Journal of Financial Economics 142 (2021) 23–45

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Hedging macroeconomic and financial uncertainty and volatility☆

Ian Dew-Becker[a], Stefano Giglio[b], Bryan Kelly[c,d,*]

[a] Northwestern University, United States
[b] Yale University, United States
[c] Yale University, United States
[d] AQR Capital Management (Greenwich, CT USA)

ARTICLE INFO

Article history:
Received 17 February 2020
Revised 29 August 2020
Accepted 26 September 2020
Available online 1 June 2021

JEL classification:
G12
G13
D83
C33

ABSTRACT

We study the pricing of shocks to uncertainty and volatility using a wide-ranging set of options contracts covering a variety of different markets. If uncertainty shocks are viewed as bad by investors, they should carry negative risk premiums. Empirically, however, uncertainty risk premiums are positive in most markets. Instead, it is the *realization* of large shocks to fundamentals that has historically carried a negative premium. In other words, we find that the return premium for gamma is negative, while that for vega is positive. These results imply that it is jumps, for which exposure is measured by gamma, not forward-looking uncertainty shocks, measured by vega, that drive investors' marginal utility. In further support of the jump interpretation, the return patterns are more extreme for deeper out-of-the-money options.

1. Introduction

1.1. Background

It is well established that a wide range of measures of economic volatility and uncertainty vary over time. Uncertainty about all features of the aggregate economy, including productivity, the level of the stock market, inflation, interest rates, and energy prices, varies substantially, often as the direct result of policy choices. It is therefore important to understand how uncertainty affects the economy, both to reveal the basic drivers of economic fluctuations and also to guide policymakers.

There are numerous theories, both in macro and finance, that explore the relation between uncertainty and real activity. This literature highlights that causation runs in both directions, so even the sign of the relation between the two is ambiguous in many cases.[1] The empirical literature studying uncertainty in macroeconomics has focused almost entirely on analyzing raw correlations or using vector autoregressions (VAR) with varying identifying assumptions, and thus far it has not resolved the question of whether uncertainty is contractionary in either the

☆ We appreciate helpful comments from the Editor, an anonymous referee, Dmitry Muravyev, Federico Gavazzoni, Nina Boyarchenko, Vito Gala, Alex Hsu, Ivan Shaliastovich, Emil Siriwardane, and seminar participants at Kellogg, CITE, Syracuse, Yale, the University of Illinois, the Federal Reserve Board, UT Austin, LBS, LSE, Columbia, Queen Mary, FIRS, WFA, INSEAD, SITE, the NBER, UIUC, the MFA, Temple, the AEA, UBC, the CBOE, the Federal Reserve Bank of Chicago, and the Macro Finance Society.

* Corresponding author.
E-mail address: bryan.kelly@yale.edu (B. Kelly).

[1] For example, see Schwert (1989), Caballero (1999), Bloom (2009), Schwert (2011), Pástor and Veronesi (2009), Bachmann and Moscarini (2012), and a summary discussion in Bloom et al. (2018) about the potentially expansionary effects of uncertainty shocks. In finance, see the finance literature on good and bad uncertainty, e.g., Bekaert et al. (2015) and Segal et al. (2015).

https://doi.org/10.1016/j.jfineco.2021.05.053
0304-405X/

short- or long-run: that is, whether uncertainty is typically good or bad.

Parallel to the macro literature, there is a long-running literature in finance that studies how uncertainty and volatility are priced in financial markets. That literature distinguishes between the pricing of shocks to uncertainty about the future – i.e., shocks to conditional variances or implied volatilities – and realized volatility, or the actual occurrence of jumps. Constantinides et al. (2013) and Cremers et al. (2015), for example, study the pricing of uncertainty and jump risk, looking at option portfolios with different vega (implied volatility) and gamma (realized volatility or jump) exposure.

1.2. Contribution and methods

This paper takes a finance approach to evaluating the effects of uncertainty shocks, building on the work of Constantinides et al. (2013), Cremers et al. (2015), and Dew-Becker et al. (2019). Instead of studying a VAR with all of the associated identification challenges, as in the macro literature, we use one of the key insights of the finance literature, that financial markets provide a direct window on how investors perceive shocks.[2] The main contribution of this paper relative to past work is to use options across a wide range of underlyings and maturities to measure the risk premiums associated with shocks to uncertainty and to realized volatility. Those premiums can furthermore be used to construct implied premiums on shocks to major macro uncertainty indexes and hence shed light on the question of how uncertainty shocks affect the real economy.

If investors are willing to accept negative average returns on portfolios that hedge uncertainty shocks, as they would on an insurance contract, that implies that they view uncertainty as being bad in that it rises in high marginal utility states. On the other hand, if the hedging portfolios have positive average returns, then investors view uncertainty as typically rising in low marginal utility (good) states. So rather than running sophisticated regressions of output on uncertainty, we follow the finance tradition of letting investors speak to the question.

While there is a large literature that estimates the risk premiums for uncertainty about the S&P 500 based on the pricing of options,[3] recent evidence in Ludvigson et al. (2015) and Baker et al. (2016) shows that aggregate uncertainty has multiple dimensions beyond the financial uncertainty captured by the S&P 500. This paper contributes to the literature by estimating risk premiums associated with uncertainty and realized volatility (jumps) in 19 different markets covering a range of features of the economy, including financial conditions, inflation, and the prices of real assets. The broad range allows the analysis to uncover consistent patterns in investors' attitudes to different types of uncertainty. We also use all the options together to construct hedging portfolios for aggregate uncertainty measures developed in the literature, specifically, the JLN indexes in Jurado et al. (2015) and the economic policy uncertainty (EPU) index of Baker et al. (2016). Fitting those indexes actually requires using more than just the S&P 500: the results show that to span uncertainty about the real economy, it is important to have implied volatilities for real assets, like energies and metals, underscoring the value of the breadth of our data set.

In each of the 19 markets, we construct straddles and strangles at maturities of one to five months, and measure two-week holding period returns. We show, both theoretically and empirically, that the different maturities have different loadings on the underlying risks, allowing estimation of risk premiums using standard factor models. We examine risk premiums for two types of shocks: to uncertainty, and to realized volatility (jumps). An uncertainty shock represents an increase in the dispersion of agents' conditional distribution for future outcomes, and an option's exposure to uncertainty shocks is measured (approximately) by its vega. The second shock is to the realization of large outcomes, i.e., exposure to realized volatility, or gamma (formally, exposure to squared returns).

Vega and gamma – exposures to implied and realized volatility – have a formal link to theoretical models. Whereas uncertainty in models is a forward-looking conditional variance, realized volatility is a contemporaneous sample variance. That is, for some shock ε, with $var_t(\varepsilon_{t+1}) = \sigma_t^2$, uncertainty is σ_t^2, while volatility is ε_t^2. Vega is literally the exposure of an option to σ_t^2, while gamma is exposure to ε_t^2. The distinction between σ_t^2 and ε_t^2 is crucial from a theoretical point of view: models in which forward-looking uncertainty matters for the economy have predictions about σ_t^2 but not about ε_t^2.

To summarize, then, the basic method in the paper is to measure risk premiums on implied and realized volatility (jumps), or vega and gamma, using a typical factor pricing model on a panel of option returns across maturities, strikes, and numerous different underlyings. The estimated premiums are then used to infer the relation of marginal utility with uncertainty and realized volatility, both for specific underlyings and also for prominent macro uncertainty indexes.

1.3. Results

The main results focus on straddles, because the options in the portfolio are initially at the money and hence most liquid. The empirical analysis yields three key findings. First, across 19 option markets, the risk premium for hedging uncertainty shocks, vega, is in the majority of cases positive. For nonfinancial underlyings and the JLN macro and inflation uncertainty indexes, the premiums are statistically and economically significantly positive, with Sharpe ratios near 0.5. The results imply that investors in these markets view periods of high uncertainty about the real economy as being good on average. For the financial sector (including the S&P 500) and the JLN financial uncer-

[2] To be clear, the analysis of risk premiums does not identify structural shocks; it only reveals the correlation of innovations in marginal utility with reduced-form innovations to uncertainty (since there is no structural identification here, we will use the terms "shock" and "innovation" interchangeably).

[3] See Egloff et al. (2010), Dew-Becker et al. (2017), Van Binsbergen and Koijen (2017), Andries et al. (2015), and Ait-Sahalia et al. (2019).

tainty and EPU index, the premium on uncertainty is not clearly distinguishable from zero.

The second empirical result runs in the opposite direction: consistently across both the financial and real sectors of the economy, portfolios that hedge realized volatility, or jumps, earn statistically and economically significantly negative returns. Investors on average therefore view periods in which shocks to fundamentals themselves are large as being bad.

It is well known that both volatility and uncertainty are countercyclical, but their overall correlation is not as high as one might expect – only about 65% on average across markets – and the average correlation between their innovations is only 0.2. The results here show that innovations in realized volatility identify the states of the world that investors view as actually negative, whereas surprise increases in implied volatility, which is high in other, mostly unrelated, states of the world, are not on average perceived as bad.

Our findings for realized volatility contribute to the growing literature studying skewness risk in the economy: if shocks to the economy (i.e., aggregate consumption) are skewed to the left, then large shocks tend to be bad.[4] An explanation for the pricing of realized volatility could then simply be that hedging realized volatility helps hedge downward jumps and disasters in aggregate consumption. If it is truly jumps that drive pricing, then we would expect that the negative returns on options would be larger for options that are farther out of the money. To test the hypothesis that the pricing is compensation for jump risk, we extend the baseline results to examine returns on strangles, which are like straddles, in holding both a put and a call, but in which both options are out of the money at inception. Relative to straddles, strangles only have positive payoffs for relatively large movements in the underlying.

Our third result is that the gamma/jump premiums for strangles are about twice as large as those for straddles, providing formal evidence for the idea that it is jumps, rather than small (or diffusive) movements in underlying prices, that investors are averse to. As with the results for straddles, the result that deeper out-of-the-money options have more negative returns is well known for the S&P 500. Our results are novel for showing that the same result appears in a wide range of markets, including those linked to the real economy.

Because the variance risk premium is robustly negative across many markets, jumps, which drive surprises in realized volatility, tend to be robustly viewed as bad events by investors, regardless of where they occur. According to asset prices, what policymakers should focus on, rather than uncertainty about the future (the possibility that something extreme might happen), is the realization of extreme (typically negative) events. For investors, the results imply that the mean-variance efficient portfolio among the assets we study is short gamma (jump risk) and either neutral to or long vega (exposure to implied volatility), and we show that large Sharpe ratios are available when buying vega and selling gamma across many markets. In the paper, we also build a simple extension of the standard long-run risk model of Bansal and Yaron (2004) that shows how our results can arise in equilibrium.

1.4. Relation to past work

The paper is related to two main strands of literature. The first studies the relation between uncertainty and the macroeconomy. Numerous channels have been proposed through which uncertainty about various aspects of the aggregate economy may have real effects, but the models do not generate a uniform prediction that uncertainty shocks are necessarily contractionary.[5] Our results are more consistent with the expansionary forces present in the models. There are also models with joint or reverse causation, such as Pástor and Veronesi, (2009) and Bachmann and Moscarini, (2012).[6] The related empirical literature tries to measure whether uncertainty does in fact have contractionary effects, finding often conflicting results.[7]

This paper builds on that work from a finance perspective by providing measures of risk premiums that indicate how investors perceive the effects of aggregate uncertainty shocks across many markets. The finance perspective of this paper means that the methods and data are very different from papers that have instead used a macroeconomic approach to the question. For example, Berger et al. (2020) estimate a structural vector autoregression, as is common in the macroeconomics literature, to try to understand the effect of uncertainty shocks on the economy. While trying to answer a similar question, this paper takes a financial economics approach, studying risk premiums, and requiring none of the VAR identifying assumptions.

As discussed above, Constantinides et al. (2013) and Cremers et al. (2015) are important precedents in the finance literature for studying the pricing of shocks to uncertainty and volatility. We build on Constantinides et al. (2013) in that we also examine factor risk premiums estimated from option returns, with the innovation that we look across a broader range of markets. Our analysis uses methods similar to that paper and also to those of Cremers et al. (2015), in that we study both a factor model and replicating portfolios. We differ from Cremers et al. (2015) in that we use option returns to measure risk premiums, rather than projecting stock returns onto uncertainty and volatility factors. Because stock returns are driven by so many different

[4] See Barro (2006), Salgado et al. (2016), Seo and Wacheter (2018); Seo and Wachter (2018), Siriwardane (2015), and Berger et al. (2020). Dew-Becker et al. (2019) provide a structural model for the source of aggregate skewness.

[5] See Basu and Bundick (2017) Berger et al. (2020), Bloom (2009), Bloom et al. (2018), Leduc and Liu (2016), Gourio (2013), Gilchrist and Williams (2005), and Bloom et al. (2018).

[6] See also Decker et al. (2016), Berger and Vavra (2013), Ilut et al. (2015), Kozlowski et al. (2016), Cesa-Bianchi et al. (2018), and Diercks et al. (2019).

[7] For example, Schwert (1989), Schwert (2011), Berger et al. (2020), Bretscher et al. (2019), Jurado et al. (2015), Jurado et al. (2015), Baker et al. (2016), Bachmann and Bayer (2013), and Baker et al. (2016); Alexopoulos and Cohen (2009). For papers on different types of uncertainty, see also Bretscher et al. (2018), Elder and Serletis (2010), Darby et al. (1999), Huizinga (1993), and Elder (2004).

risk factors, options can be useful for helping to isolate underlying risks relatively precisely. That difference can help explain differences between our results and those obtained by Constantinides et al. (2013) relative to Cremers et al. (2015).

The paper also draws on a literature in finance estimating the pricing of volatility (ε^2) risk. The past literature almost exclusively studies the S&P 500, and in general studies just the variance risk premium, which is the pricing of realized volatility (as measured by the average gap between option-implied and realized volatility).[8] In addition to studying a much broader range of markets, our contribution is to also isolate the premium on implied volatility.

The remainder of the paper is organized as follows. Section 2 describes the data and its basic characteristics. Our main results on the cost of hedging uncertainty and volatility shocks are in Section 3. We then provide a theoretical derivation of the risk exposures of the options in Section 4 and use it to construct replicating portfolios. Section 5 reports the cost of hedging macroeconomic uncertainty and realized volatility, combining all 19 markets together. Section 6 presents robustness results and Section 7 concludes.

2. Measures of uncertainty and realized volatility

This section describes our main data sources and then examines various measures of uncertainty and realized volatility.

2.1. Data

2.1.1. Options and futures

We obtain data on prices of financial and commodity futures and options from the end-of-day database from the CME Group, which reports closing settlement prices, volume, and open interest over the period 1983–2015. Each market includes both futures and options, with the options written on the futures. The futures may be cash- or physically settled, while the options settle into futures. As an example, a crude oil call option gives its holder the right to buy a crude oil future at the strike price. The underlying crude oil future is itself physically settled: if held to maturity, the buyer must take delivery of oil.[9]

To be included in the analysis, contracts are required to have least 15 years of data and maturities for options extending to at least six months, which leaves 14 commodity and 5 financial underlyings. The final contracts included in the data set have 18 to 31 years of data. A number of standard filters are applied to the data to reduce noise and eliminate outliers. Those filters are described in Appendix A.

We calculate implied volatility for all of the options using the Black and Scholes (1973) model and, technically, the Black (1976) model for the case of futures.[10] Unless otherwise specified, implied volatility is calculated at the five-month maturity. We take this value as the benchmark measure of uncertainty in each market. In general, longer maturities are naturally more tightly linked to long-lived economic decisions, like physical investments. We do not go past five months because there is less trade and liquidity at longer maturities, making prices less reliable.

Implied volatilities extracted from options reflect market uncertainty about future returns, but they also contain a risk premium, which can potentially vary over time. However, even in the presence of that risk premium, implied volatilities appear to provide very good summaries of the available information in the data for forecasting future volatility, driving out other standard uncertainty measures from forecasting regressions. Online Appendix Section OA.1 compares implied volatilities to regression-based forecasts of future volatilities and shows that they are all over 90% correlated (with an average correlation of 97%), indicating that option-implied volatility is a good, if not perfect, proxy for true (physical) uncertainty. For that reason, in what follows we refer to implied volatility and uncertainty interchangeably, with the understanding that deviations due to time-varying risk premiums are quantitatively small at the monthly frequencies we focus on.[11]

2.2. The time series of implied volatility

Fig. 1 plots option implied volatility for three major futures: the S&P 500, crude oil, and US Treasury bonds. The implied volatilities clearly share common variation; for example, all rise around 1991, 2001, and 2008. On the other hand, they also have substantial independent variation. Their overall correlations (also reported in the figure) are only in the range 0.5–0.6.

Table 1 reports pairwise correlations of implied volatility across the 19 underlyings. The largest correlations in implied volatility are among similar underlyings: crude and heating oil, the agricultural products, gold and silver, and the British Pound and Swiss Franc. Correlations outside those groups are notably smaller, in many cases close to zero. The largest principal component (PC) of the correlation matrix explains 46% of the total variation. The remaining PCs are much smaller, though: even the second largest only explains 16% of the total variation. Eight PCs are required to explain 90% of the total variation in the IVs, which is perhaps a reasonable estimate of the number of independent components in the data.

The common variation in the implied volatilities is much larger than the common variation in the underlying futures returns. The largest PC for the futures returns

[8] For example, see Ait-Sahalia et al. (2019), Bollerslev and Todorov (2011), Andersen et al. (2015, 2017), Dew-Becker et al. (2017), Constantinides et al. (2013), Cremers et al. (2015), and Farago and Tédongap (2018) for work on the S&P 500. A few papers have studied specific markets, like Bakshi et al. (2003), Choi et al. (2017), Prokopczuk et al. (2017), and Trolle and Schwartz (2010).

[9] The underlying futures in general expire in the same month as the option. Crude oil options, for example, currently expire three business days before the underlying future.

[10] The majority of the options that we study have American exercise, while the Black model technically refers to European options. We examine IVs calculated assuming both exercise styles (we calculate American IVs using a binomial tree) and obtain nearly identical results. Since there are no dividends on futures contracts, early exercise is only rarely optimal for the options studied here.

[11] See also Bekaert et al. (2013) for an analysis of the variation in risk premiums in implied volatilities.

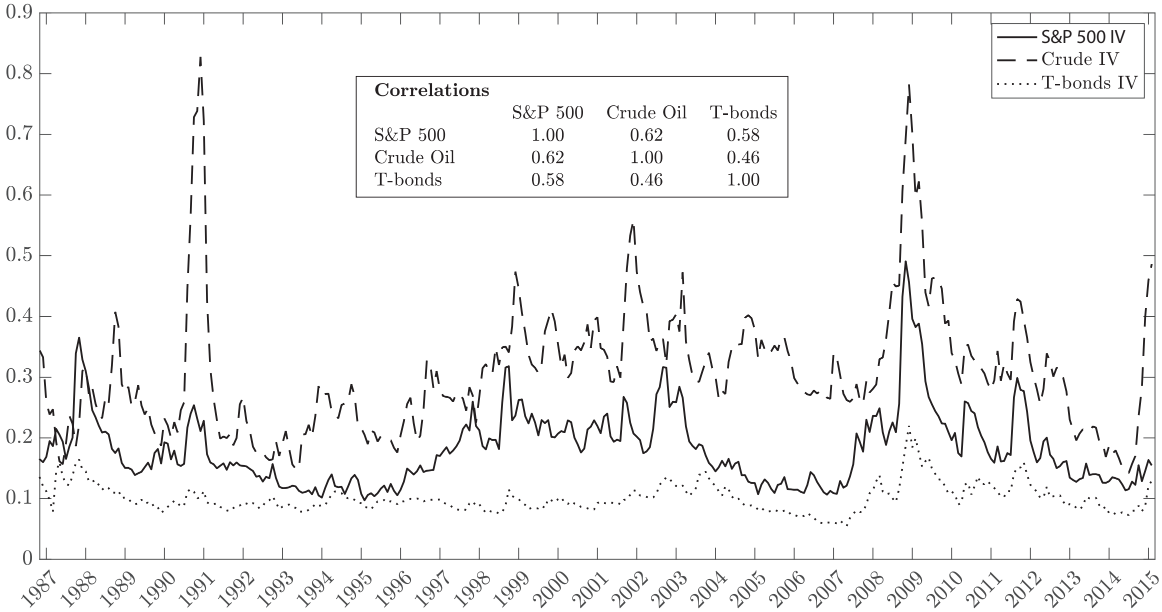

Fig. 1. Sample implied volatilities. Monthly implied volatilities calculated from three-month options using the Black-Scholes model.

Table 1
Pairwise correlations of implied volatility across markets.

IV	Treasuries	S&P 500	Swiss Franc	Yen	British Pound	Gold	Silver	Copper	Crude oil	Heating oil	Natural gas	Corn	Soybeans	Soybean meal	Soybean oil	Wheat	Lean hog	Feeder cattle
S&P 500	0.56																	
Swiss Franc	0.53	0.29																
Yen	0.40	0.56	0.48															
British Pound	0.45	0.40	0.75	0.45														
Gold	0.52	0.57	0.21	0.28	0.37													
Silver	0.42	0.34	0.19	0.29	0.34	0.78												
Copper	0.39	0.49	0.15	0.35	0.36	0.74	0.77											
Crude oil	0.42	0.63	0.25	0.39	0.27	0.54	0.31	0.48										
Heating oil	0.41	0.64	0.23	0.36	0.23	0.51	0.28	0.51	0.95									
Natural gas	0.11	0.44	-0.03	0.04	0.03	0.33	0.06	0.44	0.49	0.63								
Corn	0.25	0.37	-0.11	0.14	0.11	0.50	0.56	0.58	0.22	0.18	0.11							
Soybeans	0.22	0.35	-0.05	0.17	0.17	0.47	0.48	0.57	0.29	0.29	0.21	0.85						
Soybean meal	0.28	0.33	-0.08	0.16	0.06	0.53	0.50	0.57	0.30	0.27	0.23	0.81	0.94					
Soybean oil	0.31	0.30	0.10	0.12	0.23	0.48	0.49	0.56	0.26	0.29	0.23	0.73	0.89	0.83				
Wheat	0.38	0.42	0.01	0.19	0.10	0.62	0.62	0.60	0.34	0.31	0.17	0.84	0.77	0.75	0.64			
Lean hog	0.29	0.42	-0.03	0.28	-0.10	0.27	0.16	0.35	0.40	0.47	0.40	0.29	0.37	0.39	0.38	0.36		
Feeder cattle	0.45	0.35	0.11	0.16	0.07	0.40	0.51	0.50	0.31	0.34	0.13	0.48	0.47	0.50	0.48	0.52	0.43	
Live cattle	0.51	0.28	0.24	0.18	0.07	0.38	0.41	0.45	0.32	0.39	0.26	0.32	0.33	0.43	0.49	0.43	0.47	0.84

Note: Pairwise correlations of three-month option-implied volatility across markets. The darkness of the shading represents the degree of correlation.

explains less than half as much variation, 19% versus 46%. In other words, while the individual futures prices may be driven by idiosyncratic shocks, or their correlations with each other might change over time, masking common variation, investor uncertainty about futures returns has a substantial degree of commonality across markets that is similar to findings in Herskovic et al. (2016), showing that we are not studying uncertainty that is purely idiosyncratic and isolated to individual futures markets. The table below formalizes that result, reporting the variation explained by the first PC for implied volatility, realized volatility (discussed below), and the underlying futures returns, along with bootstrapped 95% confidence bands.

Fraction of variation explained by first principal component

			Futures
	IV	RV	return
First PC (% explained)	45.9%	28.1%	19.1%
95% Bootstrap CI	37.3% 49.5%	23.7% 41.8%	16.7% 21.2%

2.3. Relation between implied volatility and macro uncertainty indexes

Our ultimate goal is to understand the pricing of economic uncertainty. We therefore want to check whether the implied volatilities in the futures markets we study are related to other prominent measures of uncertainty. Fig. 2 quantifies how well the 19 IVs can replicate two well-known macro uncertainty indexes: the JLN indexes from Jurado et al. (2015) and the EPU index of Baker et al. (2016) (see Section 5 for a more detailed description of the indexes). Fig. 2 plots the time series of the JLN indexes and EPU index against the fitted values from their projection onto the 19 implied volatilities. The right-hand panels plot the pairwise correlations of the implied volatilities in the individual markets with the fitted uncertainty. For financials, the correlation with S&P 500 implied volatility is 97%. The next highest correlation is only 68%, for Treasury bonds. So Fig. 2 shows that fitted financial uncertainty is very nearly equivalent to S&P 500 implied volatility.[12]

The second row plots fitted uncertainty for real variables. In this case, gold, copper, crude oil, and heating oil are the most important contributors. The third row shows similar results for the price component of JLN uncertainty. Uncertainty about the real economy and inflation are therefore driven by similar factors, and those factors are notably distinct from financial uncertainty, which shows why having a broad range of IVs and looking at markets beyond the S&P 500 are important.

The bottom panels plot results for the EPU index. The highest pairwise correlations are with financial IVs, Treasuries, gold, the S&P 500, and currencies. That implies that the EPU index measures a similar type of uncertainty as other financial uncertainty measures, perhaps because news coverage often focuses on financial markets.[13]

3. The cost of hedging uncertainty and volatility

In this section we present the main results of the paper: we estimate the cost of hedging shocks to volatility and uncertainty using option portfolios.

We compute the cost of hedging a shock as the negative of the average excess return (risk premium) on the portfolio that hedges that shock. We report all risk premiums in terms of Sharpe ratios, which reveal the compensation for bearing a risk (or the cost of hedging it) per unit of risk, and are therefore more easily comparable across markets. The option returns are highly skewed, so an investor here would care about more than just the Sharpe ratio; we use it simply as a device for holding effective leverage constant across markets. For reference, the historical Sharpe ratio of US equities in our sample is 0.52.

We estimate risk premiums for implied and realized volatility using a standard linear factor model, and we use straddle returns of different maturities as test assets. Typical factor models use a small number of aggregate factors. Here, though, we are interested in the price of risk for shocks to all 19 types of uncertainty. We therefore estimate market-specific factor models. This is similar to the common practice of pricing equities with equity-specific factors, bonds with bond factors, currencies with currency factors, etc.[14]

The cost of hedging a risk has a simple but important economic interpretation: it measures the extent to which the risk is "bad" with respect to state prices or marginal utility. Consider a factor X and an asset with returns R_X that hedges it, in the sense that R_X varies one-for-one (and is perfectly correlated) with innovations to X. Then if M represents the stochastic discount factor,

$$E\left[\frac{R_{X,t+1}-R_f}{std_t(R_{X,t+1})}\right]=-cov\left(M_{t+1}-E_tM_{t+1},\frac{X_{t+1}-E_tX_{t+1}}{std_t(X_{t+1})}\right)R_f, \tag{1}$$

where R_f is the gross risk-free rate, which we treat as constant for the sake of exposition, E_t is the expectation operator, and std_t is the standard deviation conditional on date-t information. The equation says that the negative of the risk premium on a portfolio that hedges the risk X measures the covariance of innovations in X_{t+1} with state prices. More generally, when the correlation between R_X and innovations in X is less than 1, $E\left[R_X-R_f\right]$ measures the covariance of state prices with the part of innovations to X that is spanned by R_X. So if the premium $E\left[R_X-R_f\right]$ is negative, times when R_X (and hence X) rise are bad times, in which state prices are high. The factor model and subsequent analysis will deliver estimated Sharpe ratios for the various risk factors we study.

Finally, as we review in Online Appendix Section OA.2, the risk premiums estimated from linear factor models correspond to the average excess returns of portfolios that isolate each risk (that is, each portfolio has beta of 1 with respect to one risk factor, and 0 with respect to all other factors). These portfolios are precisely those portfolios that allow an investor to change risk exposure to any factor and that factor only; we refer to them as factor-hedging portfolios.

[12] The strong fit with S&P 500 implied volatility is not simply due to the fact that S&P 500 returns are included in the JLN construction. The results are similar when all variables involving the S&P 500 index (returns, dividends, etc.) are dropped.

[13] To account for possible overfitting due to the fact that we have 19 explanatory variables, we experimented with lasso and variable selection based on information criteria. The results were highly similar in all cases.

[14] The analysis is similar to those of Jones (2006) and Constantinides et al. (2013).

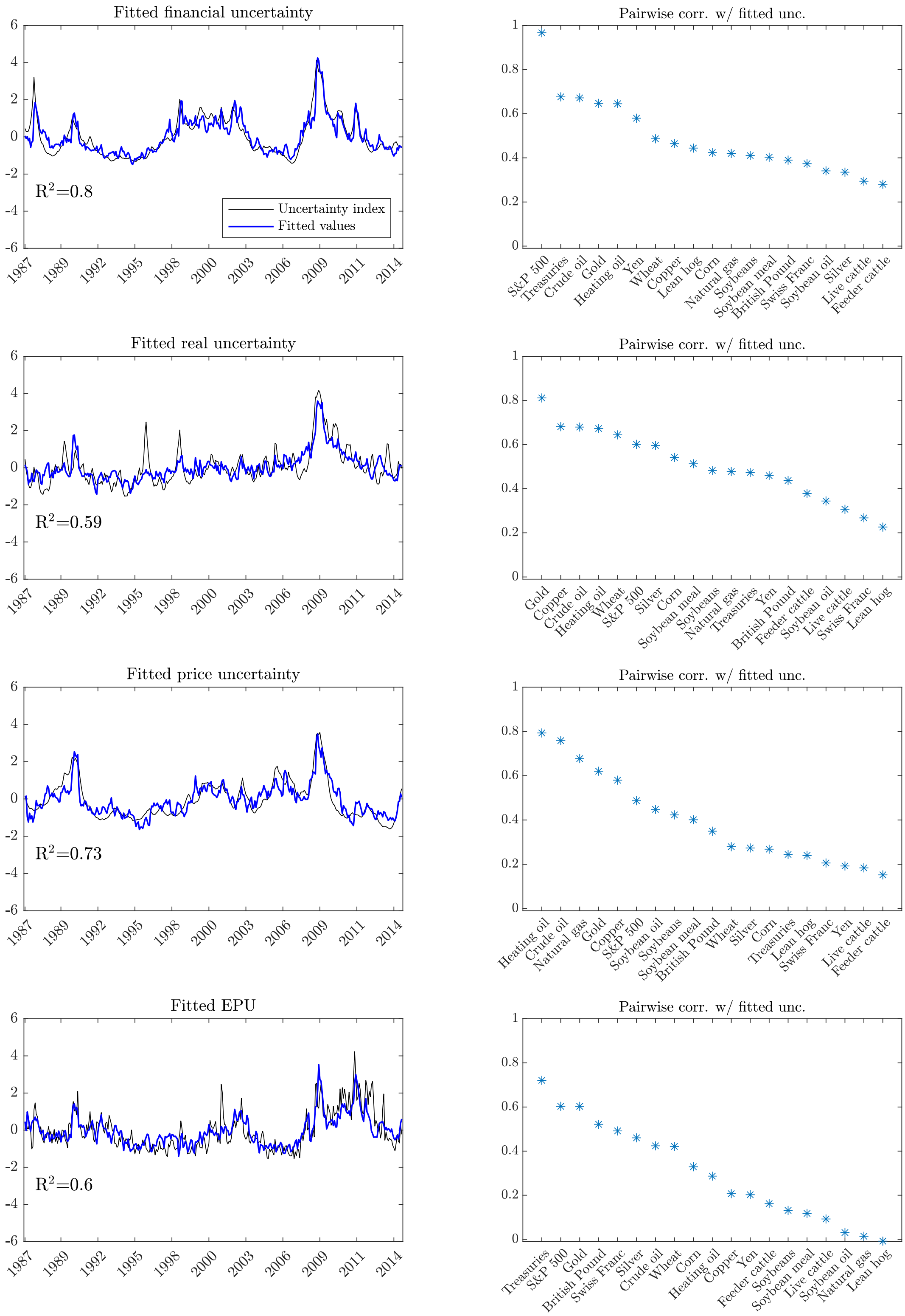

Fig. 2. Fit to uncertainty indexes. The left-hand panels plot the fitted values from the regressions of the EPU and JLN indexes on three-month implied volatility in the 19 markets. The right-hand panels plot pairwise correlations between the individual implied volatility series and the fitted values from the regressions.

3.1. Method

3.1.1. Factor model specification

For each market we estimate a time-series model of the form

$$r_{i,n,t} = a_{i,n} + \beta_{i,n}^{f} \frac{f_{i,t}}{IV_{i,t-1}} + \beta_{i,n}^{f^2} \frac{1}{2}\left(\frac{f_{i,t}}{IV_{i,t-1}}\right)^2 + \beta_{i,n}^{\Delta IV} \frac{\Delta IV_{i,t}}{IV_{i,t-1}} + \varepsilon_{i,n,t}, \quad (2)$$

where $f_{i,t}$ is the futures return for underlying i and $\Delta IV_{i,t}$ is the change in the five-month at-the-money implied volatility for underlying i. The term $r_{i,n,t}$ is a return on each of the N test assets (straddles and strangles, described in greater detail below).

The underlying futures return $f_{i,t}$ controls for any exposure of the test assets to the underlying, though in general that loading will be small, given that we use as test assets portfolios with payoffs that are symmetric in the value of the underlying. Much more important is the fact that straddles and strangles have nonlinear exposures to the futures return. The expression $\left(f_{i,t}/IV_{i,t-1}\right)^2$ captures that nonlinearity; $\beta_{i,n}^{f^2}$ will be interpreted as the exposure of the options to realized volatility.[15] Finally, the third factor is the change in the at-the-money implied volatility for the specific market at the five-month maturity, representing an uncertainty shock in that market.[16]

The three factors are scaled by lagged implied volatility for two reasons. First, this helps control heteroskedasticity. Intuitively, the factors are measuring innovations in standard deviation units, so that we are pricing based on how much the underlying moves relative to what investors expected. The second reason will be demonstrated in the next section: it is what the Black–Scholes model implies for the exposures of straddles and strangles. That is, the option portfolios yield exposure to the scaled factors used here, rather than, for example, the raw futures return (and raw futures return squared). So while the analysis in this section does not rely on Black-Scholes, this scaling will be useful for interpreting the results.

We estimate a standard linear specification for the risk premiums,

$$E[r_{i,n,t}] = \gamma_i^{f} \beta_{i,n}^{f} Std\left(\frac{f_{i,t}}{IV_{i,t-1}}\right) + \gamma_i^{f^2} \beta_{i,n}^{f^2} Std\left(\left(\frac{f_{i,t}}{IV_{i,t-1}}\right)^2\right) + \gamma_i^{\Delta IV} \beta_{i,n}^{\Delta IV} Std\left(\frac{\Delta IV_{i,t}}{IV_{i,t-1}}\right) + \alpha_{i,n}, \quad (3)$$

where $\alpha_{i,n}$ is a fitting error, using standard two-step cross-sectional regressions. The γ coefficients represent the risk premiums that are earned by investments that provide direct exposure to the factors. Due to the scaling by standard deviations, γ denotes the Sharpe ratios of the hedging portfolios for each factor constructed using the test assets.[17]

3.1.2. Test assets

Our main results are for two-week returns on straddles with maturities between one and five months.[18] A straddle is a portfolio holding a put and a call with the same maturity and strike; we specifically study zero-delta straddles, with the strike set so that the Black–Scholes delta of the portfolio is zero. The final payoff of a zero-delta straddle depends on the absolute value of the return on the underlying, meaning that they have symmetrical exposures to positive and negative returns. For the remainder of the paper, we refer to zero-delta straddles simply as straddles (that is, we only work with zero-delta straddles).

Straddles give investors exposure both to realized and implied volatility. They are exposed to realized volatility because the final payoff of the portfolio is a function of the absolute value of the underlying futures return. But when a straddle is sold before maturity (as in our case, since we use two-week holding period returns), the sale price will also depend on expected future volatility, meaning that straddles can give exposure to uncertainty shocks. Since the options in the straddle are at the money at inception, a straddle is the most liquid zero-delta portfolio we can construct.

In principle, it is also possible to estimate the factor risk premiums using other assets, like stock or bond returns, as in Cremers et al. (2015). We focus on option returns because they depend directly on realized volatility and uncertainty – which is why they are used to construct implied volatility measures – whereas for other assets the connection is less clear (many other factors affect their returns) and there could be nontrivial problems with exposures shifting over time. We show below that under the simple Black–Scholes benchmark, the factor loadings will be constant.

[15] The results are similar when the second factor is the absolute value of the futures return or when it is measured as the sum of squared daily returns over the return period.

[16] Since the IVs may be measured with error, we construct this factor by regressing available implied volatilities on maturity for each underlying and date and then taking the fitted value from that regression at the five-month maturity.

[17] While $f_{i,t}^2$ and $\Delta IV_{i,t}$ are nontradable factors, $f_{i,t}$ itself is tradable, so we include it as a test asset, yielding the additional restriction $E[f_{i,t}/IV_{i,t-1}] = \gamma_i^{f} Std(f_{i,t}/IV_{i,t-1})$. See Cochrane (2005).

[18] Past work on option returns and volatility risk premiums has examined returns at frequencies of anywhere from a day, as in Andries et al. (2015), to holding to maturity, as in Bakshi and Kapadia (2003). The precision of estimates of the riskiness of the straddles is, all else being equal, expected to be higher with shorter windows. On the other hand, shorter windows cause any measurement error in option prices (e.g., from differences between settlement prices and true fair values or trade prices, or from simple data errors) to have larger effects.Some of the existing literature, beginning with Bakshi and Kapadia (2003), examines delta-hedged returns. Even with delta hedging, the higher-order risk exposures of the straddles change substantially as the price of the underlying changes over time.Another alternative is to examine returns on synthetic variance swaps. Synthetic variance swap prices are constructed using the full range of strikes, so they require much more data than straddles. The markets we study do not all have liquid options at extreme strikes and multiple maturities, so we focus on straddles, which just require liquidity near the money.

3.2. Empirical results

3.2.1. Hedging uncertainty shocks

The dotted red series in Fig. 3 plots estimated risk premiums and confidence bands for the realized and implied volatility factors – $\gamma_i^{f^2}$ and $\gamma_i^{\Delta IV}$, respectively, using straddles as test assets. Again, the risk premiums should be interpreted as annualized Sharpe ratios, since they are scaled to measure average annualized returns per unit of annualized standard deviation. The top panel plots premiums for implied volatility and the bottom panel realized volatility. The boxes are point estimates, while the bars represent 95% confidence bands based on a block bootstrap. The bootstrap is constructed with 50-day blocks and 5000 replications. It is used to account for the fact that the returns use overlapping windows. Hansen-Hodrick type standard errors are not feasible here due to the fact that observations in the data are not equally spaced in time. The block bootstrap additionally accounts for other sources of serial correlation in the returns, such as time-varying risk premiums.

Across the top panel, implied volatility shocks carry zero or even positive premiums. For financials, the average Sharpe ratios tend to be near zero or weakly negative. The S&P 500 has a positive premium, consistent with results for variance swaps discussed extensively in Dew-Becker et al. (2017). That result is not completely robust here, however, as we discuss further below, but there is certainly no evidence of a significantly negative premium for S&P 500 uncertainty. For the nonfinancials, on the other hand, all 14 sample Sharpe ratios are actually positive, and 5 of those are individually statistically significant. Overall, for only 1 out of 19 contracts, the British Pound, do we find a significantly negative Sharpe ratio.

To formally estimate the average risk premiums across contracts, we use a random effects model, which yields an estimate of the population mean risk premium while simultaneously accounting for the fact that each of the sample Sharpe ratios is estimated with error, and that the errors are potentially correlated across contracts (see Appendix B).

For both nonfinancials and all markets overall, the estimated population mean Sharpe ratio is statistically and economically significantly positive, while for financials it is close to zero. The group-level means have the advantage of being much more precisely estimated than the Sharpe ratios for the markets individually. They show that on average, instead of there being a cost to hedging uncertainty shocks, the factor risk premium for uncertainty shocks is actually positive. For nonfinancials, the average Sharpe ratio is 0.43, and the lower end of the 95% confidence interval is 0.13. For the overall mean, the corresponding numbers are 0.32 and 0.08, so the average Sharpe ratios are significantly positive in both cases. The top panel of Table 3 reports the estimated average Sharpe ratios for financials and nonfinancials, and, in the third column, their difference, and shows that the difference in risk premiums between the two groups is not statistically significant.

The top panel of Fig. 3 contains our key results on the risk premium for uncertainty. It shows that across the board, risk premiums for uncertainty are indistinguishable from zero or, if anything, somewhat positive. The results allow us to quantify the overall correlation between uncertainty and marginal utility. For financial underlyings, including the S&P 500, the zero or very weakly negative risk premium implies that the correlation is close to zero. For the nonfinancial underlyings, which are closely linked to the JLN real and price uncertainty series, the results imply that the correlation is positive.

3.2.2. Hedging realized volatility shocks

The bottom panel of Fig. 3 reports risk premiums for realized volatility across the 19 markets, representing our second main result. The numbers are drastically different from those for IV. Whereas implied volatility has earned a zero or even positive premium, the realized volatility premiums are almost all estimated to be negative. For the S&P 500, this result is well known and is referred to as the variance risk premium. The S&P 500 realized volatility risk premium is most negative, at -1.26. That is, the premium for selling insurance against shocks to realized volatility is more than twice as large as the premium on the stock market over the same period. For the other financial underlyings, the premium on realized volatility is not statistically significantly negative. For the nonfinancials, 11 of 14 estimated premiums are negative (6 significantly).

Looking at the category means, in this case all three estimates (financials, nonfinancials, and all assets) are negative. The values are on the edge of statistically significant for the nonfinancials and the overall mean, with confidence bands just barely encompassing zero. The point estimate for the overall mean Sharpe ratio is -0.26 and the upper end of the 95% confidence interval is 0.04. Those values are almost the same as what we obtain for uncertainty, but with the opposite sign. As with uncertainty, Table 3 shows that the difference between financials and nonfinancials is not statistically significant.

In sum, in stark contrast to the results for hedging uncertainty, the bottom panel of Fig. 3 shows that there has historically been, consistently across markets, an economically significant cost to hedge realized volatility. Contracts that, rather than loading on changes in implied volatility, load on actual realized squared returns, earn negative Sharpe ratios with magnitudes up to twice as large as that for the overall stock market. So while uncertainty is viewed as neutral or even good on average, realized volatility or jumps – the realization of large squared returns – is viewed as mostly bad, for both financials and nonfinancials.

3.2.3. Goodness of fit

Fig. 4 reports a scatter plot of realized returns on the various straddle returns against the fitted returns from the model. The figure shows that there is a wide spread in realized returns that the model is able to capture. In addition, there are no large outliers. Table OA.1 in the Online Appendix reports the p-values of the χ^2 test of the model based on the squared fitting errors, bootstrapped following Constantinides et al. (2013). That test is very stringent, especially when the fitting errors are small on

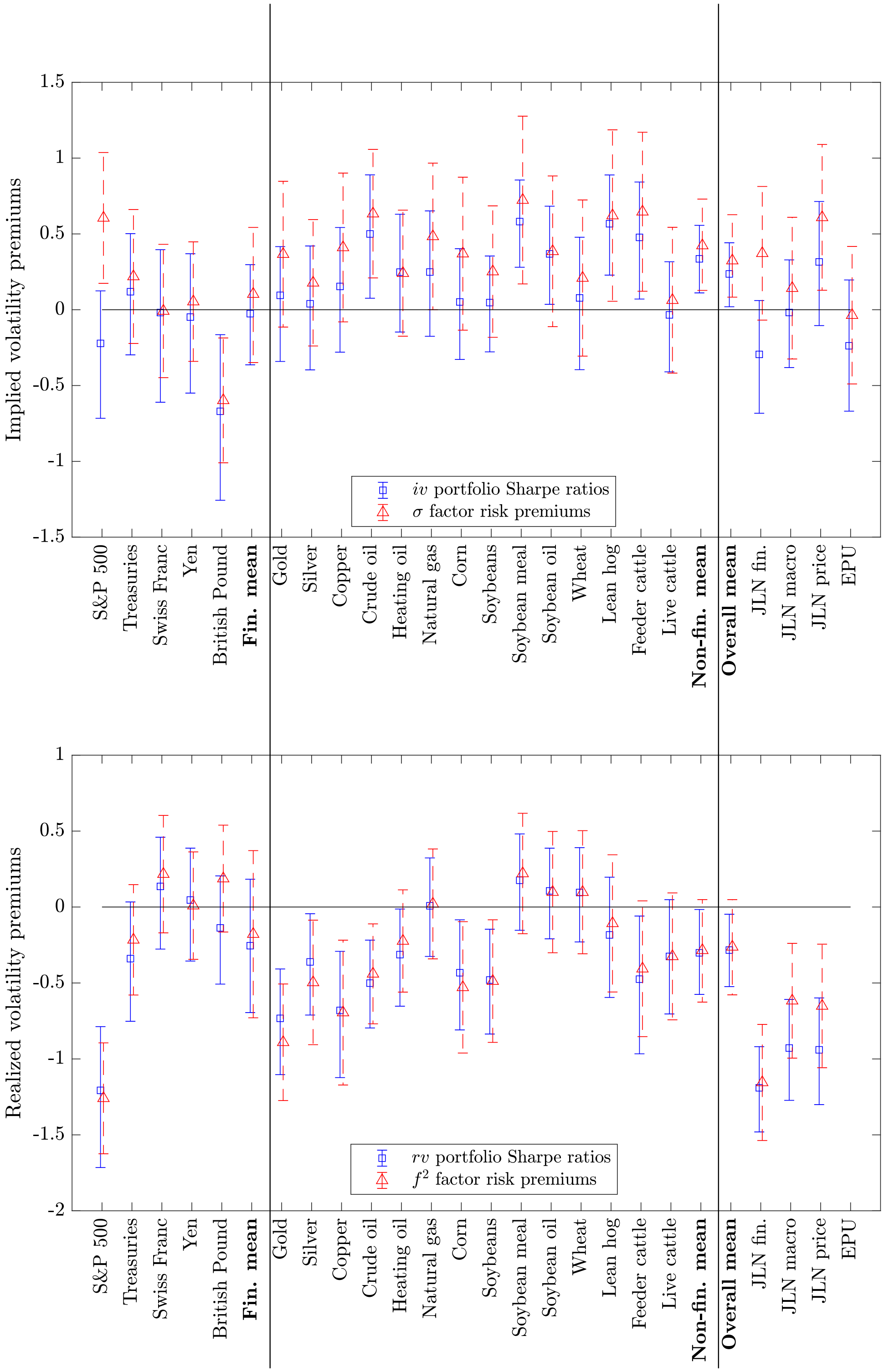

Fig. 3. RV and IV portfolio Sharpe ratios and factor risk premia: straddles. Squares are point estimates and vertical lines represent 95% confidence intervals. The solid series plots the Sharpe ratios for the *rv* and *iv* portfolios. The dotted series plots the estimated risk premiums from the factor model. In both cases, all estimation uses straddles. The confidence bands for the *rv* and *iv* Sharpe ratios are calculated through a 50-day block bootstrap, while those for the factor model use GMM standard errors with the Hansen-Hodrick (1980) method used to calculate the long-run variance. The "Fin. mean," "Non-fin. mean," and "Overall mean" points represent random effects estimates of group-level and overall means. The "JLN" and "EPU" points are for the portfolios that hedge those indexes.

Table 2
Pairwise correlations of realized volatility across markets.

RV	Treasuries	S&P 500	Swiss Franc	Yen	British Pound	Gold	Silver	Copper	Crude oil	Heating oil	Natural gas	Corn	Soybeans	Soybean meal	Soybean oil	Wheat	Lean hog	Feeder cattle
S&P 500	0.63																	
Swiss Franc	0.17	0.12																
Yen	0.31	0.32	0.15															
British Pound	0.43	0.36	0.24	0.31														
Gold	0.44	0.47	0.15	0.24	0.31													
Silver	0.42	0.43	0.15	0.22	0.27	0.65												
Copper	0.52	0.51	0.11	0.24	0.43	0.50	0.53											
Crude oil	0.24	0.24	0.13	0.20	0.20	0.32	0.14	0.24										
Heating oil	0.20	0.22	0.04	0.14	0.15	0.30	0.11	0.15	0.91									
Natural gas	0.03	0.08	0.04	-0.04	0.00	0.05	-0.06	0.00	0.08	0.18								
Corn	0.33	0.35	0.04	0.09	0.27	0.37	0.40	0.50	0.12	0.03	-0.04							
Soybeans	0.33	0.30	0.03	0.16	0.30	0.33	0.35	0.40	0.11	0.05	-0.07	0.74						
Soybean meal	0.33	0.25	0.03	0.19	0.19	0.31	0.32	0.30	0.08	0.02	-0.06	0.68	0.94					
Soybean oil	0.48	0.43	0.11	0.21	0.42	0.40	0.41	0.51	0.17	0.12	-0.04	0.67	0.88	0.72				
Wheat	0.30	0.24	0.02	0.08	0.11	0.31	0.34	0.33	0.11	0.04	-0.08	0.63	0.51	0.47	0.47			
Lean hog	0.12	0.12	0.08	0.20	-0.03	0.00	0.00	0.05	0.10	0.09	0.11	0.07	0.11	0.12	0.11	0.12		
Feeder cattle	0.22	0.20	0.03	0.04	0.07	0.10	0.16	0.30	0.10	0.07	0.12	0.35	0.32	0.32	0.27	0.22	0.26	
Live cattle	0.41	0.24	0.13	0.11	0.11	0.17	0.24	0.28	0.07	0.07	0.09	0.22	0.22	0.27	0.30	0.23	0.28	0.63

Note: Pairwise correlations of monthly realized volatility across markets. The darkness of the shading represents the degree of correlation.

Table 3
Risk premiums for financials and nonfinancials, and their difference.

		Financials	Nonfinancials	Difference
Factor model	RV	-0.18	-0.29	0.11
		[-0.63]	[-1.62]	[0.34]
	IV	0.10	0.43	-0.32
		[0.47]	[2.82]	[-1.28]
Replicating port.	rv	-0.25	-0.30	0.05
		[-1.13]	[-2.14]	[0.18]
	iv	-0.02	0.34	-0.36
		[-0.14]	[2.95]	[-2.02]

Note: The table reports the average risk premiums for RV and IV risks, across financials (first column), across nonfinancials (second column) and for the difference between the two groups (third column), with corresponding t-statistics in square brackets. The top panel estimates the risk premiums using the linear factor model; the bottom panel estimates the risk premiums as the average excess returns of the rv and iv portfolios.

average, since they are scaled by their sample variance. That said, the test rejects in only 3 of the 19 markets. The p-value for the S&P 500 is 0.22, similar to the one obtained by Constantinides et al. (2013). The fact that the model is rejected for only 1 of the 14 nonfinancials suggests that the results for nonfinancials, where the differences in the pricing of implied and realized volatility are most pronounced, should be most reliable. The test rejects for two of the five financial underlyings, which implies that they are more likely to have specification error.

3.3. Interpretation of the results

How can realized volatility have a negative price of risk, while uncertainty has a positive risk price? Key to understanding this distinction is noticing that realized volatility (which is computed by squaring shocks) is strongly dominated by large price movements like jumps, which our empirical results suggest tend to be bad for investors on average. So it is easy to see how investors might dislike realized volatility, as it captures the occurrence of a large, bad shock.

On the other hand, innovations in implied volatility are driven by changes in the perceived uncertainty about good and bad potential events: so a higher probability of a bad jump will increase uncertainty, but a higher probability of a good event (e.g., a new technology) will also increase uncertainty. Our results show that on net, investors seem to perceive increases in uncertainty as being associated with good states of the world.

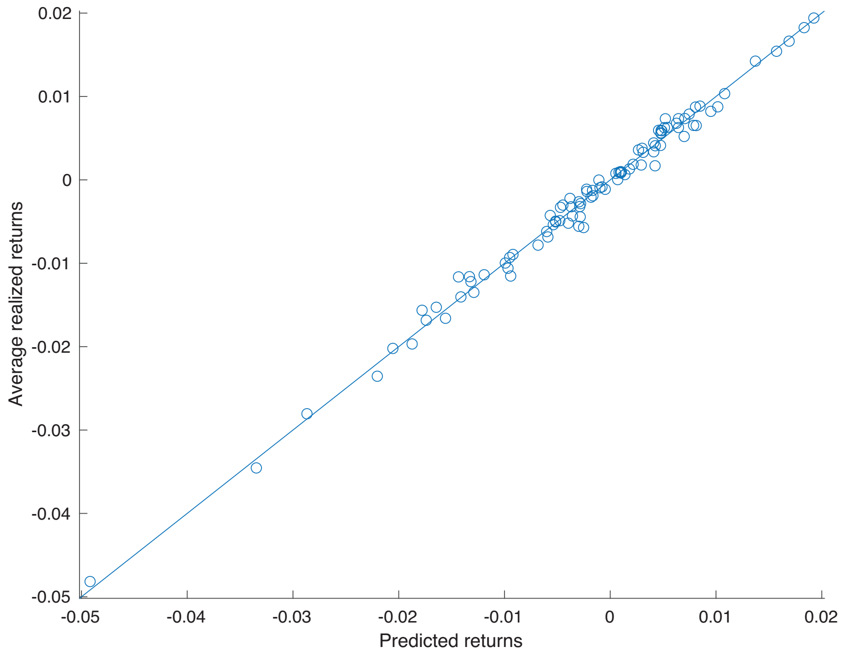

Fig. 4. Cross-sectional fit of factor models. For each straddle of maturity one to five months, and for each of the 19 markets, the figure reports the predicted risk premium against the realized average excess return. Predicted risk premiums are obtained estimating a linear factor model separately in each market.

Section OA.8 in the Online Appendix formalizes this idea, describing a simple extension of the standard long-run risk model of Bansal and Yaron (2004) that is consistent with our results on the pricing of both volatility and uncertainty shocks.

Finally, it is valuable to compare our analysis with some closely related past work. As discussed above, both Constantinides et al. (2013, CJS) and Cremers et al. (2015, CHW) also examine the pricing of uncertainty and realized volatility in the S&P 500 using factor models. While we cannot compare our full range of results with theirs, we can at least see how those for the S&P 500 compare.

The analysis of CJS is closest to us, as they also use option portfolios as test assets. In Table 8, they report a premium of approximately zero for shocks to uncertainty and a large negative premium for realized volatility for the S&P 500. So consistent with our findings, they find much stronger pricing of realized than implied volatility, though their uncertainty premium is less positive. CHW, instead, use the cross-section of equities as their test assets and find a more strongly negative premium for uncertainty. However, they also report returns on an uncertainty hedging portfolio, which aligns very closely with our analysis in the next section (see their Table 1). In that case, their results are quantitatively highly similar to ours. We discuss this observation further below.

3.4. Is realized volatility about jumps? evidence from strangles

Similar to others such as Cremers et al. (2015), we have argued thus far that the exposures to squared returns on the underlying (or gamma) represent exposure to jump risk. While CHW focuses on straddles, we further test the hypothesis that the premiums are for jumps by examining returns on strangles. A strangle is, like a straddle, a portfolio long a put and a call, with the delta set to zero here by construction. However, in the case of a strangle, the two options are out of the money, with different strikes, rather than both having the same strike. So whereas the final payoff of a straddle depends on the absolute value of the change in the underlying, a strangle only pays off if the underlying moves sufficiently far from its initial value, with that required distance being a choice variable.

We examine returns on strangles where the put and call strikes are one standard deviation unit (scaling by time to maturity) from the forward price when the portfolio is formed, so they only have positive payoffs at maturity if the underlying moves further than that. As with the straddles, we examine two-week returns.

Fig. 5 replicates Fig. 3 for the case of strangles. For the uncertainty risk premiums, the results are qualitatively and quantitatively similar to those for straddles: for financials the premium is close to zero, and for nonfinancials it is 0.42.

It is for the RV/gamma risk premiums that we find a substantial difference, representing our third main result. Across the various markets, the premiums are generally twice as large for strangles as for straddles. Every single point estimate is now negative, and only one confidence band contains zero. For financial underlyings, the average premium is now statistically significant, at -1.54. For nonfinancials and all assets combined, the means are both -1.48 and -1.5, respectively.

These results provide clear evidence that it is really the tail of the distribution that drives the RV results. The find-

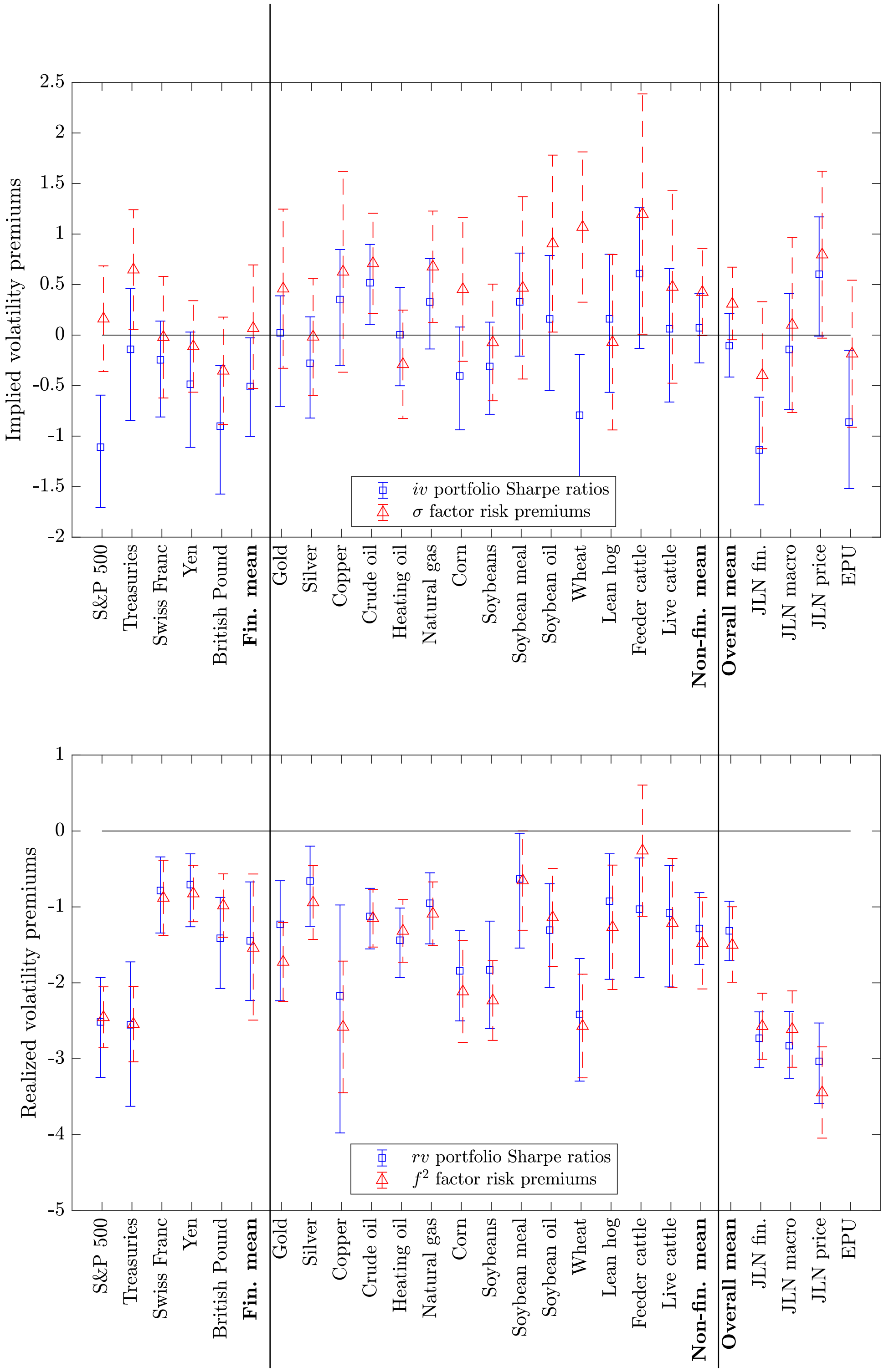

Fig. 5. RV and IV portfolio Sharpe ratios and factor risk premiums: strangles. See Fig. 3. This figure differs only in replacing the straddles with 1-sigma strangles.

ing that deep out-of-the-money options have the largest premiums is well known for the S&P 500. This paper is novel for showing that the relation of the gamma premium with moneyness in fact holds across all the markets that we study, and is strikingly different from the patterns on uncertainty.

To sum up, Figs. 3 and 5 contain our three main results. Pervasively across markets, premiums related to vega (uncertainty) are zero or positive, while premiums for gamma (jump risk) are significantly negative. Furthermore, the jump risk premiums are largest for out-of-the-money options. Economically, the results show that it is periods with extreme shocks – realized volatility or jumps – that investors are averse to, rather than simple increases in forward-looking uncertainty.

4. Theoretical risk exposures of straddles and strangles

We argued heuristically above that straddles and strangles are natural test assets for a factor model involving realized and implied volatility since they have zero delta and payoffs that are convex in the underlying return. This section formalizes that intuition by calculating the theoretical exposures of options of different maturities to those shocks, following the analysis of Cremers et al. (2015). Similar to their analysis, we then show that we can construct replicating portfolios that, under the theory, should provide direct exposure to shocks to either implied or realized volatility. Formally, under the Black-Scholes model, one portfolio has positive vega and zero gamma, and the other has positive gamma and zero vega. These portfolios give an alternative, and in some sense more direct, way of measuring the risk premiums.

4.1. Return exposures

The exposures of the portfolios studied above to the risk factors we use in our linear factor model can be approximated theoretically using the Black-Scholes model, as in Coval and Shumway (2001), Bakshi and Kapadia (2003), and Cremers et al. (2015). Online Appendix OA.3 shows that the partial derivatives of the zero-delta straddle and strangle return with respect to the underlying futures return, f, its square, and the change in volatility, can be approximated as

$$\frac{\partial r_{n,t}}{\partial f_t} \approx 0, \tag{4}$$

$$\frac{\partial^2 r_{n,t}}{\partial (f_t/\sigma_{t-1})^2} \approx n^{-1}, \tag{5}$$

$$\frac{\partial r_{n,t}}{\partial (\Delta\sigma_t/\sigma_{t-1})} \approx 1, \tag{6}$$

where $r_{n,t}$ is the return on date t of a straddle or strangle with maturity n, f_t is the return on the underlying future, σ_t is the implied volatility of the underlying, and Δ is the first-difference operator.[19]

It is perhaps surprising at first that the exposures are the same for both straddles and strangles. Intuitively, the two types of portfolios have the same exposures up to the second order. Where they differ is in their higher-order exposures, which are naturally larger for the strangles. The first partial derivative says that the straddles and strangles have close to zero local exposure to the futures return. The second line says that the exposure of the options to squared returns on the underlying (realized volatility) is approximately inversely proportional to time to maturity. The third line shows that they are also exposed to changes in expected future volatility, through $\frac{\Delta\sigma_t}{\sigma_{t-1}}$, and that exposure is approximately constant across maturities.

To see how the risk exposures differ in their higher order terms, Fig. OA.4 in the Online Appendix plots the return on a straddle and a one standard-deviation strangle as a function of the change in the price of the underlying. It is apparent that the two curves are not just tangent at zero, but that they have the same curvature, consistent with having the same second derivative, as in Eq. (5). They only begin to differ noticeably as the returns get extreme. So straddles and strangles have equal local exposures to the underlying, but in the tails, e.g., in response to jumps, strangles become more sensitive. This shows why strangle returns help isolate the extra premium earned for exposure to tail risk.

4.2. Replicating portfolios

Cremers et al. (2015) show that the implied sensitivities in (4)–(6) give a method for constructing portfolios that the Black-Scholes model says should give exposures only to realized volatility, as expressed by $(f_{n,t}/\sigma_{t-1})^2$, or implied volatility, measured by $\Delta\sigma_t/\sigma_{t-1}$. The method is to construct, for each market, two portfolios,

$$rv_{i,t} = \frac{5}{24}(r_{i,1,t} - r_{i,5,t}) \approx (f_t/\sigma_{t-1})^2, \tag{7}$$

$$iv_{i,t} = \frac{5}{4}r_{i,5,t} - \frac{1}{4}r_{i,1,t} \approx \Delta\sigma_t/\sigma_{t-1}. \tag{8}$$

where the approximations follow from Eqs. (4)–(6).[20] Throughout this section, capitalized *RV* and *IV* refer to the levels of realized and implied volatility, while lower-case rv and iv refer to the associated portfolio returns. We use the one- and five-month options to construct the portfolios, since it is exactly five-month implied volatility that is priced in the main analysis. The iv portfolio is dominated by an investment in the five-month options, with just a small short position in the one-month options. In that sense, the iv portfolio is a rather direct claim on exactly the implied volatility priced in the factor model.

The purpose of constructing these portfolios is to give a simple and direct method of measuring the premiums associated with realized and implied volatility that does not

[19] We ignore here the fact that options at different maturities have different underlying futures contracts. If that elision is important, it can be expected to appear as a deviation of the estimated factor loadings from the predictions of the approximations (4)–(6).

[20] Note that Eq. (5) gives the second derivative, which has weight 1/2 in the Taylor approximation. So the loading on the squared future return for a straddle of maturity n is $(2n)^{-1}$, which implies that the coefficient for Eq. (7) is 5/24.

require full estimation of the factor model. If the loadings used to construct the portfolios are correct, this method will also be more efficient. On the contrary, if the assumptions of the model are not correct, then the results will be biased (whereas the factor model will still be correct, as it estimates the risk exposures instead of using the ones implied by the model). There is thus a bias/variance trade-off between the factor model, which requires fewer assumptions but will have greater estimation error, and the replicating portfolios, which require stronger assumptions but will have less estimation error.

The key concern, then, is how accurate the Black-Scholes-implied loadings are. Fig. OA.2 and Table OA.2 in the Online Appendix show that the theoretical predictions for the loadings are fairly accurate (though not perfect) empirically. Online Appendix OA.3 also examines the accuracy of the Black-Scholes approximation for returns in a simulated setting.

Table OA.2 shows that the biggest deviations from the model-implied loadings are for the S&P 500 *iv* portfolio. In that case, there is a large positive loading on realized volatility – a GARCH effect – and a large negative loading on the underlying futures return – the leverage effect. Both should be expected to bias the return on the *iv* portfolio down relative to the estimated implied volatility factor loading from above. The effects are three times larger for the S&P 500 than for any other market. That suggests that for measuring pricing of S&P 500 uncertainty, in particular, it is best to use the factor model, as in Constantinides et al. (2013). For all other markets, instead, the Black-Scholes assumptions appear relatively accurate, so we would expect the results to line up well with those of the factor model.

Note that even though the *rv* and *iv* portfolios theoretically load on separate risk factors, they need not be uncorrelated. It is well known from the GARCH literature, for example Engle (1982) and Bollerslev (1986), that in many markets, innovations to realized volatility are correlated with innovations to implied volatility. Table 4 reports the correlations between the *rv* and *iv* returns in the 19 markets. GARCH effects appear most strongly for the financial underlyings and precious metals, for which the average correlation is 0.44. For the other nonfinancial underlyings, the effects are much smaller, and the correlation between the *rv* and *iv* returns is only 0.03 on average (it is 0.09 on average across all nonfinancials). So for the nonfinancials, innovations to realized and implied volatility returns are essentially independent on average. These weak correlations are valuable for the identification, since they show that surprises in realized and implied volatility are far from the same and can be hedged separately using the *rv* and *iv* portfolios.

4.3. Risk premiums

4.3.1. Straddles

The solid blue series in the two panels of Fig. 3 report annualized Sharpe ratios for the *rv* and *iv* portfolios constructed from straddles in the 19 markets. As with the factor model, we begin by focusing on the straddle returns because they use more liquid near-the-money options.

The results in Fig. 3 for the *rv* and *iv* portfolios are highly similar to those for the factor model. The *iv* portfolios earn returns close to zero on average for the financial underlyings and returns that are consistently positive for the nonfinancial underlyings. For the nonfinancials, the average Sharpe ratio for the *iv* portfolios is again statistically significantly positive. As expected, since the *iv* portfolios are formed using stronger assumptions, the standard errors for the risk premiums are tighter than for the factor model.

The bottom panel of Table 3 summarizes the estimates for the realized and implied volatility risk premiums for financials and nonfinancials computed using the *rv* and *iv* portfolios, and also reports tests for whether the two are different. In all cases, the premiums for the financials are insignificant while those for the nonfinancials are significant. However, note that there are fewer financial underlyings, limiting our statistical power. The difference between financials and nonfinancials itself is not significant, so we cannot actually say that there is strong evidence for a difference between the two in three out of four cases. The only case where the difference is statistically significant is for the Sharpe ratio on the *iv* portfolio.

That difference appears to be driven largely by the fact that the return on the S&P 500 *iv* portfolio is very different from the estimated risk premium for implied volatility from the factor model. In fact, the confidence bands do not even overlap. This result is driven by the fact that there are much stronger GARCH effects in the S&P 500 than the other underlyings that we study, creating a bias, as discussed above (see Table OA.2 showing that the S&P 500 *iv* portfolio actually loads strongly on realized volatility). We thus place relatively less trust in the results from the *rv* and *iv* portfolios (as opposed to the results from the factor model) for the S&P 500 than the other underlyings, for which there is very strong agreement between the factor model and the *iv* portfolio returns. Even in the case of the S&P 500, though, the premium for uncertainty shocks is not statistically significantly negative.

The Sharpe ratios for the *rv* portfolios are also highly similar to the estimated risk premiums on realized volatility in the factor model (even for the S&P 500). The financial underlyings other than the S&P 500 again have premiums generally close to zero, while the S&P 500 and the nonfinancials have consistently negative premiums.

The returns on the *rv* and *iv* portfolios for the S&P 500 can be compared to those reported in Table 1 of Cremers et al. (2015). For their analog to our *rv* portfolio, they obtain a Sharpe ratio of -0.9, compared to -1.2 in our case, while for their analog to the *iv* portfolio, they report a Sharpe ratio of -0.5, compared to -0.2 here. In both cases, the confidence bands for our estimates easily contain theirs. We thus obtain substantial agreement with the findings of CHW for returns on option portfolios. Our results differ from theirs in two key ways. First, we focus on factor models using options as test assets, instead of equities. We choose to use options, similar to Constantinides et al. (2013), because they have risk exposures very directly tied to uncertainty and volatility, whereas equity returns have many other risk exposures that have been explored in the literature. Second, obvi-

Table 4
Correlations between *rv* and *iv* portfolio returns in each market.

	Std(rv)	Std(iv)	Corr(rv,iv)
S&P 500	0.03	0.08	0.48
T-bonds	0.03	0.08	0.01
CHF	0.04	0.08	0.63
JPY	0.04	0.08	0.61
GBP	0.04	0.07	0.41
Gold	0.04	0.12	0.48
Silver	0.04	0.08	0.45
Copper	0.03	0.10	0.03
Crude Oil	0.04	0.09	0.05
Heating oil	0.04	0.08	0.01
Natural gas	0.04	0.08	-0.17
Corn	0.04	0.08	0.06
Soybeans	0.04	0.09	0.17
Soybean meal	0.04	0.11	0.20
Soybean oil	0.04	0.09	0.21
Wheat	0.04	0.08	0.08
Lean hog	0.05	0.10	-0.24
Feeder cattle	0.05	0.10	0.03
Live cattle	0.04	0.08	-0.12

Note: The table reports, for each underlying, the standard deviation of the two-week returns to the rv and iv portfolios, and their correlation.

ously, we explore the pricing of options in a wide range of markets, not just the S&P 500.

4.3.2. Strangles

The results for strangles are again consistent with those for straddles, but more extreme. In Fig. 5, as in Fig. 3, the point estimates and confidence bands from the factor model (red) and the *rv* and *iv* portfolios (blue) are similar, with the model-based *rv* and *iv* portfolios again having narrower confidence intervals, showing that the results are robust to the estimation method.

We again find that the strangles have much more negative jump/gamma premiums than the straddles. Since we showed above that the exposures of the strangles and straddles are the same up to second order, this section clearly indicates that it is the difference in higher order exposures of the different strategies that drives the larger premiums for strangles.

4.4. Summary

The results in this section are useful for three reasons. First, they show that our results are not driven by some hidden detail of the factor model estimation. The *rv* and *iv* portfolios are simple to construct and yield highly similar results to the factor model, both for straddles and strangles. So the three key findings, zero or positive premiums for uncertainty, substantially negative premiums for realized volatility, and even larger premiums for realized volatility for strangles, appear to be robust.

Second, the replicating portfolios help clarify exactly what the source of identification is in the factor model. The options have exposures to implied and realized volatility that differ across maturities, so including a panel of multiple maturities allows us to separately measure their premiums.

Finally, by analyzing the risk exposures of the options, we can link the factor model estimates back to widely studied and applied features of options – their greeks. The estimate of the price of shocks to implied volatility from the factor model is essentially identical to the Sharpe ratio on a portfolio with positive vega and zero gamma, while the estimate of the price of shocks to realized volatility is almost the same as the Sharpe ratio on a portfolio with positive gamma and zero vega.

4.5. Combined portfolios

As we discussed in Section 2.3, the uncertainty in our 19 markets is related to various measures of aggregate uncertainty. It is then natural to ask what the cost of hedging is for aggregate uncertainty. A simple way to do that is to buy all the *iv* or *rv* portfolios simultaneously. We focus on just the straddles here since they are most liquid and thus most feasible for an investor to hold. Since Tables 1 and 2 show that realized and implied volatility are imperfectly correlated across markets, even larger Sharpe ratios can be earned by holding portfolios that diversify across the various underlyings. Table 5 reports results of various implementations of such a strategy. Looking first at the top panel, the first row reports results for portfolios that put equal weight on every available underlying in each period, the second row uses only nonfinancial underlyings, and the third row only financial underlyings. The columns report Sharpe ratios for various combinations of the *rv* and *iv* portfolios. The first two columns report Sharpe ratios for strategies that hold only the *rv* or only the *iv* portfolios, the third column uses a strategy that is short *rv* and long *iv* portfolios in equal weights, while the final column is short *rv* and long *iv*, but with weights inversely proportional to their variances (i.e., a simple risk parity strategy).

Table 5
Portfolios of rv and iv across markets.

Panel A: Sharpe ratios			rv+iv	
	rv	iv	Equal weight	Risk-parity
All underlyings	-0.74 ***	0.49 **	1.05 ***	0.90 ***
Nonfinancials	-0.63 ***	0.62 ***	0.91 ***	0.90 ***
Financials	-0.37 **	-0.04	0.42 ***	0.13

Panel B: Skewness			rv+iv	
	rv	iv	Equal weight	Risk-parity
All underlyings	1.23 ***	1.82 ***	-0.79 ***	1.05 ***
Nonfinancials	2.11 ***	1.55 ***	-2.00 ***	0.75 ***
Financials	2.01 ***	2.91 ***	-1.40 ***	2.19 ***

Note: Sharpe ratios and skewness of portfolios combining rv and iv portfolios across markets. For each panel, the first row reports a portfolio constructed using straddles from all available markets on each date, the second row using only nonfinancial underlyings, the third row only financial underlyings. Each column corresponds to a different portfolio. The first column is an equal-weighted RV portfolio, the second is an equal-weighted IV portfolio, the third is an equal-weighted long-short IV minus RV portfolio, and the last is the same long/short portfolio but weighted by the inverse of the variance (risk-parity). *** indicates significance at the 1% level, ** the 5% level, and * the 10% level.

The Sharpe ratios reported in Table 5 are generally larger than those in Fig. 3. The portfolios that are short rv and long iv are able to attain Sharpe ratios above one. The largest Sharpe ratios come in the portfolios that combine rv and iv, which follows from the fact that they are positively correlated, so going short rv and long iv leads to internal hedging. All of that said, these Sharpe ratios remain generally plausible. Values near one are observed in other contexts, for example, Broadie et al. (2009) for put option returns, Asness et al. (2013) for global value and momentum strategies, and Dew-Becker et al., (2017) for variance swaps.

The portfolios that take advantage of all underlyings simultaneously seem to perform best, presumably because they are the most diversified. While holding exposure to implied volatility among the financials earns effectively a zero risk premium, it is still generally worthwhile to include financials for the sake of hedging.

Finally, the bottom panel of Table 5 reports the skewness of the various strategies from above. One might think that the negative returns on the rv portfolio are driven by its positive skewness, but the iv portfolio also is positively skewed and has positive average returns. So the degree of skewness does not seem to explain differences in average returns in this setting.

5. Hedging uncertainty indexes

The results so far give the cost of directly hedging shocks in commodity markets. This section examines how options can be used to hedge shocks to macro uncertainty indexes. Section 2.3 showed that the commodity IVs do a good job of spanning the macro uncertainty indexes. We now discuss those indexes in more detail and examine the cost of hedging both the implied and realized parts of macro volatility.

The JLN index is developed in a pair of papers by Jurado et al. (2015) and Ludvigson et al. (2015). We follow their construction methodology and further extend it to yield separate measures of uncertainty that pertain to financial markets, real activity, and goods prices, with the latter two also being combined into an overall macroeconomic uncertainty index.[21] The goal of the JLN framework is to estimate uncertainty on each date, σ_t^2. The method can also be extended to create a realized volatility index.[22] We refer to the JLN uncertainty indexes by *JLNU* and realized volatility indexes by *JLNRV*.

The Economic Policy Uncertainty (EPU) index of Baker et al. (2016) is constructed based on media discussion of uncertainty, the number of federal tax provisions changing in the near future, and forecaster disagreement. Unlike JLN, there is no distinction in this case between volatility and uncertainty, so we treat EPU as measuring only uncertainty.

[21] The construction involves two basic steps. First, realized squared forecast errors are constructed for 280 macroeconomic and financial time series, of which 134 macro series are from McCracken and Ng (2016), while the remaining financial indicators are from Ludvigson and Ng (2007). Our analysis uses code from the replication files of JLN. The macro price series are defined as those referring to price indexes, and the real series are the remainder of the macro time series. Denoting the error for series i as $\varepsilon_{i,t}$, there is a variance process, $\sigma_{i,t}^2 = E\left[\varepsilon_{i,t}^2\right]$. So $\varepsilon_{i,t}^2$ constitutes a noisy signal about $\sigma_{i,t}^2$. JLN then estimate $\sigma_{i,t}^2$ from the history of $\varepsilon_{i,t}^2$ using a two-sided smoother and create an uncertainty index as the first principal component of the estimated $\sigma_{i,t}^2$. For the component indexes, we take the first principal component of the $\sigma_{i,t}^2$ corresponding to the relevant group of indicators.

[22] This is done by taking the first principal component from the cross-section of the $\varepsilon_{i,t}^2$ in a given month, instead of the $\sigma_{i,t}^2$.

Fig. 2 shows that the 19 IVs span most of the variation in the JLN and EPU uncertainty indexes. We can then measure risk premiums associated with those indexes by constructing hedging portfolios using our straddles. For each index, we obtain the weights for the hedging portfolio from the coefficients of the projection we presented in Section 2.3. Specifically, for each uncertainty index j, we estimate the regression

$$JLNU_t^j = a + \sum_i b_i^j IV_{i,t} + \varepsilon_{j,t} \tag{9}$$

We then use the risk premiums estimated in the factor model to calculate a premium for hedging the *JLN* indexes. In particular, we construct a hypothetical portfolio that has exposure b_i^j to $\Delta IV_{i,t}/IV_{i,t-1}$. The mean return on that portfolio can be calculated from Eq. (3), while the standard deviation is obtained from the covariance matrix of $\Delta IV_{i,t}/IV_{i,t-1}$ across i (again weighting by b_i^j). The same method also yields a risk premium for the EPU and *JLNRV* indexes (see Online Appendix Fig. OA.1 for the analog of Fig. 2 for realized volatilities).

The right-hand section of Fig. 3 (red lines) reports the Sharpe ratios for straddle portfolios hedging the EPU and JLN indexes, computed using the estimates from the factor models. Since those hedging premiums are constructed combining the individual factor premiums, it is not surprising that they are similar. In all three cases, the risk premium for JLN indexes (financial, macro, and price uncertainty) is positive, in one case statistically significantly. Furthermore, the confidence bands rule out economically large negative premiums: the lowest confidence band only runs to -0.32. For EPU we find a point estimate of approximately zero (-0.03), though a confidence band that runs to -0.49.

The right-hand section of the bottom panel of Fig. 3 reports the returns from the JLN realized volatility hedging portfolios (again, the red lines use the risk premiums estimates from the factor model). Again, consistent with the fact that the RV risk premiums themselves are consistently negative, hedging the JLN indexes for realized volatility historically has a positive cost. For all three subindexes, the risk premiums are very negative, with the Sharpe ratios for financial, real, and price volatility at -1.15, -0.62, and -0.65, respectively, all three of which are statistically significant. So the conclusions from hedging the JLN and EPU indexes are highly similar to those in the main analysis, providing further evidence that in the macroeconomy, it is realized volatility that is priced, rather than uncertainty about the future. The blue lines in the figure, which use the estimates from the *rv* and *iv* portfolio, show similar results, with the uncertainty Sharpe ratios slightly lower but still statistically indistinguishable from zero, and the realized volatility premiums strongly negative. Fig. 5 shows that the results for straddles are again similar, with hedging realized volatility in this case again carrying a more negative premium.

6. Robustness

This section examines some potential concerns about the robustness of the results.

6.1. One-week holding period returns

Our main analysis is based on two-week holding period returns for straddles, which strike a balance between having more precise estimates of risk premiums and reducing the impact of measurement error in prices. We have repeated all of our analysis using one-week holding period returns, and find very similar results. Online Appendix Fig. OA.6 is the analog of Fig. 3, but constructed using one-week returns. The results are qualitatively and quantitatively similar to the baseline.

6.2. Split sample and rolling window results

To address the concern that the results could be driven by outliers (though note that there would need to be outliers in all 19 markets), Figs. OA.7 and OA.8 replicate the main results in Fig. 3, but splitting the sample in half (before and after June 2000). The confidence bands are naturally wider, and the point estimates vary more from market to market in the two figures. Nevertheless, the qualitative results are the same as in the full-sample case, showing that realized volatility earns a negative premium while the premium on implied volatility is positive.

To further evaluate the possibility that the results are driven by a small number of observations, Fig. OA.9 plots Sharpe ratios for the *rv* and *iv* portfolios in five-year rolling windows for each of the 19 markets, as well as for the equal-weighted portfolios of all 19 markets. The sample Sharpe ratios are reasonably stable over time. In no case do the results appear to be driven by a single outlying period or episode. Note that these results are not informative about variation in the conditional risk premium; with a five-year window, the standard error for the Sharpe ratios is 0.45, so even if the true conditional Sharpe ratios are constant, the five-year rolling estimates should display large amounts of variation over time.

6.3. Alternative maturities

Our main results use the five-month maturity for implied volatility, both in the factor model and as the second leg in the *rv* and *iv* portfolios. Fig. OA.10 in the Online Appendix replicates the analysis using two-month implied volatility instead in both cases. The results are qualitatively and quantitatively similar to the main specification. Note that the GARCH effects that bias the estimates for the *iv* portfolio risk premium (blue) in the top panel downward relative to the estimates from the factor model (red) are stronger when using two-month IV instead of five-month IV (see the loadings of the *iv* portfolio on realized volatility in Table OA.4).

To help understand why the maturity choice does not have strong effects, the top panel of Table OA.3 in the Online Appendix reports loadings of the *rv* portfolio on changes in implied volatility at maturities of one to five months. In all cases, the coefficients are close to zero – no larger than 0.1 – indicating that the exposures to implied volatility at any maturity are economically small, especially in comparison to the loading on realized volatility, which can be seen from Table OA.2 to be closer to one.

The bottom panel shows the same loadings, but for the RV-hedging portfolio built using the factor model. By construction, this portfolio has loading one on RV and zero on five-month IV, as the last column of the table highlights: see Online Appendix Section OA.2 for more details.

6.4. Weighted least squares

Johnson (2019) argues that there can be efficiency gains from weighting by implied volatility in estimating risk premiums. We explore that in Fig. OA.11 in the Online Appendix, which reports the risk premiums (computed with the factor model) with and without weighting by implied volatility. Weighting drives most of the risk premiums to be less negative or more positive, but the patterns all remain qualitatively and quantitatively similar. The premium for implied volatility shocks becomes even more strikingly positive.

6.5. Pricing the independent parts of realized and implied volatility

The main results above report returns associated with assets that hedge innovations to realized and implied volatility. Table 4 shows that those returns are positively correlated: months with increases in realized volatility also tend to have increases in implied volatility. A natural question is what would happen if we were to construct a portfolio that loaded on the independent part of those returns, e.g., an increase in implied volatility holding realized volatility fixed. Section OA.6 in the Online Appendix reports an SDF-based analysis that prices the independent components and shows that the results are similar to the main specification (see Fig. OA.12).

6.6. Oil and gas equity options

Since the stock returns of firms in the energy sector are naturally exposed to changes in energy prices, it is natural to ask whether returns on their options behave similarly to what we report for oil and gas futures options. We obtain data from Optionmetrics on firms with an Optionmetrics industry code between 120 and 125, corresponding to the energy sector. We then construct *rv* and *iv* portfolios for those firms using the same methods as for the main analysis, again with maturities of one and five months. We construct two-week returns and sum them across whatever firms are available on each date, weighting by market capitalization. The Optionmetrics data covers the period 1996–2018.

	Sharpe ratio
rv	-0.56
95% CI	[-1.02,-0.10]
iv	0.05
95% CI	[-0.42,0.52]

The Sharpe ratios for the *rv* and *iv* portfolios for oil and gas companies are shown above. Similar to the main results, we obtain a significantly negative premium on realized volatility and a marginally positive premium on implied volatility. The premium for the *iv* portfolio for oil and gas companies is less positive than for crude oil futures options, but more positive than for S&P 500 index options. In other words, the results imply that options on oil and gas companies behave as though they are a mixture of options on the S&P 500 and on crude oil, which is not an unrealistic desciption of oil and gas companies.

Because of the relatively short sample compared to the main results, this analysis has relatively low power. The point estimate for *rv* is outside the confidence band for *iv* and vice versa, but their confidence bands do overlap and the Sharpe ratios are not statistically significantly different from each other. That also illustrates the benefit in the main analysis of using information from many different markets to help increase estimation power. Nevertheless, the results in this section are consistent with our main findings, if statistically weaker. Section OA.7 further extends these results by examining options on energy sector ETFs and finds similar results.

6.7. Liquidity

If the options used here are highly illiquid, the analysis will be substantially complicated for three reasons. First, to the extent that illiquidity represents a real cost faced by investors, such as a bid/ask spread, then returns calculated from settlement prices do not represent returns earned by investors. Second, illiquidity itself could carry a risk premium that the options might be exposed to. Third, bid/ask spreads represent an added layer of noise in prices. The identification of the premiums for realized volatility and uncertainty depends on differences in returns on options across maturities, so what is most important for our purposes is how liquidity varies across maturities. This section shows that the liquidity of the straddles studied here is generally highly similar to that of the widely studied S&P 500 contracts traded on the CBOE, and the liquidity does not appear to substantially deteriorate across maturities. It is important to note that measuring trading costs is nontrivial, especially for complex orders, and bid/ask spreads are not necessarily the best measure of the true cost of liquidity. See Muravyev and Pearson (2020) for a detailed analysis.

While a long history of bid/ask spreads is not available to us, we obtained posted bid/ask spreads for the options closest to the money on Friday, 8/4/2017 for our 19 contracts plus the CBOE S&P 500 options at maturities of one, four, and seven months.[23] Those spreads are plotted in Fig. OA.13. For the majority of the options, the spreads are less than 3%, consistent with the 4.1% bid/ask spread for one-month S&P 500 options at the CBOE. Across nearly all the contracts, the posted spreads again decline with maturity, and for 10 of the 19 contracts the one-month posted spreads are nearly indistinguishable from that for the S&P 500, which is typically viewed as a highly liquid market and where incorporating bid/ask spreads generally has minimal effects on return calculations (Bondarenko, 2014).

[23] Longer histories of bid/ask spreads for options are available for purchase from the CME (at significant cost), which would enable these results to be extended.

Note that the decline with maturity is relative to the price of the options themselves, not in absolute terms.

Fig. OA.13 yields two important results. First, it shows that the liquidity of the straddles is reasonably high, in the sense that posted spreads are currently relatively narrow in absolute terms for most of the contracts and that they compare favorably with spreads for the more widely studied S&P 500 options traded at the CBOE. Second, liquidity does not appear to deteriorate as the maturity of the options grows, and in fact in many cases there are improvements with increasing maturities, again consistent with CBOE data.

Section OA.3.5 in the Online Appendix reports statistics for volume across maturities, showing that the markets are generally fairly similar. Section OA.3.6 reports an additional robustness test that measures returns using a method that is robust to certain types of measurement errors in prices, showing that the main results are essentially identical.

Finally, it is useful to note that while the liquidity of option markets changed significantly in the past 30 years, the patterns in risk premiums for the *rv* and *iv* portfolios appear stable over time (see, for example, the rolling Sharpe ratios of Fig. OA.9), suggesting that liquidity is not the main driver of our results.

Even though the liquidity is similar across many of the markets, one might still ask how trading costs affect the returns we have been studying. Any trading cost will lower the return of a portfolio, regardless of whether an investor is long or short. By studying returns based on settlement prices, we are essentially looking at the return averaged across what the buyer and seller receive. For example, if the return on a portfolio based on settlement prices is 10% and there are total trading costs to each side of 1%, then the buyer earns a return of 9% while the seller has a loss of 11%. Looking at prices is therefore natural for illustrating the return that the average investor sees.

7. Conclusion

This paper studies the pricing of uncertainty and realized volatility across a broad array of options on financial and commodity futures. Uncertainty is proxied by implied volatility, which theoretically measures investors' conditional variances for future returns, and a number of uncertainty indexes developed in the literature. Realized volatility, on the other hand, measures how large realized shocks have been. In modeling terms, if $\varepsilon_{t+1} \sim N\left(0, \sigma_t^2\right)$, uncertainty is σ_t^2, while volatility is the realization of ε_t^2.

A large literature in macroeconomics and finance has focused on the effects of uncertainty on the economy. This paper explores empirically how investors perceive uncertainty shocks. If uncertainty shocks have major contractionary effects so that they are associated with high marginal utility for the average investor, then assets that hedge uncertainty should earn negative average returns. On the other hand, the finance literature has recently argued that in many cases uncertainty can be good. For example, during the late 1990s, it may have been the case that investors were not sure about how good new technologies would turn out to be.

The contribution of this paper is to construct hedging portfolios for a range of types of macro uncertainty, including interest rates, energy prices, and uncertainty indexes. The cost of hedging uncertainty shocks reveals the relative importance of good and bad types of uncertainty. Furthermore, using a wide range of options is important for capturing uncertainty about the real economy and inflation, as opposed to just about financial markets. The empirical results imply that uncertainty shocks, no matter what type of uncertainty we look at, are not viewed as being negative by investors, or at least not sufficiently negative that it is costly to hedge them. Financial uncertainty appears to be roughly equally split between the good and bad types, while nonfinancial uncertainty is relatively more strongly driven by good uncertainty – the cost of hedging nonfinancial uncertainty shocks is negative.

What is highly costly to hedge is realized volatility. Portfolios that hedge extreme returns in futures markets and hence large innovations in macroeconomic time series earn strongly negative returns, with premiums that are in many cases one to two times as large as the premium on the aggregate stock market over the same period. So what is consistently high in bad times is not uncertainty, but realized volatility. Periods in which futures markets and the macroeconomy are highly volatile and display large movements appear to be periods of high marginal utility, in the sense that their associated state prices are high. This is consistent with (and complementary to) the findings in Berger et al. (2020), who provide VAR evidence that shocks to volatility predict declines in real activity in the future, while shocks to uncertainty do not.

Berger et al. (2020) show that the VAR evidence and pricing results for realized volatility are consistent with the view that it is downward jumps in the economy that investors are most averse to. They show that a simple model in which fundamental shocks are both stochastically volatile and negatively skewed can quantitatively match the pricing of uncertainty and realized volatility, along with the VAR evidence. Similarly, Seo and Wachter (2018a); Seo and Wachter (2018b) show that negative skewness can explain the pricing of credit default swaps and put options. This paper thus also contributes to the growing literature studying the effects of skewness. In a world where fundamental shocks are negatively skewed, the most extreme shocks – those that generate realized volatility – tend to be negative, which can explain why realized volatility would be so costly to hedge.

Appendix A. Data filters and transformations

The observed option prices very often appear to have nontrivial measurement errors. This section describes the various filters we use and then provides more information about the specifics of the data transformations we apply. Code is available on request.

First, we note that the price formats for futures and strike prices for many of the commodities change over time. That is, they will move between, say, 1/8ths, 1/16ths, and pennies. We make the prices into a consistent decimal time series for each commodity by inspecting the prices directly and then coding by hand the change dates.

We then remove all options with the following properties:

1. Strikes greater than five times the futures price
2. Options with open interest below the fifth percentile across all contracts in the sample
3. Price less then five ticks above zero
4. Maturity less than nine days
5. Maturity greater than eight months
6. Options with prices below their intrinsic value (the value if exercised immediately)

Note that in our baseline results, we do not remove options for which we have no volume information, or for which volume is zero. However, we have reproduced our main analysis (Fig. 3) including that filter and find, if anything, stronger results. We report them in Online Appendix Fig. OA.5.

We then calculate implied volatilities using the Black-Scholes formula, treating the options as though they are European. We also replicate the analysis using American implied volatilities and find nearly identical results. The reason for this is that in most cases we ultimately end up converting the IVs back into prices, meaning that any errors in the pricing formula are largely irrelevant: it is just a temporary data transformation, rather than actually representing a volatility calculation.

The data are then further filtered based on the IVs:

1. Eliminate all zero or negative IVs
2. All options with IV more than 50% (in proportional terms) different from the average for the same underlying, date, and maturity
3. We then filter outliers along all three dimensions, strike, date, and maturity, removing the following:
   (a) If the IV changes for a contract by 15% or more on a given day then moves by 15% or more in the opposite direction in a single day within the next week, and if it moves by less than 3% on average over that window, for options with maturity greater than 90 days. This eliminates temporary large changes in IVs that are reversed, which tend to be observed early in the life of the options.
   (b) If the IV doubles or falls by half in either the first or last observation for a contract
   (c) If, looking across maturities at a given strike on a given date, the IV changes by 20% or more and then reverses by that amount at the next maturity (i.e., spikes at one maturity). This is restricted to maturities within 90 days of each other.
   (d) If the last, second to last, or third to last IV is 40% different from the previous maturity
   (e) If, looking across strikes at a given maturity on a given date, the IV changes by 20% and reverses at the next strike (for strikes within 10% of each other)
   (f) If the change in IV at the first or last strike is greater than 20%, or the change at the second or second to last option is greater than 30%

At-the-money (ATM) IVs are constructed by averaging the IVs of the options with the first strike below and above the futures price. The ATM IV is not calculated for any observation where we do not have at least one observation (a put or a call) on both sides of the futures price.

To calculate ATM straddle returns for each maturity, we interpolate linearly between the IVs of the two closest out-of-the-money options on either side of the spot, and use this to compute the implied price of the ATM straddle at the beginning of the holding period; similarly, we interpolate linearly the IVs of those options at the end of the holding period, and obtain the corresponding price of the straddle at the end of the holding period. These prices are then used to compute the holding period return. Finally, to calculate returns of straddles at standardized maturities, we interpolate linearly the returns across maturities (which corresponds to a feasible portfolio). If options are not available on the maturities on both sides of the target, then we use a single straddle if it has a maturity within 35 days of the target maturity.

Appendix B. Random effects models

Denote the vector of true Sharpe ratios for the straddles in market i as sr_i. Our goal is to estimate the distribution of sr_i across the various underlyings. A natural benchmark distribution for the means is the normal distribution,

$$sr_i \sim N(\mu_{sr}, \Sigma_{sr}) \tag{B.1}$$

This section estimates the parameters μ_{sr} and Σ_{sr}, where μ_{sr} represents the high-level mean of Sharpe ratios across all the markets, and Σ_{sr} describes how the market-specific means vary. The estimates of the market-specific Sharpe ratios differ noticeably across markets, but much of that variation is likely driven by sampling error. $Theterm \Sigma_{sr}$ is an estimate of how much the true Sharpe ratios vary, as opposed to the sample estimates.

Denote the sample estimate of the Sharpe ratio in each market as $\widehat{sr}_i$, and the stacked vector of sample Sharpe ratios as $\widehat{\mathbf{sr}} \equiv \left[\widehat{sr}_1', \widehat{sr}_2', \ldots\right]'$. Similarly, denote the vector of true Sharpe ratios as $\mathbf{sr} \equiv \left[sr_1', sr_2', \ldots\right]'$. Under the central limit theorem,

$$\widehat{\mathbf{sr}} \Rightarrow N(\mathbf{sr}, \Sigma_{\widehat{\mathbf{sr}}}), \tag{B.2}$$

where $\Rightarrow$ denotes convergence in distribution and the covariance matrix $\Sigma_{\widehat{sr}}$ depends on the covariance between all the returns, across both maturities and underlyings, along with the lengths of the various samples.[24] Online Appendix OA.4 describes how we construct $\Sigma_{\widehat{\mathbf{sr}}}$.

The combination of (B.1) and (B.2) represents a fully specified distribution for the data as a function of μ_{sr} and Σ_{sr}. It is then straightforward to construct point estimates and confidence intervals for μ_{sr} and Σ_{sr} with standard methods.

To allow for the possibility that average returns differ between the financial and nonfinancial underlyings, the mean in the likelihood can be replaced by $\mu_{sr} + \mu_D I_F$,

[24] More formally, we would say that $\widehat{\mathbf{sr}}$ properly scaled by the square root of the sample size converges to a normal distribution. The expression (B.2) implicitly puts the sample size in $\Sigma_{\widehat{\mathbf{sr}}}$. The derivation of this result is a straightforward application of the continuous mapping theorem, nearly identical to the proof that a sample t-statistic is asymptotically normally distributed.

where μ_D is the difference in Sharpe ratios and I_F is a 0/1 indicator for whether the associated underlying is financial. We calculate the sampling distribution for the estimated parameters through Bayesian methods, treating the parameters as though they are drawn from a uniform prior. The point estimates are therefore identical to MLE, and the confidence bands represent samples from the likelihood.[25]

References

Ait-Sahalia, Y., Karaman, M., Mancini, L., 2019. The term structure of variance swaps, risk premia and the expectations hypothesis. Forthcoming, Journal of Econometrics.

Alexopoulos, M., Cohen, J., 2009. Uncertain times, uncertain measures. Working paper.

Andersen, T.G., Fusari, N., Todorov, V., 2015. The risk premia embedded in index options. J. Financ. Econ. 117 (3), 558–584.

Andersen, T.G., Fusari, N., Todorov, V., 2017. Short-term market risks implied by weekly options. J. Finance 72 (3), 1335–1386.

Andries, M., Eisenbach, T. M., Schmalz, M. C., Wang, Y., 2015. The term structure of the price of variance risk.

Asness, C.S., Moskowitz, T.J., Pedersen, L.H., 2013. Value and momentum everywhere. J. Finance 68 (3), 929–985.

Bachmann, R., Bayer, C., 2013. "Wait-and-see" business cycles? J. Monet. Econ. 60 (6), 704–719.

Bachmann, R., Moscarini, G., 2012. Business cycles and endogenous uncertainty. Working paper.

Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. Q. J. Econ. 131 (4), 1593–1636.

Bakshi, G., Kapadia, N., 2003. Delta-hedge gains and the negative market volatility risk premium. Rev. Financ. Stud. 16(2), 527–566.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. Rev. Financ. Stud. 16 (1), 101–143.

Barro, R.J., 2006. Rare disasters and asset markets in the twentieth century. Q. J. Econ. 121(3), 823–866.

Basu, S., Bundick, B., 2017. Uncertainty shocks in a model of effective demand. Econometrica 85 (3), 937–958.

Bekaert, G., Engstrom, E., Ermolov, A., 2015. Bad environments, good environments: a non-gaussian asymmetric volatility model. J. Econom. 186, 258–275.

Bekaert, G., Hoerova, M., Duca, M.L., 2013. Risk, uncertainty and monetary policy. J. Monet. Econ. 60 (7), 771–788.

Berger, D., Dew-Becker, I., Giglio, S., 2020. Uncertainty shocks as second–moment news shocks. Rev. Econ. Stud. 87 (1), 40–76.

Berger, D., Vavra, J., 2013. Volatility and pass-through. Working paper.

Black, F., 1976. The pricing of commodity contracts. J. Financ. Econ. 3 (1–2), 167–179.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. J. Polit. Econ. 81 (3), 637–654.

Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77(3), 623–685.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., Terry, S.J., 2018. Really uncertain business cycles. Econometrica 86 (3), 1031–1065.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31 (3), 307–327.

Bollerslev, T., Todorov, V., 2011. Tails, fears, and risk premia. J. Finance 66(6), 2165–2211.

Bondarenko, O., 2014. Why are put options so expensive? Q. J. Finance 04 (03), 1450015.

Bretscher, L., Hsu, A., Tamoni, A., 2019. The real response to uncertainty shocks: the risk premium channel. Working paper.

Bretscher, L., Schmid, L., Vedolin, A., 2018. Interest rate risk management in uncertain times. Rev. Financ. Stud. 31, 3019–3060.

Broadie, M., Chernov, M., Johannes, M., 2009. Understanding index option returns. Rev. Financ. Stud. 22(11), 4493–4529.

Caballero, R.J., 1999. Aggregate investment. Handbook Macroeconomic. 1, 813–862.

Cesa-Bianchi, A., Pesaran, M. H., Rebucci, A., 2018. Uncertainty and economic activity: A multi-country perspective. Working paper.

Choi, H., Mueller, P., Vedolin, A., 2017. Bond variance risk premiums. Rev. Financ 21 (3), 987–1022.

Cochrane, J.H., 2005. Asset pricing. Princeton University Press.

Constantinides, G.M., Jackwerth, J.C., Savov, A., 2013. The puzzle of index option returns. Rev. Asset Price Studies 3 (2), 229–257.

Cremers, M., Halling, M., Weinbaum, D., 2015. Aggregate jump and volatility risk in the cross-section of stock returns. J. Finance 70 (2), 577–614. doi:10.1111/jofi.12220.

Darby, J., Hallett, A.H., Ireland, J., Piscitelli, L., 1999. The impact of exchange rate uncertainty on the level of investment. Economic J. 109, 55–67.

Decker, R., D'Erasmo, P.N., Boedo, H.M., 2016. Market exposure and endogenous firm volatility over the business cycle. Am. Econ. J. 8 (1), 148–198.

Dew-Becker, I., Giglio, S., Le, A., Rodriguez, M., 2017. The price of variance risk. J. Financ. Econ. 123 (2), 225–250.

Dew-Becker, I., Tahbaz-Salehi, A., Vedolin, A., 2019. Skewness and stochastic volatility in a network production model. Working paper.

Diercks, A., Hsu, A., Tamoni, A., 2019. When it rains it pours: Cascading uncertainty shocks. Working paper.

Egloff, D., Leippold, M., Wu, L., 2010. The term structure of variance swap rates and optimal variance swap investments. J. Financ. Quant. Anal. 45(5), 1279–1310.

Elder, J., 2004. Another perspective on the effects of inflation uncertainty. J. Money Credit Bank. 36 (5), 911–928.

Elder, J., Serletis, A., 2010. Oil price uncertainty. J. Money Credit Bank. 42, 1137–1159.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: J. Econ. Soc. 987–1007.

Farago, A., Tédongap, R., 2018. Downside risks and the cross-section of asset returns. J. Financ. Econ. 129 (1), 69–86.

Gilchrist, S., Williams, J.C., 2005. Investment, capacity, and uncertainty: a putty–clay approach. Rev. Econ. Dyn. 8 (1), 1–27.

Gourio, F., 2013. Credit risk and disaster risk. Am. Econ. J. 5(3), 1–34. Working paper

Herskovic, B., Kelly, B., Lustig, H., Van Nieuwerburgh, S., 2016. The common factor in idiosyncratic volatility: quantitative asset pricing implications. J. Financ. Econ. 119 (2), 249–283.

Huizinga, J., 1993. Inflation uncertainty, relative price uncertainty, and investment in u.s. manufacturing. J. Money Credit Bank. 25 (3), 521–549.

Ilut, C., Kehrig, M., Schneider, M., 2015. Slow to hire, quick to fire: employment dynamics with asymmetric responses to news. NBER Working Paper Series.

Jones, C.S., 2006. A nonlinear factor analysis of s&p 500 index option returns. J. Finance 61 (5), 2325–2363.

Jurado, K., Ludvigson, S., Ng, S., 2015. Measuring uncertainty. Am. Econ. Rev. 105 (3), 1177–1216.

Kozlowski, J., Veldkamp, L., Venkateswaran, V., 2016. The tail that wags the economy: Belief-driven business cycles and persistent stagnation. Working paper.

Leduc, S., Liu, Z., 2016. Uncertainty shocks are aggregate demand shocks. J. Monet. Econ. 82, 20–35.

Ludvigson, S.C., Ma, S., Ng, S., 2015. Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response? Technical Report. National Bureau of Economic Research.

Ludvigson, S.C., Ng, S., 2007. The empirical risk-return relation: a factor analysis approach. J. Financ. Econ. 83(1), 171–222.

McCracken, M.W., Ng, S., 2016. Fred-md: a monthly database for macroeconomic research. J. Bus. Econ. Stat. 34 (4), 574–589.

Muravyev, D., Pearson, N.D., 2020. Option trading costs are lower than you think. Rev. Financ. Stud.. Forthcoming

Pástor, L., Veronesi, P., 2009. Technological revolutions and stock prices. Am. Econ. Rev. 1451–1483.

Prokopczuk, M., Symeonidis, L., Simen, C.W., 2017. Variance risk in commodity markets. J. Bank. Finance 81, 136–149.

Salgado, S., Guvenen, F., Bloom, N., 2016. Skewed business cycles. Working paper.

Schwert, G.W., 1989. Business cycles, financial crises, and stock volatility. In: Carnegie-Rochester Conference series on public policy, Vol. 31. Elsevier, pp. 83–125.

Schwert, G.W., 2011. Stock volatility during the recent financial crisis. Eur. Financ. Manag. 17 (5), 789–805.

Segal, G., Shaliastovich, I., Yaron, A., 2015. Good and bad uncertainty: macroeconomic and financial market implications. J. Financ. Econ. 117 (2), 369–397.

[25] We use Bayesian methods to calculate the sampling intervals because likelihood-based methods require inverting large second derivative matrices, which can be numerically unstable. The estimation in this section is performed using the Bayesian computation engine Stan, which provides functions that both maximize the likelihood and rapidly sample from the posterior distribution. Code is available on request.

Seo, S.B., Wacheter, J.A., 2018a. Do rare events explain CDX tranche spreads? J. Finance 0 (ja).
Seo, S.B., Wachter, J.A., 2018b. Option prices in a model with stochastic disaster risk. Manage. Sci..
Siriwardane, E., 2015. The probability of rare disasters: estimation and implications. Harvard Bus. School Finance Working Paper (16–61).
Trolle, A.B., Schwartz, E.S., 2010. Variance risk premia in energy commodities. J. Derivativ. 17 (3), 15–32.
Van Binsbergen, J.H., Koijen, R.S., 2017. The term structure of returns: facts and theory. J. Financ. Econ. 124 (1), 1–21.