

Hedging macroeconomic and financial uncertainty and volatility*

Ian Dew-Becker, Stefano Giglio, and Bryan Kelly

April 1, 2019

Abstract

We study the pricing of shocks to uncertainty and volatility using a novel and wide-ranging data set of options contracts directly related to the state of the macroeconomy and financial markets. If uncertainty shocks are viewed as bad by investors – in the sense of being associated with high marginal utility – portfolios that hedge them should earn negative premia. Empirically, however, contracts that provide protection against shocks to macroeconomic uncertainty have historically earned statistically and economically significantly positive excess returns. Instead, portfolios exposed to the realization (as opposed to the expectation) of large shocks to fundamentals have historically earned large and negative risk premia. These results imply that it is large realizations of shocks to fundamentals, not forward-looking uncertainty shocks, that drive investors' marginal utility; in turn, these implications dictate the role of volatility in macroeconomic models and indicate which shocks policymakers should aim to counteract.

1 Introduction

It is well established that a wide range of measures of economic volatility and uncertainty vary over time and with the business cycle. Uncertainty about all sides of the economy, including productivity, the level of the stock market, inflation, interest rates, and energy prices, varies substantially, and 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 that explore the relationship between uncertainty and real activity. Some models focus on contractionary effects of uncertainty, such as models with wait-and-see effects in investment (e.g. Caballero (1999), Bloom (2009)), while others argue that uncertainty can be high in periods of high growth, like the late 1990's, due to learning effects (Pastor and Veronesi (2009)). Furthermore, even theoretical work that focuses on contractionary effects of uncertainty tends to find responses of the economy to uncertainty shocks whose sign is parameter-dependent. Gilchrist

*Dew-Becker: Northwestern University. Giglio: Yale University. Kelly: Yale University. We appreciate helpful comments from Dmitry Muravyev, Federico Gavazzoni, 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, and the Federal Reserve Bank of Chicago.

and Williams (2005) and Bloom et al. (2017) extensively discuss the potentially expansionary effects of uncertainty shocks.

The empirical literature studying uncertainty has focused almost entirely on analyzing raw correlations or using vector autoregressions with varying identifying assumptions. Empirical work thus far has not resolved the question of whether uncertainty is contractionary in either the short- or long-run, with some arguing that uncertainty shocks are briefly contractionary followed by a large rebound (Bloom (2009)), others arguing that they are persistently contractionary (e.g. Alexopoulos and Cohen (2009); Leduc and Liu (2016), Caldara et al. (2016)), and a third set finding that they have little effect at all (Bachmann and Bayer (2013); Berger, Dew-Becker, and Giglio (2018)). A fourth set of papers argues the causation may run the opposite direction, with economic activity driving uncertainty (e.g. Bachmann and Moscarini (2012), Ludvigson, Ma, and Ng (2015), and Creal and Wu (2017)).¹

This paper develops a novel empirical approach to evaluate the effects of uncertainty shocks. Instead of studying a VAR with all of the associated identification challenges, we argue that financial markets provide a direct window on how investors perceive uncertainty shocks. The basic idea is to construct portfolios that directly hedge innovations in uncertainty and then measure their average returns. If investors are willing to accept negative average returns on those hedging portfolios (i.e., negative risk premia), as they would on an insurance contract, that implies that they view uncertainty as being bad in the sense 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. The magnitude of the average return moreover measures the correlation between uncertainty shocks and state prices. So rather than running sophisticated regressions of output on uncertainty, we let investors speak to the question.

To be clear, the analysis of risk premia 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). Since asset prices naturally encode investors’ conditional expectations, returns (i.e., changes in prices) measure innovations in the expectations without requiring the construction of a statistical model.² Section 4 formalizes this point.

Importantly, the correlation between innovations to uncertainty and marginal utility – which risk premia reveal – is not the same as the unconditional correlation between uncertainty and the state of the economy. For example, if uncertainty rises as an endogenous response to real contractions (e.g. Ludvigson, Ma, and Ng (2015)), then its unconditional correlation with the economy may be negative even if the correlation of its innovations with the state of the economy is zero or positive (that is, even if investors are not worried about uncertainty).

While there is a large literature that estimates the risk premia for uncertainty about the

¹For related theories, see Decker, D’erasmo and Boedo (2016), Berger and Vavra (2013), Ilut, Kehrig and Schneider (2015), Kozłowski, Veldkamp, and Venkateswaran (2016), and Cesa-Bianchi, Pesaran, and Rebucci (2018).

²This is the same as the idea of using changes in prices of Fed Funds futures to measure monetary policy shocks.

S&P 500, based on the pricing of options,³ recent evidence shows that aggregate uncertainty has multiple dimensions (Ludvigson, Ma, and Ng (2015); Baker, Bloom, and Davis (2015)). S&P 500 uncertainty is related to conditions in the financial sector, but there is good reason to think that the driving force in the economy is actually uncertainty about other features of the macroeconomy, such as interest rates, inflation, or the availability of inputs to production, like crude oil. This paper contributes to the literature by estimating risk premia associated with uncertainty in 19 different markets covering a range of different features of the economy, including financial conditions, inflation, and the prices of real assets. The broad range of assets allows the analysis to uncover consistent patterns in the attitudes of investors to different types of uncertainty.

We also use the range of contracts to construct portfolios of options that allow investors to hedge innovations in uncertainty indexes from Jurado, Ludvigson, and Ng (JLN; 2015) and the economic policy uncertainty (EPU) index of Baker, Bloom, and Davis (2015). None of those indexes necessarily gives a direct, errorless measure of whatever the true latent uncertainty is that drives the decisions of agents in the economy, so the goal is not to perfectly replicate them with option contracts. Instead, we show that the options are able to almost perfectly span the part of those indexes that is related to aggregate outcomes like industrial production, employment, and interest rates. Moreover, replicating the indexes requires using the full range of implied volatilities, and not just the S&P 500.

Together, these results confirm that hedging shocks to implied volatility in the 19 markets represents a good way to hedge various types of aggregate uncertainty shocks, both macroeconomic and financial, and they show why it is important to study more than just S&P 500 options.

We next examine the pricing of shocks. The discussion so far has focused on economic uncertainty – some measure of the dispersion of agents’ conditional distribution for future outcomes. But much of the literature also studies *volatility* – the magnitude of realized shocks to fundamentals. Whereas uncertainty in theoretical models is a forward-looking conditional variance, volatility is a backward-looking sample variance. That is, for some shock ε , with $\text{var}_t(\varepsilon_{t+1}) = \sigma_t^2$, uncertainty is σ_t^2 , while volatility is ε_t^2 . The distinction is crucial from the 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 .

Our analysis of returns on options yields hedging portfolios for both uncertainty and volatility, σ^2 and ε^2 , taking advantage of the fact that options of different maturities have different exposures to σ^2 and ε^2 . The empirical analysis yields two key findings. First, across 19 individual option markets and also when hedging the JLN and EPU indexes, portfolios that directly hedge uncertainty shocks have historically earned returns that are in almost all cases statistically and economically significantly positive. The average returns are nearly as positive as those on the aggregate US stock market. That result implies that investors in these markets view periods of high uncertainty as being good on average, rather than bad, in the same way and to the same degree that stock

³See Egloff, Leippold, and Wu (2010), Dew-Becker et al. (2017), Van Binsbergen and Koijen (2017), Andries et al. (2015), and Ait-Sahalia, Karaman, and Mancini (2015).

returns are high in good times. The second result runs in the opposite direction: portfolios that hedge realized volatility – large realized futures returns, or ε^2 – earn statistically and economically significantly negative returns, implying that investors on average view periods in which shocks to fundamentals themselves are large as being bad.

The returns on the uncertainty hedging portfolios are difficult to reconcile with the view that innovations in economic uncertainty reduce utility. If uncertainty shocks were viewed as bad in the sense of raising marginal utility, then we would find a negative premium on implied volatility – investors would be willing to accept negative average returns on assets that are hedges against high marginal utility states. Instead, the results imply that investors have historically viewed periods of high uncertainty and implied volatility as being good, in the sense that they are associated with low marginal utility, consistent with models such as that of Pastor and Veronesi (2009).

What is associated with bad outcomes, from the perspective of investors, is instead realized volatility. That finding contributes to the growing literature studying skewness risk in the economy (e.g. Barro (2006), Bloom, Guvenen, and Salgado (2016), and Seo and Wachter (2018a,b)). If shocks to the economy are skewed to the left, then large shocks tend to be negative. That is, $E[\varepsilon^3] < 0$ implies $cov(\varepsilon, \varepsilon^2) < 0$. An explanation for the pricing of realized volatility is then simply that hedging realized volatility helps hedge downward jumps and disasters.

It is well known that both volatility and uncertainty are countercyclical, but their overall correlation not as high as one might expect – only about 65 percent, 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.

The paper is related to two main strands of literature. The first studies the relationship 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.⁴ While there are contractionary forces, such as wait-and see effects and Keynesian demand channels, there are also forces through which uncertainty can be expansionary, including precautionary saving and the Oi–Hartmann–Abel effect that is extensively discussed by Gilchrist and Williams (2005) and Bloom et al. (2017).⁵ Our results are more consistent with the expansionary forces.

The related empirical literature tries to measure whether uncertainty does in fact have contractionary effects.⁶ This paper builds on that work by providing measures of risk premia that indicate how investors perceive the effects of aggregate uncertainty shocks. The basic ideas underlying the

⁴These include a Keynesian demand channel (Basu and Bundick (2017)), real options effects on investment (Bloom (2009), Bloom et al. (2017)), effects on labor search (Leduc and Liu (2015)), or through financial frictions and credit spreads (Gourio (2013)).

⁵Oi (1961), Hartman (1972), and Abel (1983).

⁶Recent examples include Berger, Dew-Becker, and Giglio (2017), Jurado, Ludvigson, and Ng (2015), Ludvigson, Ma, and Ng (2015), Baker, Bloom, and Davis (2015), Bachmann and Bayer (2013), and Alexopoulos and Cohen (2009), among many others.

paper, in particular the emphasis on the distinction between uncertainty and realized volatility, make it superficially similar to Berger, Dew-Becker, and Giglio (2018). The methods and data used in this paper are very different, though. That paper estimates a structural vector autoregression as is common in the macroeconomics literature, whereas this paper takes a financial economics approach, studying risk premia, requiring none of the VAR identifying assumptions. Furthermore, while the past literature has primarily studied S&P 500 implied volatility,⁷ an important contribution of this paper is to examine a much broader range of asset classes, showing that they have a better link to uncertainty about macroeconomic outcomes and that they behave qualitatively differently from the S&P 500.

The paper also draws on a literature in finance estimating the pricing of volatility (ε^2) risk. As in macroeconomics, that literature almost exclusively studies the S&P 500.⁸ In addition to studying a much broader range of markets, our contribution is to isolate the premium on implied volatility as opposed to just the realized variance risk premium studied in past work – the distinction between the two is crucial because it is only implied volatility, not realized volatility, that captures the forward-looking concept of uncertainty on which the theoretical models are based.

It is important to distinguish between aggregate uncertainty – uncertainty about the state of the aggregate economy – and idiosyncratic uncertainty, or uncertainty about shocks at, say, the household or firm level. Our results apply to sector-level uncertainty, since we price uncertainty shocks in areas like the stock market, interest rates, and the price of oil and other goods. The concept of uncertainty thus lies somewhere between aggregate and purely idiosyncratic. Our results do not address models based on purely firm- or household-level shocks, e.g. Christiano, Motto, and Rostagno (2014).

The remainder of the paper is organized as follows. Section 1.1 describes in more detail the paper’s key distinction between realized and implied volatility and discusses how they are distinguished conceptually and empirically. Section 2 describes the data and its basic characteristics. Section 3 discusses the construction of portfolios that hedge realized volatility and uncertainty. Section 4 reports the cost of hedging volatility and uncertainty in our data. Section 5 presents robustness results. To provide more confidence in some of the results, section 6 examines the crude oil market in detail. Finally, section 7 concludes.

1.1 The distinction between implied and realized volatility

In models of time-varying uncertainty, both structural and reduced form, there is typically a shock, say, ε_t , that has a time-varying conditional variance, $var_{t-1}(\varepsilon_t) = \sigma_{t-1}^2$. Given that structure, *realized volatility* measures the volatility of the realized shock in period t , defined as ε_t^2 . *Uncertainty*, on the other hand, is the forward-looking conditional variance, σ_t^2 , which can also be viewed as

⁷See Bloom (2009), and Basu and Bundick (2017), among many others.

⁸There are a few papers that have studied specific markets, such as individual equities (e.g. Bakshi, Kapadia, and Madan (2003)) or Treasury bonds (Mueller, Vedolin, and Yen (2017)). Prokopczuk et al. (2017) examine the variance risk premium across many of the same markets that we study (see also Trolle and Schwartz (2010)).

the expectation of *future* realized volatility ($\sigma_t^2 = E_t[\varepsilon_{t+1}^2]$). Realized volatility therefore measures the magnitude of the shock that occurred in the present period, while uncertainty measures the expected future magnitude of shocks.

The distinction between realized and implied volatility is the same as the distinction between realized technology shocks and news shocks, that has been explored in the macro literature: one affects technology on impact, while the other affects its expected future path. Whereas the news shock literature has focused on first-moment shocks, we explore the distinction between realizations and expectations for second-moment shocks.

Just like a realized productivity shock and the expectation of future productivity are related, current realized volatility and the expectation of future volatility (uncertainty) are also related. So a natural question is whether it makes theoretical and practical sense to distinguish between implied and realized volatility. Specifically, since we are studying asset prices, when uncertainty rises, current asset prices typically fall; in turn, this means that the square of the price change (realized volatility) rises. This intuition seems to imply that implied and realized volatility are mechanically connected, and cannot logically or empirically be distinguished.

That intuition has a flaw. It is true that when uncertainty *rises*, prices fall and realized volatility increase. But it is also true that when uncertainty *falls*, prices rise and realized volatility still increases: realized volatility is a *quadratic* function of price movements, so both uncertainty increases and decreases induce an increase in realized volatility. To a good approximation, uncertainty shocks will actually have *no average effect* on – and hence no mechanical correlation with – realized volatility.

Berger, Dew-Becker, and Giglio (2018) discuss these issues in greater detail, and also simulate three workhorse macroeconomic models of uncertainty shocks. They show that, consistent with that logic, uncertainty shocks have no effect on average on realized volatility, so that the two can be distinguished.⁹

As a practical matter, we show below that for most markets we study, the correlation between implied and realized volatility shocks is far from 1 – in fact averaging only 0.2 (see table 5 and section 3.2) – showing that there is independent variation that can be used to distinguish investors attitudes to them.¹⁰ Intuitively, for use to have identifying variation, there must be variation in uncertainty independent of realized volatility, which is exactly what is shown in table 5.

⁹There is also a time aggregation question. Suppose we were to measure implied and realized volatility shocks over the course of a year. An increase in implied volatility on the first day of the year will, by definition, lead to higher realized volatility over the remainder of the year, thus making the two processes correlated. That will be true even if they are unrelated at the daily level. In other words, time aggregation can lead implied and realized volatility to be mechanically related. To avoid that problem, we use relatively high-frequency data, focusing on two-week returns, where time aggregation problems should be much smaller (one can formally show that they disappear in the continuous limit).

¹⁰This implies volatility is not driven by a pure GARCH model (Engle (1982); Bollerslev (1986)): implied volatility is not a deterministic function of past realized volatility.

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 for the period 1983–2015. The CME data is important for covering a broad array of features of the economy, including stock prices, interest rates, exchange rates, and prices of metals, petroleum products, and agricultural products.

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.¹¹

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.1.

We calculate implied volatility for all of the options using the Black–Scholes (1973) model (technically, the Black (1976) model for the case of futures).¹² Unless otherwise specified, implied volatility is calculated at the three-month maturity. We treat the implied volatilities as measuring uncertainty, in the sense that they represent the conditional variance (expected future volatility) of futures returns. In addition to depending on uncertainty, implied volatilities also contain a risk premium, which can potentially vary over time. However, even in the presence of that risk premium implied volatilities appear to summarize all available information in the data for forecasting future volatility, driving out other standard uncertainty measures from forecasting regressions (see Berger, Dew-Becker, and Giglio (2018)).

¹¹The 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.

¹²The 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.

2.1.2 Alternative uncertainty measures

The implied volatilities of the CME options give direct measures of investor uncertainty, similar to the VIX. In addition to implied volatilities, we also examine two other measures of uncertainty. The first uncertainty index is developed in a pair of papers by Jurado, Ludvigson, and Ng (JLN; 2015) and Ludvigson, Ma, and Ng (2017). The construction involves two basic steps. First, realized squared forecast errors are constructed for 280 macroeconomic and financial time series.¹³ Denoting the error for series i as $\varepsilon_{i,t}$, the basic assumption is that there is a variance process, $\sigma_{i,t}^2$, such that $E[\varepsilon_{i,t}^2] = \sigma_{i,t}^2$. 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 $\sigma_{i,t}^2$. We divide the 280 series among those that pertain financial markets, real activity, and goods prices, with the latter two also being combined into an overall macroeconomy group, and take the first principal component from each group to get different subindexes.

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 by taking the first principal component from the cross-section of the $\varepsilon_{i,t}^2$ in a given month. We therefore construct both uncertainty and realized volatility under the JLN framework.

The second uncertainty index is the Economic Policy Uncertainty (EPU) index of Baker, Bloom, and Davis (2015). The EPU index is constructed based on media discussion of uncertainty, the number of federal tax provisions changing in the near future, and forecaster disagreement. Unlike the JLN framework, there is no distinction in this case between volatility and uncertainty, so we treat the EPU index as measuring only uncertainty.

2.2 The time series of uncertainty

Figure 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. The period around the 1991 Gulf War was a period of extremely high implied volatility for crude oil, but much lower uncertainty for stocks and bonds. Conversely, the Financial Crisis was associated with larger relative increases in stock and bond than crude oil implied volatility. So while they move together, 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, and also gives the first introduction to the full list of 19 markets. The various markets are sorted in this table into related categories, with the result that the largest correlations are generally along the main diagonal. Shading denotes the degree of correlation, with darker cells representing greater

¹³Following Ludvigson, Ma, and Ng (2017), we augment the original JLN data set with returns on a broad range of equity portfolios. 134 macro series are from McCracken and Ng (2016), while the 148 financial indicators are from Ludvigson and Ng (2007). Our analysis uses code from the replication files of JLN. Among the macro series from McCracken and Ng (2016), the price series are defined as those referring to price indexes, and the real series are the remainder.

correlation. 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 eigenvalues of the correlation matrix quantify the degree of common variation. The largest eigenvalue explains 43 percent of the total variation. The remaining eigenvalues are much smaller, though – even the second largest only explains 15 percent of the total variation. Eight eigenvalues are required to explain 90 percent of the total variation in the IVs, which is perhaps a reasonable estimate of the number of independent components in the data.

Importantly, the common variation in the implied volatilities is much larger than the common variation in the underlying futures returns. The largest principal component for the futures returns explains less than half as much variation – 17 percent versus 43. 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, 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 eigenvalue for implied volatility, realized volatility (discussed below), and the underlying futures returns, along with bootstrapped 95-percent confidence bands.

Fraction of variation explained by largest eigenvalue

| | IV | RV | Futures return |
|----------------------------------|-------|-------|-------------------|
| Largest Eigenvalue (% explained) | 43.3% | 31.9% | 17.4% |
| 95% Bootstrap CI | 37.1% | 23.7% | 16.0% |
| | 50.0% | 41.6% | 20.6% |

2.3 Projecting the uncertainty indexes onto the 19 IVs

Figure 2 examines how well the 19 IVs can fit the JLN and EPU indexes. These regressions are then used to construct hedging portfolios for the indexes. Figure 2 plots the time series of the JLN (top three rows) and EPU (bottom row) indexes against the fitted values from their projection onto the 19 implied volatilities and a constant. 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 95 percent. The next highest correlation is only 62 percent, for Treasury bonds. So figure 2 shows that fitted financial uncertainty is very nearly equivalent to S&P 500 implied volatility.¹⁴

¹⁴The strong fit the 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.

The second row plots fitted uncertainty for real variables. A very different set of implied volatilities now are important for driving the fitted values. Gold, copper, crude oil, and heating oil are the most important contributors. The implied volatilities capture well the lower-frequency variation, though they may miss some of the more high-frequency variation; overall, these investments provide a hedge against a substantial fraction of real (GDP, IP, etc.) risk.

The results for the price component of JLN uncertainty are reported in the third row. In this case, the highest correlations are again for heating oil, crude oil, natural gas, gold, and copper. These results suggest that uncertainty about the real economy and inflation are driven by similar factors, and that those factors are notably distinct from financial uncertainty.

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. So the fit of the IVs to the EPU index comes mostly from the financial rather than the nonfinancial options. 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.¹⁵

2.3.1 Goodness of fit

An important question to ask here is how well the implied volatilities are able to span actual economic uncertainty. If the EPU and JLN indexes were measures of some true underlying uncertainty with no error, then the goal would simply be to have the highest possible R^2 in these regressions. But in reality, they are noisy estimates, so what is more important for our purposes is that the IVs are able to span the part of the JLN and EPU indexes *that is related to actual economic outcomes*.

To measure that relationship, denote the fitted value from the regression of the $JLNU$ and EPU indexes on the IVs with a circumflex, and the residuals by ε ; we then have the following decomposition for the JLN financial index:

$$JLNU_t^{\text{Financial}} = \underbrace{\widehat{JLNU}_t^{\text{Financial}}}_{\text{Fitted value}} + \underbrace{\varepsilon_t^{JLN\text{Financial}}}_{\text{Residual}} \quad (1)$$

and naturally we have similar decompositions for each other index. Panel A of Table 2 reports the coefficients from regressions of industrial production growth, employment growth, and the Fed funds rate on their own lags and the fitted and residual uncertainty for the four indexes. That is, for industrial production and JLN financial uncertainty, we estimate the regression

$$\Delta \log IP_t = a + b_1 \Delta \log IP_{t-1} + b_2 \widehat{JLNU}_t^{\text{Financial}} + b_3 \varepsilon_t^{JLN\text{Financial}} + \mu_{IP,t} \quad (2)$$

where a and the b_j are estimated coefficients and $\mu_{IP,t}$ is a residual. Each column of the table corresponds to a different uncertainty index for this regression, and each set of two rows corresponds to a different macroeconomic variable. For example, the top left of the panel corresponds to a

¹⁵To 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.

regression of IP growth onto the fitted and residual components of the financial uncertainty index of JLN.

The relative values of b_1 and b_2 give a measure of the relative importance of the spanned part of uncertainty versus the residual. Furthermore, since the fitted values and residuals are uncorrelated, the variance of the part of IP growth explained by the uncertainty index is

$$\frac{b_2^2 \text{Var} \left(\widehat{JLNU}_t^{\text{Financial}} \right)}{b_2^2 \text{Var} \left(\widehat{JLNU}_t^{\text{Financial}} \right) + b_3^2 \text{Var} \left(\varepsilon_t^{\text{JLNFinancial}} \right)} \quad (3)$$

This variance decomposition is reported in Panel B of Table 2.

Table 2 shows that the fitted part of uncertainty is consistently associated with lower growth in employment and IP and lower interest rates across the various specifications. Moreover, the coefficient on fitted uncertainty (b_2) is almost always substantially larger than that on the unexplained residual (b_3). That has the consequence that when we examine the variance decompositions, across the various specifications, the part of the uncertainty indexes that is spanned generates on average 87 percent of the total fitted variance. That is, the options explain 87 percent of the relationship between the EPU and JLN uncertainty indexes with these three aggregate outcomes.

Recall that the R^2 s in the regressions for fitting the JLN and EPU uncertainty indexes in figure 2 are all less than 1. Table 2 shows that the unexplained residual parts of those indexes is largely not significantly associated with macro outcomes, while the fitted part is. So while the option-implied volatilities do not perfectly span the uncertainty indexes, they do span the economically relevant part – what is left does not appear to have a meaningful link to the aggregate economy and may simply represent measurement error.

As discussed above and in section 4, the fact that the coefficients in the top panel of Table 2 are negative does not mean uncertainty shocks are associated with low utility. First, these are unconditional correlations, which depend on the endogenous response of uncertainty, whereas the paper’s results on risk premia below address the relationship between *innovations* in uncertainty and marginal utility. Second, to the extent that utility depends mostly on the long-run component of income growth, associations with monthly activity measures will be a weak measure of importance for welfare. Finally, since the level of implied volatility is correlated with other factors, like realized volatility, an important question will be what the relationship of a pure shock to uncertainty with marginal utility is (see section 5.4). What the results in this section show is that our findings about the pricing of uncertainty are not due to uncertainty and implied volatility being unrelated to real activity.

2.4 Realized volatility

While implied volatility is measured based on option prices at a point in time, realized volatility must be measured over some time period. In our main analysis of option returns below, realized

volatility is measured over the two-week period over which the return is computed. In this section – and this section only – since the uncertainty indexes are measured at the monthly frequency, we also construct realized volatility at the monthly frequency. In general, realized volatility for market m in some period t is defined as

$$RV_{m,t} = \sum_{n \in t} r_{m,n}^2 \quad (4)$$

where $r_{m,n}$ is the log futures return in market m in subperiod n . For the case of this section with monthly time periods, t corresponds to months and we calculate realized volatility based on daily returns, so that each n represents a day. Realized volatility can be measured based on returns at different frequencies – e.g. with n representing hourly returns – and aggregated over different intervals – e.g. with t representing just a day or a week.

Realized volatility is a squared realization (or sum of squared realizations) of a random variable, which means that it is itself random and appears to be “noisy” relative to implied volatility, which can be thought of as expected realized volatility (in many models that is exactly true, or true up to an affine transformation).

Realized volatility therefore tends to be substantially more volatile than implied volatility, and implied and realized volatility are also naturally correlated with each other. The key difference between the two is that realized volatility isolates realizations of extreme events – price jumps – whereas implied volatility measures expectations of the probability or size of future extreme events.¹⁶

Table 3 reports the correlation matrix for realized volatility across the 19 markets. As in the IV correlation matrix, the correlations are relatively strong near the main diagonal, but they are all smaller in the RV case. The largest principal component accounts for 34 percent of the total variation, compared to 46 percent for IV, implying there is less common and more idiosyncratic variation in realized than implied volatility. However, that is still larger than the 19 percent in the case of futures returns themselves. So the realizations of *extreme* returns have a stronger factor structure than the realizations of returns themselves, which can help explain why realized volatility might be a priced risk factor.

Figure 3 replicates figure 2, but using realized instead of implied volatility. That is, it examines the ability of the RV series for the 19 futures markets to fit the three JLN RV indexes. The R^2 s in this case are smaller than for IV, which is consistent with the result from the correlation matrices that there is more common variation in IV than RV. Interestingly, S&P 500 realized volatility appears to fit better to the JLN RV indexes than in the IV case. It remains the case that for fitting real and price RV, the nonfinancial markets, including in particular the energies and copper, are especially useful.

Table 4 also shows that, similar to IV, the part of RV that is spanned by the options captures

¹⁶ As a simple example, suppose returns on date t are equal to 0 with probability $1 - p_{t-1}$ or some number r^* with probability p_{t-1} . The date- $t - 1$ conditional variance of returns on date t is then $p_{t-1}(1 - p_{t-1})(r^*)^2$, while the realized variance on date t is either 0 or $(r^*)^2$. More generally, the dependence of realized volatility on squared returns is what causes it to respond most strongly to extreme events.

nearly all the relationship of $JLNRV$ with aggregate outcomes – in this case the options explain 95 percent of the systematic variation (and, again, this does not identify the effect of *changes in* realized volatility on real activity).

3 Constructing option portfolios to hedge uncertainty

Implied volatility and the uncertainty indexes are not directly tradable – only the options themselves are. This section shows how to construct option portfolios that hedge shocks to implied and realized volatility in each of the 19 markets and also the JLN and EPU indexes.¹⁷ Whereas realized volatility in the previous section was measured over the period of a month, here realized volatility is hedged over the period that the options are held (two weeks, for the main results).

3.1 Straddle portfolios

We study two-week returns on straddles with maturities between one and six months.¹⁸ A straddle is a portfolio holding a put and a call with the same maturity and strike, with the strike set so that the Black–Scholes delta of the portfolio – the derivative of its price with respect to the value of the underlying – is zero. The final payoff of a straddle depends on the absolute value of the return on the underlying, meaning that they have symmetrical exposures to positive and negative returns, and no local directional exposure to the underlying.

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, the sale price will also depend on expected future volatility, meaning that straddles can give exposure to uncertainty shocks.

The exposures of straddles can be approximated theoretically using the Black–Scholes model, as in Coval and Shumway (2001), Bakshi and Kapadia (2003), and Cremers, Halling, and Weinbaum

¹⁷In principle, one could try to hedge uncertainty shocks with a much larger range of assets, e.g. using also equity returns. We focus on option returns for two reasons. First, they are directly linked to the implied volatilities that we are able to measure, and economic theory gives formal, quantitative predictions relationship between uncertainty shocks and option price returns. Unlike, say, stock returns, option returns depend essentially entirely on implied and realized volatility; they are not contaminated by other influences. Second, since the hedging weights must be estimated, adding more assets to the model increases estimation error and overfitting. That problem is limited by using a relatively small number of option portfolios.

¹⁸Past work on option returns and volatility risk premia has examined returns at frequencies of a day (e.g. Andries et al. (2017)), a week (Coval and Shumway (2001)), a month (Constantinides, Jackwerth, and Savov (2013); Dew-Becker et al. (2017)), and holding the options to maturity (Bakshi and Kapadia (2003)). The precision of estimates of the riskiness of the straddles is, all else equal, expected to be higher with shorter windows. On the other hand, shorter windows cause any measurement error in option prices to have larger effects. We choose two-week windows because they are within the typical range used and they are short enough to allow us to still calculate returns on relatively short-maturity options. 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. Higher-frequency returns avoid that problem. Section 5 describes alternative specifications that we have examined to check the robustness of the main results.

(2015). Appendix A.2 shows that the partial derivatives of the straddle return with respect to the underlying futures return, f , its square, and the change in volatility, can be approximated as

$$\frac{\partial r_{n,t}}{\partial f_t} \approx 0, \quad (5)$$

$$\frac{\partial^2 r_{n,t}}{\partial (f_t/\sigma_{t-1})^2} \approx n^{-1}, \quad (6)$$

$$\frac{\partial r_{n,t}}{\partial (\Delta\sigma_t/\sigma_{t-1})} \approx 1, \quad (7)$$

where $r_{n,t}$ is the return on date t of a straddle 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

The first partial derivative says that the straddles all have close to zero local exposure to the futures return, which is natural since their payoff is a symmetrical function of the underlying return. The second line says that the exposure of straddles to *squared* returns on the underlying – scaled by volatility – is approximately inversely proportional to time to maturity. Since the squared return measures realized volatility, the second line reveals loadings on (scaled) realized volatility. The third line shows that straddles 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.

Overall, then, all straddles have approximately equal exposure to proportional shifts in implied volatility, while the exposure to realized volatility decreases with maturity. Long-maturity straddle returns, for which the term n^{-1} becomes small, therefore reveal the premium associated with uncertainty shocks.

3.2 Hedging RV and IV in each market

The implied sensitivities in (5)–(7) give a method for constructing portfolios that the Black–Scholes model says should give exposures *only* to realized volatility – squared returns, measured by $(f_{n,t}/\sigma_{t-1})^2$ – and *only* implied volatility, measured by $\Delta\sigma_t/\sigma_{t-1}$ (Cremers, Halling, and Weinbaum (2015)). Specifically, we construct, for each market, two portfolios,

$$rv_{i,t} = \frac{5}{48} (r_{i,1,t} - r_{i,5,t}), \quad (8)$$

$$iv_{i,t} = \frac{5}{4} r_{i,5,t} - \frac{1}{4} r_{i,1,t}. \quad (9)$$

Throughout the paper, 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.

Given equations (5)–(7), the rv and iv portfolios will both have zero local sensitivity to f_t . The rv portfolio will have a unit sensitivity to $(f_t/\sigma_t)^2$ and zero sensitivity to $\Delta\sigma_t/\sigma_{t-1}$ in each market,

¹⁹We 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 (5)–(7).

while the *iv* portfolio will have a unit sensitivity to $\Delta\sigma_t/\sigma_{t-1}$ and zero sensitivity to the squared returns in each market. We use the one- and five-month straddles to construct the portfolios as those are the shortest and longest maturities that we consistently observe in the data.

The purpose of constructing these portfolios is to give a simple and direct method of measuring the premia associated with realized and implied volatility that does not require any complicated estimation or data transformation. One might worry, though, that they do not obtain the desired exposures in practice. Figure A.2 in the appendix shows that the loadings of the straddles fit the Black–Scholes predictions well. Furthermore table A.1 reports results of regressions, for each underlying, of the returns of the two portfolios on the underlying futures return, the squared futures return, and the change in implied volatility. The table shows that while the Black–Scholes predictions do not hold perfectly, the *rv* portfolio is nevertheless much more strongly exposed to realized than implied volatility, and the opposite holds for the *iv* portfolio. The coefficients on $(f_t/\sigma_{t-1})^2$ average 0.76 for the *rv* portfolio and 0.10 for the *iv* portfolio. Conversely, the coefficients on $\Delta\sigma_t/\sigma_{t-1}$ average 0.03 for the *rv* portfolio and 0.79 for the *iv* portfolio. The R^2 s are also large, averaging 72 percent across the various portfolios, implying that their returns are well described by the approximation (5). Appendix A.2 also examines the accuracy of the Black–Scholes approximation for returns in a simulated setting. Finally, section 5 reports estimates of hedging costs that do not rely on the Black–Scholes assumptions at all.

It is important to note that we would not necessarily expect the returns of the *rv* and *iv* portfolios to be uncorrelated. It is well known from the GARCH literature (e.g. Engle (1982) and Bollerslev (1986)) that in many markets, innovations to realized volatility are correlated with innovations to implied volatility. Table 5 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. In those cases, the average correlation is 0.47. While that shows that GARCH effects are present, only a minority of the variation in the *rv* and *iv* returns is driven by a common component. For the other nonfinancial underlyings, the effects are much smaller, and the correlation between the *rv* and *iv* returns is only 0.02 on average (it is 0.08 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 central to our identification, since they show that surprises in realized and implied volatility are far from the same and can be hedged separately.

3.3 Hedging the JLN and EPU indexes

Finally, using the results in figure 2 showing that the 19 IVs span most of the variation in the JLN and EPU uncertainty indexes, we construct portfolios that optimally hedge those indexes. For each index, we obtain the weights for the hedging portfolio from the regression coefficients in sections 2.3 and 2.4. For each uncertainty index j , we estimate the regression

$$JLNU_t^j = a + \sum_i b_i^j IV_{i,t} + \varepsilon_{j,t} \quad (10)$$

and then construct a hedging portfolio as

$$iv_t^{hedge,j} \equiv \sum_i b_i^j iv_{i,t} \quad (11)$$

the coefficients b_i^j therefore tell us the weight of the iv portfolio of market i in the hedging portfolio for index j . We create such portfolios for each of the JLN uncertainty indexes and the EPU index. We also construct similarly a hedge portfolio for the JLN realized volatility series ($JLNRV$) from the regression

$$JLNRV_t^j = a + \sum_i b_i^{RV,j} RV_{i,t} + \varepsilon_{RV,j,t} \quad (12)$$

$$rv_t^{hedge,j} \equiv \sum_i b_i^j rv_{i,t} \quad (13)$$

For the main results, the hedging weights, b_i^j , are estimated using the full-sample regression, and we show that the exact weights are unimportant in practice for the qualitative findings. Section 4.2 reports results using hedging weights estimated on a rolling basis, $b_{i,t}^j$, where $b_{i,t}^j$ is estimated with a regression using data up to date $t - 1$.

4 The cost of hedging

This section reports our main results on the price of hedging shocks to volatility and uncertainty. Given a hedging portfolio, the cost of hedging is the negative of the average excess return (risk premium) on the portfolio. For example, holding an iv portfolio represents holding insurance against increases in implied volatility, so if the iv portfolio earns, say, a -10 percent excess return on average, the cost of that insurance is 10 percent on average. A mean excess return cannot be interpreted without reference to the associated volatility – leveraging a portfolio up or down with debt will shift the mean excess return but also the volatility – so we report all risk premia in terms of Sharpe ratios: the mean excess return divided by the standard deviation. The Sharpe ratio reveals the compensation for bearing a risk (or the cost of hedging it) per unit of risk, and is therefore more easily comparable across markets. For reference, the historical Sharpe ratio of US equities in our sample is 0.52.

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. Formally, 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 (i.e. the Arrow–Debreu state prices divided by state probabilities), then

$$E[R_{X,t+1} - R_f] = -cov(M_{t+1} - E_t M_{t+1}, X_{t+1} - E_t X_{t+1}) R_f, \quad (14)$$

where R_f is the gross risk-free rate, which we treat as constant for the sake of exposition and E is the expectation operator (with E_t conditioning on date- t information). The equation says that the negative of the risk premium on a portfolio that hedges the risk X captures the covariance of innovations in X_{t+1} with state prices.²⁰ More generally, when the correlation between R_X and innovations in X is less than 1, $E[R_X - R_f]$ measures the covariance of state prices with the part of innovations to X that is spanned by R_X . So if the premium $E[R_X - R_f]$ is negative, times when R_X , and hence X , rise are bad times, in which state prices are high (in consumption-based models, these are the times when consumption is low and the marginal utility of consumption – which drives M – is high).

Note also that this shows that studying risk premia naturally isolates innovations of the form $X_{t+1} - E_t X_{t+1}$, rather than either levels, X_{t+1} , or structural shocks (intuitively, because the expected value, $E_t X_{t+1}$ is incorporated into the price at time t). That is the sense in which our analysis reveals how investors perceive the effects of uncertainty *shocks*, rather than just the raw correlation of the level of uncertainty with the state of the economy. In other words, hedging costs correspond intuitively to the welfare effects of reduced-form VAR innovations, as opposed to identified structural shocks.²¹

4.1 Hedging uncertainty shocks

The solid series in figure 4 plots sample Sharpe ratios and confidence bands for the various *rv* and *iv* portfolios, which hedge realized and implied volatility in the individual markets. The top panel plots results for *iv* and the bottom panel *rv*. The boxes are point estimates while the bars represent 95-percent confidence bands based on a block bootstrap.²²

Across the top panel, the *iv* portfolios tend to earn zero or even positive returns. For financials, the average Sharpe ratios tend to be near zero, while for the nonfinancials, all 14 sample Sharpe ratios are actually positive. To formally estimate the average Sharpe ratios, we use a random effects model, which yields an estimate of the population mean Sharpe ratio 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. The procedure is described in detail in appendix A.3. The estimated mean Sharpe ratios for just the financial and nonfinancial groups are reported in their respective sections, and the estimated population mean across both groups is in the right-hand section (“overall mean”).

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

²⁰The last term on the right, R_f , is close to 1, and is the same for all assets and all risk factors, so it plays no significant role in interpreting this equation.

²¹In section 5.4, we instead show the results using rotated portfolios that correspond intuitively to identified structural VAR shocks.

²²The 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 premia.

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 (in the form of a negative return) to hedging uncertainty shocks, *uncertainty-hedging portfolios actually earn positive returns*. In particular, for nonfinancials, the average Sharpe ratio is 0.48, and the lower end of the 95-percent confidence interval is 0.25. For the overall mean, the corresponding numbers are 0.39 and 0.17. These are not just statistically but economically significant – portfolios hedging uncertainty shocks have earned average returns nearly as high as the overall stock market. But whereas the stock market is risky, in the sense that it rises in good times and falls in bad, the *iv* portfolios are actually *hedges*, by construction giving positive returns when uncertainty rises. Even for financials, the point estimate for the average Sharpe ratio is positive, though the confidence band runs below zero.

The right-hand section of figure 4 reports the Sharpe ratios for the portfolios hedging the EPU and JLN indexes. Since those hedging portfolios are constructed combining the individual *iv* portfolios (weighting them across the 19 markets to obtain the best hedge for the JLN and EPU indexes), it is not surprising that they are all near zero or positive. The hedging portfolios for JLN financial uncertainty and the EPU index both place relatively more weight on the financials, which have Sharpe ratios close to zero or even slightly negative, so they have overall lower Sharpe ratios. The portfolios hedging macro and price uncertainty, though, since they have larger weights on markets like crude oil, heating oil, and copper, have statistically and economically significantly positive Sharpe ratios, with point estimates both near 0.50, similar to the overall mean for the *iv* portfolios.

As discussed above, even if an investor did not know precisely what weight to put on the various financial or nonfinancial underlyings, it is clear from the figure that almost any set of weights would yield similar results, in the sense that the Sharpe ratios within the nonfinancial and financial categories are all similar, and 18 of the 19 are positive. That fact makes uncertainty about the coefficients in the hedging portfolios unlikely to have important quantitative effects.

The top panel of figure 4 contains all of our key results on the cost of hedging different types of uncertainty shocks. It shows that in our sample spanning almost 30 years, the cost of hedging shocks to uncertainty, whether it is uncertainty in a specific commodity or financial market or a more general macro uncertainty index, has been zero or even negative (the risk premium has been zero or positive).

If uncertainty was perceived to be bad by investors, hedging uncertainty shocks would be costly, and the point estimates in the top panel of figure 4 would be negative – the graph would be the opposite of what we actually see. But at most, some of the *iv* portfolios and hedging portfolios for JLN and EPU have very slightly negative Sharpe ratios. In the majority of the cases – and in particular for uncertainty about the nonfinancial macroeconomy – the Sharpe ratios are statistically and economically significantly positive. In other words, investors have been able to purchase portfolios that directly hedge them against uncertainty shocks and simultaneously earn returns as large as those on the overall stock market.

4.2 Hedging realized volatility shocks

The bottom panel of figure 4 reports analogous results for the cost of hedging realized volatility shocks. The numbers are drastically different. Whereas the *iv* portfolios have historically earned positive returns, the *rv* portfolios have almost all historically earned negative returns. For the S&P 500, this result is well known and is referred to as the variance risk premium. The S&P 500 *rv* portfolio has the most negative Sharpe ratio, at -0.99 – the return to selling insurance against shocks to realized volatility is *twice* as large as the average return on the stock market over the same period. Treasuries also have a significantly negative return, but the other financials in our sample – all currencies – have Sharpe ratios slightly above zero. For the nonfinancials, 10 of 14 estimated Sharpe ratios are negative. So whereas the cost of hedging uncertainty shocks with the *iv* portfolios is consistently negative in the top panel, the cost of hedging realized volatility shocks using the *rv* portfolios is positive in the bottom panel.

As with the *iv* portfolios, we use a random effects model to calculate the population mean Sharpe ratios and report them in the three sections of the figure. In this case, all three estimates – financials, nonfinancials, and all assets – are negative. The values are again statistically and economically significant. The point estimate for the overall mean Sharpe ratio is -0.33 and the upper end of the 95-percent confidence interval is -0.08. Those values are almost the same as what we obtain for the *iv* portfolios, but with the opposite sign.

Finally, the right-hand section of the bottom panel of figure 4 reports the returns from the JLN *rv* hedging portfolios – those that hedge the realized volatility of the JLN macro series.²³ Again, consistent with the fact that the *rv* portfolios themselves consistently earn negative returns, hedging the JLN indexes for realized volatility – as opposed to uncertainty – historically has a positive cost. For all three subindexes, the hedging portfolios earn extremely negative returns, with the Sharpe ratios for financial, real, and price volatility at -1.02, -0.84, and -0.82.

The result that portfolios hedging the uncertainty indexes (JLN and EPU) have positive Sharpe ratios, while those hedging the realized volatility indexes (JLN) have negative Sharpe ratios is robust to the exact weights of used to build the mimicking portfolios. For example, we can construct the weights using an expanding window as opposed to the full sample. The following table reports the annualized Sharpe ratios of the *iv* and *rv* portfolios hedging the indexes with weights estimated with expanding windows. The results are similar to those reported in Figure 4: negative for the *rv* portfolio and positive for the *iv* portfolio.

Sharpe ratios for portfolios hedging the JLN and EPU indexes with rolling weights

| | Financial Unc. | Real Unc. | Price Unc. | EPU |
|-----------|----------------|-----------|------------|------|
| <i>iv</i> | 0.20 | 0.48 | 0.54 | 0.11 |
| <i>rv</i> | -1.16 | -0.19 | -0.83 | |

In sum, in stark contrast to the results for hedging uncertainty, the bottom panel of figure 4

²³There is no realized volatility equivalent of the EPU index, so here we only look at the JLN ones, for which both the uncertainty and the realized volatility versions can be constructed, as discussed above.

shows that there has historically been an extremely large cost to hedge realized volatility. Contracts that, rather than loading on changes in implied volatility, load on actual realized squared returns – which the analysis above shows directly hedge extreme events in the macroeconomy – earn negative Sharpe ratios with magnitudes up to twice as large as that for the overall stock market.

In summary, across both individual markets and also the hedging portfolios for the JLN and EPU indexes, exposure to realized volatility has consistently earned a negative premium, while exposure to implied volatility has earned a zero or positive premium. Investors have therefore historically paid money – by accepting negative returns – to hedge surprise realizations of large shocks, while hedging surprises in uncertainty has had a zero or even negative cost. Those results hold across a wide range of markets that provide hedges against uncertainty in both real activity and aggregate prices.

The finding that there is a negative cost to hedge uncertainty shocks is inconsistent with the view that uncertainty shocks are major drivers of economic declines. If they were – that is, if they were associated with periods of high marginal utility – the equilibrium cost of hedging would be positive. If anything is associated with high marginal utility in our data, it is not periods when investors are particularly uncertain about the future, but periods of high realized volatility, when large movements occur in stock, bond, and commodity markets.

4.3 Hedging average rv and iv

An alternative way to hedge aggregate uncertainty is simply to buy all the iv or rv portfolios simultaneously. Since tables 1 and 3 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 6 reports results of various implementations of such a strategy. 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).

The Sharpe ratios reported in table 6 are generally larger than those in figure 4. The portfolios that are short rv and long iv are able to attain Sharpe ratios well above 1. 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 1 are observed in other contexts (e.g. Broadie, Chernov, and Johannes (2009) for put option returns, Asness and Moskowitz (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 a relatively small premium, it is still generally worthwhile to include financials for the sake of hedging.

Finally, the bottom panel of table 6 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 Robustness

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

5.1 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 premia 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. Appendix figure A.3 is the analog of figure 4, but constructed using one-week returns. The results are qualitatively and quantitatively very similar, confirming the robustness of our analysis to the period considered.

5.2 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), figures A.4 and A.5 replicate the main results in figure 4, 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, figure A.7 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 Sharpe ratios are reasonably stable over time (they should not be constant, even if just for sampling variability – the standard deviation should be theoretically approximately 0.45). In no case do the results appear to be driven by a single outlying period or episode.

5.3 Linear factor models

The evidence presented on the pricing of implied and realized volatility risk relies on the Black–Scholes model to give an approximation for the risk exposures of the portfolios. Appendix A.2 provides evidence that those predictions are an accurate description of the data, but our findings

are not actually dependent on Black–Scholes holding with perfect accuracy. To estimate the price of risk for realized and implied volatility purely empirically, with no appeal to exposures from a theoretical model, we now estimate standard factor specifications which estimate risk exposures freely from the data.

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.²⁴

5.3.1 Specification

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

$$r_{i,n,t} = a_{i,n} + \beta_{i,n}^f \frac{f_{i,t}}{IV_{i,t-1}} + \beta_{i,n}^{f^2} \frac{1}{2} \left(\frac{f_{i,t}}{IV_{i,t-1}} \right)^2 + \beta_{i,n}^{\Delta IV} \frac{\Delta IV_{i,t}}{IV_{i,t-1}} + \varepsilon_{i,n,t}, \quad (15)$$

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 underlying futures return controls for any exposure of the straddles to the underlying, though the Black–Scholes model predicts that effect to be small.

Much more important is the fact that straddles have a nonlinear exposure to the futures return. $(f_{i,t}/IV_{i,t-1})^2$ captures that nonlinearity. Consistent with the construction and interpretation of the rv portfolio, $\beta_{i,n}^{f^2}$ will be interpreted as the exposure of the straddles to realized volatility, since realized volatility is calculated based on squared returns of the underlying.²⁵ Finally, the third factor is the change in the at-the-money implied volatility for the specific market at the five-month maturity.²⁶

We estimate a standard linear specification for the risk premia,

$$\begin{aligned} E[r_{i,n,t}] &= \gamma_i^f \beta_{i,n}^f Std\left(\frac{f_{i,t}}{IV_{i,t-1}}\right) + \gamma_i^{f^2} \beta_{i,n}^{f^2} Std\left(\left(\frac{f_{i,t}}{IV_{i,t-1}}\right)^2\right) + \gamma_i^{\Delta IV} \beta_{i,n}^{\Delta IV} Std\left(\frac{\Delta IV_{i,t}}{IV_{i,t-1}}\right) + \alpha_{i,n}, \\ E[f_{i,t}/IV_{i,t-1}] &= \gamma_i^f Std(f_{i,t}/IV_{i,t-1}). \end{aligned} \quad (16)$$

where $\alpha_{i,n}$ is a fitting error. The γ coefficients represent the risk premia that are earned by investments that provide direct exposure to the factors. That is, the γ 's are estimates of what the Sharpe ratios on the factors would be if it were possible to invest in them directly (neither $f_{i,t}^2$ nor

²⁴The analysis is similar to those of Jones (2006) and Constantinides, Jackwerth, and Savov (2013).

²⁵There are obviously numerous closely related specifications of that second term that could be substituted. We obtain similar results when the second factor is the absolute value of the futures return instead of its square, for example, or when it is measured as the sum of squared daily returns over the return period (recall that the straddle returns cover two weeks, so the factor in that case is the two-week daily realized volatility). We focus on the squared return because it can be interpreted as a second-order term in the pricing kernel and also because it allows a direct link to the gamma of the straddles.

²⁶Since 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.

$\Delta IV_{i,t}$ is an asset return that one can directly purchase in our data; $f_{i,t}$ itself is tradable, though, which is why we impose the second equality). The difference between the method here and the *rv* and *iv* portfolios discussed above is that the factor model does not require assumptions about the risk exposures of the straddles – instead estimating them from (15) – whereas the *rv* and *iv* portfolios rely on the Black–Scholes model. So the results using the factor models should be more robust, but also have more estimation error.

5.3.2 Results

The dashed series in figure 4 plots the estimated risk premia across the various markets along with 95-percent confidence bands. The top panel plots $\gamma_i^{\Delta IV}$, while the bottom panel plots $\gamma_i^{f^2}$. Simple inspection shows that the results are nearly identical to those for the *iv* and *rv* portfolios. The $\gamma_i^{\Delta IV}$ estimates are almost all positive, while the $\gamma_i^{f^2}$ are almost all negative. As before, we produce a random effects estimator of the mean of the risk premia in various groups. The random effects estimates of the means in the various groups are also similar, both in magnitude and statistical significance, to the main results in the solid series. The main difference between the two series is that the confidence bands are wider for the factor model estimates, which is consistent with the fact that the factor model estimates impose less structure and must estimate the factor loadings of the individual straddles.

5.4 Pricing the independent parts of *RV* and *IV*

The main results above report returns associated with assets that hedge innovations to realized and implied volatility. Table 5 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. Equivalently, how does marginal utility respond to an increase in implied volatility, conditional on the level of realized volatility? Given that table 5 shows that the correlation between the *rv* and *iv* returns is small, one should expect this orthogonalization to have small effects on the results.

The marginal effects of realized and implied volatility can be estimated using the stochastic discount factor representation of the factor model estimated in the previous section. Specifically, given the set of straddle returns in each market, one can construct a pricing kernel M_t of the form

$$M_t = \bar{M} - m_i^f \frac{f_{i,t}}{IV_{i,t-1}} - m_i^{f^2} \left(\frac{f_{i,t}}{IV_{i,t-1}} \right)^2 - m_i^{\Delta IV} \frac{\Delta IV_{i,t}}{IV_{i,t-1}} \quad (18)$$

where M_t represents state prices (or marginal utility) and $1 = E_{t-1} M_t R_t$ for any return priced by M . The difference between this specification and that in the previous section is that the coefficients $m^{\cdot\cdot}$ represent the marginal impact of each term on marginal utility, whereas the $\gamma^{\cdot\cdot}$ coefficients represent the premium for total exposure to the factors. Cochrane (2001) discusses the distinction

extensively.

Denoting the covariance matrix of the factors in market i by Σ_i , the m coefficients can be recovered as

$$\begin{bmatrix} m_i^f, m_i^{f^2}, m_i^{\Delta IV} \end{bmatrix}' = \Sigma_i^{-1} \begin{bmatrix} \gamma_i^f, \gamma_i^{f^2}, \gamma_i^{\Delta IV} \end{bmatrix}' \quad (19)$$

Intuitively, the m 's now represent Sharpe ratios on portfolios with exposure to each of the individual factors, orthogonalized to the other two. That is, $m_i^{\Delta IV}$ is the Sharpe ratio for a portfolio exposed to the part of $\frac{\Delta IV_{i,t}}{IV_{i,t-1}}$ that is orthogonal to $\frac{f_{i,t}}{IV_{i,t-1}}$ and $\left(\frac{f_{i,t}}{IV_{i,t-1}}\right)^2$.

Figure A.6 reports the results of this exercise. The findings are qualitatively consistent with the main results in figure 4 and in fact even stronger quantitatively. The marginal effect of an increase in uncertainty on marginal utility, holding realized volatility fixed, is consistently negative, while an increase in realized volatility increases marginal utility. The fact that these results are close to the benchmark case is a consequence of the weak correlation between innovations in realized and implied volatility, so that the rotation by Σ_i^{-1} has small effects.

Figure A.6 also reports premia on orthogonalized versions of the rv and iv portfolios.²⁷ Again, the results are similar to the main analysis.

5.5 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 – e.g. 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 premia 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.

We measure liquidity using two methods. First, since our data set does not include posted bid/ask spreads, we estimate the standard Roll (1984) effective spread using the daily returns, which is a monotone transformation of negative autocorrelation in returns.²⁸ The top panel of appendix figure A.8 plots the effective bid/ask spreads for straddles at maturities of 1, 3, and 5 months for the 19 contracts that we study. The average posted bid/ask spreads for CBOE S&P 500 straddles, for which we have data since 1996, are also reported in the figure. At the one-month

²⁷These are constructed simply through a rotation. The rv_{\perp} portfolio has a positive correlation with rv and zero correlation with iv , while the iv_{\perp} portfolio has zero correlation with rv and a positive correlation with iv .

²⁸The Roll model assumes that there is an unobservable mid-quote that follows a random walk in logs and that observed prices have equal probability of being from a buy or sell order. Bid-ask bounce then induces negative autocorrelation in returns, from which the spread can be inferred (when the autocorrelation is positive, we set the spread to zero).

maturity, the effective spreads are approximately 6 percent on average, which is similar to the 6.6-percent average posted spreads for one-month CBOE S&P 500 straddles since 1996. More importantly, the spreads actually decline at longer maturities indicating that there is less observed negative autocovariance in returns for options at those maturities. For the three- and five-month options, the spreads are smaller by about half, averaging 2 to 3 percent. This is again consistent with posted spreads for CBOE S&P 500 contracts, which decline to 4.0 percent on average for 6-month options.²⁹

As a second measure of liquidity, 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 1, 4, and 7 months. Those spreads are plotted in the bottom panel of figure A.8. For the majority of the options, the spreads are less than 3 percent, consistent with the 4.1-percent bid/ask spread for one-month S&P 500 options at the CBOE. More importantly, though, across nearly all the contracts, the posted spreads again decline with maturity, consistent with the effective spreads. That said, for some of the contracts, there were no available bids or asks at the 4- and 7-month maturities on 8/4/2017. Note also, again similar to the effective spreads, 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)). For crude oil, which is studied in detail in the next section, the spreads at all three maturities are essentially identical to those for the S&P 500.

Figure A.8 yields two important results. First, it shows that the liquidity of the straddles is reasonably high, in the sense that effective and posted spreads are both 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.

Figure A.9 reports the average daily volume of all of the option contracts across maturities 1 to 6 months. For crude oil, which is the focus of the more in-depth study in the next section, the figure reports average daily volume in dollars; for all other contracts, it reports the average daily volume relative to crude oil. Empirically, crude oil options have volume numbers of the same order of magnitude as the S&P 500, while there is more heterogeneity across the other markets. Looking across maturities, the general pattern is that dollar volume declines by about a factor of three in almost all the markets between the 1- and 6-month maturities – so the 6-month maturity has less volume, but far from zero.

Given the decline in volume across maturity, one might worry that prices would be noisier at longer than short maturities (though, again, the bid/ask spreads in figure A.8 belie that view). Since the returns have prices in their denominators, measurement error in prices can potentially

²⁹Even though postead spreads growth in absolute terms with maturity, straddle prices grow by more (approximately with the square root of maturity), causing the percentage spreads to decline.

lead to an upward bias in returns, which could vary across maturities and bias our results. To account for that scenario, section A.2.3 of the appendix shows how returns can be constructed in such a way that they all share the same denominator, and hence the same bias from measurement error. Figure A.10 replicates the main results from figure 4 with these alternative returns and shows that they are little changed, giving further evidence that differences in liquidity across maturities that cause measurement error do not drive the results.

Finally, it is useful to note that while the liquidity of option markets changed significantly in the last 30 years, the patterns in risk premia for the RV and IV portfolios appeared very stable over time (see, for example, the rolling Sharpe ratios of figure A.7), suggesting that liquidity is not the main driver of our results.

6 Case study: crude oil

It is worthwhile to briefly delve more deeply into one market to build confidence in the robustness of the paper’s results. We choose the crude oil market for this exercise because it has one of the longest time series available with the most maturities of any of the markets that we study, it is highly liquid (e.g. Gibson and Schwartz (1990) and Trolle and Schwartz (2010)), and it has a strong link to the macroeconomy.

Figures 5 and 6 contain several plots that help illustrate the historical behavior of the crude oil market. Panel A of figure 5 plots the history of total volume for one- and five-month options (specifically, average daily dollar volume of all contracts with maturities between 15 and 45 or 135 and 165 days to maturity, respectively). The volume of contracts in both maturity bins has risen over time, peaking in 2008, with a subsequent decline. On average, there is about 3 times more volume in the one- than the five-month option, though the volume in the five-month option has been trending upward, reaching as high as 75 percent of the volume for the one-month option.

Panel B of figure 5 plots 5-year rolling sample Sharpe ratios for the *iv* and *rv* portfolios. The left-hand section plots results for crude oil, while, for reference, the right-hand panel plots results for the S&P 500. For crude oil, the *rv* portfolio had negative average returns in almost all five-year periods in our data, while the *iv* portfolio had positive returns in almost all five-year periods. The *rv* returns trend down over time, implying that the variance risk premium may have been growing. The *iv* returns are more consistent, though the returns are close to zero or even negative for short periods at the beginning and end of the sample.

The right-hand side of panel B gives further context to those results by plotting the *rv* and *iv* returns for the S&P 500 options. For the S&P, the *rv* portfolio has relatively more negative returns than for crude, while the *iv* portfolio has average returns that are generally centered on zero, rather than staying consistently positive as we observe for crude oil.

Panel C of figure 5 is similar to panel B, except instead of plotting returns on the *rv* and *iv* portfolios, it plots their constituents, the returns on the one- and five-month straddles. For crude oil, the five-month straddle has consistently *positive* returns, unlike the S&P 500, for which the five-

month straddle tends to have negative returns. In both cases the one-month straddle has negative returns, though that effect is stronger for the S&P 500.

Overall, panels B and C have two uses. First, they show that the returns that we observe on the *iv* and *rv* portfolios are not driven by a small number of outliers; rather, they are consistent over time. Second, they provide further detail on the divergences between the behavior of straddle returns for the S&P 500 compared to crude oil.

Next, to help understand how crude oil volatility relates to macroeconomic uncertainty, the top panel of figure 6 plots one-month at-the-money implied volatility for crude oil along with the JLN financial and price uncertainty series. The correlation of oil price uncertainty with the two series is immediately apparent. The various spikes upward in crude oil volatility are all traceable to spikes in either price or financial uncertainty. This figure thus underscores the utility to an investor of buying five-month crude oil straddles: they provide good protection against increases in the JLN uncertainty indexes and at the same time earn positive average returns.

Because the crude oil market is so large, it has relatively more traded maturities than the other underlyings. At any given time, the CME currently has trading in the next 12 monthly expirations and also December expirations for a number of years into the future. Panel B of figure 6 plots average returns for crude oil straddles with maturities between 1 and 11 months (not 12 because of how we interpolate to construct the monthly portfolios); panel C reports Sharpe ratios. The figure shows that the behavior at longer maturities remains similar, and returns continue to rise slightly beyond the five months examined in the main analysis, though they eventually flatten. When we calculate the *iv* portfolio using the 11- instead of the five-month maturity, we also obtain similar results.

Because crude oil prices are such a widely followed indicator, there are also exchange traded funds (ETFs) that track oil prices, and those ETFs have options traded on them. Appendix A.6 examines the returns on those options and shows that the results are consistent with those we obtain for the CME options, though with more noise because the ETF options were introduced in the 2000's.

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 $\varepsilon_{t+1} \sim N(0, \sigma_t^2)$, 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.

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. Using a wide range of options is important for capturing uncertainty about the real economy and inflation. 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.

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 premia 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 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 the findings in Berger, Dew-Becker, and Giglio (2018), who provide VAR evidence that shocks to volatility predict declines in real activity in the future, while shocks to uncertainty do not.

Berger, Dew-Becker, and Giglio (2018) 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,b) 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.

References

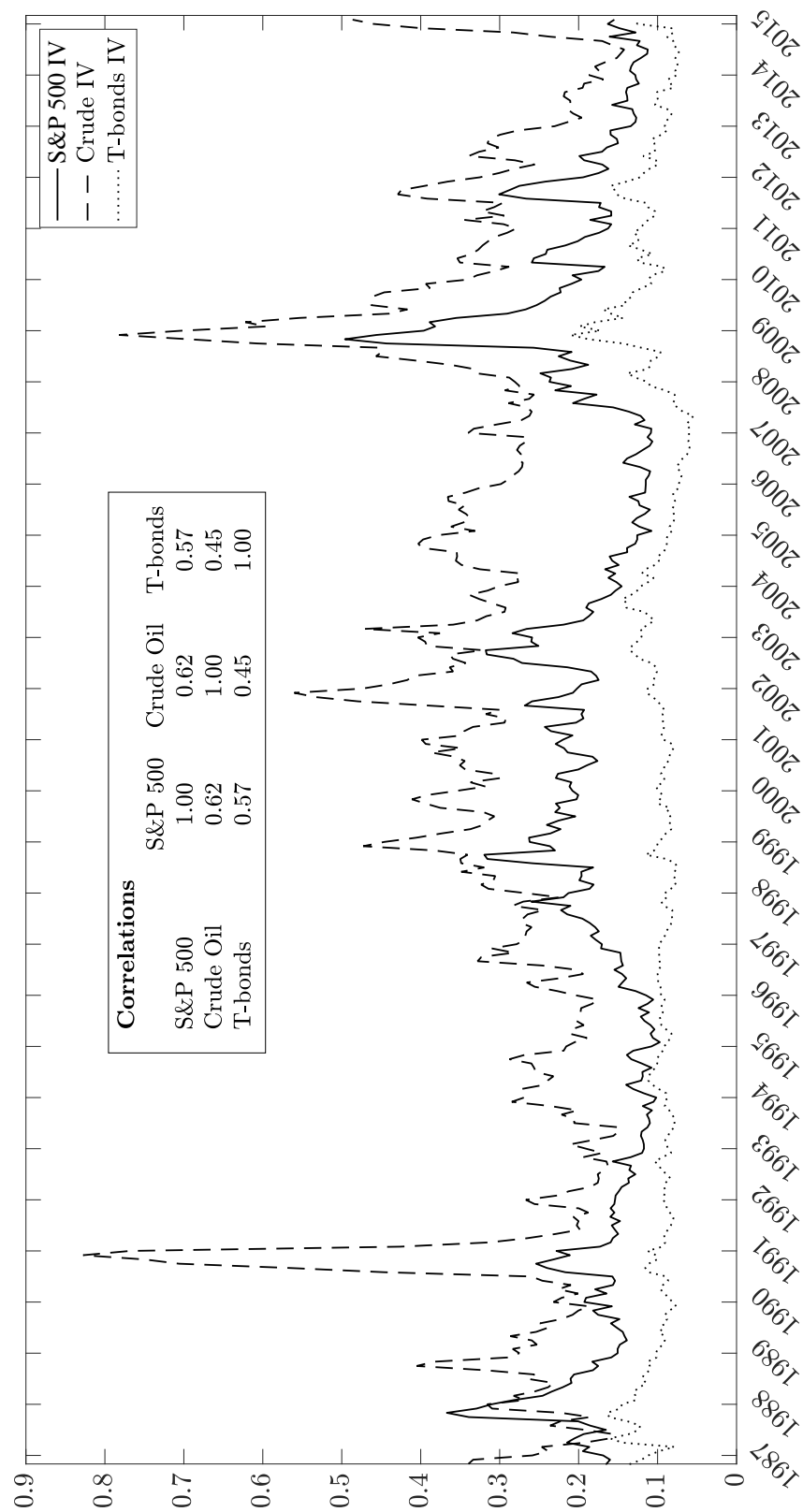
- Abel, Andrew B., “Optimal Investment Under Uncertainty,” *American Economic Review*, 1983, 73 (1), 228–233.
- Ait-Sahalia, Yacine, Mustafa Karaman, and Lorian Mancini, “The Term Structure of Variance Swaps, Risk Premia and the Expectations Hypothesis,” 2015. Working paper.
- Alexopoulos, Michelle and Jon Cohen, “Uncertain Times, Uncertain Measures,” 2009. Working paper.
- Andries, Marianne, Thomas M Eisenbach, Martin C Schmalz, and Yichuan Wang, “The term structure of the price of variance risk,” 2015.
- Asness, Clifford S, Tobias J Moskowitz, and Lasse Heje Pedersen, “Value and momentum everywhere,” *The Journal of Finance*, 2013, 68 (3), 929–985.
- Bachmann, Rudiger and Christian Bayer, ““Wait-and-See” business cycles?,” *Journal of Monetary Economics*, 2013, 60 (6), 704–719.
- and Giuseppe Moscarini, “Business Cycles and Endogenous Uncertainty,” 2012. Working paper.

- Baker, Scott, Nicholas Bloom, and Steven J. Davis**, “Measuring Economic Policy Uncertainty,” 2015. Working paper.
- Bakshi, Gurdip and Nikunj Kapadia**, “Delta-Hedge Gains and the Negative Market Volatility Risk Premium,” *The Review of Financial Studies*, 2003, *16*(2), 527–566.
- , —, and **Dilip Madan**, “Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options,” *Review of Financial Studies*, 2003, *16* (1), 101–143.
- Barro, Robert J.**, “Rare Disasters and Asset Markets in the Twentieth Century,” *Quarterly Journal of Economics*, 2006, *121*(3), 823–866.
- Basu, Susanto and Brent Bundick**, “Uncertainty Shocks in a Model of Effective Demand,” 2015. Working paper.
- Berger, David and Joseph Vavra**, “Volatility and Pass-through,” 2013. Working paper.
- , **Ian Dew-Becker, and Stefano Giglio**, “Uncertainty shocks as second-moment news shocks,” 2017. Working paper.
- Binsbergen, Jules H. Van and Ralph S.J. Koijen**, “The Term Structure of Returns: Facts and Theory,” *Journal of Financial Economics*, 2017, *124* (1), 1–21.
- Black, Fischer**, “The pricing of commodity contracts,” *Journal of Financial Economics*, 1976, *3* (1-2), 167–179.
- and **Myron Scholes**, “The Pricing of Options and Corporate Liabilities,” *The Journal of Political Economy*, 1973, *81* (3), 637–654.
- Bloom, Nicholas**, “The Impact of Uncertainty Shocks,” *Econometrica*, 2009, *77*(3), 623–685.
- , **Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry**, “Really Uncertain Business Cycles,” 2014. Working paper.
- Bollerslev, Tim**, “Generalized autoregressive conditional heteroskedasticity,” *Journal of econometrics*, 1986, *31* (3), 307–327.
- Bondarenko, Oleg**, “Why Are Put Options So Expensive?,” *Quarterly Journal of Finance*, 2014, *04* (03), 1450015.
- Broadie, Mark, Mikhail Chernov, and Michael Johannes**, “Understanding Index Option Returns,” *The Review of Financial Studies*, 2009, *22*(11), 4493–4529.
- Caballero, Ricardo J**, “Aggregate investment,” *Handbook of macroeconomics*, 1999, *1*, 813–862.
- Caldara, Dario, Cristina Fuentes-Albero, Simon Gilchrist, and Egon Zakrajšek**, “The macroeconomic impact of financial and uncertainty shocks,” *European Economic Review*, 2016.
- Cesa-Bianchi, Ambrogio, M. Hashem Pesaran, and Alessandro Rebucci**, “Uncertainty and Economic Activity: A Multi-Country Perspective,” 2018. Working paper.
- Choi, Hoyong, Philippe Mueller, and Andrea Vedolin**, “Bond Variance Risk Premiums,” *Review of Finance*, 2017, *21* (3), 987–1022.
- Christiano, Lawrence J., Roberto Motto, and Massimio Rostagno**, “Risk Shocks,” *American Economic Review*, 2014, *104*(1), 27–65.
- Constantinides, George M, Jens Carsten Jackwerth, and Alexi Savov**, “The puzzle of index option returns,” *Review of Asset Pricing Studies*, 2013, *3* (2), 229–257.
- Coval, Joshua D. and Tyler Shumway**, “Expected Option Returns,” *The Journal of Finance*, 2001, *56*(3), 983–1009.
- Creal, Drew D and Jing Cynthia Wu**, “Monetary policy uncertainty and economic fluctuations,” *International Economic Review*, 2017, *58* (4), 1317–1354.

- Cremers, Martijn, Michael Halling, and David Weinbaum**, “Aggregate Jump and Volatility Risk in the Cross-Section of Stock Returns,” *The Journal of Finance*, 2015, 70 (2), 577–614.
- Decker, Ryan, Pablo N. D’Erasmus, and Hernan Moscoso Boedo**, “Market exposure and endogenous firm volatility over the business cycle,” *American Economic Journal: Macroeconomics*, 2016, 8 (1), 148–198.
- Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez**, “The price of variance risk,” *Journal of Financial Economics*, 2017, 123 (2), 225 – 250.
- Egloff, Daniel, Markus Leippold, and Liuren Wu**, “The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments,” *Journal of Financial and Quantitative Analysis*, 2010, 45(5), 1279–1310.
- Engle, Robert F**, “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation,” *Econometrica: Journal of the Econometric Society*, 1982, pp. 987–1007.
- Gibson, Rajna and Eduardo S Schwartz**, “Stochastic convenience yield and the pricing of oil contingent claims,” *The Journal of Finance*, 1990, 45 (3), 959–976.
- Gilchrist, Simon and John C Williams**, “Investment, capacity, and uncertainty: a putty–clay approach,” *Review of Economic Dynamics*, 2005, 8 (1), 1–27.
- Gourio, Francois**, “Credit Risk and Disaster Risk,” *American Economic Journal: Macroeconomics*, 2013, 5(3), 1–34. Working paper.
- Hartman, Richard**, “The effects of price and cost uncertainty on Investment,” *Journal of Economic Theory*, 1972, 5 (2), 258–266.
- Ilut, Cosmin, Matthias Kehrig, and Martin Schneider**, “Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News,” *NBER Working Paper Series*, 2015.
- Jones, Christopher S**, “A nonlinear factor analysis of S&P 500 index option returns,” *The Journal of Finance*, 2006, 61 (5), 2325–2363.
- Jurado, Kyle, Sydney Ludvigson, and Serena Ng**, “Measuring Uncertainty,” *American Economic Review*, 2015, 105 (3), 1177–1216.
- Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran**, “The Tail that Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation,” 2016. Working paper.
- Leduc, Sylvain and Zheng Liu**, “Uncertainty shocks are aggregate demand shocks,” *Journal of Monetary Economics*, 2016, 82, 20–35.
- Ludvigson, Sydney C, Sai Ma, and Serena Ng**, “Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?,” Technical Report, National Bureau of Economic Research 2015.
- McCracken, Michael W and Serena Ng**, “FRED-MD: A monthly database for macroeconomic research,” *Journal of Business & Economic Statistics*, 2016, 34 (4), 574–589.
- Oi, Walter**, “The Desirability of Price Instability Under Perfect Competition,” *Econometrica*, 1961, 29 (1), 58–64.
- Pástor, L’uboš and Pietro Veronesi**, “Technological Revolutions and Stock Prices,” *The American Economic Review*, 2009, pp. 1451–1483.
- Prokopczuk, Marcel, Lazaros Symeonidis, and Chardin Wese Simen**, “Variance risk in commodity markets,” *Journal of Banking & Finance*, 2017, 81, 136–149.
- Salgado, Sergio, Fatih Guvenen, and Nicholas Bloom**, “Skewed Business Cycles,” 2016. Working paper.
- Seo, Sang Byung and Jessica A. Wacheter**, “Do Rare Events Explain CDX Tranche Spreads?,” *The Journal of Finance*, 0 (ja).

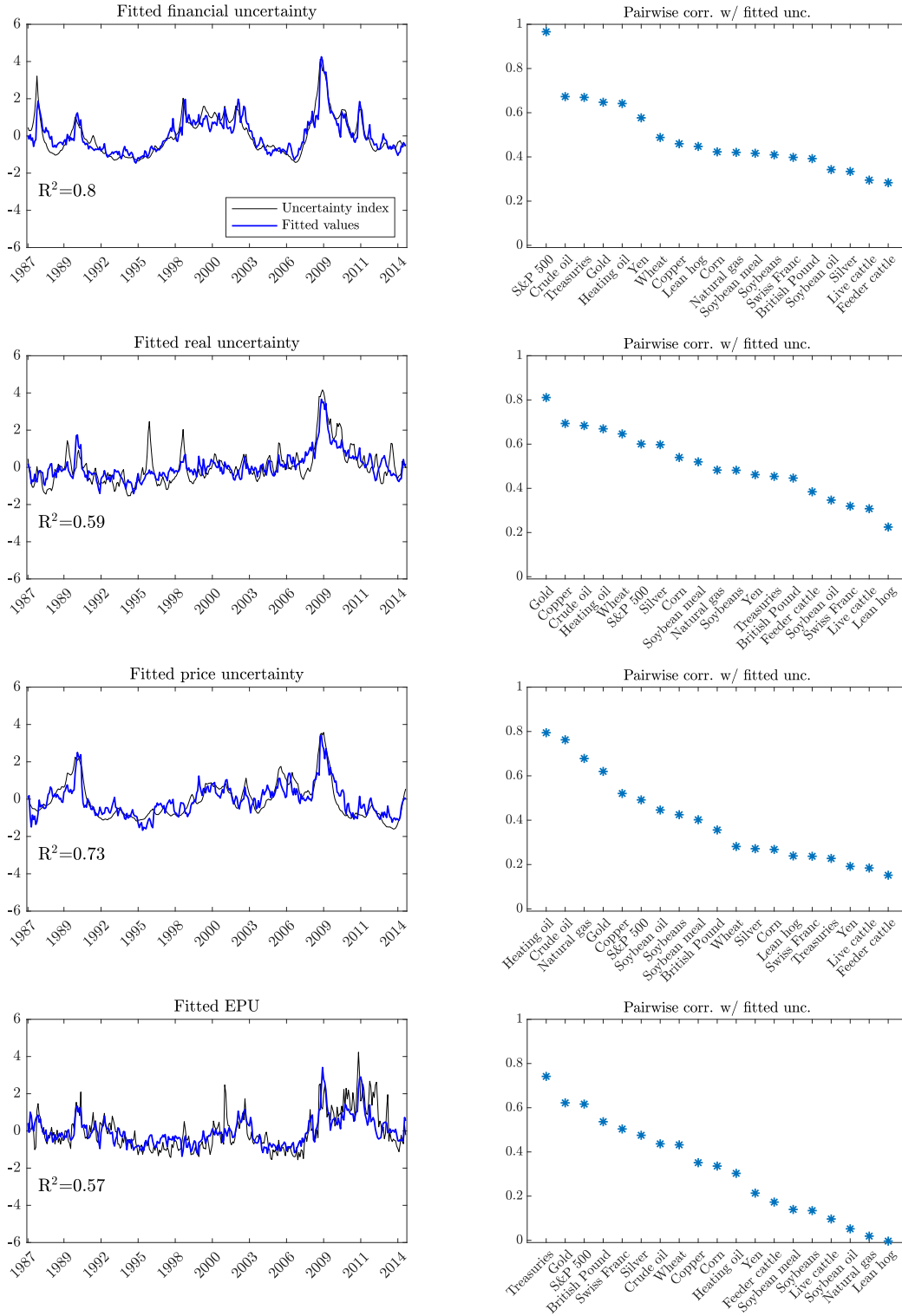
- and **Jessica A Wachter**, “Option prices in a model with stochastic disaster risk,” *Management Science*, 2018.
- Trolle, Anders B and Eduardo S Schwartz**, “Variance risk premia in energy commodities,” *The Journal of Derivatives*, 2010, *17* (3), 15–32.

Figure 1: Sample implied volatilities



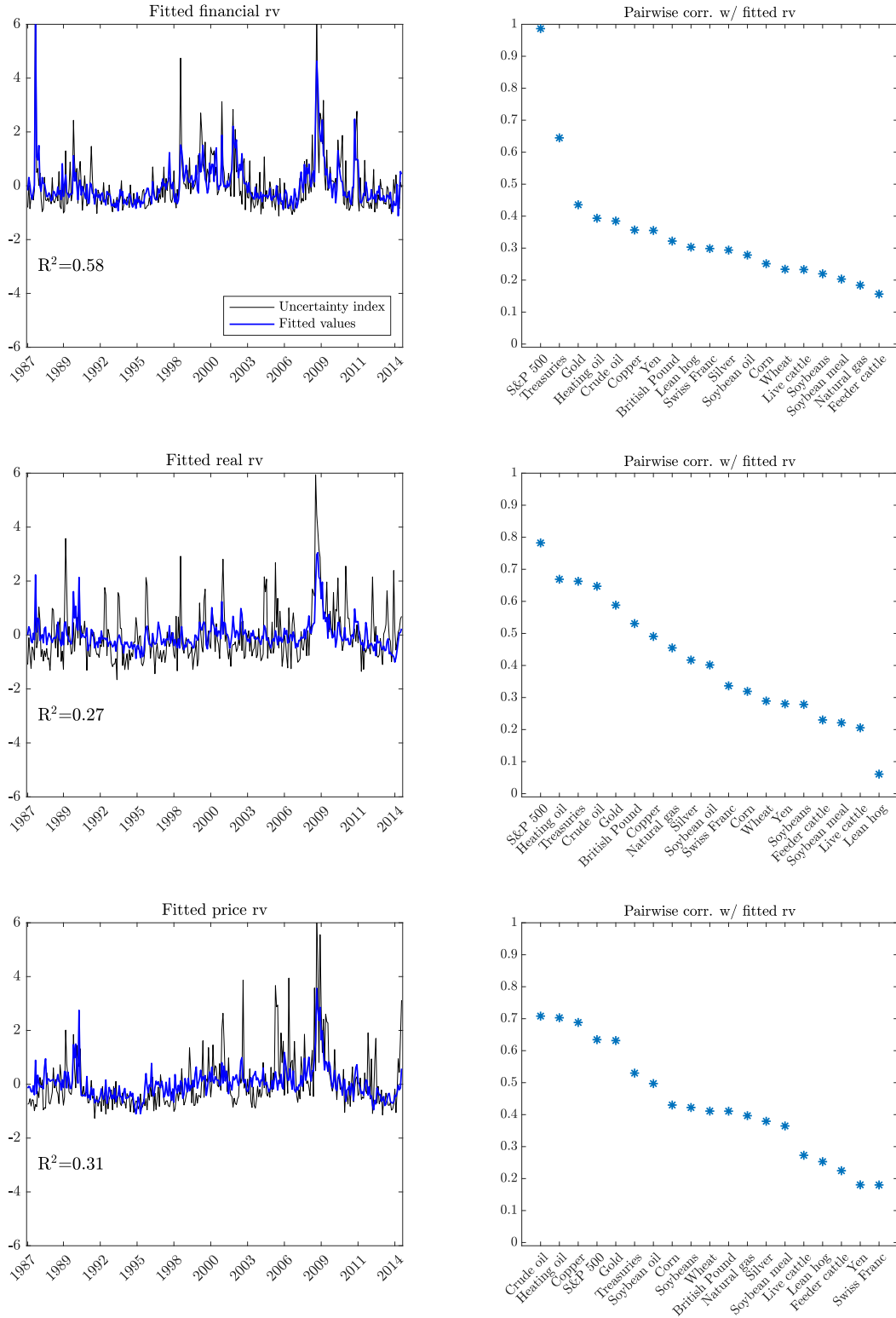
Note: Monthly implied volatilities calculated from three-month options using the Black–Scholes model.

Figure 2: Fit to uncertainty indexes



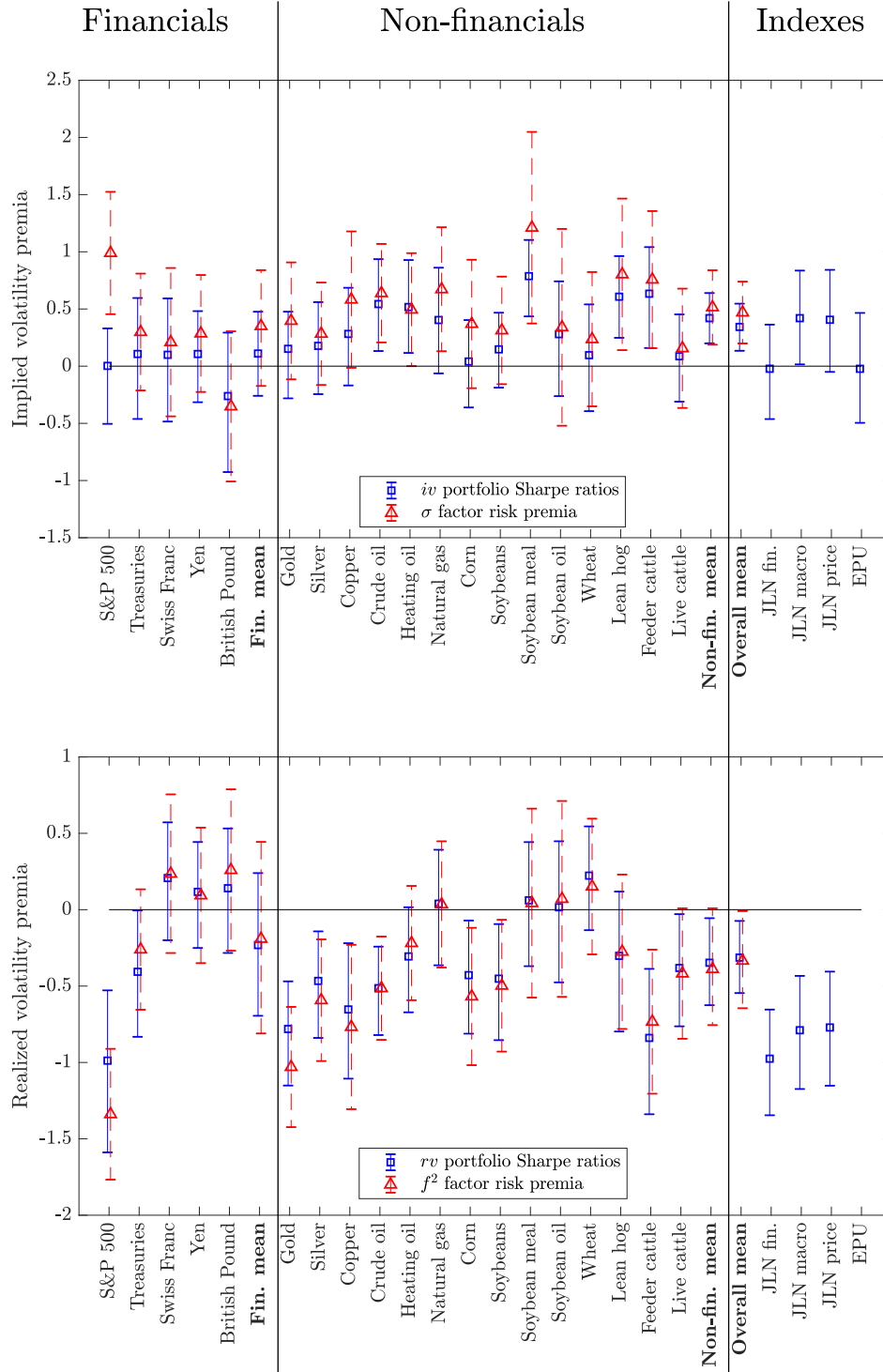
Note: The left-hand panels plot the fitted values from the regressions of the EPU and JLN indexes on three-month implied volatility in the 19 markets. The right-hand panels plot pairwise correlations between the individual implied volatility series and the fitted values from the regressions.

Figure 3: Fit to realized volatility indexes



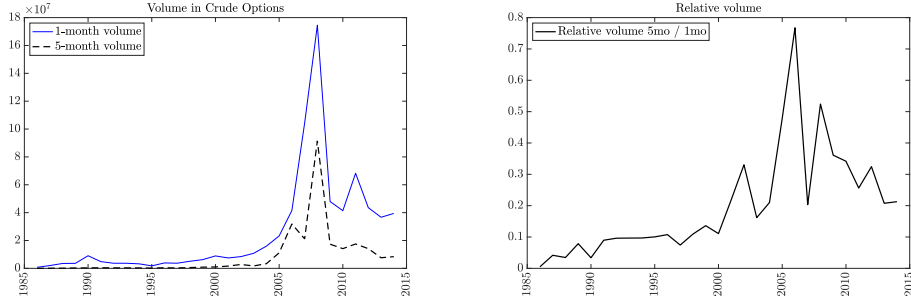
Note: See figure 2. This figure uses the JLN realized volatility series instead of uncertainty.

Figure 4: RV and IV portfolio Sharpe ratios and factor risk premia

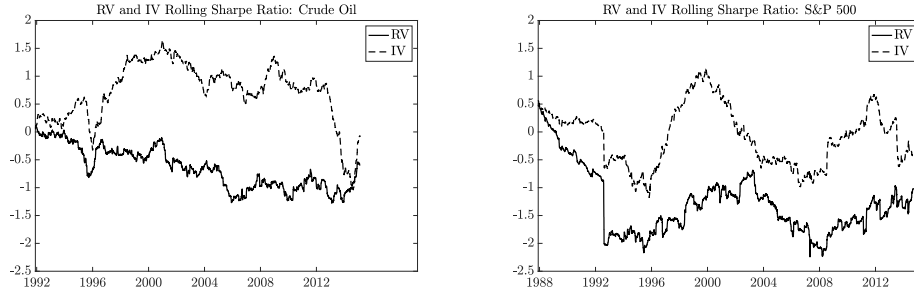


Note: Squares are point estimates and vertical lines represent 95-percent confidence intervals. The solid series plots the Sharpe ratios for the rv and iv portfolios. The dotted series plots the estimated risk premia from the factor model. The confidence bands for the rv and iv Sharpe ratios are calculated 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.

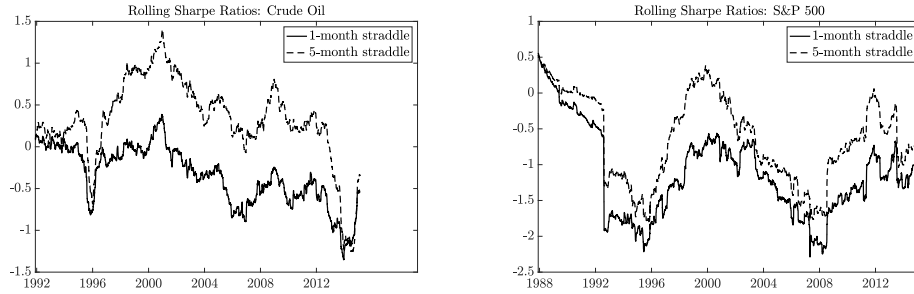
Figure 5: Case study: crude oil (I)



(a) Volume



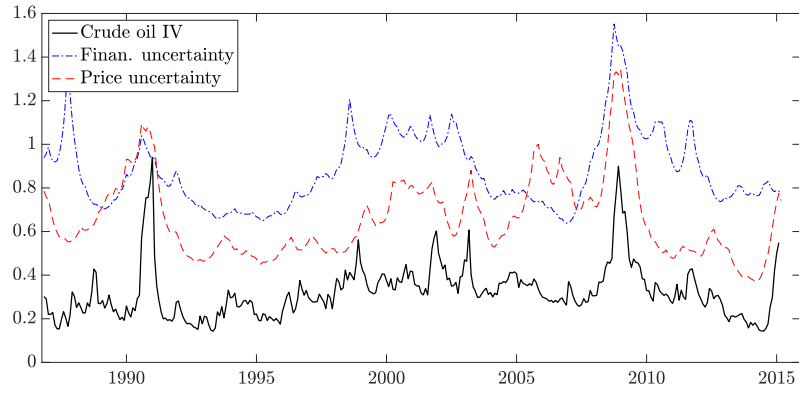
(b) Five-year rolling Sharpe ratios, RV and IV



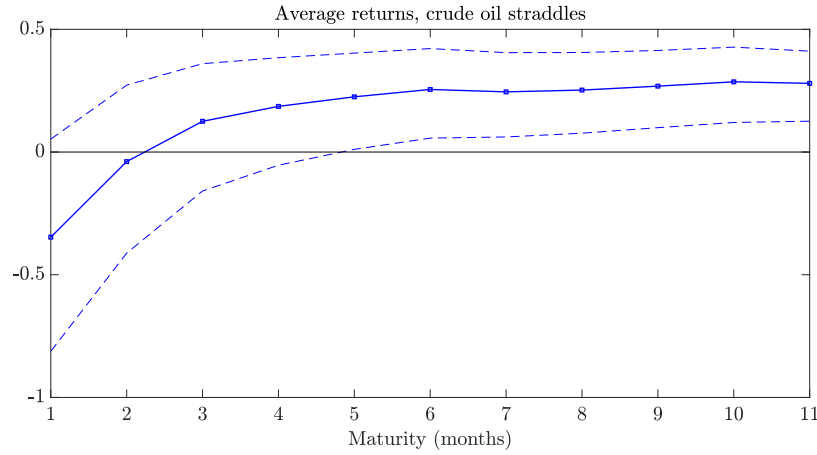
(c) Five-year rolling Sharpe ratios, 1mo and 5mo straddles

Note: The top row reports dollar volume for crude oil options. The middle row reports rolling sharpe ratios for the *rv* and *iv* portfolios for crude oil on the left and the S&P 500 on the right. The bottom panel reports rolling Sharpe ratios for the 1-month and 5-month straddles.

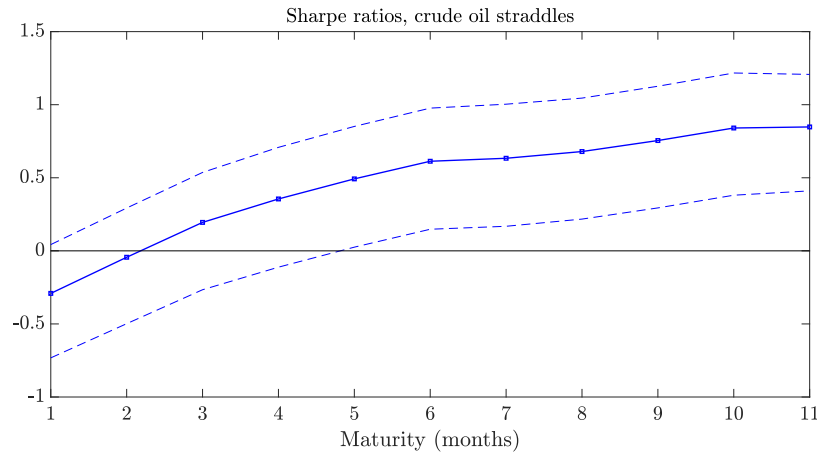
Figure 6: Case study: crude oil (II)



(a) Crude IV and Macro Uncertainty



(b) Straddle average returns



(c) Straddle Sharpe ratios

Note: The top panel reports the Jurado, Ludvigson, and Ng (2015) financial uncertainty series and the macroeconomic price uncertainty series together with the implied volatility for crude oil. The middle and bottom panels plot average returns and Sharpe ratios for straddles along with block-bootstrapped 95-percent confidence intervals.

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 | 0.55 | | | | | | | | | | | | | | | | | |
| Swiss Franc | 0.55 | 0.31 | | | | | | | | | | | | | | | | |
| Yen | 0.39 | 0.56 | 0.49 | | | | | | | | | | | | | | | |
| British Pound | 0.45 | 0.42 | 0.76 | 0.45 | | | | | | | | | | | | | | |
| Gold | 0.51 | 0.57 | 0.24 | 0.27 | 0.37 | | | | | | | | | | | | | |
| Silver | 0.41 | 0.35 | 0.22 | 0.29 | 0.34 | 0.78 | | | | | | | | | | | | |
| Copper | 0.38 | 0.50 | 0.12 | 0.35 | 0.33 | 0.79 | 0.82 | | | | | | | | | | | |
| Crude oil | 0.42 | 0.64 | 0.28 | 0.39 | 0.29 | 0.55 | 0.33 | 0.51 | | | | | | | | | | |
| Heating oil | 0.41 | 0.65 | 0.23 | 0.36 | 0.23 | 0.54 | 0.30 | 0.51 | 0.96 | | | | | | | | | |
| Natural gas | 0.11 | 0.44 | -0.03 | 0.03 | 0.04 | 0.33 | 0.07 | 0.47 | 0.49 | 0.62 | | | | | | | | |
| Corn | 0.25 | 0.37 | -0.09 | 0.14 | 0.12 | 0.51 | 0.57 | 0.61 | 0.23 | 0.20 | 0.13 | | | | | | | |
| Soybeans | 0.21 | 0.35 | -0.03 | 0.16 | 0.17 | 0.47 | 0.48 | 0.62 | 0.30 | 0.29 | 0.22 | 0.85 | | | | | | |
| Soybean meal | 0.29 | 0.35 | -0.05 | 0.17 | 0.07 | 0.55 | 0.52 | 0.64 | 0.32 | 0.30 | 0.24 | 0.82 | 0.93 | | | | | |
| Soybean oil | 0.32 | 0.31 | 0.12 | 0.12 | 0.24 | 0.50 | 0.50 | 0.62 | 0.28 | 0.30 | 0.24 | 0.73 | 0.89 | 0.83 | | | | |
| Wheat | 0.38 | 0.42 | 0.04 | 0.20 | 0.12 | 0.62 | 0.63 | 0.64 | 0.35 | 0.32 | 0.18 | 0.84 | 0.77 | 0.75 | 0.64 | | | |
| Lean hog | 0.28 | 0.41 | -0.03 | 0.28 | -0.09 | 0.27 | 0.16 | 0.37 | 0.41 | 0.46 | 0.41 | 0.29 | 0.37 | 0.40 | 0.39 | 0.36 | | |
| Feeder cattle | 0.45 | 0.36 | 0.14 | 0.17 | 0.09 | 0.40 | 0.51 | 0.55 | 0.33 | 0.34 | 0.13 | 0.49 | 0.47 | 0.50 | 0.48 | 0.53 | 0.42 | |
| Live cattle | 0.50 | 0.28 | 0.25 | 0.18 | 0.08 | 0.37 | 0.40 | 0.46 | 0.33 | 0.38 | 0.27 | 0.31 | 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.

Table 2: Economic activity and option-fitted *uncertainty* indexes**Panel A: regression coefficients**

| Uncertainty measure: | | Financial | Real | Price | EPU |
|----------------------|----------|------------|------------|------------|------------|
| Empl | Fitted | -0.207 *** | -0.321 *** | -0.276 *** | -0.143 ** |
| | Residual | -0.088 | -0.063 | -0.194 *** | -0.029 |
| FFR | Fitted | -0.019 *** | -0.023 *** | -0.012 * | -0.022 *** |
| | Residual | -0.015 | -0.010 | -0.018 * | -0.023 *** |
| IP | Fitted | -0.321 *** | -0.469 *** | -0.439 *** | -0.296 ** |
| | Residual | 0.088 | 0.036 | -0.278 ** | 0.051 |

Panel B: variance decomposition

| Uncertainty measure: | | Financial | Real | Price | EPU |
|----------------------|----------|-----------|-------|-------|-------|
| Empl | Fitted | 0.957 | 0.974 | 0.845 | 0.969 |
| | Residual | 0.043 | 0.026 | 0.155 | 0.031 |
| FFR | Fitted | 0.869 | 0.894 | 0.548 | 0.564 |
| | Residual | 0.131 | 0.106 | 0.452 | 0.436 |
| IP | Fitted | 0.982 | 0.996 | 0.870 | 0.978 |
| | Residual | 0.018 | 0.004 | 0.130 | 0.022 |

Note: Panel A reports the coefficients from regressions of industrial production growth, employment growth, and the Fed funds rate on their own lags and the fitted and residual uncertainty for the four indexes. * indicates significance at the 10-percent level, ** 5-percent, and *** 1-percent. Fitted uncertainty is obtained by projecting each uncertainty index on the 19 IVs. Panel B reports the corresponding variance decomposition.

Table 3: 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 | 0.51 | 0.51 | 0.11 | 0.24 | 0.43 | 0.50 | 0.52 | | | | | | | | | | | |
| 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 | 0.49 | 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 | 0.41 | 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 | 0.06 | 0.11 | 0.12 | 0.11 | 0.12 | | |
| Feeder cattle | 0.22 | 0.20 | 0.03 | 0.04 | 0.06 | 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 4: Economic activity and option-fitted *realized volatility* indexes**Panel A: regression coefficients**

| Uncertainty measure: | | Financial | Real | Price |
|----------------------|----------|------------|------------|------------|
| Empl | Fitted | -0.178 *** | -0.431 *** | -0.389 *** |
| | Residual | -0.077 | -0.010 | -0.109 *** |
| FFR | Fitted | -0.028 *** | -0.040 *** | -0.026 *** |
| | Residual | -0.006 | -0.003 | -0.001 |
| IP | Fitted | -0.267 ** | -0.597 *** | -0.610 *** |
| | Residual | 0.018 | 0.000 | -0.156 |

Panel B: variance decomposition

| Uncertainty measure: | | Financial | Real | Price |
|----------------------|----------|-----------|-------|-------|
| Empl | Fitted | 0.880 | 0.999 | 0.852 |
| | Residual | 0.120 | 0.001 | 0.148 |
| FFR | Fitted | 0.968 | 0.988 | 0.998 |
| | Residual | 0.032 | 0.012 | 0.002 |
| IP | Fitted | 0.997 | 1.000 | 0.873 |
| | Residual | 0.003 | 0.000 | 0.127 |

Note: Same as table 2, but using realized volatilities.

Table 5: Correlations between rv and iv portfolio returns in each market

| | Std(rv) | Std(iv) | Corr(rv,iv) |
|---------------|-------------|-------------|-----------------|
| S&P 500 | 0.03 | 0.09 | 0.56 |
| T-bonds | 0.03 | 0.06 | 0.10 |
| CHF | 0.04 | 0.08 | 0.56 |
| JPY | 0.04 | 0.09 | 0.62 |
| GBP | 0.04 | 0.07 | 0.50 |
| Gold | 0.04 | 0.11 | 0.49 |
| Silver | 0.04 | 0.08 | 0.41 |
| Copper | 0.03 | 0.10 | 0.05 |
| Crude Oil | 0.04 | 0.08 | 0.03 |
| Heating oil | 0.04 | 0.08 | -0.07 |
| Natural gas | 0.05 | 0.08 | -0.20 |
| Corn | 0.04 | 0.08 | 0.12 |
| Soybeans | 0.04 | 0.09 | 0.19 |
| Soybean meal | 0.04 | 0.11 | 0.19 |
| Soybean oil | 0.04 | 0.09 | 0.14 |
| Wheat | 0.04 | 0.08 | 0.08 |
| Lean hog | 0.05 | 0.10 | -0.22 |
| Feeder cattle | 0.04 | 0.10 | 0.02 |
| Live cattle | 0.04 | 0.08 | -0.13 |

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

Table 6: Portfolios of rv and iv across markets

| Panel A: Sharpe ratios | | | rv+iv | | | |
|------------------------|-----------|----------|--------------|--|-------------|--|
| | rv | iv | Equal weight | | Risk-parity | |
| All underlyings | -0.89 *** | 0.65 *** | 1.28 *** | | 1.13 *** | |
| Nonfinancials | -0.80 *** | 0.66 *** | 1.10 *** | | 1.02 *** | |
| Financials | -0.50 *** | 0.10 | 0.61 *** | | 0.33 | |

| Panel B: Skewness | | | rv+iv | | | |
|-------------------|----------|----------|--------------|--|-------------|--|
| | rv | iv | Equal weight | | Risk-parity | |
| All underlyings | 1.11 *** | 1.96 *** | -0.67 *** | | 1.15 *** | |
| Nonfinancials | 1.25 *** | 1.51 *** | -0.91 *** | | 0.80 *** | |
| Financials | 1.82 *** | 3.33 *** | -0.70 | | 3.37 *** | |

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-percent level, ** the 5-percent level, and * the 10-percent level.

A.1 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 proceeds to provide 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 5 times the futures price
2. Options with open interest below the 5th percentile across all contracts in the sample
3. Price less than 5 ticks above zero
4. Maturity less than 9 days
5. Maturity greater than 8 months.
6. Volume equal to zero or missing
7. Options with prices below their intrinsic value (the value if exercised immediately)

We then calculate implied volatilities using the Black–Scholes formula, treating the options as though they are European. We have also replicated the analysis using American implied volatilities and find nearly identical results (the reason 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 percent (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 percent or more on a given day then moves by 15 percent or more in the opposite direction in a single day within the next week, and if it moves by less than 3 percent on average over that window, for options with maturity greater than 90 days (this eliminates temporary large changes in IVs that are reversed that 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 percent 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 percent different from the previous maturity.
- (e) If, looking across strikes at a given maturity on a given date, the IV changes by 20 percent and reverses at the next strike (for strikes within 10 percent of each other).
- (f) If the change in IV at the first or last strike is greater than 20 percent, or the change at the second or second to last option is greater than 30 percent.

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, we first construct returns for straddles with all observable strikes. We calculate ATM straddle returns by averaging across the two closest strikes above and below the current futures price as long as they are less than 0.5 ATM standard deviations from the futures price. Denote the returns on the four straddles in order of increasing strike as R_1 to R_4 , with associated strikes S_1 to S_4 . The interpolated return is then

$$\frac{1}{2} \left(R_2 \frac{S_3 - F}{S_3 - S_2} + R_3 \frac{F - S_2}{S_3 - S_2} \right) + \frac{1}{2} \left(R_1 \frac{S_4 - F}{S_4 - S_1} + R_4 \frac{F - S_1}{S_4 - S_1} \right) \quad (\text{A.1})$$

That is, we linearly interpolate pairwise through R_2 and R_3 and then R_1 and R_4 and average across those two interpolations. The reason to use four straddles instead of two is to try to reduce measurement error. The linear interpolation ensures that the portfolio has an average strike equal to the forward price F . If there is only one straddle available on either side of the forward price, we then interpolate using just a single pair of options, the nearest to the money on either side of the forward price.

To calculate returns at standardized maturities, we again interpolate. If there are options available with maturities on both sides of the target maturity and they both have maturities differing from the target by less than 60 days, then we linearly interpolate. If options are not available on both sides of the target, then we use a single option if it has a maturity within 35 days of the target. This does mean that the maturity of the option used for a portfolio at a desired maturity can deviate from the target.

A.2 Approximating straddle return sensitivities

This section describes the approximation of option returns used to obtain the *rv* and *iv* portfolios. P denotes the price of an at-the-money straddle. σ is the Black-Scholes volatility, n is the time to maturity, F is the forward price, and K is the strike. N denotes the standard Normal cumulative distribution function.

The general formula for the price of a straddle is

$$P(F, \sigma) = e^{-rn} \left(\begin{aligned} &FN \left(\frac{1}{\sigma\sqrt{n}} \left[\log F/K + \frac{\sigma^2}{2}n \right] \right) - KN \left(\frac{1}{\sigma\sqrt{n}} \left[\log F/K - \frac{\sigma^2}{2}n \right] \right) \\ &- FN \left(\frac{-1}{\sigma\sqrt{n}} \left[\log F/K + \frac{\sigma^2}{2}n \right] \right) + KN \left(\frac{-1}{\sigma\sqrt{n}} \left[\log F/K - \frac{\sigma^2}{2}n \right] \right) \end{aligned} \right) \quad (\text{A.2})$$

We calculate the straddles at the strike such that $d_1 = 0$, which is

$$K = F \exp \left(\frac{\sigma^2}{2}n \right) \quad (\text{A.3})$$

The price is then

$$P = e^{-rn} F \exp\left(\frac{\sigma^2}{2}n\right) (-N(-\sigma\sqrt{n}) + N(\sigma\sqrt{n})) \quad (\text{A.4})$$

The derivative of the price with respect to the underlying at that point is zero. The second derivative of the straddle's price is

$$P_{FF}(F_t, \sigma_t) = 2e^{-rn} \frac{N'(0)}{F_t \sigma_t \sqrt{n}} \quad (\text{A.5})$$

The sensitivity to volatility is

$$P_\sigma(F_t, \sigma_t) = 2e^{-rn} F_t N'(0) \sqrt{n} \quad (\text{A.6})$$

The local approximation for returns that we use is

$$\frac{\partial r_{t+1}}{\partial x_{t+1}} = \frac{\partial}{\partial x_{t+1}} \frac{P(F_{t+1}, \sigma_{t+1})}{P(F_t, \sigma_t)} \quad (\text{A.7})$$

and we evaluate the derivatives at the point $F_{t+1} = F_t$, $\sigma_{t+1} = \sigma_t$.

We have

$$\frac{\partial^2 r_{t+1}}{\partial F_{t+1}^2} = \frac{P_{FF}}{P} \quad (\text{A.8})$$

$$= 2e^{-rn} \frac{N'(0)}{F_t \sigma_t \sqrt{n}} \frac{1}{e^{-rn} F \exp\left(\frac{\sigma^2}{2}n\right) (-N(-\sigma\sqrt{n}) + N(\sigma\sqrt{n}))} \quad (\text{A.9})$$

We then use the approximation

$$N(\sigma\sqrt{n}) - N(-\sigma\sqrt{n}) \approx 2N'(0) \sigma_t \sqrt{n} \quad (\text{A.10})$$

and use $\exp\left(\frac{\sigma^2}{2}n\right) \approx 1$, yielding

$$\frac{\partial^2 r_{t+1}}{\partial F_{t+1}^2} \approx \frac{1}{F_t^2 \sigma_t^2 n} \quad (\text{A.11})$$

Since $\partial F_{t+1}^2 / F_t^2 \sigma_t^2 = \partial (f_{t+1} / \sigma_t)^2$, where f_t is the log futures return, we have

$$\frac{\partial^2 r_{t+1}}{\partial (f_{t+1} / \sigma_t)^2} \approx \frac{1}{n} \quad (\text{A.12})$$

For the sensitivity to σ , we have, following similar steps,

$$\frac{\partial r_{t+1}}{\partial \sigma_{t+1}} = \frac{P_\sigma(F_t, \sigma_t)}{P(F_t, \sigma_t)} \quad (\text{A.13})$$

$$\approx \frac{1}{\sigma_t} \quad (\text{A.14})$$

$$\frac{\partial r_{t+1}}{\partial (\Delta \sigma_{t+1} / \sigma_t)} \approx 1 \quad (\text{A.15})$$

A.2.1 Accuracy

To study how effective the above approximation is, we examine a simple simulation. We assume that options are priced according to the Black–Scholes model. We set the initial futures price to 1 and the initial volatility to 30 percent per year. We then examine instantaneous returns (i.e. through shifts in σ and S) on the IV and rv portfolios defined exactly as in the main text, allowing the futures return to vary between $+/- 23.53$ percent, which corresponds to variation out to four two-week standard deviations. We allow volatility to move between 15 and 60 percent – falling by half or doubling.

The top two panels of figure A.1 plot contours of returns on the rv and iv portfolios defined in the main text, while the middle panels plot the contours predicted by the approximations for the partial derivatives. For the iv portfolio, except for very large instantaneous returns – 15–20 percent – the approximation lies very close to the truth. The bottom-right panel plots the error – the middle panel minus the top panel – and except for cases where the underlying has an extreme movement and the implied volatility falls – the exact opposite of typical behavior – the errors are all quantitatively small, especially compared to the overall return.

For the rv portfolio, the errors are somewhat larger. This is due to the fact that we approximate the rv portfolio using a quadratic function, but its payoff has a shape closer to a hyperbola. Again, for underlying futures returns within two standard deviations (where the two-week standard deviation here is 5.88 percent), the errors are relatively small quantitatively, especially when σ does not move far. Towards the corners of the figure, though, the errors grow somewhat large.

These results therefore underscore the discussion in the text. The approximations used to construct the iv and rv portfolios are qualitatively accurate, and except in more extreme cases also hold reasonably well quantitatively. But they are obviously not fully robust to all events, so the factor model estimation, which does not rely on any approximations, should be used in situations where the nonlinearities are a concern.

A.2.2 Empirical return exposures

To check empirically the accuracy of the expressions for the risk exposures of the straddles, figure A.2 plots estimated factor loadings for straddles at maturities from one to five months for each market from time series regressions of the form

$$r_{i,n,t} = a_{i,n} + \beta_{i,n}^f \frac{f_{i,t}}{IV_{i,t-1}} + \beta_{i,n}^{f^2} \frac{1}{2} \left(\frac{f_{i,t}}{IV_{i,t-1}} \right)^2 + \beta_{i,n}^{\Delta IV} \frac{\Delta IV_{i,t}}{IV_{i,t-1}} + \varepsilon_{i,n,t} \quad (\text{A.16})$$

The prediction of the analysis above is that $\beta_{i,n}^f = 0$, $\beta_{i,n}^{f^2} = 1/n$, and $\beta_{i,n}^{\Delta IV} = 1$.

Across the panels, the predictions hold surprisingly accurately. The loadings on $f_{i,t}$ are all near zero, if also generally slightly positive. The loadings on the change in implied volatility are all close to 1, with little systematic variation across maturities. And the loadings on the squared futures return tend to begin near 1 (though sometimes biased down somewhat) and then decline monotonically, consistent with the predicted n^{-1} scaling.

Table A.1 reports results of similar regressions for each underlying of the returns on the rv and iv portfolios on the underlying futures return, the squared futures return, and the change in implied volatility. The table shows that while the Black–Scholes predictions do not hold perfectly, it is true that the rv portfolio is much more strongly exposed to realized than implied volatility,

and the opposite holds for the *iv* portfolio. The coefficients on $(f_t/\sigma_{t-1})^2$ average 0.76 for the *rv* portfolio and 0.10 for the *iv* portfolio (though that average masks some variation across markets). Conversely, the coefficients on $\Delta\sigma_t/\sigma_{t-1}$ average 0.03 for the *rv* portfolio and 0.79 for the *iv* portfolio. Furthermore, the R^2 s are large, averaging 72 percent across the various portfolios, implying that their returns are well described by the approximation (5).

A.2.3 Alternative scaling for returns

Because returns have a price in the denominator, if that price is measured with error, returns can be biased upwards. The *iv* portfolio is net long the straddles, while the *rv* portfolio has a total weight of zero, so measurement error in prices would bias *iv* returns up but not *rv* returns. To account for that possibility, this section examines results when all the straddle returns are scaled by the price of the one-month straddle, instead of the price of a straddle with the same maturity.

Specifically, denoting $P_{n,t}$ the price of a straddle of maturity n on date t , the return on an n -month straddle used in the main results is

$$R_{n,t} = \frac{P_{n-1,t+1} - P_{n,t}}{P_{n,t}} \quad (\text{A.17})$$

We consider returns on a portfolio that puts weight $\frac{P_{n,t}}{P_{1,t}}$ on the n -month straddle and weight $1 - \frac{P_{n,t}}{P_{1,t}}$ on the risk-free asset (which is a tradable portfolio), which is

$$r_{n,t+1}^{rescaled} = \frac{P_{n-1,t+1} - P_{n,t}}{P_{n,t}} \frac{P_{n,t}}{P_{1,t}} + \left(1 - \frac{P_{n,t}}{P_{1,t}}\right) r_{f,t} \quad (\text{A.18})$$

$$= \frac{P_{n-1,t+1} - P_{n,t}}{P_{1,t}} + \left(1 - \frac{P_{n,t}}{P_{1,t}}\right) r_{f,t} \quad (\text{A.19})$$

This portfolio is useful for two reasons. First, the one-month maturity has the highest volume in most markets we study, and it is typically considered to be the most accurate. Second, this eliminates differences in bias across maturities since in this specification, the denominator is the same for all n .

For $r_{n,t+1}^{rescaled}$, similar calculations to those above yield the results that

$$\frac{\partial^2 r_{n,t+1}^{rescaled}}{\partial (f_{t+1}/\sigma_t)^2} \approx \frac{1}{\sqrt{n}} \quad (\text{A.20})$$

$$\frac{\partial r_{n,t+1}^{rescaled}}{\partial (\Delta\sigma_{t+1}/\sigma_t)} \approx \sqrt{n} \quad (\text{A.21})$$

We then calculate alternative *rv* and *iv* portfolios as

$$iv_t^{rescaled} = \frac{3}{\sqrt{12}} \left(\sqrt{5/12} r_{5,t}^{rescaled} - \sqrt{1/12} r_{1,t}^{rescaled} \right) \quad (\text{A.22})$$

$$rv_t^{rescaled} = \frac{5/48}{\sqrt{12}} \left(\sqrt{12} r_{1,t}^{rescaled} - \sqrt{12/5} r_{5,t}^{rescaled} \right) \quad (\text{A.23})$$

Figure A.10 replicates 4 with the rescaled returns. The results are nearly identical to the baseline for both the Sharpe ratios on the *iv* and *rv* portfolios and the estimated factor risk premia. The

price of risk for implied volatility is slightly less positive, but otherwise there is no quantitatively notable difference. These results show that when we correct for the potential bias induced by low liquidity and measurement error at longer maturities, the estimates are essentially unchanged.

A.3 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,

$$sr_i \sim N(\mu_{sr}, \Sigma_{sr}) \quad (\text{A.24})$$

This section estimates the parameters μ_{sr} and Σ_{sr} . 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 is variation is likely driven by sampling error. Σ_{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 \hat{sr}_i , and the stacked vector of sample Sharpe ratios as $\hat{\mathbf{sr}} \equiv [\hat{sr}'_1, \hat{sr}'_2, \dots]'$. Similarly, denote the vector of true Sharpe ratios as $\mathbf{sr} \equiv [sr'_1, sr'_2, \dots]'$. Under the central limit theorem,

$$\hat{\mathbf{sr}} \Rightarrow N(\mathbf{sr}, \Sigma_{\hat{\mathbf{sr}}}) \quad (\text{A.25})$$

where \Rightarrow denotes convergence in distribution and the covariance matrix $\Sigma_{\hat{\mathbf{sr}}}$ depends on the covariance between all the returns, across both maturities and underlyings, along with the lengths of the various samples.¹ Appendix A.4 describes how we construct $\Sigma_{\hat{\mathbf{sr}}}$.

The combination of (A.24) and (A.25) 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 $\mu_{sr} + \mu_D I_F$, 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.²

A.4 Calculating the covariance of the sample mean returns

There are two features of our data that make calculating covariance matrix of sample means difficult: we have an unbalanced panel and the covariance matrix is either singular or nearly so. We deal with those issues through the following steps.

¹More formally, we would say that $\hat{\mathbf{sr}}$ properly scaled by the square root of the sample size converges to a normal distribution. The expression (A.25) implicitly puts the sample size in $\Sigma_{\hat{\mathbf{sr}}}$. 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.

²We 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.

1. For each market, we estimate the two largest principal components, therefore modeling straddle returns for underlying i and maturity n on date t as

$$r_{i,n,t} = \lambda_{1,i,n} f_{1,i,t} + \lambda_{2,i,n} f_{2,i,t} + \theta_{i,n,t} \quad (\text{A.26})$$

where the λ are factor loadings, the f are estimated factors, and θ is a residual that we take to be uncorrelated across maturities and markets (it is also in general extremely small).

2. We calculate the long-run covariance matrix of all $J \times 2$ estimated factors. The covariance matrix is calculated using the Hansen–Hodrick method to account for the fact that the returns are overlapping (we use daily observations of 2-week returns). The elements of the covariance matrix are estimated based on the available nonmissing data for the associated pair of factors. That means that the covariance matrix need not be positive semidefinite. To account for that fact, we set all negative eigenvalues of the estimated covariance matrix to zero.

Given the estimated long-run covariance matrix of the factors, denoted Σ_f , and given the (diagonal) long-run variance matrix of the residuals θ , denoted Σ_θ , the long-run covariance matrix of the returns is then

$$\Sigma_r \equiv \Lambda \Sigma_f \Lambda' + \Sigma_\theta \quad (\text{A.27})$$

where Λ is a matrix containing the factor loadings λ .

3. Finally, it is straightforward to show that the covariance matrix of the sample mean returns is

$$\Sigma_{\hat{r}} = M \odot \Sigma_r \quad (\text{A.28})$$

where \odot denotes the elementwise product and M is a matrix where the element for a given return pair is equal to the ratio of the number of observations in which both returns are available to the product of the number of observations in which each return is available individually (if all returns had the same number of observations T , then we would obtain the usual T^{-1} scaling). We then have the asymptotic approximation that

$$\hat{r} \Rightarrow N(\bar{r}, \Sigma_{\hat{r}}) \quad (\text{A.29})$$

where \hat{r} is a vector that stacks the \hat{r}_i and \bar{r} stacks the \bar{r}_i and \Rightarrow denotes convergence in distribution.

To construct $\Sigma_{\hat{r}}$, we simply divide the i, j element of $\Sigma_{\hat{r}}$ by the product of the sample standard deviations of r_i and r_j .

A.5 Calculating risk prices with unbalanced panels and correlations across markets

In estimating the factor models, we have two complications to deal with: the sample length for each underlying is different, and returns are correlated across underlyings. This section discusses how we deal with those issues.

We have the model

$$E_{T_i} [R_i] = \lambda_i \beta_i + \alpha_i \quad (\text{A.30})$$

where E_{T_i} denotes the sample mean in the set of dates for which we have data for underlying i , R_i is the vector of returns of the straddles, λ_i is a vector of risk prices, β_i is a vector of risk prices, and α_i is a vector of pricing errors. Note that these objects are all population values, rather than estimates. In order to calculate the sampling distribution for the estimated counterparts, we need to know the covariance of the pricing errors. Note that there is also a population cross-sectional

regression with

$$E_{T_i} [R_i] = a_i + \beta_i E_{T_i} [f_i] + E_{T_i} [\varepsilon_i] \quad (\text{A.31})$$

where ε_i is a vector of residuals and f_i is a vector of pricing factors. That formula can be used to substitute out returns and obtain

$$\alpha_i = a_i + \beta_i E_{T_i} [f_i] + E_{T_i} [\varepsilon_i] - \lambda_i \beta_i \quad (\text{A.32})$$

Since a_i , λ_i , and β_i are fixed in the true model, the distribution of α_i depends only on the distributions of the sample means $E_{T_i} [f_i]$ and $E_{T_i} [\varepsilon_i]$. Denoting the long-run (i.e. Hansen–Hodrick) covariance matrix of f_i as Σ_{f_i} and that of ε_i as Σ_{ε_i} , we have

$$\text{var}(\alpha_i) = \beta_i T_i^{-1} \Sigma_{f_i} \beta_i' + T_i^{-1} \Sigma_{\varepsilon_i} \quad (\text{A.33})$$

Since the λ_i are estimated from a regression, if we denote their estimates as $\hat{\lambda}_i$, we obtain the usual formula for the variance of $\hat{\lambda}_i - \lambda_i$

$$\text{var}(\hat{\lambda}_i - \lambda_i) = (\beta_i' \beta_i)^{-1} \beta_i' \text{var}(\alpha_i) \beta_i (\beta_i' \beta_i)^{-1} \quad (\text{A.34})$$

$$= \Sigma_f + (\beta_i' \beta_i)^{-1} \beta_i' \Sigma_{\varepsilon_i} \beta_i (\beta_i' \beta_i)^{-1} \quad (\text{A.35})$$

Beyond the variance of $\hat{\lambda}_i$, we also need to know the covariance of any pair of estimates, $\hat{\lambda}_i$ and $\hat{\lambda}_j$. Using standard OLS formulas, we have

$$\begin{bmatrix} \hat{\lambda}_i - \lambda_i \\ \hat{\lambda}_j - \lambda_j \end{bmatrix} = \begin{bmatrix} (\beta_i' \beta_i)^{-1} \beta_i' \alpha_i \\ (\beta_j' \beta_j)^{-1} \beta_j' \alpha_j \end{bmatrix} \quad (\text{A.36})$$

$$= \begin{bmatrix} (\beta_i' \beta_i)^{-1} \beta_i' (\beta_i E_{T_i} [f_t] + E_{T_i} [\varepsilon_{j,t}]) \\ (\beta_j' \beta_j)^{-1} \beta_j' (\beta_j E_{T_j} [f_t] + E_{T_j} [\varepsilon_{j,t}]) \end{bmatrix} \quad (\text{A.37})$$

The covariance between $\hat{\lambda}_i$ and $\hat{\lambda}_j$ is then

$$\frac{T_{12}}{T_1 T_2} \left(\Sigma_{f,i,j} + (\beta_1' \beta_1)^{-1} \beta_1' \Sigma_{\varepsilon,i,j} \beta_2 (\beta_2' \beta_2)^{-1} \right) \quad (\text{A.38})$$

where $\Sigma_{f,i,j}$ and $\Sigma_{\varepsilon,i,j}$ are now long-run covariance matrices (again from the Hansen–Hodrick method). Using these formulas, we then have estimates of risk prices in each market individually along with a full covariance matrix of all the estimates.

A.6 Robustness: ETF options

This section provides an alternative check on the results for crude oil options by examining returns on straddles for options on two exchange traded funds. The first is the United States Oil Fund (USO), which invests in short-term oil futures. USO has existed since 2006, and Optionmetrics reports quotes for options beginning in May, 2007. The second fund is the Energy Select Sector SPDR fund (XLE), which tracks the energy sector of the S&P 500. XLE has existed since 1998 and Optionmetrics reports data since December, 1998.

We eliminate observations using the following filters:

1. Volume less than 10 contracts
2. Time to maturity less than 15 days
3. Bid-ask spread greater than 20 percent of bid/ask midpoint
4. Initial log moneyness – log strike divided by the futures price – greater than 0.75 implied volatility units in absolute value (where implied volatility is scaled by the square root of time to maturity).

We then calculate straddle returns as in the main text over two-week periods and average across the two straddles nearest to the money for each maturity, weighting them by the inverse of their absolute moneyness.

The top section of table A.6.1 reports the number of (potentially overlapping) two-week straddle return observations across maturities for USO, XLE, and the CME Group futures options used in the main analysis. Since the CME data goes back to 1983, there are far more observations for that series than the other two. More interestingly, though, the number of observations only declines by about 10 percent between the 1- and 6-month maturities, while it falls by more than 2/3 for the XLE and USO samples. The CME data therefore has superior coverage at longer horizons, which justifies its use in our main analysis.

The bottom section of table A.6.1 reports the correlations of the USO and XLE straddle returns with those for the CME on the days where they overlap. The correlations are approximately 90 percent at all maturities for USO and 50 percent for XLE. The 90-percent correlations for USO and the CME sample provide a general confirmation of the accuracy of the CME straddle returns, since we would expect the USO and CME options to be highly similar as USO literally holds futures. The lower correlation for XLE is not surprising given that it holds energy sector stocks rather than crude oil futures.

Table A.6.1.

| | Maturity: | 1 | 2 | 3 | 4 | 5 | 6 |
|----------|-----------|------|------|------|------|------|------|
| # obs. | USO | 1640 | 1616 | 1721 | 1679 | 1118 | 525 |
| | XLE | 2612 | 2545 | 2454 | 1928 | 1134 | 369 |
| | CME | 6762 | 6645 | 6817 | 6801 | 6606 | 5998 |
| Corr. w/ | USO | 0.93 | 0.96 | 0.95 | 0.92 | 0.89 | 0.83 |
| | XLE | 0.43 | 0.48 | 0.50 | 0.49 | 0.50 | 0.53 |

In the main text, the RV and IV portfolio returns are calculated using 5- and 1-month straddles. Since the number of observations drops off substantially between 4 and 5 months for both XLE and USO, though, here we examine returns on RV and IV portfolios using both 5- and 4-month straddles for the long-maturity side.

Figure A.11 plots estimated annualized Sharpe ratios along with 95-percent confidence bands for the RV and IV portfolios using 4- and 5-month straddles for the three sets of options. In all four cases, the three confidence intervals always overlap substantially. The fact that the sample for the CME options is far larger is evident in its confidence bands being much narrower than those for the other two sources. For the IV portfolios, USO has returns that are close to zero, but its confidence bands range from -1 to greater than 0.5, indicating that it is not particularly informative about the Sharpe ratio.

Table A.6.2 reports confidence bands for the *difference* between the IV and RV average returns constructed with the CME data and the same portfolios constructed using USO and XLE. The top panel shows that the differences for the IV portfolios are negative for USO and positive for XLE, but only the difference for USO constructed with the 4-month straddle is statistically significant. The bottom panel similarly shows mixed results for the point estimates for the differences for the

RV portfolios, with none of the differences being statistically significant.

Table A.6.2. Differences between CME and USO, XLE mean returns

| | USO minus CME, 4mo. | USO minus CME, 5mo. | XLE minus CME, 4mo. | XLE minus CME, 5mo. |
|-----------|---------------------|---------------------|---------------------|---------------------|
| IV return | -2.2 [-3.9,-0.2] | -2.2 [-4.8,0.4] | -0.8 [-2.5,4.1] | -1.4 [-4.1,6.3] |
| RV return | 0.43 [-0.6,1.4] | 0.47 [-0.6,1.4] | -0.27 [-1.8,1.3] | 0.67 [-1.5,2.6] |

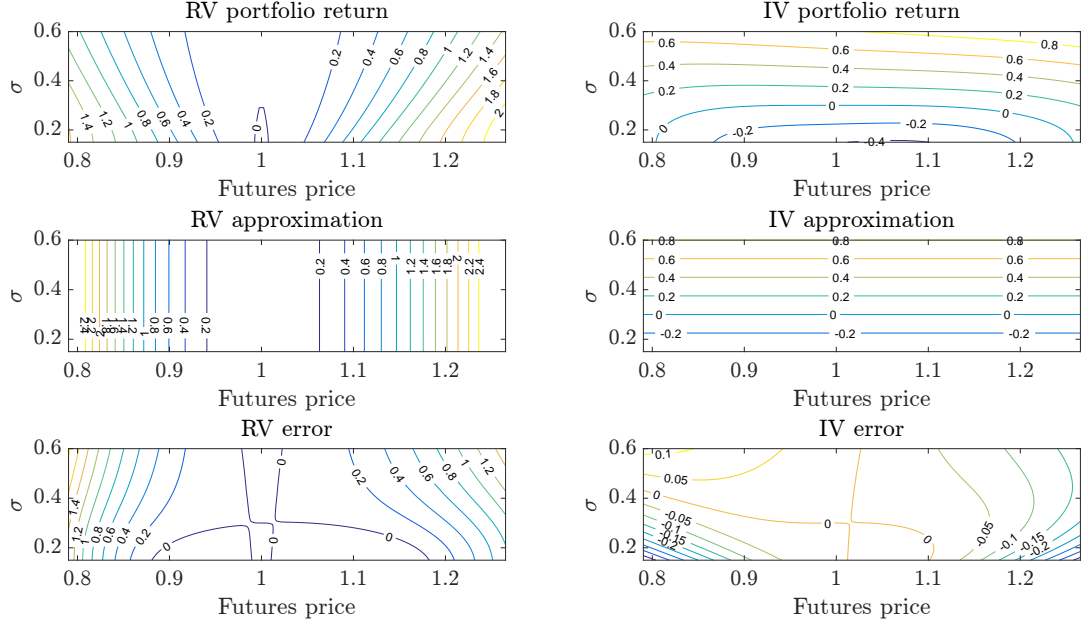
Notes: the table reports percentage (two-week) returns on USO and XLE minus returns on CME RV and IV portfolios. 95-percent confidence intervals are reported in brackets.

The fact that the USO and CME straddle returns are highly correlated does not necessarily mean that the CME data is accurate for the mean return on the straddles. To check whether the difference in the means observed in the USO and XLE data would affect our main results, we ask how the Sharpe ratios of the RV and IV portfolios in the CME data would change if we shifted their means by the average differences reported in table A.6.2. The bars labeled “CME, USO adj.” and “CME, XLE adj.” show how the confidence bands would change if we shifted them by exactly the point estimates from table A.6.2. Note that this is not the same as shifting the Sharpe ratio for the CME data to match that for the XLE or USO data. The reason is that the difference in table A.6.2 is calculated only for the returns on matching dates, whereas the Sharpe ratio calculated in figure A.11 is calculated using the full sample for the CME data. So the two adjusted bands take the full-sample band and then shift it by the mean difference calculated on the dates that overlap between the CME data and XLE or USO.

Figure A.11 shows that the economic conclusions drawn for the crude oil straddles are not changed if the mean returns are shifted by the differences observed in table A.6.1. The RV portfolio returns remain statistically significantly negative in all four cases, the changes in the point estimates are well inside the original confidence intervals. The top panel shows that the IV returns using 5-month straddles are similarly unaffected. For the 4-month straddles, the only difference is that with the USO options, the estimated Sharpe ratio falls by about half and is no longer statistically significantly greater than zero. So, again, out of eight cases – IV and RV with 4- and 5-month straddles – in only one is there a nontrivial change in the conclusions, and even there the Sharpe ratio on the IV portfolio does not become negative, it is simply less positive.

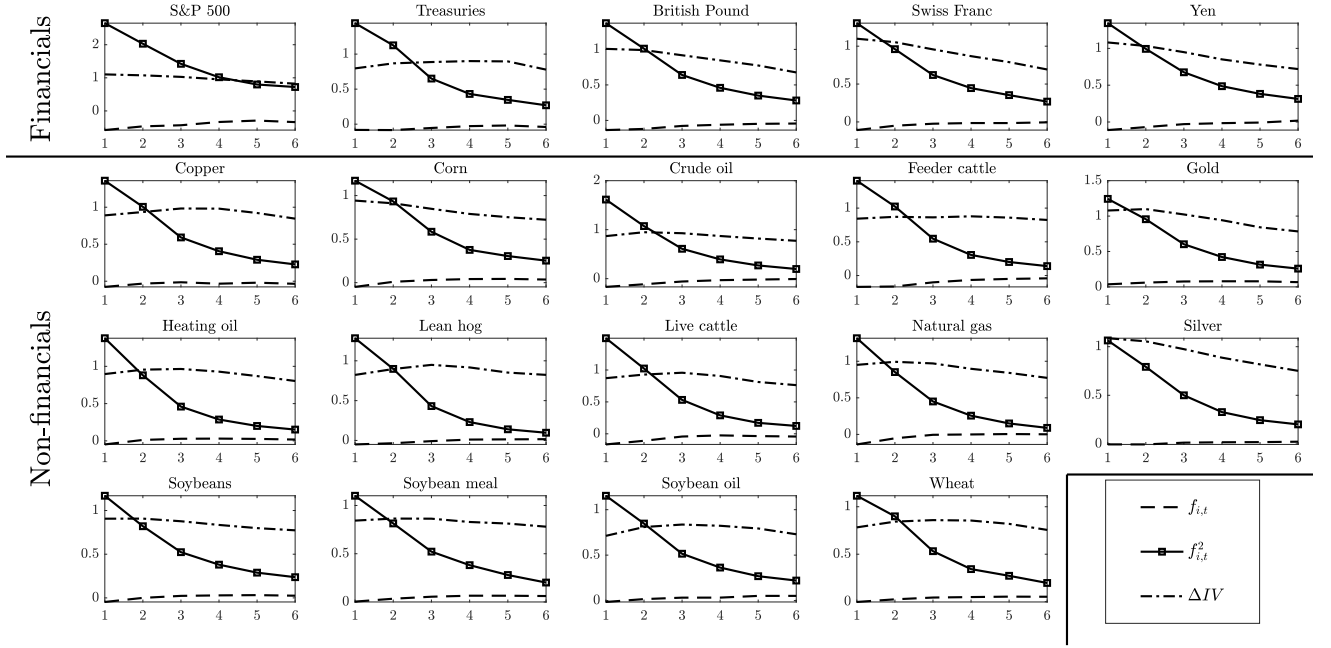
Overall, the period in which the USO and XLE options are traded is too short to use them for our main analysis. This section shows that the USO straddle returns are highly correlated with the CME returns. The mean returns on the XLE and CME straddles are highly similar, while they differ somewhat more for CME and USO. However, shifting the means used for the CME options in the main analysis by the observed difference between the CME and USO options does not substantially change any of the conclusions.

Figure A.1: rv and iv portfolio approximation errors



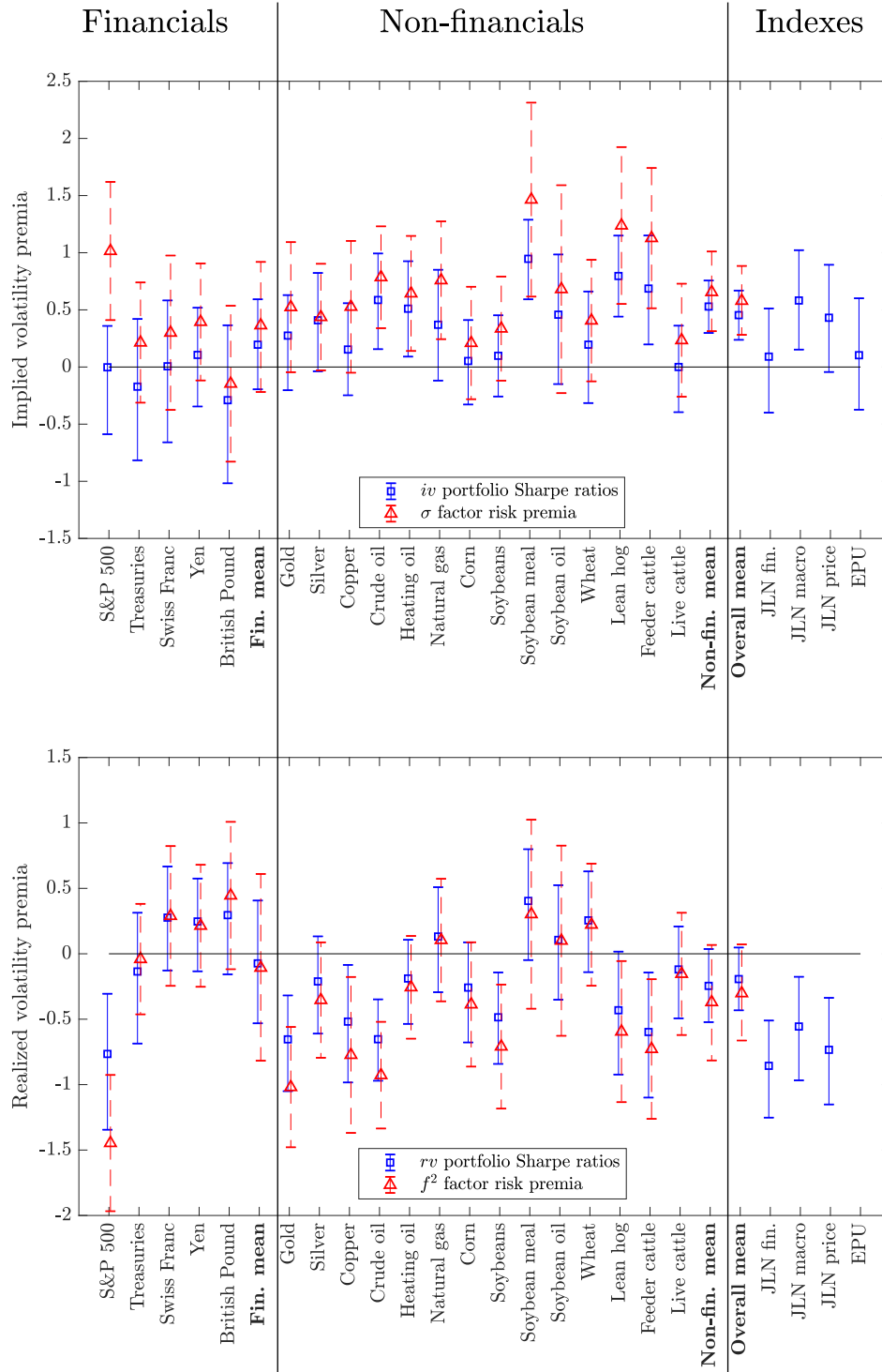
Note: The initial futures price is 1 and the initial volatility, σ , is 0.3. The top panels calculate the return on the rv and iv portfolios given an instantaneous shift in the futures price and volatility to the values reported on the axes under the assumption that the Black-Scholes formula holds. The middle panels plot returns under the approximations used in the text. The bottom panels are equal to the middle minus the top panels. All returns and errors are reported as decimals.

Figure A.2: Factor loadings



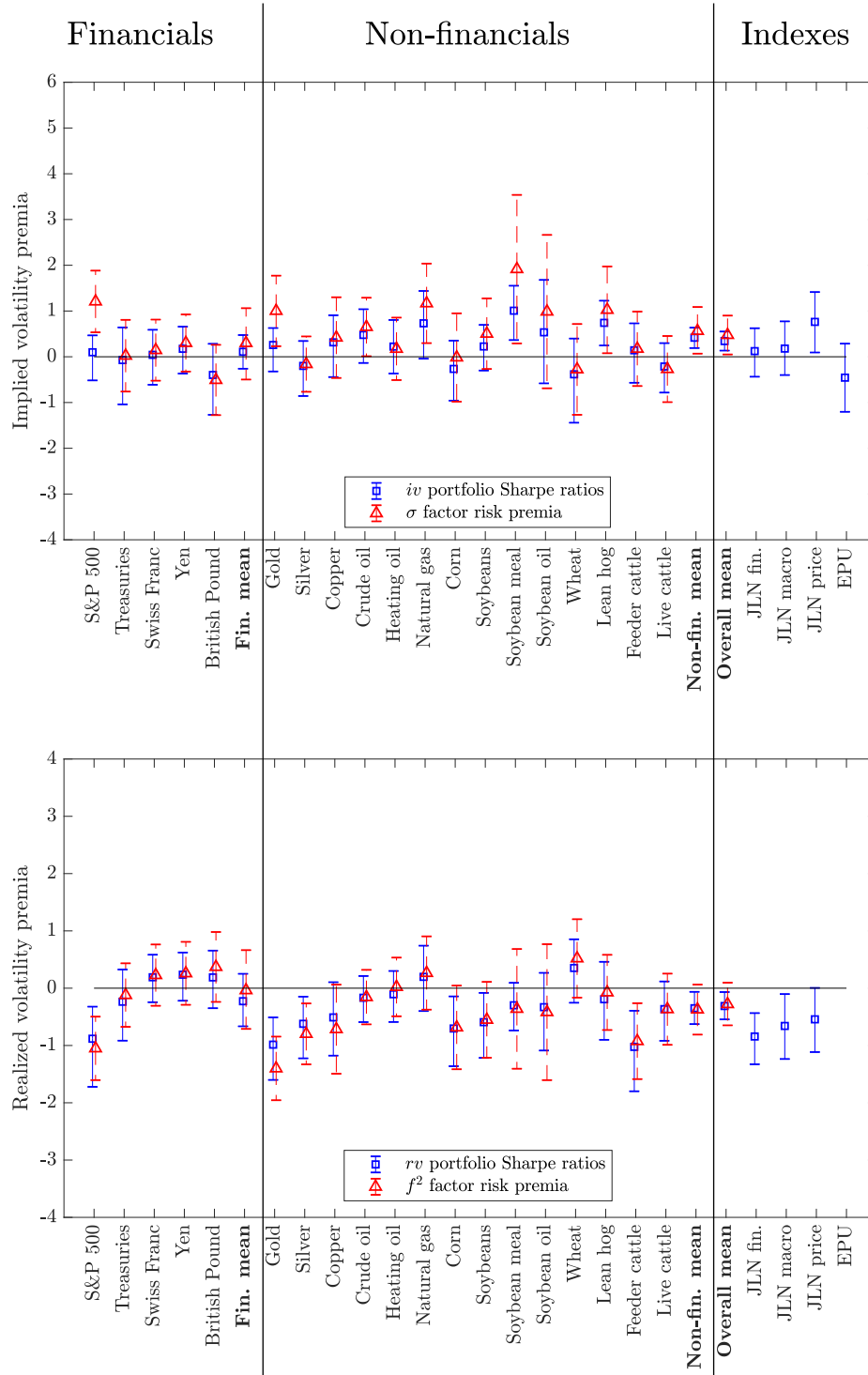
Note: Loadings of two-week straddle returns on the three risk factors. The factors are all scaled by current IV , as in equation 15.

Figure A.3: RV and IV portfolio Sharpe ratios and factor risk premia, one-week holding period



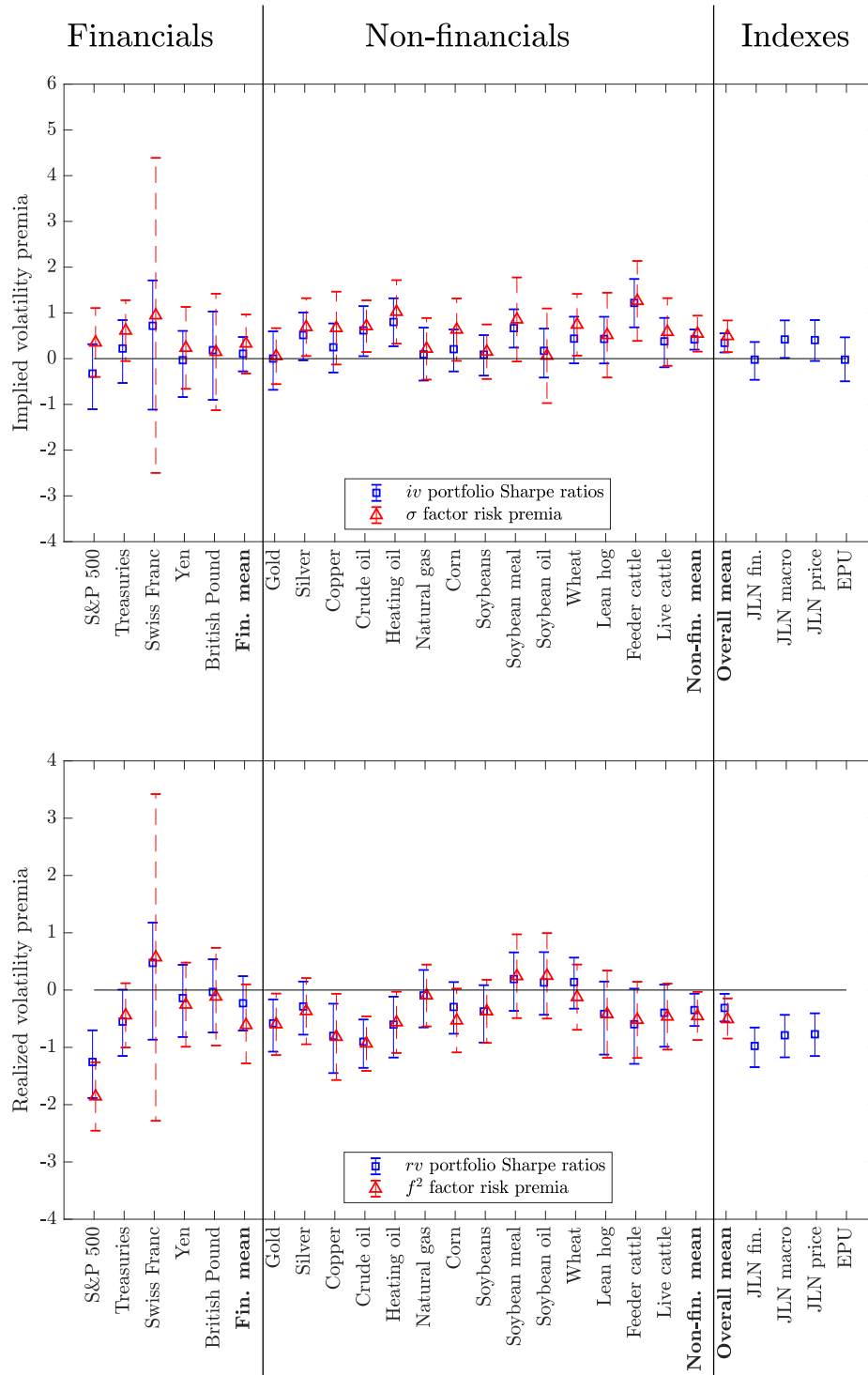
Note: Same as figure 4, but using one-week holding periods.

Figure A.4: RV and IV portfolio Sharpe ratios and factor risk premia (first half of the sample)



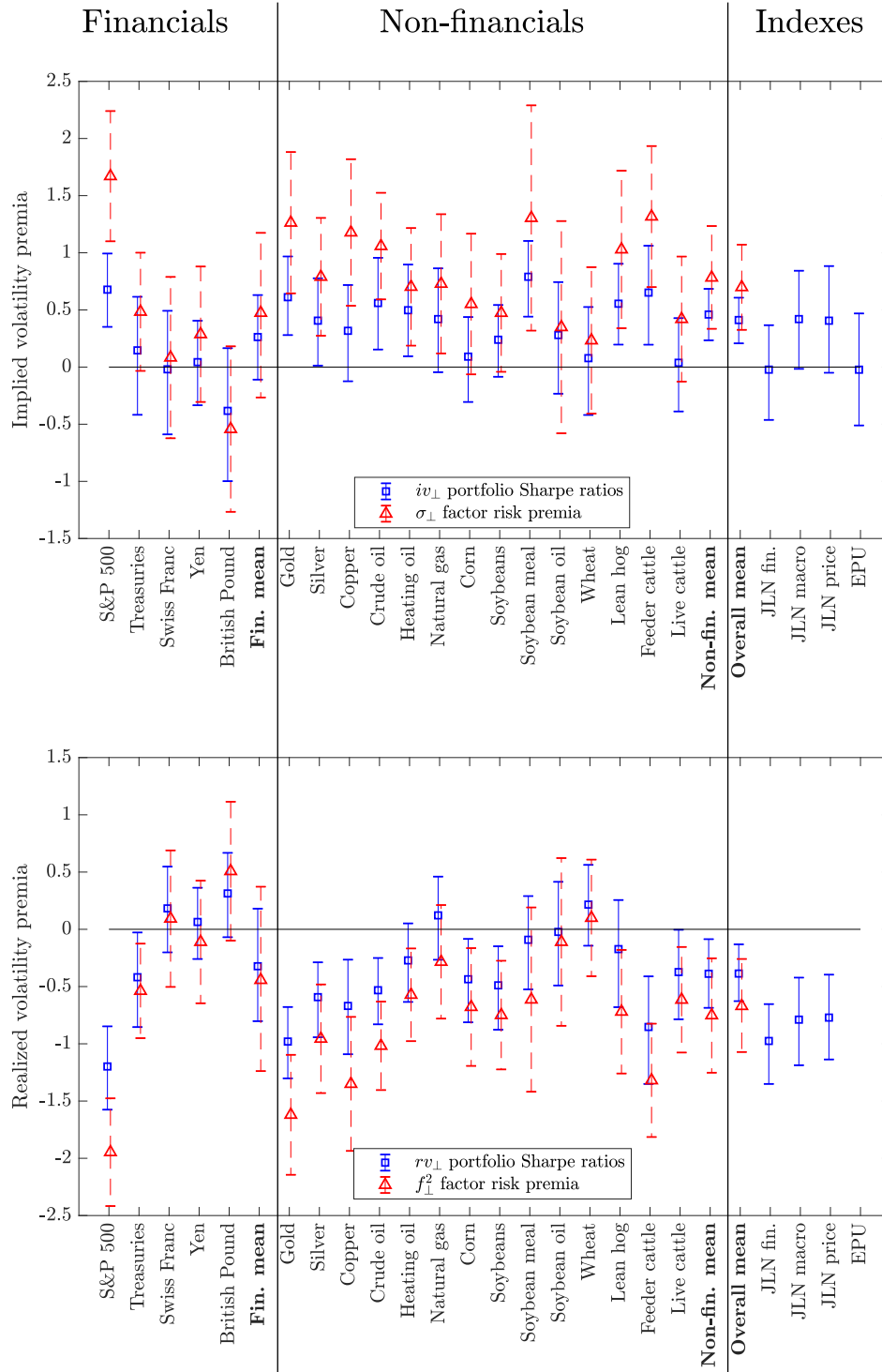
Note: Same as Figure 4, but only using the first half of the sample (up to June 2000).

Figure A.5: RV and IV portfolio Sharpe ratios and factor risk premia (second half of the sample)



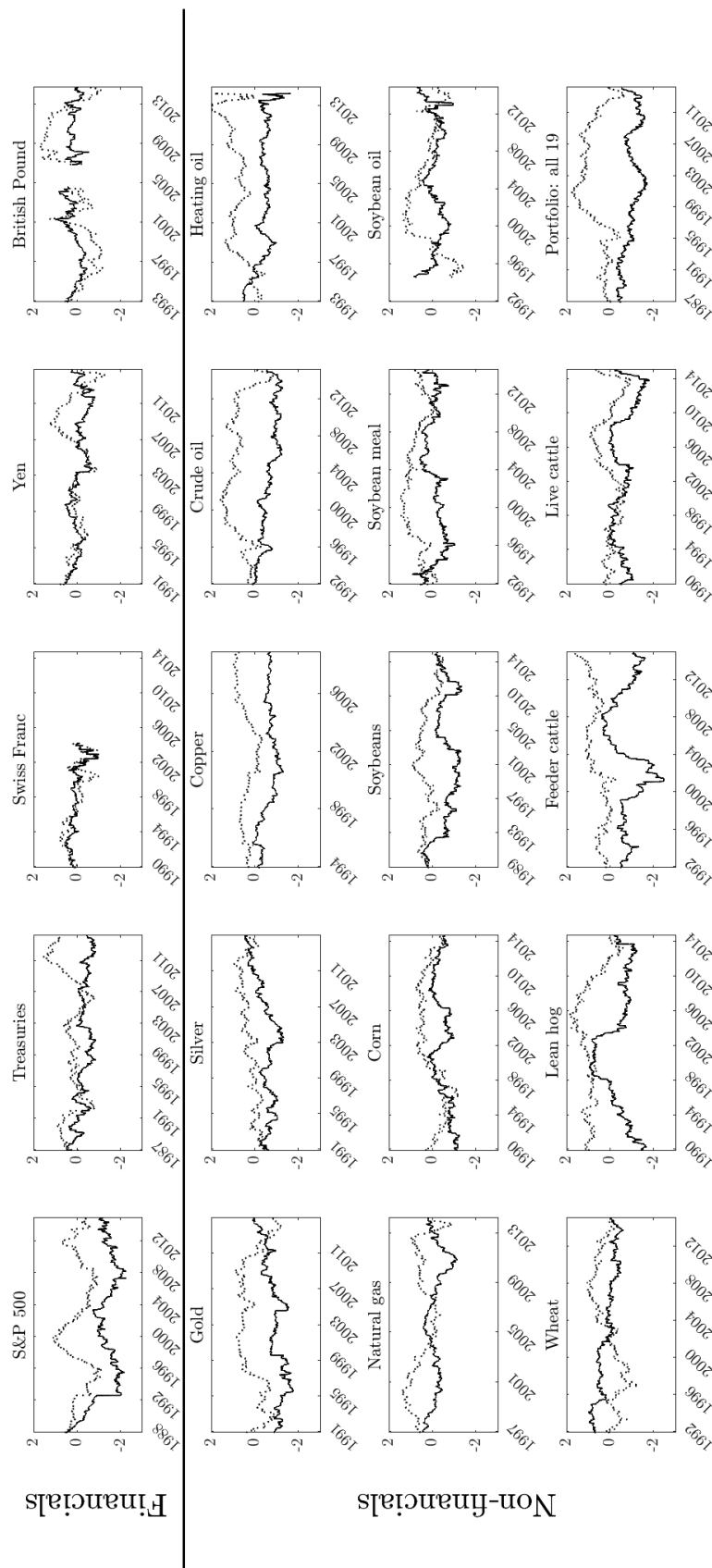
Note: Same as Figure 4, but only using the second half of the sample (after June 2000).

Figure A.6: SDF loadings on RV and IV (Sharpe ratios)



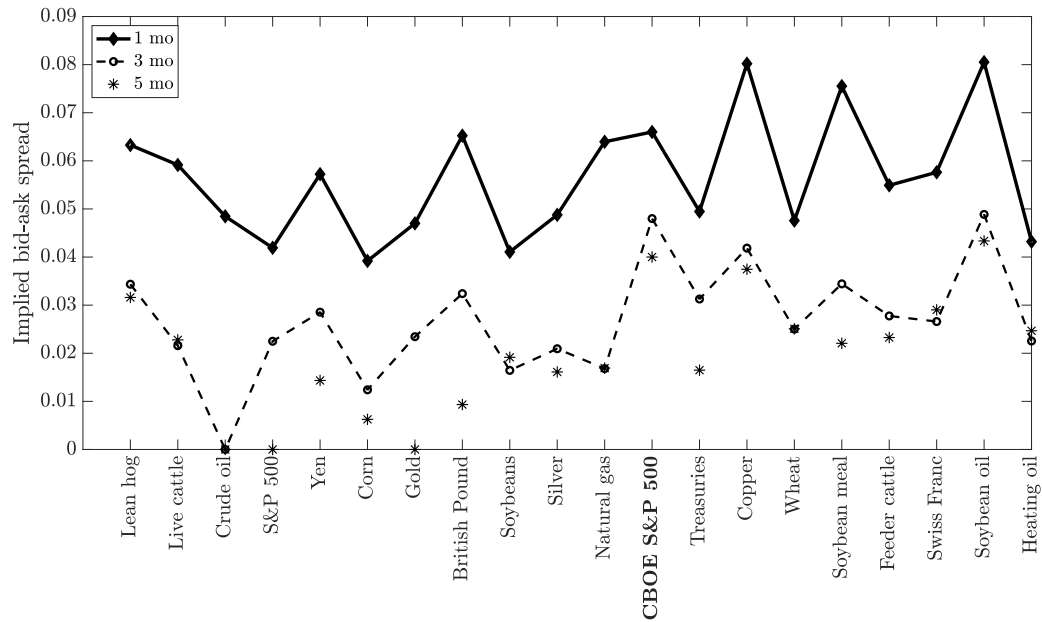
Note: The figure reports the stochastic discount factor (SDF) loadings on IV and RV. The loadings are scaled to correspond to Sharpe ratios of orthogonalized RV and IV portfolios, whose risk premia is equal to the corresponding SDF loading.

Figure A.7: Rolling Sharpe ratios of RV and IV portfolios

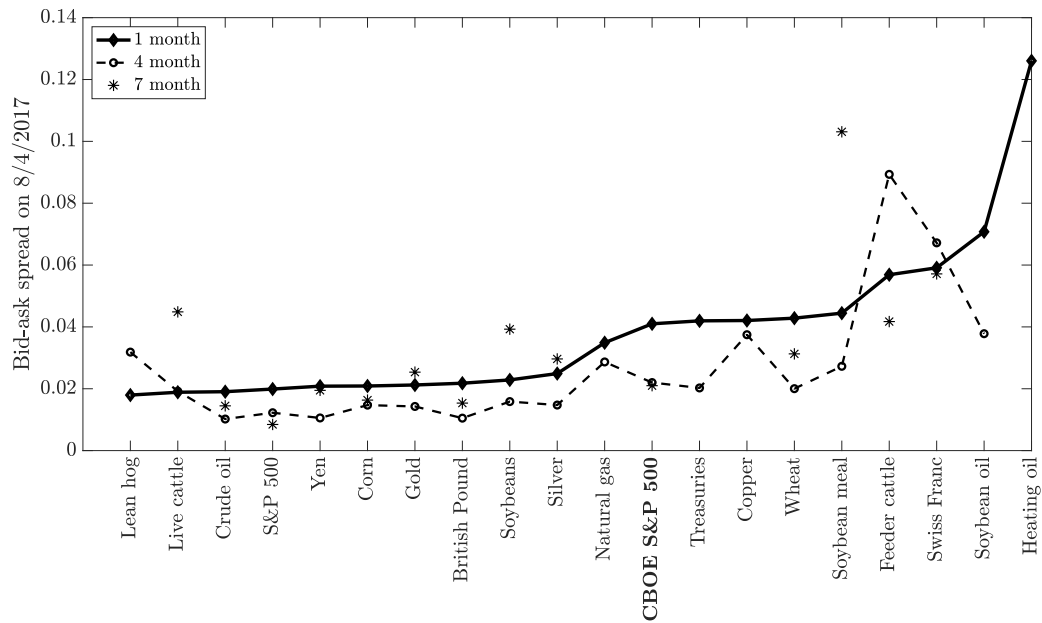


Note: 5-year rolling Sharpe ratios for RV and IV portfolios. Missing values correspond to cases when the IV and RV portfolio are not available for more than 90 days over the 5-year window. The bottom-right panel reports the rolling Sharpe ratio for RV and IV portfolio of all available markets.

Figure A.8: Bid-ask spreads



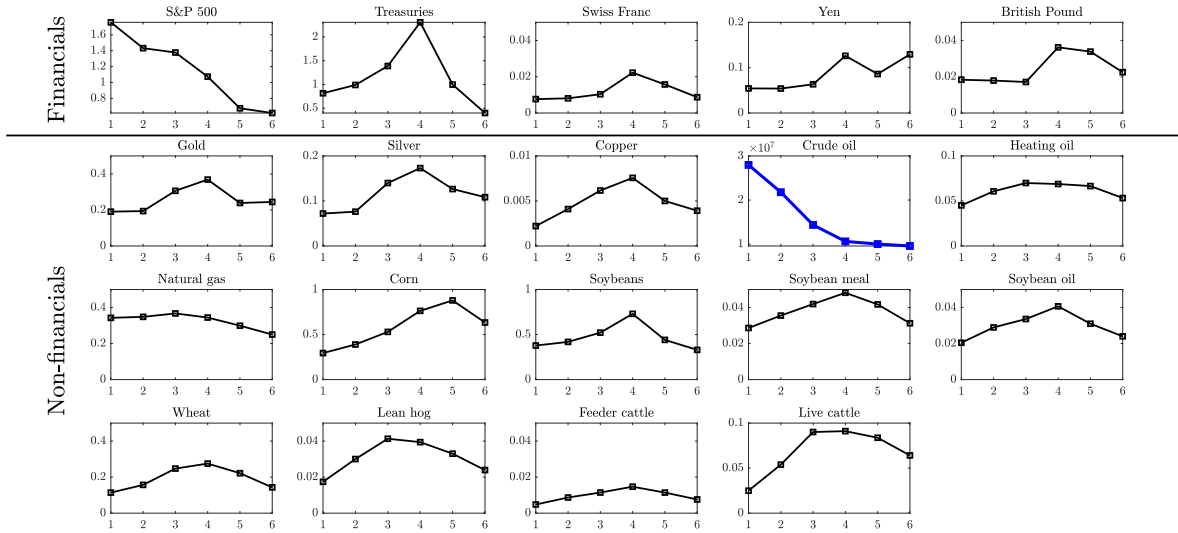
(a) Implied bid-ask spreads



(b) Observed bid-ask spreads on 8/4/2017

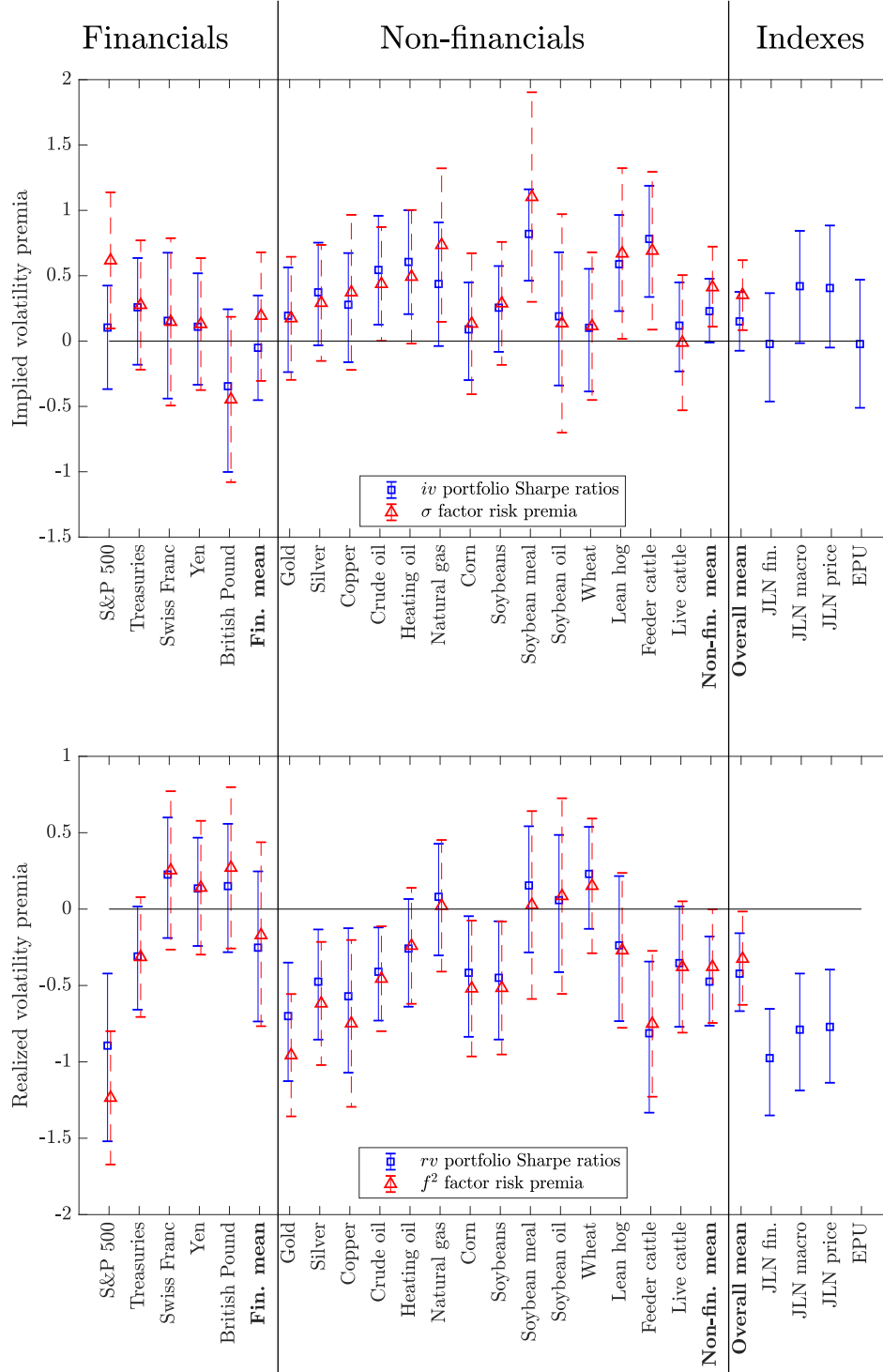
Note: The top panel plots for each market the effective bid-ask spread computed from observed option returns, calculated as in Roll (1984). The spread reported for the CBOE S&P 500 options is based on the historical mean available from Optionmetrics covering the period 1996–2015. The bottom panel reports posted bid-ask spreads for at-the-money straddles obtained from Bloomberg on of August 4, 2017 (the CBOE S&P 500 spreads on that date are also obtained from Optionmetrics).

Figure A.9: Volume across markets and maturities



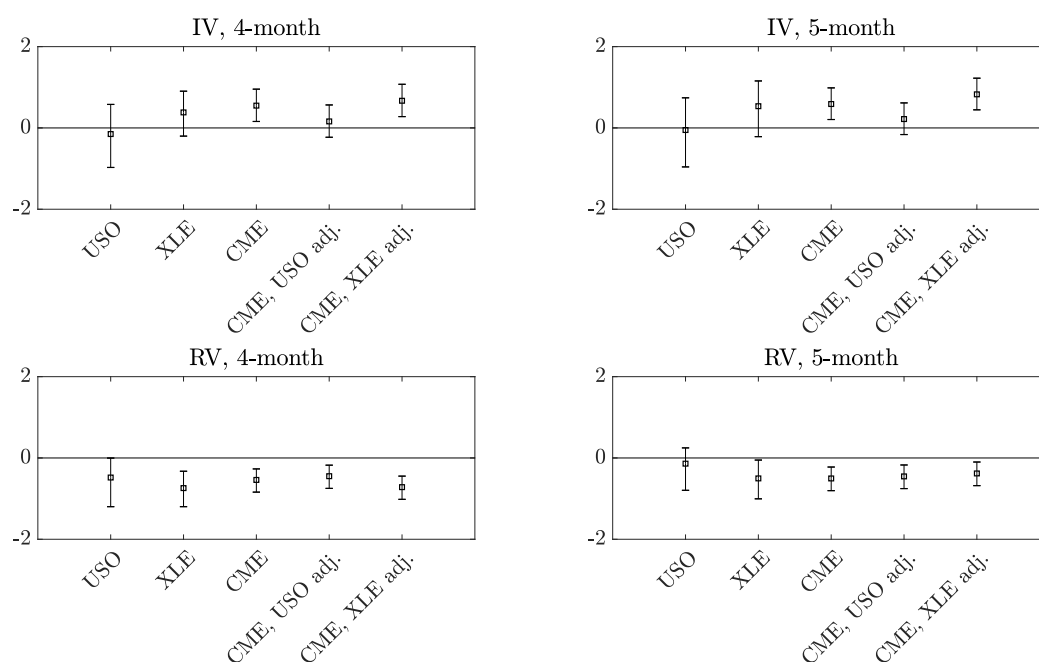
Note: Average daily volume of options in different markets. The panel corresponding to crude oil reports values in dollars. All other panels show values relative to the volume in the crude oil market, matched by maturity.

Figure A.10: RV and IV portfolio Sharpe ratios and factor risk premia (robust to measurement error)



Note: Same as Figure 4, but returns are computed using the same denominator at all maturities, to provide robustness with respect to measurement error in the prices (see section A.2.3).

Figure A.11: Options on crude futures vs ETFs



Note: Sharpe ratios on *rv* and *iv* portfolios using straddles for CME crude oil futures and the XLE and USO exchange traded funds. “4-month” and “5-month” refers to the longer of the two maturities used to construct each portfolio (the short maturity is always one month). The squares are point estimates based on the full sample available for each series. The lines are 95-percent confidence bands constructed with a 50-day block bootstrap. “CME, USO adj.” and “CME, XLE adj.” are identical to the “CME” numbers but with the mean return in the denominator of the Sharpe ratio shifted by the point estimate for the mean difference from table A.6.2.

Table A.1: Risk exposures of *rv* and *iv* portfolios

| rv portfolio | | | | | iv portfolio | | | | | Corr(rv,iv) |
|---------------|-------|----------------|-------------|----------------|---------------|-------|----------------|-------------|----------------|-------------|
| | f | f ² | ΔIV | R ² | | f | f ² | ΔIV | R ² | |
| S&P 500 | -0.06 | 1.31 | 0.05 | 0.71 | S&P 500 | -0.21 | 1.20 | 0.83 | 0.68 | 0.53 |
| T-bonds | -0.01 | 0.80 | -0.02 | 0.81 | T-bonds | 0.00 | 0.24 | 0.92 | 0.70 | 0.22 |
| GBP | -0.02 | 0.73 | 0.06 | 0.84 | GBP | -0.03 | 0.34 | 0.71 | 0.75 | 0.49 |
| CHF | -0.02 | 0.68 | 0.07 | 0.84 | CHF | 0.01 | 0.40 | 0.71 | 0.76 | 0.55 |
| JPY | -0.02 | 0.70 | 0.07 | 0.81 | JPY | 0.02 | 0.48 | 0.70 | 0.84 | 0.62 |
| Copper | -0.01 | 0.75 | -0.01 | 0.62 | Copper | -0.01 | 0.10 | 0.93 | 0.70 | 0.10 |
| Corn | -0.02 | 0.62 | 0.04 | 0.73 | Corn | 0.07 | 0.31 | 0.70 | 0.63 | 0.10 |
| Crude oil | -0.03 | 0.97 | 0.01 | 0.75 | Crude oil | 0.02 | -0.23 | 0.81 | 0.66 | 0.03 |
| Feeder cattle | -0.02 | 0.83 | 0.00 | 0.74 | Feeder cattle | -0.03 | -0.28 | 0.86 | 0.70 | 0.04 |
| Gold | -0.01 | 0.66 | 0.05 | 0.71 | Gold | 0.09 | 0.29 | 0.79 | 0.63 | 0.48 |
| Heating oil | -0.02 | 0.84 | 0.01 | 0.77 | Heating oil | 0.05 | -0.31 | 0.86 | 0.61 | -0.05 |
| Lean hog | -0.01 | 0.80 | 0.01 | 0.76 | Lean hog | 0.02 | -0.48 | 0.84 | 0.48 | -0.21 |
| Live cattle | -0.03 | 0.95 | 0.01 | 0.74 | Live cattle | 0.00 | -0.55 | 0.80 | 0.64 | -0.14 |
| Natural gas | -0.03 | 0.84 | 0.03 | 0.80 | Natural gas | 0.04 | -0.49 | 0.81 | 0.50 | -0.13 |
| Silver | 0.00 | 0.59 | 0.06 | 0.76 | Silver | 0.02 | 0.14 | 0.75 | 0.78 | 0.42 |
| Soybeans | -0.02 | 0.64 | 0.03 | 0.73 | Soybeans | 0.05 | 0.24 | 0.77 | 0.72 | 0.21 |
| Soybean meal | -0.02 | 0.58 | 0.02 | 0.72 | Soybean meal | 0.08 | 0.30 | 0.79 | 0.64 | 0.21 |
| Soybean oil | -0.02 | 0.62 | 0.01 | 0.76 | Soybean oil | 0.07 | 0.19 | 0.79 | 0.65 | 0.16 |
| Wheat | -0.01 | 0.61 | 0.00 | 0.77 | Wheat | 0.07 | 0.22 | 0.83 | 0.70 | 0.19 |
| Average | -0.02 | 0.76 | 0.02 | 0.76 | Average | 0.02 | 0.11 | 0.80 | 0.67 | |

Note: The table reports regression coefficients of the *rv* and *iv* portfolios for each market onto three market-specific factors: the futures return, the squared futures return, and the change in IV. The column on the right reports the correlation between the *rv* and *iv* portfolio returns.