This definition says that the SDF is the change of measure between the physical and risk-neutral distributions and that prices are expectations with respect to the risk-neutral measure.

Definition 1. Asset Pricing Equation

$$P_t = \mathbb{E}[M_{t,t+1}P_{t+1} | \mathcal{F}_t] \tag{1}$$

Risk prices measure the level of compensation that the representative investor demands to face additional risk. Consequently, they determine the amount that the stochastic discount factor (SDF) twists the distribution of P_t . Our model for P_t that implies $M_{t,t+1}$ is a known function of these risk prices, parameters that govern the price and volatility dynamics, the volatility — $\sigma_t^2, \sigma_{t+1}^2$, and the log excess returns — r_{t+1} . Having done that, we place this $M_{t,t+1}$ into [definition 1](#) and estimate the model by matching the reduced-form and model-implied dynamics of the returns and volatilities.

Throughout we assume that the two risks that command nonzero prices are the market risk price and the volatility risk price as discussed in the introduction. Consequently, we only use variation in

the first two moments of the data to estimate these parameters. If higher moments, such as skewness and kurtosis are also priced factors, as in Harvey and Akhtar (2000), Conrad, Dittmar, and Ghysels (2012), and Chang, Christoffersen, and Jacobs (2013), and we used higher sample moments to determine the price of our risk-factors our resulting estimates would be biased, likely substantially so. Attempting to match the risk prices of high moments with only the risk prices considered here would misattribute compensation for bearing risk associated with these higher moments to the market and volatility risk prices. Consequently, our model must be parameterized in terms of the first two moments. We use the conditional autoregressive CAR(1) model here, which we take from Serge, Christian, and Joann (2006) and Han, Khrapov, and Renault (2018).

In particular, we adopt the model that Han, Khrapov, and Renault (2018) develop. This model assumes that the P_t and σ_t^2 are first-order Markov and there is no Granger-causality from the return to the volatility and that returns are independent given the volatility. In other words, the volatility drives all of the dynamics of the process. We do allow σ_{t+1}^2 and r_{t+1} to be contemporaneously correlated, as they are in the data. We construct the model in terms of a series of Laplace transforms that we parameterize using some functions $A(x)$, $B(x)$, $C(x)$, $D(x)$, and $E(x)$ for all x in its domain. We collect the parameters in a vector ω .

3.2 The Stochastic Discount Factor

We start by parameterizing the stochastic discount factor. Let π be the price of volatility risk and θ be the price of market risk.

Definition 2. The Stochastic Discount Factor

$$M_{t,t+1}(\pi, \theta) = \exp \left(m_0(\pi, \theta) + m_1(\pi, \theta) \sigma_t^2 - \pi \sigma_{t+1}^2 - \theta r_{t+1} \right) \quad (2)$$

Since $M_{t,t+1}$ must integrate to 1 for all σ_t^2 , $m_0(\pi, \theta)$ and $m_1(\pi, \theta)$ are integration constants. This will be crucial to our derivation of the relationship between the SDF and physical measure functions below.

3.3 Parameterizing the Physical Measure Dynamics

We now introduce the data generating process for the volatility. We use a conditional autoregressive gamma process as in Christian and Joann (2006) and Han, Khrapov, and Renault (2018) for the volatility. This implies A and B are as follows.

Definition 3. Volatility Dynamics Functions

$$A(x) = \frac{\rho x}{1 + cx}, \quad (3)$$

$$B(x) = \delta \log(1 + cx), \quad (4)$$

with $\rho \in [0, 1)$, $c > 0$, $\delta > 0$.

In this specification, ρ is a persistence parameter, and c is a scaling parameter. To see this, consider the following moment conditions.

Remark 1 (Volatility Moment Conditions).

$$\mathbb{E} [\sigma_{t+1}^2 \mid \sigma_t^2] = \rho\sigma_t^2 + c\delta \quad (5)$$

$$\mathbb{V}\text{ar} [\sigma_{t+1}^2 \mid \sigma_t^2] = 2c\rho\sigma_t^2 + c^2\delta \quad (6)$$

Since these two moment conditions are sufficient to derive the unconditional moments, all of the parameters are identified as long as they are in the interior of their appropriately specified domains. We are using linear regression to estimate the slope and intercept parameters.

3.3.1 Return Dynamics

We now compute the return distribution's moments. This computation is more subtle than computing the moments of the volatility dynamics because we must relate the dynamics of the returns' dynamics to those of the volatility as given by the unobservable SDF. We use the conditional autoregressive CAR(1) model here, which we take from Serge, Christian, and Joann (2006) and Han, Khrapov, and Renault (2018). This model specifies the conditional Laplace transform of the return as a function of r_{t+1} given $\sigma_{t+1}^2, \sigma_t^2$. We start by specifying a parametric form for $C(x)$, $D(x)$, and $E(x)$. The parametric form we give $C(x)$ implies that returns are conditionally Gaussian.

Definition 4. Reduced-Form Functions

$$C(x) := \psi x + \frac{1 - \phi^2}{2} x^2, \quad (7)$$

$$D(x) := \beta x, \quad (8)$$

$$E(x) := \gamma x. \quad (9)$$

As these functions parameterize the Laplace transform, they equal zero when $x = 0$. It might seem as first that imposing linearity for $D(x)$ and $E(x)$ is a strong restriction. This is not actually the case. Since we assumed that σ_{t+1}^2 is an integrated volatility and hence greater than $\mathbb{V}\text{ar}(r_{t+1} \mid \sigma_{t+1}^2, \sigma_t^2)$ and variances are positive, the model implies they are linear.¹

To ease understanding of how r_{t+1} and σ_t^2 are related, consider the first two moments:

$$\mathbb{E} [r_{t+1} \mid \sigma_t^2, \sigma_{t+1}^2] = \psi\sigma_{t+1}^2 + \beta\sigma_t^2 + \gamma \quad (10)$$

and

$$\mathbb{V}\text{ar} [r_{t+1} \mid \sigma_t^2, \sigma_{t+1}^2] = (1 - \phi^2)\sigma_{t+1}^2. \quad (11)$$

These moments identify ψ , β , and ϕ . The ψ parameter governs the change in the conditional mean when we condition on the volatility. The ϕ parameter governs the reduction of the returns'

1. We show this in [Lemma 7](#) which is located in [Section A](#).

innovation variance. If $\phi = 0$, the variance does not decrease, and there is no leverage effect.

We derive the following by plugging in the Laplace transforms for σ_{t+1}^2 and r_{t+1} into [definition 2](#).

Lemma 1 (Characterizing the SDF Integration Constants). *Let the SDF be given as in [definition 2](#), and the model be parameterized as in [definition 3](#) and [definition 4](#). Then the SDF constants follow the following equations.²*

$$\begin{aligned} m_0(\theta, \pi) &= E(\theta) + B(C(\theta) + \pi) \\ m_1(\theta, \pi) &= D(\theta) + A(C(\theta) + \pi) \end{aligned} \tag{12}$$

3.3.2 Deriving the Conditional Mean

This section solves for γ , β , and ψ in terms of $\{\rho, c, \delta\}$ and $\{\pi, \theta\}$. We then use the difference in expected returns between the distribution given σ_t^2 and the distribution given σ_{t+1}^2 and σ_t^2 to separately identify the two risk prices. We focus on ψ because it controls this difference in means and hence identifies of the volatility risk price when $\phi \neq 0$.

As alluded to in the introduction, ψ combines three different effects. First, it has a Jensen's inequality term — the mean will shift by a value proportional to the variance. Second, the reduction in variance shifts the price because investors are risk-averse. Third, since σ_{t+1}^2 and r_{t+1} are correlated, the drift will change directly — the leverage effect. The goal is to separate the three components. The way we do this is by constructing a pseudo-return — \tilde{r}_{t+1} — whose mean is not affected by the change in variance resulting from the addition of σ_{t+1}^2 to the information set.

We exploit the fact that $\mathbb{E}[M_{t,t+1}(\theta, \pi) \exp(r_{t+1}) | \sigma_t^2] = 1$ by relating $\mathbb{E}[\exp(r_{t+1}) | \sigma_{t+1}^2, \sigma_t^2]$ to this expression. Clearly, there are two differences between these expressions. First, the first expression contains the SDF. Second, the conditioning information differs between them. These differences are two measure changes. Since prices here are conditionally log-Gaussian, the measure change can be parameterized in terms of the covariance between the log SDF — $m_{t,t+1} := \log M_{t,t+1}$ — and the return.

Applying the logarithm to [definition 2](#) gives

$$m_{t,t+1}(\pi, \theta) = m_0(\pi, \theta) + m_1(\pi, \theta)\sigma_t^2 - \pi\sigma_{t+1}^2 - \theta r_{t+1}. \tag{13}$$

Except for θr_{t+1} , the terms on the right are constant given $\sigma_t^2, \sigma_{t+1}^2$, and so they do not affect the conditional covariance between $m_{t,t+1}$ and r_{t+1} . Hence,

$$\text{Cov}(m_{t,t+1}, r_{t+1} | \sigma_{t+1}^2, \sigma_t^2) = \text{Cov}(-\theta r_{t+1}, r_{t+1} | \sigma_{t+1}^2, \sigma_t^2) = -\theta \text{Var}(r_{t+1} | \sigma_{t+1}^2). \tag{14}$$

Plugging in [Equation \(11\)](#) gives

$$\text{Cov}(m_{t,t+1}, r_{t+1} | \sigma_{t+1}^2, \sigma_t^2) = -\theta(1 - \phi^2)\sigma_{t+1}^2. \tag{15}$$

2. This is Equation 3.4 in Han, Khrapov, and Renault ([2018](#), 3.4).

This expression is the change in the mean driven by investors' risk aversion, and is controlled by their market risk price. To see this, note it equals zero if $\theta = 0$.

The second term we need is the Jensen's effect term. The mean of a conditionally log-Gaussian price depends on both the mean and variance of the underlying return. It has the standard form: minus one-half the conditional variance $\left(-\frac{1-\phi^2}{2}\sigma_{t+1}^2\right)$.

We now define a pseudo-return that only changes by an amount proportional to the leverage effect and is not affected by the changes in the return arising from the reduction in risk. From [Equation \(10\)](#), we know that $\mathbb{E}[r_{t+1} | \sigma_t^2, \sigma_{t+1}^2] = \psi\sigma_{t+1}^2$ plus some function of σ_t^2 . However, this ψ contains all three effects. We want a pseudo-return \tilde{r}_{t+1} where $\mathbb{E}[\tilde{r}_{t+1} | \sigma_t^2, \sigma_{t+1}^2] = \tilde{\psi}\sigma_{t+1}^2$ for some $\tilde{\psi}$ that only contains the direct effect, not the Jensen or any risk-compensation. We can then solve for $\tilde{\psi}$ in terms of the parameters governing the dynamics of r_{t+1} .

Lemma 2 (Separating the Leverage Effect from the Measure Changes). *Let $\tilde{r}_{t+1} := r_{t+1} - \frac{1-\phi^2}{2}\sigma_{t+1}^2 + (1-\phi^2)\theta\sigma_{t+1}^2$, and let the m_{t+1} have the form given by [Equation \(13\)](#). Then*

$$\text{Var}[\tilde{r}_{t+1} | \sigma_t^2] = \mathbb{E}[\sigma_{t+1}^2 | \sigma_t^2]. \quad (16)$$

This is not the only way we can compute $\text{Var}[\tilde{r}_{t+1} | \sigma_t^2]$. We can also use the law of total variance:

$$\text{Var}[\tilde{r}_{t+1} | \sigma_t^2] = \mathbb{E}[\text{Var}[\tilde{r}_{t+1} | \sigma_{t+1}^2] | \sigma_t^2] + \text{Var}[\mathbb{E}[\tilde{r}_{t+1} | \sigma_{t+1}^2] | \sigma_t^2]. \quad (17)$$

Since σ_{t+1}^2 is constant given $\sigma_{t+1}^2, \sigma_t^2$, the conditional variance of \tilde{r}_{t+1} is the same as r_{t+1} . In addition, $\mathbb{E}[\tilde{r}_{t+1} | \sigma_{t+1}^2] = \psi\sigma_{t+1}^2 - \frac{1-\phi^2}{2}\sigma_{t+1}^2 + (1-\phi^2)\theta\sigma_{t+1}^2$ plus σ_t^2 -measurable terms. We now fill in the values for both of the inside variables on the right-hand side using the formulas for r_{t+1} , and we use [Lemma 2](#) on the left-hand side. We then take unconditional expectations of both sides, giving us

$$\mathbb{E}[\sigma_{t+1}^2] = \mathbb{E}[\mathbb{E}[(1-\phi^2)\sigma_{t+1}^2]] + \mathbb{E}\left[\text{Var}\left[\left(\psi - \frac{1-\phi^2}{2} + (1-\phi^2)\theta\right)\sigma_{t+1}^2 \mid \sigma_t^2\right]\right]. \quad (18)$$

Rearranging terms implies

$$\left(\psi - \frac{1-\phi^2}{2} + (1-\phi^2)\theta\right) = \phi\sqrt{\frac{\mathbb{E}[\sigma_{t+1}^2]}{\mathbb{E}[\text{Var}(\sigma_{t+1}^2 | \sigma_t^2)]}}. \quad (19)$$

We can now solve for ψ and replace the moments on the right hand side by substituting in the expectations from [Remark 1](#) and some elementary calculations, which gives

$$\psi = \frac{\phi}{\sqrt{c(1+\rho)}} + \frac{1-\phi^2}{2} - (1-\phi^2)\theta. \quad (20)$$

Now that we have a formula for ψ , we can derive the formulas for β and γ in terms of the other reduced-form parameters and the risk prices.

Lemma 3 (Reparamaterizing the Physical Distribution). *Let the SDF be given as in [definition 2](#), and the model be parameterized as in [definition 3](#) and [definition 4](#). Then the following holds*

$$\psi(\pi, \theta) = \frac{\phi}{\sqrt{c(1+\rho)}} + \frac{1-\phi^2}{2} - (1-\phi^2)\theta. \quad (21)$$

Proof. This is a straightforward implication of [Lemma 2](#) and the derivation in the text. \square

3.4 Identification of the Risk Prices

The goal is to identify the risk-prices θ and π . In the previous sections, we derived a series of moment conditions in terms of the parameters. We now need to analyze when these moment conditions identify the risk prices. The asset pricing equation entirely encapsulates the information that return data contain about pricing. The definition of $M_{t,t+1}$ as a change of measure means that the following holds:

$$\mathbb{E}[M_{t,t+1}(\theta, \pi) \exp(r_{t+1}) | \mathcal{F}_t] = 1. \quad (22)$$

We now characterize, the information in this set of moment conditions regarding the risk prices. The difficult part is identifying the volatility risk price — π , which is the primary object of interest. Substituting in the SDF formula from above and replacing all of the dependence on \mathcal{F}_t with σ_t^2 , yields

$$\mathbb{E}[\exp(-\pi\sigma_{t+1}^2 - (\theta - 1)rx_{t+1}) | \sigma_t^2] = \exp(-m_0(\theta, \pi) - m_1(\theta, \pi)\sigma_t^2). \quad (23)$$

Similar to above, we use the law of iterated expectations to substitute in the conditional Laplace transforms of r_{t+1} and σ_{t+1}^2 , obtaining

$$\mathbb{E}[\exp(-A(\pi + C(\theta - 1))\sigma_t^2 - B(\pi) - D(\theta - 1)\sigma_t^2 - E(\theta - 1)) | \sigma_t^2] = \exp(-m_0(\theta, \pi) - m_1(\theta, \pi)\sigma_t^2). \quad (24)$$

Lemma 4 (Reparamterizing β and γ). *Let the SDF be given as in [definition 2](#), and the model be paramterized as in [definition 3](#) and [definition 4](#) and assume that the risk-neutral measures have distributions with the same parametric forms.*

$$\gamma = B(\pi + C(\theta - 1)) - B(\pi + C(\theta)) \quad (25)$$

$$\beta = A(\pi + C(\theta - 1)) - A(\pi + C(\theta)) \quad (26)$$

Proof. This comes from matching the coefficients of the \mathcal{F}_t -measurable variables and using the linearity of $D(x)$ and $C(x)$ from [Lemma 7](#). We substitute in the formulas for $m_0(\theta, \pi)$ and $m_1(\theta, \pi)$ that we derived in [Lemma 1](#) which gives us to equations that we solve for γ and β . \square

We can characterize the identification restrictions in [Lemma 4](#) in two different cases.³

Case 1: The price of market risk θ satisfies the following equations.

$$C(\theta - 1) = C(\theta) \quad (27)$$

3. These equations are Equation 3.7 in Han, Khrapov, and Renault ([2018](#)).

If Equation (27) holds, then some simple algebra shows that $E(\theta) = E(\theta - 1)$ and $D(\theta) = D(\theta - 1)$. In this situation, any value of π satisfies Lemma 4. Since these are the only places π appears in the model, π is not identified from equity data alone. As noted by Han, Khrapov, and Renault (2018), this is in line with the common belief that the econometrician needs options data to be able to identify the price of volatility risk.

Case 2: In general, there is no reason to expect the Equation (27) to hold. If it does not, it may be possible to identify both θ and π . In this case, π does affect the affect risk premia in the equity data. Hence, we can, in principle, identify π from the difference between the functions in the previous case when evaluated at $\theta - 1$ and θ .

We now show that if $\phi = 0$, then Equation (27) holds, but it does not hold if $\phi \neq 0$. In other words, as mentioned in Han, Khrapov, and Renault (2018, 13), a leverage effect will allow us to separately identify θ and π . To see this, we apply the definition of $C(x)$ from definition 4 and the formula for $\psi(\pi, \theta)$ from Lemma 3. Some simple algebra gives

$$C(\theta) - C(\theta - 1) = C'(0) + C''(0) \left(\theta - \frac{1}{2} \right) = \psi + (1 - \phi^2) \left(\theta - \frac{1}{2} \right), \quad (28)$$

$$= \frac{\phi}{\sqrt{c(1+\rho)}} + \frac{1 - \phi^2}{2} - (1 - \phi^2)\theta + (1 - \phi^2) \left(\theta - \frac{1}{2} \right) = \frac{\phi}{\sqrt{c(1+\rho)}}. \quad (29)$$

This shows the leverage effect creates a wedge between $C(\theta)$ and $C(\theta - 1)$ that is proportional to $\frac{\phi}{\sqrt{c(1+\rho)}}$. Clearly, this wedge equals zero if and only if $\phi = 0$. Consequently, if $\phi \neq 0$, we can separately identify π and θ .

4 The Estimation Criterion in Population

We use the moment conditions derive above to estimate the model. To complete the model, we derive moments that identify ϕ and ψ . They are straightforward implications of the discrete-time moment conditions governing the volatility and return dynamics.

4.1 Identified Set

In the previous sections, we have eliminated ψ , β , and γ from the model by solving for them in terms of the other parameters. We now collect all of the moment conditions.

Definition 5. Equilibrium Moment Conditions

$$\mathbb{E} [\sigma_{t+1}^2 \mid \sigma_t^2] = \rho \sigma_t^2 + c\delta \quad (30)$$

$$\mathbb{V}\text{ar} [\sigma_{t+1}^2 \mid \sigma_t^2] = 2c\rho\sigma_t^2 + c^2\delta \quad (31)$$

$$\mathbb{E} [r_{t+1} \mid \sigma_t^2, \sigma_{t+1}^2] = \gamma(\pi, \theta, \phi) + \beta(\pi, \theta, \phi)\sigma_t^2 + \psi(\pi, \theta, \phi)\sigma_{t+1}^2 \quad (32)$$

$$\mathbb{V}\text{ar} [r_{t+1} \mid \sigma_t^2, \sigma_{t+1}^2] = (1 - \phi^2)\sigma_{t+1}^2 \quad (33)$$

Now that we have a series of moment conditions, we can specify when the parameters are identified.

Lemma 5 (Identified Set). *Assume that the moment conditions specified in [definition 5](#) have the correct form and that the instruments we are using satisfy the standard exogeneity and relevant conditions. Let the true parameter vector $\omega := (\rho, c, \delta, \phi, \theta, \pi) \in [-1 + \epsilon_1, 1 - \epsilon_2] \times [M_1, M_2] \times [\epsilon_4, M_4] \times [M_5, M_6] \times [-1 + \epsilon_4, 1 - \epsilon_5] \times [M_7, M_8] \times [M_9, M_{10}]$, where the M_* are some large (in magnitude) known constants and the ϵ_* are some small positive constants. Let $Q_T(\omega, X)$ be the GMM objective function with moment conditions given in [definition 5](#).*

If there exists a ϵ such that $|\phi| > \epsilon > 0$, then all of the parameters are identified. If $\phi = 0$, the objective function is independent of π . Hence, π is not identified. However, even in that case all of remaining parameters are still identified.

We do need to run a constrained optimization though, because only certain values of the parameters are valid. As noted in the introduction, we develop a minimum distance criterion. Such a criterion has two parts. First, it requires a series of conditions that relate the structural parameters to the reduced-form parameters. Second, it needs a series of conditions that identifies the reduced-form parameters. Before we specify these criteria, we introduce a reduced-form parameter $\zeta = 1 - \phi^2$, which is the transformation of ϕ which we can easily estimate.

We group the structural parameters we are estimating into a vector $\xi := (\phi, \pi, \theta)'$. We also group the reduced-form parameters into a vector $\omega := (\rho, c, \delta, \psi, \zeta, \beta, \gamma)'$. We define the link function as follows, where we suppress the reduced-form parameters in the notation.

Definition 6. Link Function

$$g(\omega, \xi) := \begin{pmatrix} \gamma - \gamma(\pi, \theta, \phi) \\ \beta - \beta(\pi, \theta, \phi) \\ \psi - \psi(\pi, \theta, \phi) \\ \zeta - (1 - \phi^2) \end{pmatrix} \quad (34)$$

5 Estimating the Reduced-Form Parameters

This section gives the asymptotic distribution of the reduced-form parameter. Write $\omega = (\omega_1, \omega_2, \omega_3)$, where $\omega_1 = (\rho, c)$, $\omega_2 = (\gamma, \beta, \psi)$, and $\omega_3 = \zeta$. Below we describe the estimator $\hat{\omega} := (\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3)$ and provide its asymptotic distribution. We estimate these parameters separately because ω_1 only shows up in the conditional mean and variance of r_{t+1} , ω_2 only shows up in the conditional mean of σ_{t+1}^2 , and ϕ only shows up in the conditional variance of σ_{t+1}^2 .

We estimate ω_1 by GMM based on the moment condition:

$$\mathbb{E}[h_t(\omega_{1,0})] = 0, \text{ where}$$

$$h_t(\omega_1) := \begin{pmatrix} \sigma_{t+1}^2 - (c\delta + \rho\sigma_t^2) \\ \sigma_t^2 (\sigma_{t+1}^2 - (c\delta + \rho\sigma_t^2)) \\ \sigma_{t+1}^4 - \left(c^2\delta + 2c\rho\sigma_t^2 + \left(c\delta + \sigma_{t+1}^2 - (c\delta + \rho\sigma_t^2)^2 \right) \right) \\ \sigma_t^2 \left(\sigma_{t+1}^4 - \left(c^2\delta + 2c\rho\sigma_t^2 + \left(c\delta + \sigma_{t+1}^2 - (c\delta + \rho\sigma_t^2)^2 \right) \right) \right) \\ \sigma_t^4 \left(\sigma_{t+1}^4 - \left(c^2\delta + 2c\rho\sigma_t^2 + \left(c\delta + \sigma_{t+1}^2 - (c\delta + \rho\sigma_t^2)^2 \right) \right) \right) \end{pmatrix}. \quad (35)$$

The optimal GMM estimator is

$$\begin{aligned} \hat{\omega}_1 &= \underset{\omega_1 \in \Lambda_1}{\operatorname{argmin}} \bar{h}_T(\omega_1)' W_T \bar{h}_T(\omega_1), \text{ where} \\ \bar{h}_T(\omega_1) &= T^{-1} \sum_{t=1}^T h_t(\omega_1), \\ W_T &= T^{-1} \sum_{t=1}^T (h_t(\tilde{\omega}_1) h_t(\tilde{\omega}_1)' - \bar{h}_T(\tilde{\omega}_1) \bar{h}_T(\tilde{\omega}_1)'), \end{aligned} \quad (36)$$

and $\tilde{\omega}_1$ is the preliminary GMM estimator based on the identity covariance matrix.

We estimate ω_2 using generalized least squares (GLS) because γ, β, ψ are the intercept and linear coefficients of the conditional mean function and the conditional variance is proportional to σ_{t+1}^2 . Define $x_t = \sigma_{t+1}^{-1}(1, \sigma_t^2, \sigma_{t+1}^2)$ and $y_t = \sigma_{t+1}^{-1}r_{t+1}$. The GLS estimator for ω_2 is

$$\hat{\omega}_2 = \left(\sum_{t=1}^T x_t x_t' \right)^{-1} \sum_{t=1}^T x_t y_t. \quad (37)$$

We estimate ω_3 by the sample variance estimator. Define

$$\hat{y}_t = x_t \hat{\omega}_2 = \sigma_{t+1}^{-1}(\hat{\gamma} + \hat{\beta}\sigma_t^2 + \hat{\psi}\sigma_{t+1}^2). \quad (38)$$

The estimator of ω_3 is

$$\hat{\omega}_3 = T^{-1} \sum_{t=1}^T (y_t - \hat{y}_t)^2. \quad (39)$$

The next lemma provides the asymptotic distribution of the estimator $\hat{\omega}$. Let $h_{\omega,t}(\omega_1) \in \mathbb{R}^{5 \times 2}$ denote the derivative of $h_t(\omega_1)$ w.r.t. ω_1 . Define

$$\begin{aligned} \Sigma_1 &= \left\{ \mathbb{E} [h_{\omega,t}(\omega_{1,0})]' \mathbb{E} [h_t(\omega_{1,0}) h_t(\omega_{1,0})']^{-1} \mathbb{E} [h_{\omega,t}(\omega_{1,0})] \right\}^{-1}, \\ \Sigma_2 &= \mathbb{E} [x_t x_t']^{-1} \mathbb{E} [(y_t - x_t' \omega_{2,0})^2], \\ \Sigma_3 &= \mathbb{V}\text{ar} \left[(y_t - x_t' \omega_{2,0})^2 \right]. \end{aligned} \quad (40)$$

Lemma 6. *Suppose Assumptions *** hold. Then*

$$T^{1/2} \begin{pmatrix} \hat{\omega}_1 - \omega_{1,0} \\ \hat{\omega}_2 - \omega_{2,0} \\ \hat{\omega}_3 - \omega_{3,0} \end{pmatrix} \rightarrow_d \xi_\omega = \begin{pmatrix} \xi_{\omega 1} \\ \xi_{\omega 2} \\ \xi_{\omega 3} \end{pmatrix} \sim N \begin{pmatrix} \Sigma_1 & 0 & 0 \\ 0, & 0 & \Sigma_2 & 0 \\ 0 & 0 & \Sigma_3 \end{pmatrix}. \quad (41)$$

Proof. Will be added later. \square

5.1 Robust Inference for Risk Price

Our estimate in the previous section does not involve the structural parameters θ , π , and ϕ . We do not plug in β, γ, ψ as functions of θ and π . Instead, we treat β, γ, ψ just as linear coefficients and estimate them by GLS.

We estimate the structural parameters θ and π using $\hat{\omega}$ and the link functions specified below. First, we know that

$$\psi_0 = \phi_0 (c_0 (1 + \rho_0))^{-1/2} - (1 - \phi_0^2) / 2 - (1 - \phi_0^2) \theta_0 \quad (42)$$

when all parameters are evaluated at the true values. This equation strongly identifies θ_0 because ϕ_0 is assumed to be negative and bounded away from 1 in magnitude. It follows from (Equation (42)) that

$$\theta_0 = L(\omega_0) = -(1 - \phi_0^2)^{-1} \left[\psi_0 - \phi_0 (c_0 (1 + \rho_0))^{-1/2} - (1 - \phi_0^2) / 2 \right]. \quad (43)$$

Thus, we estimate θ_0 by

$$\hat{\theta} = L(\hat{\omega}). \quad (44)$$

By the delta method, we know that

$$T^{1/2}(\hat{\theta} - \theta_0) \rightarrow_d L_\omega(\omega_0)' v_\omega, \quad (45)$$

where $L_\omega(\omega) \in R^{d_\omega}$ denote the derivative of $L(\omega)$ w.r.t. ω . The inference for θ is standard. A confidence interval for θ can be obtained by inverting the t -statistic with a critical value obtained from the standard normal distribution.

Next, we consider inference for the structural parameter π . This is a non-standard problem because π is potentially weakly identified. Define

$$g(\pi, \omega) = \begin{pmatrix} \gamma - [B(\pi + C(\theta_L - 1)) - B(\pi + C(\theta_L))] \\ \beta - [A(\pi + C(\theta_L - 1)) - A(\pi + C(\theta_L))] \end{pmatrix}, \text{ where } \theta_L = L(\omega). \quad (46)$$

We know

$$g(\pi_0, \omega_0) = 0 \in \mathbb{R}^2. \quad (47)$$

Inference on π is based on the function $g(\pi, \hat{\omega})$ because $\hat{\omega}$ is a consistent estimator of ω_0 . By the

consistency of $\hat{\omega}$,

$$T^{1/2} [g(\pi, \hat{\omega}) - g(\pi, \omega_0)] \Rightarrow v(\pi) = G(\pi, \omega_0)' v_\omega, \quad (48)$$

where $G(\pi, \omega)$ denote the derivative of $g(\pi, \omega)$ w.r.t. to ω . The Gaussian process $v(\pi)$ has covariance kernel

$$\Sigma(\pi_1, \pi_2) = G(\pi_1, \omega_0)' V G(\pi_2, \omega_0). \quad (49)$$

We can estimate $\Sigma(\pi_1, \pi_2)$ by

$$\hat{\Sigma}(\pi_1, \pi_2) = G(\pi_1, \hat{\omega})' \hat{V} G(\pi_2, \hat{\omega}), \quad (50)$$

where \hat{V} is a consistent estimator of V .

We construct a confidence interval for π by inverting tests $H_0 : \pi = \pi_0$ vs $H_0 : \pi \neq \pi_0$. The test statistic is the QLR statistic:

$$QLR = Tg(\pi_0, \hat{\omega})' \hat{\Sigma}(\pi_0, \pi_0)^{-1} g(\pi_0, \hat{\omega}) - \min_{\pi \in \Pi} Tg(\pi, \hat{\omega})' \hat{\Sigma}(\pi, \pi)^{-1} g(\pi, \hat{\omega}). \quad (51)$$

To obtain the critical value, we follow the conditional inference approach in Andrews and Mikushcheva (2016). To this end, first construct a projection residual process:

$$h(\pi, \hat{\omega}) = g(\pi, \hat{\omega}) - \hat{\Sigma}(\pi, \pi_0) \hat{\Sigma}(\pi_0, \pi_0)^{-1} g(\pi_0, \hat{\omega}). \quad (52)$$

By construction, $h(\pi, \hat{\omega})$ and $g(\pi_0, \hat{\omega})$ are independent asymptotically. Conditional on $h(\pi, \hat{\omega})$, we obtain the $1 - \alpha$ quantile of the QLR statistic, denoted $c_\alpha(h)$, by sampling from the asymptotic distributions of $g(\pi_0, \hat{\omega})$ under the null. Specifically, we take independent draws $v^* \sim N(0, \Sigma(\pi_0, \pi_0))$ and produce simulated process:

$$g^*(\pi, \hat{\omega}) = h(\pi, \hat{\omega}) + \hat{\Sigma}(\pi, \pi_0) \hat{\Sigma}(\pi_0, \pi_0)^{-1} v^*. \quad (53)$$

We then calculate

$$QLR^* = Tg^*(\pi_0, \hat{\omega})' \hat{\Sigma}(\pi_0, \pi_0)^{-1} g^*(\pi_0, \hat{\omega}) - \min_{\pi \in \Pi} Tg^*(\pi, \hat{\omega})' \hat{\Sigma}(\pi, \pi)^{-1} g^*(\pi, \hat{\omega}), \quad (54)$$

which is a random drawn from the conditional distribution of the QLR statistic given $h_T(\pi, \hat{\omega})$, when $g(\pi_0, \hat{\omega})$ is drawn from its asymptotic distribution. In practice, we repeat this process for a large number of times and obtain $c_\alpha(h)$ by simulation.

We reject the null $H_0 : \pi = \pi_0$ if $QLR \geq c_\alpha(h)$. The confidence interval for π is the collection of null values that are not rejected as the null value. Note that the construction of this CI does not involve estimation of π .

6 Simulations

In this section, we verify that the asymptotic approximations we derived work well in practice. Since we have a parametric model, we can simulate from it directly. To investigate the robustness of the procedure with respect to various identification strengths, we vary both T and ϕ . We took the parameters in [Table 1](#) directly from the estimates in Han, Khrapov, and Renault (2018). We consider $\phi \in \{-0.40, -0.10, -0.01\}$ and $T \in \{3700, 37000\}$.

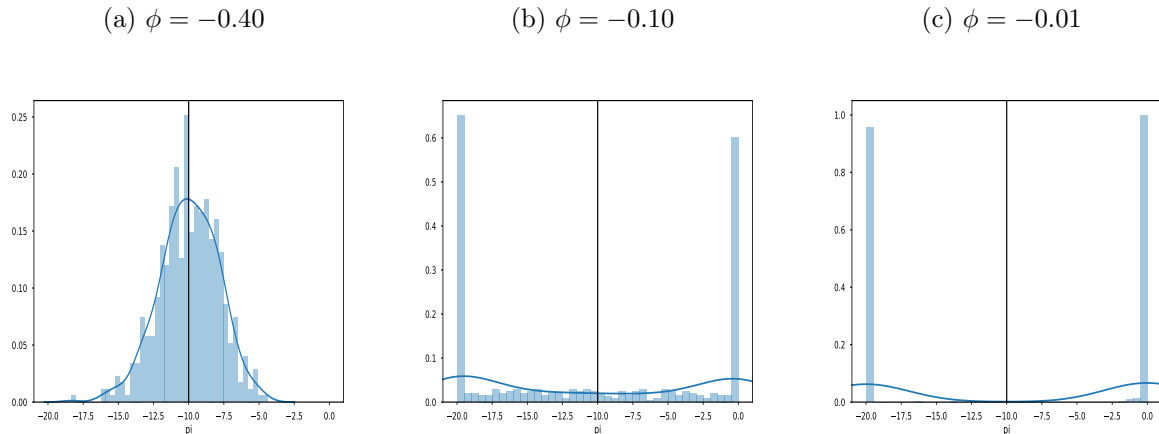
Table 1: Simulation Parameters

δ	ρ	c	π	θ
Parameter Values used by Han, Khrapov, and Renault (2018)				
0.6475	0.95	3.94128×10^{-3}	-10	1.7680

Since we are simulating the quantiles of the asymptotic distribution, we need to choose the number of simulations to use. We use 250. The results are qualitatively similar if we use 1000 draws. Instead of simulating the confidence intervals, which is computationally intensive, we simulate the test statistic, which is equivalent. We report averages over 500 simulations.

To show the identification strength varies with ϕ , we plot the distribution of estimates, $\hat{\pi}$ as reported by simulation using 37 000 observations, 500 simulations, and the parameters used in Han, Khrapov, and Renault (2018). When ϕ is far from 0 and the standard asymptotic experiments approximate the true distribution, the statistics' distribution should be close to Gaussian. Conversely, as ϕ approaches 0, the Gaussian limiting distribution does not provide a good approximation. The black lines in the middle of the figures in [Figure 1](#) are the true parameter values. We truncate the parameter space for π at $[-20, 0]$, and so the limiting distribution piles up at the boundaries as is common in this class of simulations.

Figure 1: Parameter Estimates



We now report the size of the tests in finite-sample using the standard Gaussian Quasi-Likelihood Ratio (QLR) statistic, the Anderson-Rubin (AR) statistic, and the conditional-QLR statistic pro-

posed here. Because we have four link functions, the AR-statistic is asymptotically $\chi^2(df = 4)$ distributed.

Table 2: Finite-Sample Size

ϕ	QLR	AR	CLR	QLR	AR	CLR
Parameter Values used by Han, Khrapov, and Renault (2018)						
	$T = 37000$			$T = 3700$		
-0.01						
-0.10	0.032	0.010	0.054			
-0.40	0.048	0.020	0.050			

Consider the right-hand side of Table 2. The AR-statistic is robust, but conservative, as we would expect. The standard QLR-statistic appears to be conservative. The CLR test, on the other hand, is robust and has the proper size regardless of the value of ϕ . This behavior of the CLR test is precisely what the data predicts should happen.

7 Data and Empirical Results

7.1 Data

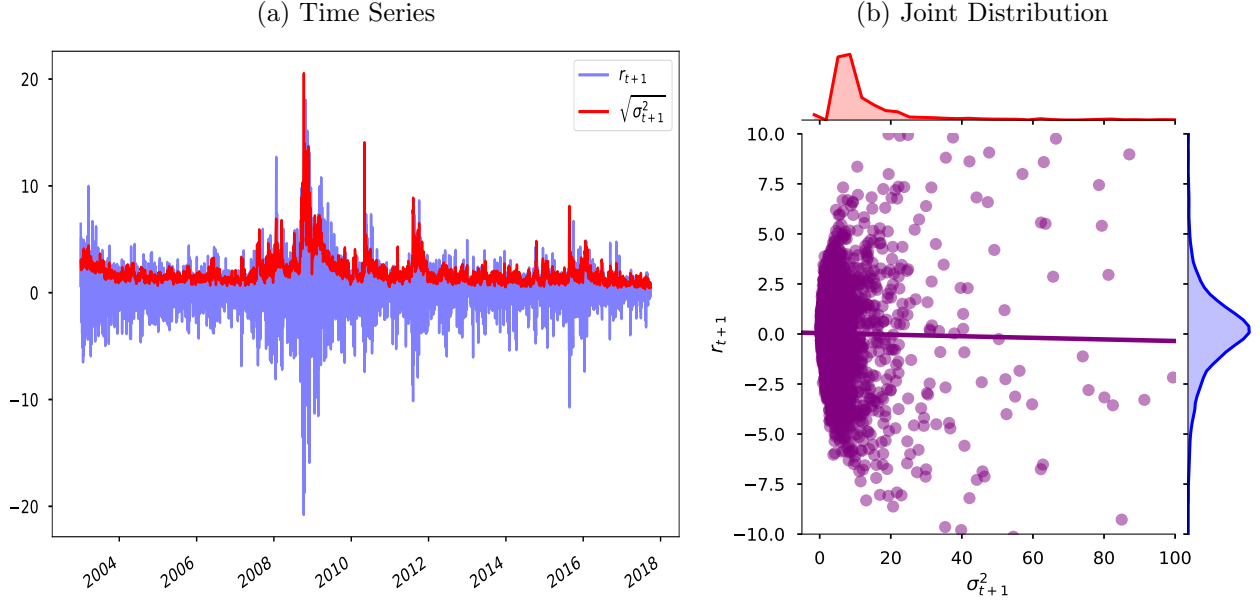
The two series we need to estimate the model are r_{t+1} and σ_{t+1}^2 , and σ_{t+1}^2 must be the volatility of r_{t+1} . That is, we need the appropriately defined variance of r_{t+1} to equal σ_{t+1}^2 in expectation as discussed in Section 3.3.2. Since this condition is automatically satisfied by the integrated volatility and daily return, we use high-frequency data to estimate σ_{t+1}^2 and use the associated daily return for r_{t+1} .

We need a market index so that the risk prices we estimate are not prices of risk investors face in practice. We do not want the asset’s risk to be easily diversifiable. We use SPY, (SPDR S&P 500 ETF Trust), which is an exchange-traded fund that mimics the S&P 500. We use the procedure Sangrey (2018) develops to estimate the integrated total volatility, which is the instantaneous expectation of the price variance, i.e., the time-derivative of the predictable quadratic variation. This measure reduces to the integrated diffusion volatility if prices have continuous paths. He shows in simulations that his method works well in the presence of market microstructure noise.

Since this paper only use one asset, and SPY is one of the most liquid assets traded, we can essentially choose the frequency at which we want to observe the underlying price. In order to balance market-microstructure noise, computational cost, and efficiency of the resultant estimators we sample at the 1 s frequency. The data starts in 2003 and ends in September 2017. Since the asset is only traded during business hours, this leads to 3713 days of data with an average of $\approx 24\,000$ observations per day. We compute r_{t+1} as the daily return from the open to close of the market because this is the interval over which we can estimate the volatility. Doing this avoids needing to specify 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). This procedure is known not to affect the consistency of the estimation procedure. The basic idea is rather simple. We average the price over a small interval to remove the noise. If we pick the rates at which we shrink the interval appropriately and thereby balance smoothing away the noise and allowing the volatility to vary intraday, the estimators are consistent.

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 are related. The volatility has a long-right tail (it has the typical gamma-type distribution) and the returns have a bell-shaped distribution. They are also slightly negative related as can be seen by the regression line in the joint plot.

We also report a series of summary statistics. The volatility and returns are weakly negatively correlated, corroborating the work by Bandi and Renò (2012) and Aït-Sahalia, Fan, and Li (2013).

Table 3: Summary Statistics

	r_{t+1}	σ^2_{t+1}
Mean	0.02	5.62
Standard Deviation	2.35	14.46
Skewness	-0.31	12.21
Kurtosis	10.06	240.40
Correlation	-0.02	

7.2 Estimates

We now report the reduced-form parameters estimates and associated confidence intervals. We only report the reduced-form parameters that true determinants of the model, i.e., they are necessary to simulate the model. The other reduced-form parameters are known functions of these parameters and the structural parameters. They have no economic meaning and are only useful in the estimation procedure. The confidence intervals reported here use the Gaussian limiting theory, (± 1.96 standard errors) but we truncate at the edge of the parameter space if necessary. Although, these reduced-form parameters can be estimated directly from the data, disentangling c and δ is difficult, and so the estimation procedure responds appropriately by reporting relatively large standard errors.

Table 4: Reduced-Form Parameter Estimates

	Point Estimate	Confidence Interval
c	3.11	(0.00, 7.38)
δ	0.27	(0.00, 3.01)
ρ	0.81	(0.40, 1.00)

We are interested in not just the reduced-form parameters. We also want estimates for the ϕ and the risk prices. Since $\phi < 0$, our estimator is consistent and the point estimate is a meaningful number. We report the 95 % confidence intervals by taking computing the joint confidence interval above, and then projecting down to the each of the components.

Table 5: Structural Parameter Estimates

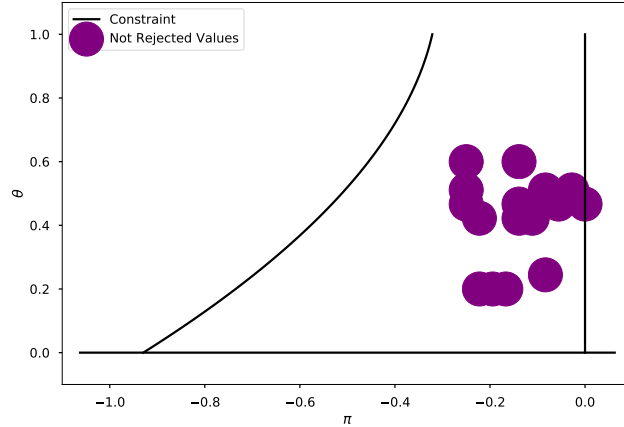
	Point Estimate	Confidence Interval
ϕ	-0.30	(0.30, -0.20)
π	-0.12	(-0.25, 0.00)
θ	0.52	(0.20, 0.60)

The parameters in [Table 5](#) have a few notable features. First, we can reject $\phi = 0$ at the 95 % level. We cannot, however, reject that $\pi = 0$. Interestingly, the values for ϕ, π , and θ reported by Han, Khrapov, and Renault (2018) lie outside our confidence regions for each of the parameters. They calibrate $\pi = -10$ and then estimate the remaining parameters. Since their calibration of π is badly misspecified, their resulting estimates for θ is biased in finite sample. Conversely, our estimates do not lie outside their confidence intervals. Even though our estimation procedure for the parameters individually is conservative, as induced by the projection above, our estimates are still more precise in practice. This difference is likely a result of our estimation procedure more effectively using the information about the various parameters jointly, which their partially calibrated model is incapable of doing.

The main issue with the estimates in [Table 5](#) is that they are conservative because we go from joint inference to sub-vector inference by using an inherently conservative projection based

procedure. Plots of confidence region in three dimensions are not particularly easy to understand, and so we report a joint confidence region for the two prices. This is still conservative as we project out ϕ , but not as conservative as the univariate analysis.

Figure 3: Confidence Set for Risk Prices



The confidence region is approximately rectangular, and so the projection inference we reported in [Table 5](#) is likely not particularly conservative. Furthermore, since we plotted the constraint on the potential value of the parameters, we can see that except for $\pi < 0$, the other constraints do not bind in practice. It is worth noting that there no reason to assume that confidence region has to be connected. The joint likelihood appears to be multi-modal but since the asymptotic distribution is not Gaussian, this not surprising.

8 Conclusion

In structural stochastic volatility models such as the one developed here, changes in the volatility affect returns through two channels. First, the investor’s willingness to tolerate high volatility in order to get high expected returns as measured by the price of market risk. Standard economic models imply there is a static trade-off between risk and expected return. Consequently, when the risk changes the expected return must change to restore equilibrium. However, investors may also be directly averse to changes in future volatility, and the empirical evidence shows that they are. Han, Khrapov, and Renault (2018) shows that we can disentangle these two channels through their differing relationships to the leverage effect. However, estimating the leverage effect is quite tricky, (Aït-Sahalia, Fan, and Li 2013). This difficulty poses a subtle identification that invalidates standard inference. When the data’s signal-to-noise ratio is small, as it is here, standard tests and confidence intervals provide misleading results. We adopt the discrete-time exponentially-affine of Han, Khrapov, and Renault (2018) and adapt weak identification methods to this framework to ensure that the resulting confidence intervals are uniformly valid.

In particular, we develop a minimum distance criterion that links the market risk price, the volatility risk price, and the leverage effect to a series of well-behaved reduced-form parameters

which govern the return and volatility’s joint distribution. We do this by adapting the conditional quasi-likelihood ratio test (CLR) Andrews and Mikusheva (2016) develop in a GMM framework to a minimum distance framework. The resulting CLR test is uniformly valid. We invert this test to derive a robust confidence set. We then apply this methodology to data on the S&P 500 and show that the market risk price lies in $(0.20, 0.60)$ yearly percentage points, and the volatility risk price lies in $(-0.20, 0.00)$ yearly percentage points. These estimates are both substantially smaller in magnitude than the values chosen by Han, Khrapov, and Renault (2018).

References

- Aït-Sahalia, Yacine, Jianqing Fan, and Yingying Li. 2013. “The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency.” *Journal of Financial Economics* 109 (1): 224–249.
- Aït-Sahalia, Yacine, Jean Jacod, and Jia Li. 2012. “Testing for Jumps in Noisy High Frequency Data.” *Journal of Econometrics* 168:207–222.
- Andrews, Donald W. K., and Xu Cheng. 2012. “Estimation and Inference With Weak, Semi-Strong, and Strong Identification.” *Econometrica* 80 (5): 2153–2211.
- Andrews, Isaiah, and Anna Mikusheva. 2016. “Conditional Inference With a Functional Nuisance Parameter.” *Econometrica* 84 (4): 1571–1612.
- Bali, Turan G., and Lin Peng. 2006. “Is there a risk–return trade-off? Evidence from high-frequency data.” *Journal of Applied Econometrics* 21 (8): 1169–1198.
- Bandi, Federico M., and Roberto Renò. 2012. “Time-varying Leverage Effects.” *Journal of Econometrics* 169 (1): 94–113.
- Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron. 2014. “Volatility, the Macroeconomy, and Asset Prices.” *The Journal of Finance* 69 (6): 2471–2511.
- Bates, David S. 2008. “The Market for Crash Risk.” *Journal of Economic Dynamics and Control* 32 (7): 2291–2321.
- Bollerslev, Tim, Robert F. Engle, and Jeffrey M. Wooldridge. 1988. *Journal of Political Economy* 96 (1): 116–131.
- Bollerslev, Tim, Tzuo Hann Law, and George Tauchen. 2008. “Risk, Jumps and Diversification.” *Journal of Econometrics* 144:234–256.
- Brandt, Michael W., and Qiang Kang. 2004. “On the Relationship Between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach.” *Journal of Financial Economics* 72 (2): 217–257.
- Breen, William, Lawrence R. Glosten, and Ravi Jagannathan. 1989. “Economic Significance of Predictable Variations in Stock Index Returns.” *The Journal of Finance* 44 (5): 1177–1189.
- Campbell, John Y. 1987. “Stock Returns and the Term Structure.” *Journal of Financial Economics* 18 (2): 373–399.
- Chang, Bo Young, Peter Christoffersen, and Kris Jacobs. 2013. “Market Skewness Risk and the Cross Section of Stock Returns.” *Journal of Financial Economics* 107 (1): 46–68.
- Christian, Gouriéroux, and Jasiak Joann. 2006. “Autoregressive Gamma Processes.” *Journal of Forecasting* 25 (2): 129–152.

- Christoffersen, Peter, Steven Heston, and Kris Jacobs. 2013. "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel." *The Review of Financial Studies* 26 (8): 1963–2006.
- Conrad, Jennifer, Robert F. Dittmar, and Eric Ghysels. 2012. "Ex-Ante Skewness and Expected Stock Returns." *The Journal of Finance* 68 (1): 85–124.
- Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez. 2017. "The Price of Variance Risk." *Journal of Financial Economics* 123 (2): 225–250.
- Drechsler, Itamar, and Amir Yaron. 2011. "What's Vol Got to Do with It." *The Review of Financial Studies* 24 (1): 1–45.
- Feunou, Bruno, Jean-Sébastien Fontaine, Abderrahim Taamouti, and Roméo Tédongap. 2014. "Risk Premium, Variance Premium, and the Maturity Structure of Uncertainty*." *Review of Finance* 18 (1): 219–269.
- Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov. 2005. "There is a Risk-Return Trade-off after All." *Journal of Financial Economics* 76 (3): 509–548.
- Han, Hyojin, Stanislav Khrapov, and Eric Renault. 2018. *The Leverage Effect Puzzle Revisited: Identification in Discrete Time*. Working Paper. Brown University.
- Hansen, Lars Peter, and Thomas J. Sargent. 2001. "Robust Control and Model Uncertainty." *The American Economic Review* 91 (2): 60–66.
- Harvey, Campbell R. 1989. "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *Journal of Financial Economics* 24 (2): 289–317.
- Harvey, Campbell R., and Siddique Akhtar. 2000. "Conditional Skewness in Asset Pricing Tests." *The Journal of Finance* 55 (3): 1263–1295.
- Heston, Steven L. 1993. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *The Review of Financial Studies* 6 (2): 327–343.
- Ju, Nengjiu, and Jianjun Miao. 2012. "Ambiguity, Learning, and Asset Returns." *Econometrica* 80 (2): 559–591.
- Lettau, Martin, and Sydney C. Ludvigson. 2010. "Measuring and Modeling Variation in the Risk-Return Tradeoff." Chap. 11, edited by Yacine Aït-Sahalia and Lars Peter Hansen, 1:617–690. *Handbook of Financial Econometrics*. Elsevier.
- Lintner, John. 1965. "Security Prices, Risk, and Maximal Gains From Diversification." *The Journal of Finance* 20 (4): 587–615.
- Ludvigson, Sydney C., and Serena Ng. 2007. "The Empirical Risk-Return Relation: A Factor Analysis Approach." *Journal of Financial Economics* 83 (1): 171–222.

- Pagan, Adrian R., and Y.S. Hong. 1991. “Nonparametric Estimation and the Risk Premium.” Chap. 2 in *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, edited by James Barnett William .A. Powell and George Tauchen, 51–75. Cambridge University Press.
- Podolskij, Mark, and Mathias Vetter. 2009. “Bipower-type Estimation in a Noisy Diffusion Setting.” *Stochastic Processes and Their Applications* 119 (9): 2803–2831.
- Sangrey, Paul. 2018. *Jumps, Realized Densities, and News Premia*. Working Paper. University of Pennsylvania.
- Serge, Darolles, Gouriéroux Christian, and Jasiak Joann. 2006. “Structural Laplace Transform and Compound Autoregressive Models.” *Journal of Time Series Analysis* 27 (4): 477–503.
- Sharpe, William F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk.” *The Journal of Finance* 19 (3): 425–442.
- Whitelaw, Robert F. 1994. “Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns.” *The Journal of Finance* 49 (2): 515–541.

Appendix A Model Characterization

Lemma 7 (Linearity of β and γ). *Letting σ_{t+1}^2 be the integrated volatility of a process with return r_{t+1} . Assume that σ_{t+1}^2 and r_{t+1} follow a bivariate CAR(1) process parametrized as in [definition 3](#) and [definition 4](#). Then $\beta''(0)$ and $\gamma''(0)$ both equal zero.*

Proof. By the Itô Isometry, and the definition of r_{t+1} as an integrated variance, the following holds for the returns' predictable information set \mathcal{F}_{t-}^r .

$$\mathbb{V}\text{ar}(r_{t+1} | \mathcal{F}_{t-}^r) \leq \mathbb{E}(r_{t+1}^2 | \mathcal{F}_{t-}^r) = \mathbb{E}(\sigma_{t+1}^2 | \mathcal{F}_{t-}^r) \quad (55)$$

The integrated volatility is predictable, and so σ_{t+1}^2 is contained in the return's predictable σ -algebra.

Consequently, $\mathbb{V}\text{ar}(r_{t+1} | \sigma_t^2, \sigma_{t+1}^2) \leq \sigma_{t+1}^2$. In addition, variance must always be positive, and so $\mathbb{V}\text{ar}(r_{t+1} | \sigma_t^2, \sigma_{t+1}^2) \geq 0$.

Since the second derivative of the log-cumulant function evaluated at zero equals the variance, we have the following set of inequalities.

$$0 \leq (1 - \phi^2)\sigma_{t+1}^2 - \beta''(0)\sigma_t^2 - \gamma''(0) \leq \sigma_{t+1}^2 \quad (56)$$

Dividing through by σ_{t+1}^2 and pulling the first term outside

$$\implies \phi^2 - 1 \leq -\frac{1}{\sigma_{t+1}^2} (\beta''(0)\sigma_t^2 + \gamma''(0)) \leq \phi^2 \quad (57)$$

On the outside of the two inequalities we have constants, and the distribution of σ_{t+1}^2 given σ_t^2 is not bounded away from zero. Consequently, the only way for [Equation \(57\)](#) to hold for all σ_{t+1}^2 is if the term inside the parentheses equals zero.

$$0 = \beta''(0)\sigma_t^2 + \gamma''(0) \quad (58)$$

However, the only way for this to hold is for both $\gamma''(0)$ and $\beta''(0)$ to equal zero. This plus the conditional Gaussianity of the returns implied by the parametric model implies that γ and β are both linear.

□

Lemma 1 (Characterizing the SDF Integration Constants). *Let the SDF be given as in [definition 2](#), and the model be parametrized as in [definition 3](#) and [definition 4](#). Then the SDF constants follow the following equations.⁴*

4. This is Equation 3.4 in Han, Khrapov, and Renault ([2018](#), 3.4).

$$\begin{aligned} m_0(\theta, \pi) &= E(\theta) + B(C(\theta) + \pi) \\ m_1(\theta, \pi) &= D(\theta) + A(C(\theta) + \pi) \end{aligned} \quad (12)$$

Proof.

$$\mathbb{E} [\exp (m_0(\theta, \pi) + m_1(\theta, \pi)\sigma_t^2 - \pi\sigma_{t+1}^2 - \theta r_{t+1}) \mid \mathcal{F}_t] = 1 \quad (59)$$

We can use [Equation \(59\)](#) to relate $m_0(\theta, \pi)$ and $m_1(\theta, \pi)$ to the physical measure functions.

$$\mathbb{E} [\exp (m_0(\theta, \pi) + m_1(\theta, \pi)\sigma_t^2 - \pi\sigma_{t+1}^2 - \theta r_{t+1}) \mid \mathcal{F}_t] = 1 \quad (60)$$

By the law of iterated expectations.

$$\mathbb{E} [\mathbb{E} [\exp (m_0(\theta, \pi) + m_1(\theta, \pi)\sigma_t^2 - \pi\sigma_{t+1}^2) \exp (-\theta r_{t+1}) \mid \mathcal{F}_t, \sigma_{t+1}^2]] = 1 \quad (61)$$

The second term is the Laplace transform of r_{t+1} .

$$\mathbb{E} [\exp (m_0(\theta, \pi) + m_1(\theta, \pi)\sigma_t^2 - \pi\sigma_{t+1}^2) \exp (-C(\theta)\sigma_t^2 - D(\theta)\sigma_t^2 - E(\theta_2)) \mid \mathcal{F}_t] = 1 \quad (62)$$

Reorganizing terms.

$$\mathbb{E} [\exp (m_0(\theta, \pi) + m_1(\theta, \pi)\sigma_t^2 - D(\theta)\sigma_t^2 - E(\theta_2)) \exp (-(\pi + C(\theta))\sigma_{t+1}^2) \mid \mathcal{F}_t] = 1 \quad (63)$$

Substituting in the Laplace transform for σ_{t+1}^2 .

$$\mathbb{E} [\exp(m_0(\theta, \pi) + m_1(\theta, \pi)\sigma_t^2 - D(\theta)\sigma_t^2 - E(\theta_2) - A(\pi + C(\theta)) - B(\pi + C(\theta))) \mid \mathcal{F}_t] = 1 \quad (64)$$

□

Lemma 2 (Separating the Leverage Effect from the Measure Changes). *Let $\tilde{r}_{t+1} := r_{t+1} - \frac{1-\phi^2}{2}\sigma_{t+1}^2 + (1-\phi^2)\theta\sigma_{t+1}^2$, and let the m_{t+1} have the form given by [Equation \(13\)](#). Then*

$$\mathbb{V}\text{ar} [\tilde{r}_{t+1} \mid \sigma_t^2] = \mathbb{E} [\sigma_{t+1}^2 \mid \sigma_t^2]. \quad (16)$$

Proof. We start by considering the expectation of \tilde{r}_{t+1} and show that it equals zero.

$$\mathbb{E} [\tilde{r}_{t+1} \mid \sigma_t^2] = \mathbb{E} \left[\mathbb{E} \left[r_{t+1} - \frac{1-\phi^2}{2}\sigma_{t+1}^2 + (1-\phi^2)\theta\sigma_{t+1}^2 \mid \sigma_t^2, \sigma_{t+1}^2 \right] \mid \sigma_t^2 \right] \quad (65)$$

By the conditional Gaussianity of r_{t+1} , we can absorb the convexity correction into an exponential. Normally, it would introduce a $1/2$ variance term, but that cancels. The second term is a measurable with respect to the conditioning information and so is not affected.

$$= \mathbb{E} \left[\log \mathbb{E} \left[\exp(r_{t+1}) \exp((1 - \phi^2)\theta\sigma_{t+1}^2) \mid \sigma_t^2, \sigma_{t+1}^2 \right] \mid \sigma_t^2 \right] \quad (66)$$

We note that normally, the covariance term would cause the mean to fall, but again that term cancels with the second term. We can divide through by $\exp(r_t)$ and $m_{t-1,t}$ because they are measurable with respect to \mathcal{F}_t , and the remainder of the SDF is also contained within the information set, and so we can add it as well (Equation (13)).

$$\propto \mathbb{E} \left[\log \mathbb{E} \left[\frac{\exp(r_{t+1})}{\exp(r_t)} \frac{M_{t,t+1}}{M_{t-1,t}} \mid \sigma_t^2, \sigma_{t+1}^2 \right] \mid \sigma_t^2 \right] \quad (67)$$

We can pull the log outside of the outer expectation while adding at most a σ_t^2 -measurable term.

$$\propto \log \mathbb{E} \left[\frac{\exp(r_{t+1})}{\exp(r_t)} \frac{M_{t,t+1}}{M_{t-1,t}} \mid \sigma_t^2 \right] \quad (68)$$

This is the log expectation of a price change in the price discounted by the change in the SDF.

$$= 0 \quad (69)$$

Because the mean of \tilde{r}_{t+1} given σ_t^2 does not change in expectation when we condition on σ_{t+1}^2 we can apply the Itô Isometry.

$$\mathbb{E} [\text{Var}(\tilde{r}_{t+1} \mid \sigma_t^2)] = \mathbb{E}[\sigma_{t+1}^2] \quad (70)$$

Intuitively, volatilities are squared returns, and so they are variances. The tricky part here is that variances are centered second moments, not the second moments themselves. The volatilities are also centered second moments, but the centering is not the same in general. In continuous time, they would only differ by a drift term, which can be ignored, which is why the Ito Isometry usually is used in that context. Using discrete-time return, as we do here, we first have to appropriately recenter the variables, which is why it applies to \tilde{r}_{t+1} but not to r_{t+1} . □

Appendix B Identification Proofs

Lemma 5 (Identified Set). *Assume that the moment conditions specified in definition 5 have the correct form and that the instruments we are using satisfy the standard exogeneity and relevant conditions. Let the true parameter vector $\omega := (\rho, c, \delta, \phi, \theta, \pi) \in [-1 + \epsilon_1, 1 - \epsilon_2] \times [M_1, M_2] \times [\epsilon_4, M_4] \times [M_5, M_6] \times [-1 + \epsilon_4, 1 - \epsilon_5] \times [M_7, M_8] \times [M_9, M_{10}]$, where the M_* are some large (in magnitude) known constants and the ϵ_* are some small positive constants. Let $Q_T(\omega, X)$ be the*

GMM objective function with moment conditions given in [definition 5](#).

If there exists a ϵ such that $|\phi| > \epsilon > 0$, then all of the parameters are identified. If $\phi = 0$, the objective function is independent of π . Hence, π is not identified. However, even in that case all of remaining parameters are still identified.

Proof. Since Q_T is a quadratic in terms of deviations between sample and population moment conditions as long as the population moment can be inverted to solve for the parameters, the Q_T process identifies them as well. In addition, since we have a sufficient number of exogenous valid instruments the conditioning implied by projecting on the instruments does not affect the arguments above.

Identifying the four parameters that govern the volatility dynamics is not particularly complicated. We have four parameters and four non-redundant moment conditions. The first two equations in [definition 5](#) identify ρ and $c\delta$. [Equation \(31\)](#) identifies ρc and $c^2\delta$. This allows us to separately identify c and δ .

Identifying ϕ is also relatively straightforward. Since r_{t+1} and σ_{t+1}^2 are known, as long as we know the conditional mean of r_{t+1} , then identifying ϕ is identified by [Equation \(33\)](#). Identifying the conditional mean of r_{t+1} is straightforward because we observe volatility and the conditional mean is a linear equation in these variables.

Identifying the risk prices π and θ is more complicated. We have to identify both parameters off of [Equation \(32\)](#). This is in principle possible because we now have two non-redundant sources of variation in the data — σ_t^2 and σ_{t+1}^2 .

The only place that [definition 5](#) that the risk-prices occurs in [Equation \(32\)](#). We showed that γ, β and ψ functions are independent of π if $\phi = 0$ in the discussion leading up [Equation \(29\)](#). We further showed that they are not independent if $\phi \neq 0$. In addition the dependence of the identification in terms of the non-singularity of the derivatives equilibrium conditions in terms of π and θ depends smoothly on ϕ . This will be important later.

□