Recent Developments in Financial Risk and the Real Economy

Ian Dew-Becker and Stefano Giglio*

November 8, 2023

Abstract

This paper reviews recent developments in macro and finance on the relationship between financial risk and the real economy. We focus on three specific topics: the term structure of uncertainty, time variation – and specifically the long-term decline – in the variance risk premium, and time variation in conditional skewness. We also introduce two new data series: implied volatility from one-day options on grains for the period 1906-1936, and on cliquet options, which provide insurance against single-day crashes on the S&P 500, both of which give some context to the recent rise in trade in extremely short-dated options. Finally, we discuss new avenues for future research.

1 Introduction

Financial risk plays an important role in both macroeconomics and finance, with a large literature investigating both how it affects the macroeconomy and how it is priced in financial markets. This paper surveys recent developments in asset pricing and macroeconomics that have implications about our understanding of the relationship between financial risk – by which we mean variation in financial volatility and skewness – and the real economy.

There are two main reasons why research in asset pricing plays a crucial role for this goal: first, while risk/uncertainty is a multi-faceted concept (e.g. sector-specific uncertainty, uncertainty about inputs, uncertainty about prices, and so on), financial uncertainty is the one that is most easily and precisely measured at high frequency and for a variety of levels of aggregation, and because of this it has most often been used as a proxy for uncertainty more

*Dew-Becker: Visiting University of Chicago and NBER. Giglio: Yale University and NBER.
generally. Second, financial markets give direct insights into how people perceive risk that can help us better understand how it affects the economy: for example, in macro-finance models there is a link between the effects of conditional volatility on macroeconomic aggregates (e.g., consumption) and the risk premium that investors require for bearing volatility risk.

Early research in this area of macroeconomics typically studied the relation between risk and the macroeconomy by relating, typically through a vector autoregression (VAR), one proxy for uncertainty (e.g., the VIX) to various macroeconomic outcomes; in finance, a corresponding literature studied the risk premium associated with the VIX itself (which captures how investors perceive the risk associated with short-term realized variance). The core of this review focuses on three new insights developed in the recent literature in this area:

1. **A focus on the term structure of volatility.** Volatility and uncertainty can differ markedly across horizons: in finance, it is well known that investors may perceive long-term and short-term risks very differently; in macroeconomics, new research has also begun to distinguish their effects, often emphasizing the importance of financial frictions.

2. **Observed declines in the variance risk premium and options premia more generally.** The variance risk premium – the large (negative) compensation for exposure to variance, implying that investors are highly averse to high-volatility states – has long been viewed as a robust empirical regularity in the finance literature, with a range of models developed to understand it. More recent research, however, has begun to question whether this premium still exists, with papers arguing that it has fallen to nearly zero. One interpretation of these results is that the measured variance risk premium may not have been due to investors’ aversion to uncertainty and volatility, but instead it reflected the effects of intermediation frictions.

3. **The recognition of variation in conditional skewness as a driving force in the economy.** While there is some evidence that suggests that financial uncertainty might not have large real effects (and, consistent with this, that the risk premia associated to it have decreased dramatically), recent research has shown that measures of higher-order risk – *skewness* in particular – may be strongly associated with macroeconomic outcomes. Specifically, while there is a long history of studying unconditional skewness, more recent work has focused on changes in skewness over time. Understanding skewness, even more so than volatility, requires tackling nonlinearity in the economy.

**New data.** In addition, this paper presents two new data series that are novel to the

---

1 For a variety of other measures, see the recent review by Cascaldi-Garcia et al. (2023).
literature and may be of interest in future work. One gives measures of implied volatility over
the period 1906–1936 and helps shed light on the historical cyclicality of risk. The second
is the time-series of prices of options that isolate one-day crash probabilities. Both are of
potential interest additionally because they give historical context for the recent rise in trade
in one-day (and even zero-day) options.

After tackling these three points in sections 2-4, we conclude with a discussion of im-
portant new avenues for research. As suggested by the above summary, one path is to
understand better how financial intermediaries affect the transmission of risk between the
financial and real sides of the economy. It is well known that the link between measures
of financial risk, like the VIX, and GDP is weak at best, but in some episodes they move
together extremely strongly. The question is what drives the conditional variation in that
relationship.

The second path is to better understand nonlinearity in both macro and finance. Coun-
tercyclical volatility is pervasive through both the real and financial sectors, but standard
models usually do not have a channel through which it can arise. Capturing that feedback re-
quires allowing nonlinearity. We discuss some recent work on that idea along with promising
avenues for future research, but view it as still very much open to new analyses.

2 The term structure of financial risk

Risk is not the same at all horizons, and it changes dynamically over time; understanding the
dynamics and maturities of risk is important both for asset pricing and for macroeconomics.
Borovicka and Hansen (2016), for example, provide a thorough review. Both macro and
finance have recently started studying in greater depth the term structure of risk in the
economy.

We begin by thinking of risk as being measured by volatility (and move on to higher-
order risk subsequently). When thinking about the dynamics of volatility and its pricing and
real effects, it is useful to start by distinguishing realized volatility from uncertainty. They
have often been used interchangeably in existing research, especially in macro, but they
correspond to very different concepts with potentially very different real effects. Consider
some economic variable \( x_t \), perhaps representing stock returns or GDP growth. Ignoring
variation in the conditional mean, we could write

\[
x_t = \sigma_{t-1} \varepsilon_t
\]

where \( \varepsilon_t \) is some mean-zero i.i.d. innovation. We refer to realized volatility as the squared
realization of the shock, $x_t^2 = \sigma_{t-1}^2 \epsilon_t^2$, a measure of the size of the realized shock. Realized volatility is, therefore, backward looking – it describes how large the shocks were during period $t$. We refer to uncertainty or risk at some horizon $j$ as $\text{var}_t[x_{t+j}]$. That captures the forward-looking uncertainty about the future values of $x_t$ (in a very similar analysis, Rossi, Sekhposyan, and Soupre (2020) refer to the forward-looking uncertainty as the “ex ante” component and realized volatility as the “ex post” component).

An immediate but key result that links the two is that

$$\text{var}_t[x_{t+j}] = E_t x_{t+j}^2$$

That is, uncertainty is the same as the expectation of future realized volatility.

To see the distinction empirically, figure 1 plots monthly realized volatility for the S&P 500 against 30- and 365-day at-the-money implied volatility for S&P 500 options. While the series are naturally strongly correlated, they also have significant independent variation. Their correlations are below:

---

Note: Figure plots monthly realized volatility, together with 30-day and 365-day at-the-money implied volatility for S&P 500 options.
The two IVs are very strongly, though imperfectly, correlated, while realized volatility has significant independent variation from the two IVs. It is therefore not surprising that the conclusions drawn from the various series might be different. Of course, all the above can also be extended to other moments – e.g. realized, short-run, and longer-run conditional skewness, which we study in section 4. The remainder of this section reviews recent work on the term structure of uncertainty, in terms both of risk premia and effects on the macroeconomy.

2.1 Finance literature

2.1.1 Empirical analyses

While the variance risk premium has been studied since the 1990’s, and work on the dynamics of volatility – implying a term structure – extends at least to the 1980’s, studies of how the variance risk premium varies across the term structure only began to arise relatively recently. Egloff, Leippold, and Wu (2010) is perhaps the earliest contribution. They estimate a two-factor model of volatility dynamics and focus on optimal investment, finding that premia are isolated at the short end of the term structure (see also, later, Feunou et al. (2014), Filipovic, Gourier, and Mancini (2016), Eraker and Wu (2017), Johnson (2017), and Ait-Sahalia, Karaman, and Mancini (2020)). Choi, Mueller, and Vedolin (2017) find a similar result for the term structure of bond variance risk premia.

Dew-Becker et al. (2017) study the term structure of volatility risk premia implied by variance swaps. The main innovation in that paper relative to the other related work on the term structure is to distinguish premia for shocks to realized volatility from shocks to implied future volatility (again, date-\(t\) shocks to \(x_t^2\) versus to \(E_t [x_{t+j}^2]\)). Dew-Becker et al. (2017) find that only shocks to realized volatility carry a significant risk premium. Claims to forward volatility (that is, expectations of future volatility) carry no risk premium at all. In the context of the figure and correlations from the previous section, the results imply that it is the part of realized volatility that is uncorrelated with the two IVs that investors have historically been averse to.

---

3For example, see Jackwerth and Rubinstein (1996), Coval and Shumway (2001), Bakshi and Kapadia (2003), Broadie, Chernov, and Johannes (2009), and Carr and Wu (2009) on the pricing of volatility, and Engle (1982) and many subsequent papers on volatility dynamics, reviewed in several places, such as Bollerslev, Engle and Nelson (1994) and Engle (2004).
Dew-Becker et al. (2017) argue that that fact is difficult to reconcile with standard consumption-based asset pricing models (though they also present a disaster model that can fit the data). A potential alternative explanation is that segmentation in the derivatives market might cause the models that fit the variance term structure (and options markets) to differ from those that fit prices in other markets like equities. That idea is discussed further below.

Dew-Becker, Giglio, and Kelly (2021) extend Dew-Becker et al.’s (2017) results to show that large negative premia for shocks to realized volatility and small or insignificant premia for shocks to implied volatility or uncertainty are pervasive across a number of other markets, including bonds, currencies, and commodities (see also Hollstein, Prokopczuk, and Wursig (2020)).

Johnson (2017) extends the results in Dew-Becker et al. (2017) looking at variation over time in variance risk premia. See also Cheng (2019) and Andries, Eisenbach, and Schmalz (2023). Time variation in the variance risk premium is discussed further in section 3.

### 2.1.2 Models

Numerous papers have developed models to fit the variance risk premium at short maturities. However, because shocks to realized volatility (which is the payoff of the derivatives from which the variance risk premium is computed) are correlated with uncertainty shocks (shocks to expected future volatility), the variance risk premium could reflect investors’ aversion to either of the two shocks. Therefore, the fact that a model successfully matches the VRP is no guarantee that the model correctly prices both realized volatility and uncertainty shocks. Only a few studies draw out the implications for the two types of shocks for the entire term structure of volatility risk, which allows separate identification of the two risk premia. As noted above, Dew-Becker et al. (2017) both compare their empirical results to predictions of earlier models (like those of Drechsler and Yaron (2011), Wachter (2013), Du (2011), and Christoffersen, Du, and Elkamhi (2015)) and also propose a disaster-based model that can fit the evidence. Eraker and Wu (2017) also present an equilibrium model to match the behavior of variance claims.

The literature has also developed models based on deviations from standard preference specifications to try to match the term structure of uncertainty risk premia. Babiak (2020) shows that a model in which agents have generalized disappointment aversion can fit the variance term structure. While they do not report detailed estimates, Andries, Eisenbach, and Schmalz (2023) suggest that their model of horizon-dependent risk aversion would explain the variance term structure.
This work complements a much larger literature that tries to explain the empirical term structures of risk premia in the equity markets (documented starting from van Binsbergen et al., 2012 and 2013)), and, of course, the vast literature on the term structure of bond yields.

2.2 Financial uncertainty and the macroeconomy

Financial market volatility has been widely used as proxy for uncertainty in macroeconomic models. The empirical evidence on the potential role of uncertainty in driving macroeconomic activity comes first of all from the observed negative correlation between measures of financial uncertainty (like the VIX) and macroeconomic downturns. Additional empirical evidence in the literature has come from trying to identify the causal effects of uncertainty shocks and the macroeconomy (mostly via vector autoregressions). In parallel with this empirical work, the theoretical literature has focused on understanding the potential mechanisms through which uncertainty could have real economic effects (a seminal paper in this literature is Bloom (2009)).

In this section we highlight several recent advances in the empirical study of the relation between financial uncertainty and the macroeconomy. We focus on three main points: 1. we document novel historical evidence supporting the negative correlation between uncertainty and the business cycle; 2. we discuss the challenges to the identification of uncertainty shocks stemming from the endogeneity of uncertainty to macroeconomic shocks, and show how term structure information described in the previous section can be used for this purpose; 3. we highlight the different cyclical properties of aggregate as opposed to idiosyncratic volatility, suggesting that they interact with economic activity through different mechanisms.

2.2.1 The correlation of financial uncertainty and the business cycle: new evidence from the early 20th century

The most widely used measure of financial uncertainty is the VIX, which is computed from S&P 500 index option prices and is available since 1987, which is when monthly expirations for these options were introduced. Empirically, panel (b) of Figure 2 plots the VIX together with indicators of the NBER recessions, showing the well-known countercyclicality of aggregate uncertainty.

That evidence is limited to a relatively short period of time, though, due to the available sample for the VIX. Several attempts have been made to extend the data before then. One approach that part of the literature has followed is to use realized volatility instead of the VIX for the time when the VIX is not available. But as discussed above, measures

\footnote{For example, Bloom (2009) splices together the two series. Baker et al. (2021) examine the causes of...}
of realized volatility are a poor proxy for forward-looking uncertainty, and the fact that uncertainty and realized volatility are priced differently in financial markets suggests that their relationship with the economy has important differences. Another approach is that of Manela and Moreira (2017), who extend the VIX back in time based on text from front-page articles in the Wall Street Journal.

Here we provide novel evidence on historical uncertainty based on prices of options on grain futures from the early 20th century. We essentially build an analogue to the VIX based on information from commodity markets – which were, at the time, of similar importance to the stock market now.

While options on the stock market index were not traded before 1987, options on other macroeconomically important underlyings were traded in the early 20th century. Among them, very short-term options on commodity futures – called “privileges”, or “bids” and “offers” – were traded at the Chicago Board of Trade. Mehl (1934) and Lurie (1979) give extensive descriptions of the CBOT options market. To get a sense of magnitudes, aggregate dividends from stocks between 1925 and 1932 averaged $2.5 billion per year, while aggregate farm income (not including receipts of government payments) averaged $8.6 billion. GDP between 1929 and 1933 averaged $78 billion. Farm income was therefore about 11% of aggregate income in the period this data covers, compared to 5% for corporate dividends since 2013.

The CBOT options market structure was different from the current structure in two important ways. First, the options traded had maturities of either one day or one week, with the dailies being traded more consistently over time. So the recent rise of daily options is actually not unprecedented historically. Second, instead of the strike of an option being fixed, it was the price (or premium) that was fixed, at $5 per 5,000 bushels (of wheat or corn), and what varied was the strike. On any given day, there was a put and a call available, each with the same price, and their strikes were then reported in newspapers along with futures prices.

We obtained data on privilege prices from scans of the Chicago Tribune over the period 1906–1936 (with some gaps when trading was outlawed). While there is data for both daily and weekly options, we focus just on dailies here for simplicity. The top panel of figure 2 plots implied volatility for corn and wheat over the sample period with NBER-dated recessions in gray, while as mentioned above the bottom panel plots, for context, the VIX and recessions.

\[ \text{large jumps since 1900 in the United States and in more recent samples in a range of other countries. Jumps are directly related to realized volatility, but correlate with uncertainty via GARCH effects, since jumps are typically followed by more volatility. See additional discussion of this in the next section.} \]

\[ \text{Both are from the NBER Macrohistory database; series m08079c and m12031a, respectively.} \]
Figure 2: Financial Uncertainty and Recession

(a) Implied Volatility of Corn and Wheat

(b) VIX

Note: Panel (a) plots the implied volatility for corn and wheat over the period 1906-1936. Panel (b) plots the VIX for the period 1987-2023. The gray shaded areas are NBER-dated recessions.
for the period 1987–2023. For each of these series, there is evidence that option-implied volatility is higher during recessions (with the statistical strength being weaker for the earlier sample).\footnote{Formally, the graphs plot the 1-month moving average of the implied volatility series. The point estimates are of higher volatility in recessions for all three series, with Newey-West t-statistics of 2.3 for corn, 0.7 for wheat, and 7.9 for the VIX.}

The focus of this paper is not economic history, so we do not take the analysis any further. The data in figure 2 is novel to the literature, though, and part of a broader dataset that includes prices of futures and options from the CBOT over a thirty-year period. We would be willing to share it with other researchers for future use.

### 2.2.2 Expected volatility and its macroeconomic effects

Going beyond simple correlations and trying to understand the causal effects of uncertainty on the macroeconomy requires thinking carefully about the dynamics of volatility, so that shocks to uncertainty can be measured and its causal effects identified. Early macroeconomic models often simplified the treatment of volatility – typically assuming that it follows an AR(1) process. In this case, news about volatility at all horizons is perfectly correlated and hence empirically indistinguishable. In addition, existing literature often ignored the distinction between realized volatility and future uncertainty that Dew-Becker et al. (2017) show is central to understanding risk premia.

Some early theoretical work (e.g. Hassler (1996)) looked at how responses to uncertainty shocks might depend on their persistence, but there was little empirical analysis initially. More recently, a number of papers have looked at responses at different horizons, allowing a distinction of the real effects of realized and expected volatility. Berger, Dew-Becker, and Giglio (BDG; 2020) use a standard VAR setup to understand the interaction of volatility and the macroeconomy, but with the innovation of treating an uncertainty shock as news about future volatility. This is closely related to the analysis above of the differential pricing of implied and realized volatility: recall from above that uncertainty is fundamentally equivalent to expectations of future realized volatility. So a shock to uncertainty is really news – it is a shock to $E_t x_{t+j}^2$. BDG build on a large literature in macroeconomics on estimating the response of the economy to \textit{news shocks} (see e.g., Beaudry and Portier (2006), Barsky and Sims (2011), and Barsky et al. (2014)), and apply these techniques to understanding the effects of news about future volatility onto the economy.

The key empirical finding of BDG is that after controlling for realized volatility – thus isolating the purely forward-looking component of uncertainty, as opposed to the part relating to shocks that have already been realized – uncertainty does \textit{not} have a significant impact
on the behavior of the economy. Interestingly, this result is consistent with those obtained in the finance literature and described in section 2, in that uncertainty shocks, which appear not to have negative macroeconomic effects, also are not priced by investors in financial markets.

Rossi, Sekhposyan, and Soupre (2020) extend the results in BDG studying various types of uncertainty. They refer to BDG’s expected and realized volatility as “ex ante” and “ex post” uncertainty and find, consistent with BDG, but across a wider range of variables and using somewhat different methods, that it is the ex post component that seems to drive the economy, as opposed to the ex ante uncertainty component.

2.2.3 Aggregate versus idiosyncratic uncertainty

Another important distinction that matters both empirically and theoretically is between aggregate and firm-level uncertainty. A recent paper focusing on the latter, and also emphasizing the importance of the term structure of volatility expectations in macro models, is Christiano, Motto, and Rostagno (CMR; 2014). Specifically, CMR take a standard New Keynesian business cycle model with financial frictions and allow for variation in firm-level risk over time. When firms face a broader distribution of shocks, they have a higher risk of bankruptcy, which essentially acts like a tax on capital due to bankruptcy costs. Increases in firm uncertainty can thus generate recessions in the model due to declines in investment demand. But CMR also show that in order to match the data it is important to have not just contemporaneous shocks to firm risk but also news about variation in future firm risk.

Building on CMR, Dew-Becker and Giglio (2023) measure firm-level uncertainty based on implied volatility from options on individual stocks over the period 1980–2020 (that paper also gives a discussion of other related work in macro on cross-sectional uncertainty shocks). They find that in only some recessions – primarily 2008 and 2020 – has firm-level uncertainty risen. Otherwise, it appears to be largely acyclical, and it was actually high during the late 1990’s boom. From this point of view, firm-level uncertainty looks very different from aggregate uncertainty in the post-1980s sample, but does resemble the early 20th century data in that the relationship between uncertainty and the business cycle is somewhat weak and ambiguous overall.

2.2.4 Other related work

A number of other papers explore different aspects of the relation between term structures of uncertainty and the business cycle. Jurado, Ludvigson, and Ng (2015) and Ludvigson, Ma
and Ng (2021) construct uncertainty indexes capturing uncertainty at different horizons. Their methods involve forward-looking data, though, making them difficult to use when trying to understand causation. Barrero, Bloom, and Wright (2017) study how different types of investment respond to short- and long-run uncertainty, using firm-level option-implied volatilities as in Dew-Becker and Giglio (2023). Other papers have exploited other identification schemes to identify the effects of uncertainty shocks and address the endogeneity of uncertainty to the state of the economy (see Bachmann, Elstner, and Sims (2013), Cesa-Bianchi, Pesaran, and Rebucci (2020), Ferrer, Rogers and Xu (2021), and Alessandri, Gazzani, and Vicondoa (2023)).

3 The decline in the variance risk premium

Most of the existing literature on the pricing of realized volatility and uncertainty has focused on estimating – and explaining with theoretical models – the unconditional risk premia associated with them (starting from Fleming (1998), Coval and Shumway (2001) and Bakshi and Kapadia (2003)). Studying conditional moments is relatively harder in these markets given the short time series available. Recent work, however, explores the time variation in the variance risk premium (VRP) and asks which theoretical models can explain the observed patterns.

3.1 Business-cycle and high-frequency variation in the VRP

Two approaches have been taken in the literature to explore time variation in the VRP. The first is based on reduced-form predictive regressions of returns to variance-related derivatives (option portfolios, variance swaps) using a variety of predictors. A second approach estimates no-arbitrage models, which model directly the pricing kernel and its dynamics as a function of observable variables or latent factors; the time variation in the variance risk premium in these models arises from changes in the quantity and/or price of volatility risk. Among these papers, Todorov (2010) shows how the variance risk premium is higher following jumps in the market. Corradi, Distaso, and Mele (2013) use a no-arbitrage model to document variation in the VRP with macroeconomic factors. Barras and Malkhozov (2016) document that the VRP depends on volatility itself, various financial and macroeconomic indicators, and the financial health of intermediaries. Finally, Johnson (2017) shows that that the slope of the term structure of the VIX predicts variation in the VRP.

See also Binder et al. (2022) and Clark et al. (2022).
A subset of papers in this literature specifically exploits information in the term structure of variance derivatives to estimate its physical and risk-neutral dynamics; implicitly or explicitly, these models also have implications about the dynamics of the conditional VRP. Among them, see Egloff et al. (2010), Feunou, Fontaine, Taamouti, and Tedongap (2014), Filipovic, Gourier, and Mancini (2016), Ait-Sahalia, Karaman and Mancini (2020), and Dew-Becker, Giglio, Le and Rodriguez (2017).

3.2 Longer-term trends in the VRP

Dew-Becker and Giglio (2023) point out that beyond the higher-frequency variation in the VRP discussed so far, there also appear to be longer-term trends likely due to changes in the markets where variance risks are traded. The starting point of their analysis is that the variance risk premium can be studied, rather than in options markets, by using synthetic options, which are dynamic portfolios of the underlying asset that replicate the payoff of the option (Black and Scholes (1973) and Merton (1973)). While replication is always imperfect in practice, it works surprisingly well, with an R\(^2\) for options of over 85% at a monthly maturity using daily rebalancing. Dew-Becker and Giglio (2023) apply the replication to the CRSP total market return since 1926 and study the long-term trends in the premium associated with synthetic options, comparing it with trends in the option-based VRP over the last decades.

The focus of the paper is on the alpha of options and synthetic options (as in Heston, Jacobs, and Kim (2023), Coval and Shumway (2001), and many others) in order to control for the “leverage effect”, the fact that innovations in S&P 500 volatility are very strongly negatively correlated with S&P 500 returns and hence have a large negative beta. Obviously, vanilla options do not give direct exposure to variance risk. However, since their payoffs are convex functions of the market return, they have volatility exposure – higher ex post volatility is generally associated with higher option returns, which is also easy to confirm with the synthetic options.

Dew-Becker and Giglio (2023) obtain two main results. First, the alpha of synthetic options was never significantly negative at any time since 1926: the market downturns and high-volatility states hedged by synthetic options were priced essentially in line with the CAPM for the last century. Second, they replicate the negative alpha of traded options since 1987 that earlier literature has found, but they also show that this premium has trended down significantly in the last two decades. In addition, other measures of the VRP, such as the gap between the VIX and realized volatility, have also trended to or even passed across zero.
Figure 3: Cumulative CAPM Alphas from Put Options

Note: Top panel plots cumulative CAPM alphas from traded and synthetic put options since 1926. Bottom panel plots 10-year rolling Information Ratios for 95% OTM puts, delta-hedged 95% OTM puts, delta-hedged ATM straddles, and RV-VIX.
These patterns are clearly visible in the top panel of Figure 3, which replicates results in Dew-Becker and Giglio (2023), showing cumulative CAPM alphas from traded and synthetic put options. The cumulative alphas for traded options show a clear flattening around 2010. Bates (2022) finds generally similar results (though with different methods). As noted by Heston and Todorov (2023) (who find zero alpha for S&P 500 variance since 2006), the market crash associated with Covid in March, 2020, has a major impact on the overall mean returns, but the cumulative returns reveal that even before Covid, the premium had flattened out substantially compared to what was observed previously. Naturally, given the data in figure 3, studies that use more recent data samples will tend to find weaker evidence of a variance risk premium – Heston and Todorov (2023) and Heston, Jacobs, and Kim (2023) being two recent examples.

The bottom panel of figure 3 plots rolling ten-year information ratios for traded options and various other investments exposed to the VRP – delta-hedged puts, delta-hedged straddles, and RV minus the VIX. In all four cases, the premium is trended significantly towards zero, and in many cases has now switched signs.

3.3 Models

Consumption-based asset pricing models have strong implications for the variance risk premium and option prices more generally, and several specifically target their moments. While a number of mechanisms can explain the unconditional variance risk premium (e.g., correlation of realized variance with long-run shocks in Drechsler and Yaron (2011) or disaster-risk exposure as in Seo and Wachter (2019)), matching the time variation in the variance risk premium documented in the data is harder. The empirical studies reviewed above show that conditional risk premia depend in a rich way on financial variables, macroeconomic conditions, as well as intermediation frictions, and these relations are hard to match for typically stylized models, though some models can match some of them. For example, the variance risk premium varies with the aggregate jump intensity in Drechsler and Yaron (2011), with the skewness of consumption shocks in the habit formation model of Bekaert & Engstrom (2017), with beliefs about the dynamics of volatility (Lochstoer and Muir 2022), and with cyclical changes in hedging demand in Cheng (2019).

Longer-term trends like the secular decline in the variance risk premium documented by Dew-Becker and Giglio (2023) are even harder to generate for structural models simply because such models are typically stationary by construction. Dew-Becker and Giglio (2023) propose, consistent with the demand-driven work of Garleanu, Pedersen, and Poteshman (2008), Cheng (2019), Constantinides and Lian (2021), and others, that shifts in the variance
risk premium and on option returns more generally might be driven by changes in the cost intermediaries face when hedging options exposure. The basic economic mechanism comes from a long line of work (see Bates (2003)), and operates as follows (in the paper, it is derived as an extension of Garleanu, Pedersen, and Potheshman (2008)). The market for options sees the participation of some retail investors who want to buy S&P 500 options, even at relatively high prices. Due to preferences or constraints, that demand is not satisfied by other retail investors, but rather by a relatively small number of intermediaries. Those intermediaries then charge a premium for bearing the risk associated with being short options, causing the options to be overpriced, consistent with the data. This explains the large negative CAPM alpha in traded options and hence the VRP. At the same time, the average equity investor does not have particularly strong aversion to market downturns and has no preference for buying (overpriced) options; this average investor’s preferences are reflected in the market price, and therefore, indirectly, in the premia associated with synthetic options. As reported above, synthetic options are priced consistently with the CAPM.

In this model, therefore, the discrepancy between the zero alpha of synthetic options and the large negative alpha of traded options is due to intermediation frictions and demand pressures in the options market. The model has one important additional prediction: the degree of overpricing of traded options should be related to the costs and risks to the intermediaries of holding the short options positions. Those costs and risks, in standard models, e.g. Garleanu, Pedersen, and Potheshman (2008), depend on the cost of trading and the degree to which the returns on options can be replicated by dynamically trading futures. Past work has provided evidence on the importance of these hedging costs empirically – see also Jackwerth (2000), Bollen and Whaley (2004), Han (2008), Jurek and Stafford (2015), and Frazzini and Pedersen (2022). Dew-Becker and Giglio (2023) show that the downward trend in the traded-option premium lines up well with the decline in these hedging costs of intermediaries as measured by bid-ask spreads and basis risk in the futures market.

In addition to those factors, over time the demand asymmetry that must be borne by dealers alone may have also shrunk. It is well known that hedge fund investment strategies produce returns that are very similar to a short options position (Jurek and Stafford (2015)). As the hedge fund sector has grown over time, then, there is effectively an increased supply of option-like exposures, which offsets the retail demand. Similarly, the rise of exchange-traded products giving exposure to the VIX (such as short-volatility ETPs) also provides a source of supply of these exposures. Overall, then, a range of factors have trended over time towards shrinkage of the VRP, as observed empirically.
4 The importance of skewness

Skewness – typically negative skewness – is pervasive in returns on financial assets, and so there is a long history of studying it in finance. Negative skewness is also pervasive in the macroeconomy, though it gets somewhat less attention. Both the finance and macroeconomic literatures have made significant advances in measuring and modeling skewness, aimed at understanding skewness both as a source of risk and fluctuations, and as an endogenous variable that depends on the state of the economy. Interestingly, the macroeconomic and finance models interact on both of those topics. In terms of causation, finance gives a way of measuring conditional skewness – again, using option prices – that can be used to evaluate whether it is a driver of the business cycle. And in terms of the source of skewness, macroeconomic models provide new tools that help to model its dependence on other variables. We review these recent advances separately in finance and in macroeconomics.

4.1 Recent work on skewness in finance

There are two main channels through which skewness in returns can arise: jump/tail risk, and stochastic volatility (with volatility increasing when returns are more negative). The two are theoretically distinct concepts in a continuous-time setting, though in discrete time they cannot be fully distinguished from each other (because the part that is originated from within-period variation is not observable). A large literature in asset pricing has proposed measurement and models of jumps and tail risk, as well as stochastic volatility, in equity and option markets. This literature has been extensively reviewed (e.g. Embrechts et al. (2014) and Aït-Sahalia and Hansen (2010)). Another strand of the literature has focused on forecasting various moments, including skewness, and again it has been reviewed extensively (see, e.g., Christoffersen, Jacobs, and Chang (2013)).

4.1.1 Measurement and pricing of skewness

There have been a number of improvements in measurement of both realized and implied skewness, tail risk, and volatility asymmetry over the past decade. Some papers have focused directly on the measurement and pricing of realized skewness; most notably, Neuberger (2012) shows how measurement of realized skewness can be improved relative to just using the sample third moment, and Amaya, Christoffersen, Jacobs, and Vasquez (2015) study the pricing of realized skewness in equity markets.

Other studies have separately focused on the two potential sources of skewness: asymmetric volatility and tail risk. Starting from the former (i.e., the idea that volatility is higher
in bad compared to good times), the asset pricing literature has made some progress incorporating it in standard consumption models. For example, Segal, Shaliastovich, and Yaron (2015) study a version of the long-run risk model of Bansal and Yaron (2004) that allows for differential upside and downside volatility, and Bekaert and Engstrom (2017) allow for differential upside and downside volatility in a model with external habit formation. Several studies have focused on the econometric properties of upward and downward volatility, like Bekaert, Engstrom, and Ermolov (2015), Patton and Sheppard (2015), and Barunik, Kocenda, and Vacha (2016) (see also the recent review in Bollerslev (2022)). Pricing in equity markets has been studied, among others, by Bollerslev, Li, and Zhao (2020), and pricing in option markets (decomposing the variance risk premium into an upward component and a downward component) has been explored in Feunou, Jahan-Parvar, and Okou, (2018) and Kilic and Shaliastovich (2019) (see also Muravyev and Ni (2020)).

A second possible source of skewness is price jumps. Whereas an earlier literature in asset pricing focused on very large but rare tail events (e.g. as a consequence of economic disasters, as in Rietz (1988), Barro (2006) and Martin (2013)), the recent literature has focused on smaller, but more frequent jumps, that appear to be more aligned with observed return behavior (see the discussion in Backus, Chernov and Martin (2011)). The most important innovation in measuring and studying the pricing of tail risks has come from exploiting information in high-frequency data and options, especially short-dated ones (see Bollerslev and Todorov (2011; 2014), Bollerslev, Todorov, and Xu (2015), and Andersen, Fusari, and Todorov (2015)). In general, option markets are especially informative about skewness and its pricing because of the asymmetric nature of option payoffs. One point to keep in mind, that parallels the distinction between realized variance and uncertainty from above, is that given a certain maturity (say, a month), option prices do not allow researchers to distinguish across the different sources of skewness in returns (volatility asymmetry and jumps occurring within the month). That said, just like studying the term structure of options can help disentangle the pricing of realized variance and that of uncertainty risk, looking at options with different maturities can help disentangle jump risk from asymmetric volatility risk. In particular, as the time to maturity shrinks to zero, out of the money option prices will be especially informative about the former.

4.2 Tail risk expectations from cliquet options

Plain-vanilla equity options provide protection against a drop in the underlying price of a certain size over a fixed horizon. That drop can come cumulatively over the course of many days, or via a single large decline. As discussed above, we might be interested in separating
the two and understand the relative pricing. Here we introduce some novel evidence from a derivative, the *cliquet option*, whose payoff relates directly to a single large drop in the underlying.

Specifically, we obtained data from a participant in the inter-dealer broker market on crash cliquets on the S&P 500. While we have a few different specifications available, the best data appears to be for put spreads that protect against a 10–20 percent decline in the index *in a day*. That is, their payout on a given day $t$ is

$$X_{t}^{cliquet} = \max\{-10\% - r_{s&p500,t}, 0\} - \max\{-20\% - r_{s&p500,t}, 0\}$$

(2)

So if in a day $t$ the return of the S&P 500 is above -10%, the payoff is zero. It then provides linear protection up to -20% (again, of the return in that day), and then pays no more for declines larger than 20%. The cliquet options in our dataset have 6-month maturity, and are knockout options. That means that the cliquet will pay the payoff $X_{t}^{cliquet}$ every day until $X_{t}^{cliquet} > 0$ (i.e., the option is triggered), then the contract terminates. The price of the option therefore technically reflects the risk-neutral expectation of the first jump over the next 6 months. Assuming that a single-day decline of more than 10 percent happening twice in a given 6-month period is vanishingly unlikely (which might actually be too strong, but we can assume it just for simplicity), the price of the cliquets we are studying will be the risk-neutral expectation of $X_{t}^{cliquet}$ conditional on a drop of at least 10%, multiplied by the risk-neutral probability of such an event.

Figure 4 plots prices of the cliquets between 8/2013 and 12/2016 against the VIX in the top panel, and against a measure of left tail probability (probability of a 10% drop in the S&P 500 over the next week under the risk-neutral probability, from Bollerslev, Todorov and Xu (2015)) in the middle panel. The bottom panel shows cliquets on the European STOXX index against the VSTOXX (the analogous of VIX for the STOXX index).

All the three panels show high correlation between the series for most of the period. The top and bottom panel suggest that a very large part of the variation in the implied volatility indexes is due to jump risk. The middle panel indicates the (risk neutral) expectation of jumps over the following 6 months is also highly correlated with the short-term (1-week) jump risk estimated using short-maturity options.

These contracts provide a novel and interesting measure of financial risk. While our sample is short, future work that expands the series to a longer time period can use this data to study market’s perceptions on the quantity and prices of tail risk in financial markets.
Note: Figure plots cliquet price between Aug 2013 and Dec 2016. The top panel plots against the VIX, the middle panel plots against a measure of left tail probability, and the bottom panel plots cliquets on the European STOXX index against the VSTOXX.
4.3 Time-varying skewness in macroeconomics

Skewness, for both aggregate time series and in the cross-section, has received growing attention in macroeconomics. As in finance, there is a relatively long literature trying to understand unconditional skewness (e.g., Sichel (1993)), but the literature on time-variation in skewness is relatively more recent and active. This section first reviews recent empirical advances and then describes some theoretical work on the topic.

4.3.1 Measurement

At the micro level, Guvenen, Ozkan, and Song (2014) and Guvenen et al. (2021) study income data from the Social Security Administration and find significant procyclicality in cross-sectional skewness in individual income growth (Schmidt (2022) links this risk back to asset prices). Salgado, Guvenen, and Bloom (2023) show that a wide range of economic variables are not just skewed, but that the skewness varies procyclically: skewness is more negative in bad times.

Dew-Becker (2023) studies the link between financial measures of skewness, like those discussed in the previous section, and the macroeconomy. Consistent with the work on labor income, he finds that option-implied skewness for individual firms – measuring idiosyncratic risk – is significantly procyclical, becoming clearly more negative in recessions. In contrast, though, option-implied skewness for the S&P 500 – measuring aggregate risk – is actually countercyclical – it becomes less negative during recessions. That is true even though overall volatility seems to rise. One simple intuition is that in bad times, the Gaussian risks become relatively larger, while non-Gaussian jump risks do not grow by as much, with the overall result that asymmetry in the shape of the option-implied distribution shrinks toward zero. Gormsen and Jensen (2022) show that countercyclical skewness holds more generally across a sample of stock markets in 17 different countries.

In contrast to Dew-Becker (2023), and more in line with the micro skewness evidence, Iseringhausen, Petrella, and Theodoridis (2022) study conditional skewness in the macroeconomy based on the McCracken–Ng (2016) panel dataset of macro indicators. They find much stronger evidence for procyclicality in skewness. Their findings again highlight that financial markets and the macroeconomy are in general relatively weakly linked. While financial risk – whether measured by volatility or skewness – is in some recessions and by some measures higher, the link is very much mixed. As we discuss below, understanding when macro and financial risks are linked and when they are not is an important area of work moving forward.
4.3.2 Models

Salgado, Guvenen, and Bloom (2023) present a model in which time-varying skewness is an exogenous driver of the business cycle. Kozlowski, Veldkamp, and Venkateswaran (2020) try to understand where agents’ beliefs about conditional skewness come from. They develop a model in which agents learn about the distribution of shocks over time, with their conditional distributions varying as they learn. After a particularly negative shock, agents will believe such shocks to be more likely going forward, which naturally means that skewness becomes more negative following bad shocks, and hence in bad times. Orlik and Veldkamp (2022) make a similar point in the context of countercyclical uncertainty.

Ilut, Kehrig, and Schneider (2018) study a model in which firms respond to shocks in a concave manner. They give conditions under which such concave responses can lead to procyclical skewness. The paper is not meant to endogenize those concave responses, but rather to show their consequences.

Dew-Becker and Vedolin (2023) endogenize that concavity. They study a production network in which sectors produce output from inputs purchased from other sectors. The sector production functions display complementarity across inputs, causing them to respond in a concave manner to shocks. The effect of concavity is that when shocks are more dispersed or skewed to the left, aggregate output is lower. This is in many ways highly similar to Ilut, Kehrig, and Schneider (2018), except with an endogenous mechanism for concave responses to shocks that relies on the network structure of the economy. Additionally, though, the concavity is not part of a given firm or sector’s decision function. Rather, it arises out of interactions across economic units. Dew-Becker and Vedolin (2023) emphasize that those interactions are critical for matching the empirical fact that skewness is much more negative at the aggregate than at the firm level.

Finally, Jovanovic and Ma (2022) develop a model in which uncertainty and skewness vary together endogenously due to the adoption of new technologies.

5 Conclusions and avenues for future research

The last decade has seen many advances in understanding the relationship between financial uncertainty and the real economy. This review highlights progress in three specific areas: the term structure of uncertainty (both its pricing and the relation with the macroeconomy); variation of the variance risk premium over time, both at high frequency and in its longer-

8Earlier work, such as Chalkley and Lee (1998), Veldkamp (2005), and Fajgelbaum et al. (2017) also study skewness via learning, but more in order to get unconditional rather than conditional skewness.
term trends; and the dynamics of conditional skewness.

The review suggests a number of avenues for future research. First as to volatility and the business cycle, the relationship is clearly mixed. While some recessions are associated with very high aggregate and cross-sectional uncertainty, others are not. Why? Does the comovement depend on observable state variables?

Second, as to the term structure, there are clearly different shocks to conditional moments at different horizons. Realized volatility – which is either contemporaneous or even backward-looking, is most strongly associated with recessions and most strongly priced. In addition, much of the trade in derivatives markets has shifted to very short maturities, emphasizing their importance. What makes these highly transitory shocks more important than more persistent shocks?

Third, all of the results here are fundamentally about nonlinear processes. These are not simple ARMA models. Volatility changes over time, both in the aggregate and the cross-section, and does so (sometimes) cyclically. That time-variation can itself be a source of unconditional skewness, and there is additionally evidence of variation in conditional skewness in the economy. The vast majority of research is about linear or linearized models, but understanding the data discussed in this review – which represents a major aspect of the business cycle itself – fundamentally requires a nonlinear approach. While there is work on nonlinear models, very little of it has made its way into the canonical models analyzed in the literature and for policy. Being able to tractably incorporate nonlinearity and in a way that captures the most important features would make a valuable contribution.

References


24


Mehl, Paul, *Trading in Privileges on the Chicago Board of Trade* Circular No. 323, United States Department of Agriculture, 1934.


