Contractionary volatility or volatile contractions?

David Berger, Ian Dew-Becker, and Stefano Giglio*

May 8, 2016

Abstract

There is substantial evidence that the volatility of the economy is countercyclical. This paper provides new empirical evidence on the relationship between aggregate volatility and the macroeconomy. We aim to test whether that relationship is causal: do increases in uncertainty about the future *cause* recessions? We measure volatility expectations using market-implied forecasts of future stock return volatility. According to both simple cross-correlations and a wide range of VAR specifications, shocks to *realized* volatility are contractionary, while shocks to *expected* volatility in the future have no clear effect on the economy. Furthermore, investors have historically paid large premia to hedge shocks to realized volatility, but the premia associated with shocks to volatility expectations are have not been statistically different from zero. We argue that these facts are inconsistent with models in which increases in expected future volatility cause contractions, but they are in line with the predictions of a simple model in which aggregate technology shocks are negatively skewed. The latter view is also consistent with evidence that equity returns and real activity are negatively skewed.

1 Introduction

Volatility in financial markets and the real economy appears to be countercyclical, and a large body of recent macroeconomic research explores the effects of shocks to volatility and uncertainty on the economy.[1] The theoretical literature has focused on mechanisms through which shocks to volatility drive business-cycle fluctuations.[2] But it is also possible that there is no causal relationship and instead volatility is simply high when other shocks to the economy are negative. The simplest

*Northwestern University, Northwestern University and the University of Chicago. We appreciate helpful comments from Larry Christiano and seminar participants at CITE and the Chicago Fed.

[1] Gilchrist, Sim, and Zakrajsek (2014) use the same fact as a starting point for an analysis of volatility, irreversible investment, and financial frictions. See Campbell et al. (2001) (equity volatility at the index, industry, and firm level is countercyclical); Storesletten, Telmer, and Yaron (2004) and Guvenen, Ozkan, and Song (2014) (household income risk is countercyclical); Eisfeldt and Rampini (2006) (dispersion in industry TFP growth rates is countercyclical); Alexopoulos and Cohen (2009) and Baker, Bloom, and Davis (2015) (news sources use uncertainty-related language countercyclically); among many others, some of which are discussed below.

[2] Bloom (2009) and Bloom et al. (2014) study models with fixed adjustment costs in investment; Gourio (2012, 2013) examines time-varying disaster risk and its effects through precautionary saving and financial frictions; Basu and Bundick (2012, 2015) and Leduc and Liu (2015) explore the effects of volatility shocks in new-Keynesian DSGE models.

example of that scenario is when shocks are negatively skewed: then large realized values tend to be negative ones, so realized volatility – driven by extreme observations – tends to be high in bad times.

This paper provides novel evidence on the question of whether expectations of future volatility cause downturns, or whether, instead, volatility is simply high in bad times. That is, does volatility lead to contractions, or are contractions simply volatile? The key distinction that we draw is between *realized* and *expected* volatility. When we refer to volatility in this paper, what we mean is dispersion in the continuous flow of shocks that hits the economy. Models are typically clear: in some, increases in the variance of agents' subjective distributions over future outcomes induce recessions; in others, recessions are periods when volatility is high.

In the data, though, realized and expected volatility are not easy to disentangle. The central empirical exercise in Bloom (2009) highlights that difficulty. That paper estimates a vector autoregression (VAR) where "uncertainty" is measured as a single time series that splices together option-implied variances since 1986 with the history of actual squared returns on the S&P 100 in the period prior to the availability of the option-implied series. The option-implied variance measures investor expectations of future volatility, while squared returns measure realized volatility.[3] The question we ask is whether those two factors have the same effect on the economy, using novel data to distinguish between realized and expected future volatility.

We examine a sample of S&P 500 index options that has been little studied in the literature, but that allows us to measure volatility expectations since 1983 at horizons of up to 6 months. We thus have a sample spanning 32 years, three recessions, and a wide range of financial and real conditions in the economy. We measure realized volatility as the sum of daily squared returns on the market. This concept differs from that used by some other recent papers, but has a number of advantages, including directly measuring the concept of interest (the realized dispersion in aggregate shocks) and being linked in a well-defined way to our volatility forecasting variable. While there are many different potential measures of volatility and uncertainty in the economy, if one wants to cleanly distinguish between expectations and realizations, equity indexes are the only source that is available.

Using both cross-correlations and VARs, we find that increases in realized volatility are associated with declines in output, consumption, investment, and employment, consistent with findings in Bloom (2009) and Basu and Bundick (2015). More surprisingly, though, we find that increases in six-month *expected* volatility have no significant or consistent effect on the real economy across a range of specifications – some methods imply expected volatility shocks are mildly contractionary, others that they are actually expansionary, but we rarely obtain statistically significant results in either direction. Moreover, the difference in the effects of shocks to realized and expected volatility is statistically significant. So in other words, increases in uncertainty about future stock market performance do not appear to reduce output, even though realized volatility tends to be high in bad times. The results are consistent in finding significant effects for realized volatility and insignifi-

[3] Basu and Bundick (2015) make a similar observation.

cant effects for volatility expectations across estimation methods and they are robust to alternative identification schemes, changing the data sample and state variables, choices about detrending, and estimating the model at the quarterly or monthly frequency.

The simplest theoretical explanation for our results is that fluctuations in economic activity are negatively skewed. That could either be because the fundamental shocks are skewed, or because symmetrical shocks are transmitted to the economy asymmetrically (e.g. because constraints, such as financial frictions, bind more tightly in bad times). Skewness literally says that the squared value of a variable is correlated with the variable itself, which is essentially what we find when realized volatility is associated with contractions. There are two important pieces of evidence in favor of this hypothesis. First, changes in a wide variety of measures of real activity are negatively skewed, as are stock returns. Second, the asset prices we examine imply that investors have paid large premia for insurance against high realized volatility and extreme negative stock returns (known as the variance risk premium and the option skew or put premium, respectively) in the last 30 years, whereas the premium paid for protection against increases in *expected* volatility has been near zero or even positive.

Two important caveats apply to our results. First, our analysis is of the effects of fluctuations in *aggregate* uncertainty. We do not measure variation in cross-sectional uncertainty. There are obviously many dimensions along which uncertainty can vary, and we try to understand one here. Second, there are other measures of aggregate uncertainty that may still cause contractions even after controlling for realized volatility. The key contribution of this paper is to show that when a single concept of volatility can be split into components coming from realizations and expectations, it is the realization component that appears to drive the results. That does not imply, though, that no other measures of uncertainty (which do not distinguish between expectations and realizations) can affect the economy.

Our work is related to a large empirical literature that studies the relationship between volatility and the macroeconomy. A number of papers use VARs, often with recursive identification, to measure the effects of volatility shocks on the economy.[4] Ludvigson, Ma, and Ng (2015), like us, distinguish between different types of uncertainty. They show that variation in uncertainty about macro variables is largely an endogenous response to business cycles, whereas shocks to financial uncertainty cause recessions. The key distinction between our work and theirs is that we focus on the distinction between uncertainty expectations and realizations, while they distinguish between uncertainty in different sectors of the economy.

Our empirical analysis uses option prices to infer investor expectations of future volatility in the economy. We therefore draw from and build on a large literature in finance examining the pricing and dynamics of volatility.[5]

[4]See Bloom (2009) and Basu and Bundick (2015), who study the VIX; and Baker, Bloom, and Davis (2015) and Alexopoulos and Cohen (2009), who study news-based measures of uncertainty. Jurado, Ludvigson, and Ng (2015) and Ludvigson, Ma, and Ng (2015) measure uncertainty based on squared forecast errors for a large panel of macroeconomic time series (using a two-sided filter to extract a latent volatility factor).

[5]There is a long history of studies of fluctuations in aggregate (i.e. stock index) volatility, including, among many

The remainder of the paper is organized as follows. Section 2 lays out a simple reduced-form model of financial markets and the real economy that helps elucidate the distinction between realized and expected volatility. Sections 3 and 4 describe our data and examines its basic characteristics. Section 5 describes our main analysis of the relationship between realized volatility, expected volatility, and the real economy. Finally, section 6 examines implications of the contractionary volatility and volatile contractions hypotheses and argues that our evidence is consistent with the view that economic shocks are negatively skewed and that contractions are inherently risky. Section 7 concludes.

2 Theory

We consider a stylized model of output and stock returns to help elucidate the VARs that we estimate and their source of identification. We think of the data as being driven by a continuous-time process that is sampled discretely, e.g. on a monthly or quarterly basis. Suppose monthly (or quarterly) samples of stock returns follow the process

$$r_t = \sigma_{r,t-1}\varepsilon_{r,t} + \varepsilon_{J,t} \tag{1}$$

where $\varepsilon_{r,t}$ is a standard normal innovation (representing the accumulation of an orthogonal increment process, for example). $\sigma_{r,t-1}$ determines the volatility of $\varepsilon_{r,t}$ (i.e. the diffusive volatility of returns *during* period t), and we assume that it is known at the end of period $t-1$.

$\varepsilon_{J,t}$ is a jump shock with $\varepsilon_{J,t} = J \times (N_t - \lambda)$ and N_t is a draw from a Poisson distribution with intensity λ. $\varepsilon_{J,t}$ is independent of the other shocks and has zero mean by construction. r_t is the cumulative return during period t.[6]

Under the assumption that returns are driven by a continuous-time process, we can calculate realized volatility as the sum of squared returns sampled at high (e.g. daily) frequency. Under the assumptions that $\varepsilon_{r,t}$ is a discretely sampled diffusion and $\varepsilon_{J,t}$ is a Poisson process, total realized volatility during period t is

$$\begin{aligned} rv_t &= \sigma^2_{r,t-1} + N_t J^2 && (2) \\ &= \sigma^2_{r,t-1} + J\left(\varepsilon_{J,t} + J\lambda\right) && (3) \end{aligned}$$

rv_t here is calculated by summing squared stock returns sampled at high frequency during period t.[7] The result that the part of realized volatility associated with the diffusion is non-random is

others, Adrian and Rosenberg (2008), Bollerslev et al. (2009), Heston (1993), Ang et al. (2006), Carr and Wu (2009), Bakshi and Kapadia (2003), Egloff, Leippold, and Wu (2010), and Ait-Sahalia, Karaman, and Mancini (2013) (see Dew-Becker et al. (2016) for a review).

[6]Specifically, during period t, the continuous return process is $r^{cont} = \sigma_{t-1}dZ + J\left(dN - \lambda dt\right)$, where dZ is an orthogonal increment process and dN is a Poisson process with intensity λ. r_t is then the integral of r^{cont} over the period t.

[7]Technically, our definition of rv_t obtains when the sampling frequency for realized volatility becomes infinitely

typical of continuous-time models. Conditional on $\sigma^2_{r,t-1}$, the only uncertainty in rv_t comes from the number of jumps realized during period t.

We assume that $\sigma^2_{r,t}$ follows an ARCH-type process (Engle (1982); Bollerslev (1986)) where realized volatility in period t may affect volatility in the future,

$$\begin{aligned} \sigma^2_{r,t} &= \bar{\sigma}^2 + k_{\sigma,rv} rv_t + \varepsilon_{\sigma,t} \quad (4) \\ &= \bar{\sigma}^2 + k_{\sigma,rv}\sigma^2_{r,t-1} + k_{\sigma,rv} J\left(\varepsilon_{J,t} + J\lambda\right) + \varepsilon_{\sigma,t} \quad (5) \end{aligned}$$

$\varepsilon_{\sigma,t}$ is a mean-zero innovation that represents news about future volatility that is independent of current realized volatility. If news about future volatility affects discount rates and hence stock prices, then ε_σ and ε_r will be correlated. $k_{\sigma,rv}$ is a coefficient determining the extent to which realized volatility in period t feeds back into volatility in future periods. There is a very large finance literature on the persistence of volatility and how it can feed back on itself.[8] The idea here is that expected volatility during period t depends on the realization of volatility in period $t-1$ (e.g. high realized volatility in January is expected to lead to above average volatility in February also).

Finally, we assume that output follows the process

$$Y_t = k_{Y,J}\varepsilon_{J,t} + \varepsilon_{Y,t} + k_{Y,\sigma}\sigma_{r,t} \quad (6)$$

where $\varepsilon_{Y,t}$ is a mean-zero innovation that is uncorrelated with all other shocks and represents the part of output unrelated to financial markets (Y_t can also depend on its own lag without affecting the conclusions below). We assume that the shocks to output and financial markets that are shared are the large skewed shocks, $\varepsilon_{J,t}$, but allowing for correlations driven also by diffusive shocks does not change the conclusions. Finally, volatility expectations, $\sigma_{r,t}$, can directly affect output if $k_{Y,\sigma} \neq 0$. $k_{Y,\sigma}$ is interpreted as the structural effect of volatility expectations (uncertainty) on output.

$k_{Y,J}$ is a coefficient determining the impact of the jump shock on output. If $k_{Y,J} > 0$ and $J < 0$, we can think of this as a model where there are infrequent but large negative shocks that induce negative skewness in output. So it can be interpreted as a setting with sharp contractions (slow recoveries could be modeled by including an autoregressive term). We discuss the evidence on sharp contractions and negative skewness in output further below.

Now note that r_t is redundant for the purpose of explaining innovations to output, since the part of it connected to output, $\varepsilon_{J,t}$, is already spanned by rv_t. So for the moment we ignore it (if r_t is correlated with the other shocks, it can simply be included last in the VAR with no effect on our identification). Consider a first-order VAR in rv_t, $\sigma^2_{r,t}$, and Y_t. Denote the vector of state variables

high. In terms of the continuous-time specification, rv_t is the integral of $\left(r^{cont}\right)^2$ over the period t.

[8] See Engle (1982), Bollerslev (1986), and the literature following them.

as $X_t \equiv \left[rv_t, \sigma^2_{r,t}, Y_t\right]'$ and the feedback matrix Φ. The VAR yields

$$X_t = C + \Phi X_{t-1} + \begin{bmatrix} J & 0 & 0 \\ k_{\sigma,rv}J & 1 & 0 \\ k_{Y,J}J + k_{Y,\sigma}k_{\sigma,rv}J & k_{Y,\sigma} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{J,t} \\ \varepsilon_{\sigma,t} \\ \varepsilon_{Y,t} \end{bmatrix} \tag{7}$$

where C is a vector of constants. The matrix multiplying the shocks is already in its Cholesky form, with rv moving first, σ^2_r second, and Y_t last. The structural "shock" to rv identifies the jump, $\varepsilon_{J,t}$. That is, the only source of uncertainty in realized volatility conditional on information available at the end of period $t-1$ is the number of jumps realized in period t. So the estimated impact of shocks to rv_t on output is $k_{Y,J} + k_{Y,\sigma}k_{\sigma,rv}$, which includes both the structural effect of $\varepsilon_{J,t}$ on output and also the fact that $\varepsilon_{J,t}$ raises expectations of future volatility, which can further reduce output.[9]

The estimated "volatility shock" (or uncertainty shock), after controlling for rv_t, is just $\varepsilon_{\sigma,t}$. So the effect of this shock on output perfectly identifies the structural impact of volatility expectations on output, $k_{Y,\sigma}$. The identification comes from the fact that $\varepsilon_{\sigma,t}$ is the part of volatility expectations that is uncorrelated with realized volatility, so that is the shock that can disentangle $k_{Y,J}$ from $k_{Y,\sigma}$. That is, the VAR allows us to ask whether output and volatility expectations are simply correlated – through the jumps – or if volatility actually has causal effects on output.

We will estimate a VAR of exactly the form in (7), in which realized volatility moves first and volatility expectations second. One might worry that by controlling for realized volatility first, we will be identifying only the response of output to a small shock to volatility. What (7) shows, though, is that the estimated response of output to a unit shock to expected volatility, after controlling for RV, identifies $k_{Y,\sigma}$. That is, it identifies precisely the object of interest – the response of output to volatility expectations. Moreover, we show that ε_σ is not particularly small in economic terms.

An alternative would be to order $\sigma^2_{r,t}$ first in the Cholesky factorization (but leaving X_t defined in the same way, $X_t = \left[rv_t, \sigma^2_{r,t}, Y_t\right]'$). We then obtain

$$X_t = C + \Phi X_{t-1} + \begin{bmatrix} JQ & J & 0 \\ 1 & 0 & 0 \\ k_{Y,\sigma} + Qk_{Y,J} & k_{Y,J} & 1 \end{bmatrix} \begin{bmatrix} k_{\sigma,rv}J\varepsilon_{J,t} + \varepsilon_{\sigma,t} \\ \varepsilon_{J,t} - Q\left(k_{\sigma,rv}J\varepsilon_{J,t} + \varepsilon_{\sigma,t}\right) \\ \varepsilon_{Y,t} \end{bmatrix} \tag{8}$$

where Q is the coefficient from a regression of $\varepsilon_{J,t}$ on the innovation to $\sigma^2_{r,t}$, $(k_{\sigma,rv}J\varepsilon_{J,t} + \varepsilon_{\sigma,t})$. In this ordering, we still find that output is negatively related to the innovation to rv_t as long as $var\left(\varepsilon_{\sigma,t}\right) \neq 0$. The reason is that the innovation to $\sigma^2_{r,t}$ is imperfectly correlated with $\varepsilon_{J,t}$. So in either ordering, rv_t should have a negative estimated impact on the economy. The effect of $\sigma^2_{r,t}$ is

[9] It is important to note that Bloom (2009) controls for market returns in the VAR first – replacing rv in our setting (though Basu and Bundick (2015) do not). Since the market return depends on shocks other than $\varepsilon_{J,t}$, controlling for the return does not completely control for ε_J. That means that the Cholesky-identified "shock" to σ^2_t in that case still contains information about $\varepsilon_{J,t}$, potentially leading to a spurious correlation with output.

now estimated to be robustly negative, though. That is, even if $k_{Y,\sigma} = 0$, shocks to $\sigma^2_{r,t}$ still appear to affect output since they include information about $\varepsilon_{J,t}$. We would obtain the same result if rv_t were not included in the VAR.

This simple model illustrates two key mechanisms that can affect the identification in our VAR. First, since ARCH effects cause high realized volatility to affect volatility expectations, controlling for rv_t before σ^2_t is critical in testing the hypothesis that volatility shocks can affect the economy. True uncertainty shocks – changes in volatility expectations that are orthogonal to shocks to other variables, $\varepsilon_{\sigma,t}$ – are identified in the VAR as the part of $\sigma^2_{r,t}$ that cannot be explained by its own lag or by rv_t. Past work has typically included only either variables analogous to rv_t or $\sigma^2_{r,t}$, so they cannot distinguish whether volatility causes declines in output or is simply correlated with negative shocks to the economy.

For the identification to have any power, the shocks $\varepsilon_{\sigma,t}$ must have non-trivial variance. So it will be important for us to show that there are in fact shocks to expected volatility that are uncorrelated with current realized volatility – i.e. $std\left(\varepsilon_{\sigma,t}\right) > 0$. We will do so by showing that equity index options have forecasting power for future volatility even after controlling for current and past realized volatility.

The second important mechanism in the model is that realized volatility can be correlated with output because it isolates the large shock, $\varepsilon_{J,t}$, that is common to output and financial markets. In the model, there are small shocks, $\varepsilon_{Y,t}$ and $\varepsilon_{r,t}$, that are isolated to the real and financial sides of the economy, respectively, and obscure the relationship between output and stock returns (explaining their low unconditional correlation). But when big shocks hit the economy, they tend to affect both output and stock returns. rv_t, by looking at *squared* returns, isolates the large shocks in stock returns, essentially stripping out the diffusive noise (especially after controlling for lagged volatility expectations).

The results we report below will be consistent with a version of the model here in which there is no structural relationship between volatility and output: $k_{Y,\sigma} = 0$. Specifically, when $\sigma^2_{r,t}$ is ordered first (or when rv_t is not included in the VAR), it will be estimated to have a negative impact on output, consistent with past findings (Bloom (2009), Basu and Bundick (2015), etc.). But when rv_t is ordered first, only rv_t will have negative effects; the estimated impact of $\sigma^2_{r,t}$ will be economically and statistically insignificant. Again, though, this does not imply that realized volatility has a *causal* negative impact on the economy. The estimated response of output to "rv shocks" is simply interpreted as coming from large negative shocks that simultaneously affect the real economy and stock markets. We show in section 6.5 that innovations in real activity and stock returns are both negatively skewed.

The model here thus illustrates two key mechanisms that may be driving the data. First, if there are large and rare shocks that can affect both stock returns and output, then it will appear that volatility drives output, even though there is actually no structural relationship between the two: they are simply driven by a common shock. Second, due to ARCH effects, volatility expectations can appear to have an effect on output if one does not properly distinguish between changes

in volatility expectations due to realized volatility itself (the ARCH effects) and pure shocks to volatility expectations ($\varepsilon_{\sigma,t}$).

3 Data

The past literature has examined a number of different measures of risk and uncertainty in the economy. There is work that examines volatility in productivity and household income, text-based measures of uncertainty, and many other concepts of risk. The concept of volatility studied in this paper is aggregate equity return volatility. We focus on stock market volatility for a number of reasons.

The feature of the data that we desire to measure is the variance of the flow of aggregate shocks that hit the economy. We thus do not aim to measure cross-sectional dispersion in shocks or even forecast uncertainty. Jurado, Ludvigson, and Ng (2015), for example, construct a monthly measure of forecast uncertainty for a wide range of macroeconomic variables. Our goal, on the other hand, is to measure the variance of the *common* shocks to measures of activity, rather than the total dispersion of each measure.

In principle, realized volatility can be measured with data sampled at many different frequencies, and much past work has focused on monthly or quarterly data. We use data on daily stock returns for two reasons. First, having more observations allows one to measure the dispersion of shocks more accurately (a standard deviation can be measured with even a single observation – if the mean is known – but with very little accuracy). Second, by using high-frequency data, we eliminate the need to estimate the mean of the distribution. With sufficiently high-frequency data, the observed squared values reveal the variance, with the conditional mean having no effects.[10]

So stock market volatility both measures what we are interested in – the variance of the flow of shocks – and also does so accurately. This is not the only paper to study stock market volatility, and for good reason. Equity prices summarize information about the future path of the economy, so volatility in the economy should be expected to be related to volatility in the stock market. One would expect that almost any factor that affects risk in the economy would affect the riskiness of firms, since the revenue and profitability of firms ultimately depend on all the features of the economy. For example, in standard investment theories, stock prices are closely related to the discounted present value of the marginal product of capital (in q theory, that link is exact). So volatility in stock prices measures volatility in the stream of future marginal products that determines incentives to invest (as opposed to, for example, uncertainty about just a single month of profits or income).

[10]Specifically, suppose that there is some fundamental unit of time, dt, indexed by t, and shocks are drawn in each period. In each period, shocks are independently drawn from a distribution with mean μ_{dt} and standard deviation σ_{dt}. The expected squared value of a variable is equal to the mean squared plus the variance. The mean squared is proportional to dt^2, while the variance is proportional to dt. So if we measure the realized variance of a series as the sum of squared realizations using data sampled at a sufficiently high rate – with dt small – then the conditional mean of the series eventually has no impact on the realized variance.

That said, the stock market is obviously not the only way to measure volatility, and there are certainly aspects of the economy that it will not capture. But while alternative volatility concepts, like survey and text-based measures (like the measure of Baker, Bloom, and Davis (2015)), are important for providing insights into how consumers feel about the future, they are sometimes difficult to interpret quantitatively. Stock market volatility, on the other hand, is clearly defined and has a direct link to economic activity.

Finally, by measuring stock market volatility, we are able to draw a clear distinction between expectations and realizations of volatility ($\sigma^2_{r,t}$ and rv_t in the model above). None of the other measures of uncertainty that we are aware of has such a feature. Some sources clearly measure volatility realizations, some clearly measure expectations, and some do both. None, though, yield both expectations and realizations of volatility for the *same* underlying variable. And not only can we construct expectations, but those expectations are available at a range of different maturities. News shocks – i.e. shocks to expectations of future rather than current values of state variables – play an important role in many recent macroeconomic studies (e.g. Beaudry and Portier (2006); Alexopoulos (2011); Christiano, Motto, and Rostagno (2014)). Since we have measures of expectations at multiple horizons, we are able to measure these news shocks for volatility directly.

The most important drawback of our data is that it does not measure variation in purely idiosyncratic risk, which plays an important role in some recent models. This paper is not able to shed light on that class of models, but in future work the methods we study here may be able to be extended to examine fluctuations in idiosyncratic risk.

3.1 Model-free implied volatility (EV_n)

Equity return volatility is measured in this paper as the volatility of returns on the S&P 500 index. In particular, realized volatility in each month is measured as the sum of squared daily returns. We denote realized volatility in month t as RV_t. We obtain data on daily stock returns from the CRSP database, which has coverage since 1926.

We construct measures of expected future volatility, $\sigma^2_{r,t}$, using option prices. The prices of S&P 500 options are obtained from the Chicago Mercantile Exchange (CME), with traded maturities from 1 to at least 6 months since 1983.

The most widely used measure of expected volatility is the VIX. The VIX is an index, based on option prices, of implied volatility over the next 30 days. There are two main issues with using the VIX as a measure of expected future volatility in our empirical analysis. First, since realized volatility is persistent, one-month expectations are very highly correlated with realizations (e.g. due to the ARCH effects in the model above). But at longer horizons, we observe larger differences between realized and expected volatility, allowing for their effects to be separately identified (i.e. allowing us to measure $\varepsilon_{\sigma,t}$).

Moreover, the very short maturity of the VIX makes it less than ideal to capture the kind of uncertainty about the future that might matter to firms and consumers, whose horizons are

typically much longer (in principle it is possible to calculate a one-day implied volatility, but it is hard to believe it would be relevant for typical economic decisions).

The problems with the VIX are solved by examining implied volatilities at longer maturities. From a theoretical point of view, the expectation of future volatility 6 or 12 months ahead lines up more closely with firms' decision frequency in macroeconomic models. In addition, as we will argue below, the correlation of shocks to expectations of future volatility with shocks to realized volatility are low enough that they allow separate identification of the effects of the two shocks.

In this section, we describe how we construct a version of the VIX (the price of a claim to future realized volatility), extending it beyond the usual one-month horizon. We will refer to the extension of the VIX to arbitrary maturities as *model-free implied volatility*, or EV_n (expected volatility), where n refers to the horizon in months (so that the standard, 30-day VIX is EV_1).

The VIX and EV_n are claims to future realized volatility. The fundamental time period that we analyze in this paper is a single month, which is indexed by t. Within any month, the realized volatility of equity returns is calculated as

$$RV_t \equiv \left(\sum_{days \in t} r_i^2 \right)^{1/2} \tag{9}$$

where r_i is the log return on the S&P 500 on day i in month t.

There is a large literature in finance that examines model-free implied variances for various assets. Under very general conditions (Jiang and Tian (2005) and Carr and Wu (2009)), the price of a claim to realized volatility between dates $t+1$ and $t+n$ can be written as a function of time-t option prices:

$$EV_{n,t} = \left(2 \int_0^\infty \frac{O_t(n,K)}{B_t(n) K^2} dK \right)^{1/2} \approx \left(E_t \left[\sum_{m=1}^{n} \frac{M_{t,t+m} RV_{t+m}^2}{E_t M_{t,t+m}} \right] \right)^{1/2} \tag{10}$$

where EV represents market implied volatility expectations and $M_{t,t+m}$ represents the stochastic discount factor between dates t and $t+m$. The 1/2 exponent in all these formulas implies that all the quantities we look at are in volatility (standard deviation) rather than variance terms, the same units in which the VIX is expressed.

Model-free implied volatility EV_n is calculated as an integral over option prices, where K denotes strikes, $O_t(n,K)$ is the price of an out-of-the-money option with strike K and maturity n, and $B_t(n)$ is the price at time t of a bond paying one dollar at time $t+n$. $EV_{n,t}$ is approximately equal to total expected volatility over the next n months (to be most precise, model-free implied volatility is the square root of expected variance; RV^2 is used since variances are additive over time). The approximate equality is related to the discretization in the calculation of realized variance.

The model-free implied volatility makes extremely minimal assumptions about the dynamics of stock returns, which is why $EV_{1,t}$ is used by the CBOE as its definition of the VIX. Crucially,

unlike the Black–Scholes (1973) implied volatility, it does not require that volatilities be constant over time.[11]

Expectations involving the term $\frac{M_{t,t+n}}{E_t M_{t,t+n}}$ are often referred to as *risk-neutral*, or *risk-adjusted expectations*,

$$E_t^Q [X_{t+1}] \equiv E_t \left[\frac{M_{t,t+1}}{E_t M_{t,t+1}} X_{t+1} \right] \tag{11}$$

The model-free implied volatility (and hence the VIX) is a risk-neutral or market-implied expectation of future realized volatility. A risk-neutral expectation depends on both the physical expectation of future volatility and also any risk adjustment due to covariation with the pricing kernel (state prices). While we generally expect risk-neutral expectations to be biased on average (for volatility, risk premia will tend to push the price above the physical expectation), to the extent that risk premia are stable, changes over time in risk-neutral expectations are equivalent to changes in physical expectations. That is why the past literature on the effects of uncertainty shocks has used the VIX (a risk-neutral expectation) as a measure of expected future volatility.

Our main results use the model-free implied volatility as a direct measure of expectations. There are a number of potential justifications for that choice. First, if risk premia are constant or proportional to model-free implied volatility itself, then no bias is introduced to our results by using the model-free implied volatility instead of a physical expectation (risk premia proportional to model-free implied volatility simply induce a rescaling of the coefficient on it). Second, many economic decisions are made under the risk-neutral measure, making the model-free implied volatility the relevant concept of uncertainty.[12] For example, since firms take state prices as given, a firm's investment decision depends on state probabilities weighted by the stochastic discount factor. Finally, in order to directly address the possibility that risk premia could contaminate our results, we also will examine below an identification strategy (following Barsky and Sims (2011)) which only requires EV to be correlated with physical expectations of future volatility, rather than directly reveal them.

We construct the model-free implied volatility, EV_n, for the S&P 500 at maturities between 1 and 6 months. We calculate it using data on option prices from the Chicago Mercantile Exchange (CME), which allows us to construct a time series dating back to 1983. Throughout the paper and appendix, we examine a range of robustness tests, including constructing volatility expectations using alternative sources of option prices and using variance swaps, whose payoffs are directly linked to realized volatility (but for which the sample is only half as long). Computing the model-free implied volatility with real-world data requires several steps; the appendix provides an extensive description of our calculation methods and analyzes the accuracy of the data.

Finally, in the remainder of the paper we focus on the logs of realized volatility (RV) and 6-month expected volatility (EV_6). Given the high skewness of realized variance, the log transfor-

[11] Martin (2013) provides an alternative implied variance that involves even fewer restrictions than that we use here. We have replicated our analysis with his measure and find nearly identical results. We focus on the VIX-type measure simply because it has been widely studied in the past.

[12] Feldman et al. (2016) make a similar argument in the context of optimal monetary policy.

mation makes the results less dependent on the occasional volatility spikes. We refer to variables in logs everywhere with lower-case letters, e.g. $rv_t = \log RV_t$.

3.2 The time series of realized volatility and its expectations

Figure 1 plots the history of (annualized) realized volatility along with 6-month market expectations (EV_6). Both realized volatility and volatility expectations vary considerably over the sample. The two most notable jumps in volatility are the financial crisis and the 1987 market crash, which both involved realized volatility above 60 annualized percentage points and rises of the 6-month expectations to 40 percent. At lower frequencies, the periods 1997–2003 and 2008–2012 are associated with persistently high volatility expectations, while expectations are lower in other periods, especially the early 1980's, early 1990's, and mid-2000's. There are also distinct spikes in expected volatility in the summers of 2010 and 2011, likely due to concerns about the stability of the Euro and the willingness of the United States to continue to pay its debts.

Panel A of Table 1 reports descriptive statistics for the series in figure 1. The means of volatility expectations is substantially higher than that of realized volatility, which is due to the risk-adjustment mentioned above. Specifically, there is a negative risk premium on volatility (Coval and Shumway (2001) and Dew-Becker et al. (2016)), which causes the prices of financial claims on volatility to be upward biased estimates of future volatility. As we would expect, the standard deviation of expectations is smaller than that of realized volatility.

Panel B of Table 1 reports raw correlations of the logs of realized and 6-month expected volatility with measures of real economic activity – capacity utilization, the unemployment rate, and returns on the S&P 500 (correlations are similar in levels). Volatility is correlated with all three macroeconomic variables, most strongly with capacity utilization; what the rest of this paper will explore is whether this association is due to simple correlation with realized volatility, or whether expected future volatility is causally linked to macroeconomic activity.

4 The dynamics of volatility expectations

We now examine the dynamics of realized volatility and its market expectations more formally. We first estimate simple forecasting regressions to measure how well realized volatility is predicted by lagged market expectations. It is obviously critical that we show that model-free implied volatility actually forecasts volatility, otherwise we cannot claim to identify volatility news, $\varepsilon_{\sigma,t}$.

4.1 Forecasting regressions for realized volatility

Under the assumption that risk premia are constant, six-month volatility expectations, EV_6, should correspond to the expectation of realized volatility over the next six months plus a constant. If risk premia are time-varying, EV_6 may not be entirely driven by expectations of future volatility, but

we should still expect it to have forecasting power for future RV even after controlling for other predictors like lagged RV.

The top panel of Figure 2 plots the coefficient β_h in the univariate regression

$$rv_t = \alpha + \beta_h ev_{6,t-h} + \varepsilon_t \tag{12}$$

for different lags h (on the x-axis). The figure shows that market-implied expected volatility indeed has highly significant forecasting power for future realized volatility, with a statistically and economically significant coefficient even 12 months ahead. The coefficients are well below 1, though, which implies that ev_6 is not a pure statistical expectation of future realized volatility (which could be due to both measurement error and time-varying risk premia).

Since rv is a persistent process, and since there are well known ARCH effects in the data, a critical question is whether ev_6 contains any information about future rv after controlling for rv itself. We therefore estimate the multivariate specification,

$$rv_t = \alpha + \beta_h ev_{6,t-h} + \gamma_h rv_{t-h} + \varepsilon_t \tag{13}$$

β_h now measures whether market expectations have any predictive content for future realized volatility after controlling for lagged volatility. In order to β_h to be non-zero, there must be shocks to volatility expectations that are independent of ARCH effects, i.e. $std\left(\varepsilon_{\sigma,t}\right) \neq 0$.

The second row of Figure 2 shows the two sets of coefficients, β_h and γ_h, for different lags h. The left panel shows the coefficient on lagged rv, γ_h. As expected, lagged rv forecasts future rv, with a coefficient declining with the horizon. More interestingly, though, the right panel shows that ev does in fact have significant predictive power for future volatility at all horizons, even after controlling for lagged rv. So Figure 2 shows that there are in fact volatility news shocks, $\varepsilon_{\sigma,t}$, and they can be identified from option prices.

Table 2 expands the analysis by looking at realized volatility over 6-month horizons and adding a set of additional predictors. In particular, columns (1) and (2) repeat the analysis of Figure 2, demonstrating the predictive power of ev_6 for future 6-month realized volatility. Column (3) adds three macroeconomic variables (industrial production, employment, and the federal funds rate) that we will use in our VAR analysis below. The table shows that the forecasting power of ev_6 for future realized volatility survives adding many variables. The macroeconomic indicators do not have any significant forecasting power for future volatility.

Overall, figure 2 and table 2 show that market-implied volatility ev contains by itself a large fraction of the total expectation of future volatility; this will allow us to identify the effects of volatility expectations on the real economy.

4.2 Decomposing the variance of realized volatility

The results above show that model-free implied volatility can forecast future realized volatility. The next natural question is how much news there actually is about future volatility, and to what extent fluctuations in realized volatility are surprises. Exploiting properties of expectations, the variance of rv_t can be decomposed into three components,

$$V(rv_t) = V\left(rv_t - E_{t-1}\left[rv_t\right]\right) + V\left(E_{t-1}\left[rv_t\right] - E_{t-6}\left[rv_t\right]\right) + V\left(E_{t-6}\left[rv_t\right]\right) \tag{14}$$

where $V(X)$ denotes the unconditional variance of a variable X. The first term is the variance of the surprise in rv conditional on information available in the previous period. The second component is the news about rv_t that occurs between months $t-6$ and $t-1$. That is, it is the variance of the innovations in the expectation of rv_t over those months. It thus measures how much investors learn about rv_t on average in the five months before it is realized. Finally, $V\left(E_{t-6}\left[rv_t\right]\right)$ is the variance of expectations of rv six months ahead.

To implement this decomposition, we must construct expectations of future volatility at 1- and 6-month horizons. In particular, to account for the possibility that risk premia vary, to measure $E_t\left[rv_{t+n}\right]$, we project rv_{t+n} on $ev_{1,t}$ and $ev_{6,t}$ in an unrestricted regression (we obtain similar results if we add other predictors to the forecasting regression, consistent with the view that option prices contain all or nearly all available information about future volatility).

Panel C of Table 1 reports results for the variance decomposition. The table shows that approximately 45 percent of the variance of rv is due to the purely unexpected component, $rv_t - E_{t-1}rv_t$. That is, almost half of the entire variance of rv_t is a surprise, whether we form expectations using univariate models or taking into account information from asset prices that directly reflect investor expectations of future equity market variance.

About 33 percent of the variance of rv is instead due to the news that investors gain between months $t-6$ and $t-1$. Finally, the variance of $E_{t-6}\left[rv_t\right]$ accounts for 22 percent of the total variance of rv_t.

The results in table 1 show that when modeling realized variance, it is important to take explicitly into account the fact that almost half of the variation in rv is unpredictable, and only about 20 percent of it is predictable at horizons 6 months or longer. That is, there appears to be a high-frequency term, like a low-order moving average, that must be accounted for.

5 Variance shocks and the real economy

We now examine the relationship between shocks to volatility and the real economy and estimate the VAR described in the theoretical analysis.

5.1 Data

We focus on monthly data to maximize statistical power, especially since fluctuations in both expected and realized volatility are rather short-lived. We confirm that our results are highly similar in quarterly data, though.

We measure real activity using the Federal Reserve's measure of industrial production for the manufacturing sector. Employment and hours worked are measured as that of the total private non-farm economy. Inflation is measured using the CPI.

All of the variables except volatility are non-stationary, so we detrend them with a one-sided HP filter with a smoothing parameter of 1.296×10^7.

As we discuss further below, the results are robust to alternative specifications of the data. In particular, they are highly similar when we set the HP smoothing parameter to 1.296×10^5 or drop it entirely; examine volatility in levels or logs; use the 6-month, 3-month, or 1-month market expectation for volatility; or use alternative methods of calculating market volatility expectations (at-the-money implied volatility, variance swaps, or Martin's (2013) simple variance swap measure).

5.2 Cross-correlations

We begin by examining the raw correlations between measures of real activity and leads and lags of changes in realized and expected volatility. Rather taking a position on a causal interpretation of the dynamic relationship between volatility shocks and macroeconomic outcomes, in this section we simply report reduced-form empirical patterns of these variables.

We estimate a bivariate VAR with rv and ev_6 as state variables. We then regress four macroeconomic variables – industrial production, hours worked per employee, employment, and inflation – on 12 leads and lags of the two residuals from the VAR. The residuals are not orthogonalized to each other – the regression coefficients simply represent the conditional covariance of the macroeconomic variables with leads and lags of innovations to rv and ev_6. Specifically, denoting the residuals from the VAR as $\varepsilon_{rv,t}$ and $\varepsilon_{6,t}$, we estimate regressions of the form

$$Y_t = \sum_{j=-12}^{12} (b_{rv,j}\varepsilon_{rv,t-j} + b_{6,j}\varepsilon_{6,t-j}) + \mu_t \tag{15}$$

where $b_{rv,j}$ and $b_{6,j}$ are coefficients and μ_t is a residual. Y_t here denotes any of the four macroeconomic series used. We rescale both the residuals and the macroeconomic variables to have unit standard deviations. Figure 3 reports the coefficients, $b_{rv,j}$ and $b_{6,j}$.

The figure shows a very consistent pattern. For all macroeconomic variables except the CPI, an increase in $\varepsilon_{rv,t}$ is followed by a significant decline in activity within the next year. Furthermore, $\varepsilon_{rv,t}$ is weakly positively predicted by positive lagged economic conditions. That is, a strong economy is associated with low realized volatility in the past and higher volatility in the future. On the other hand, $\varepsilon_{6,t}$ does not predict declines in activity, and if anything predicts expansions. That is,

increases in expected volatility appear to be associated with *high* future output and employment. The relationship between the innovations and the CPI can be viewed as something of a placebo test. There is little reason to expect that volatility shocks would have substantial effects on inflation, so it is good to see that the estimated coefficients for inflation are close to zero.

While these graphs only reveal cross-correlations between different types of volatility shocks and the macroeconomy, and have no direct causal interpretation, they indicate a substantive difference between increases in realized and expected volatility.

5.3 Vector autoregressions

We now examine more standard vector autoregressions (VARs) to measure the impact of shocks to realized volatility (rv) and expected volatility (ev_6) on the economy. For all the VARs that we run, we include four lags, as suggested by the Akaike information criterion for our main specification.

Following the implications of the model in section 2, our main analysis identifies structural shocks in the VAR using a Cholesky factorization with realized volatility ordered first and expected volatility second. So what we refer to as an identified "realized volatility shock" is simply the residual from a regression of realized volatility on the lagged variables in the VAR. Under the model, this identifies the joint jumps in equity returns and output. As in the model, since expected volatility is ordered second, "expected volatility shocks" here represent changes in volatility expectations that are orthogonal to realized volatility – i.e. separate from ARCH effects. Again, we do not interpret the Cholesky ordering here as a statement about the timing of the ordering of shocks. Rather, we interpret the volatility expectations shock as what it literally is in an econometric sense: the change in six-month volatility expectations that cannot be explained by news about realized volatility in the same period.[13]

5.3.1 Coefficient estimates

We begin by examining the sums of the coefficients on lagged values of rv and ev_6 in the VAR. These results are useful to understand because they turn out to drive our key results and they are not dependent on any SVAR identification. Table 3 reports the sum of the coefficients for a range of different VARs. The first row reports the coefficients from the regression of log employment on rv and ev_6 from a VAR that includes only those three variables. The sum of the coefficients on rv is negative, while the sum of the coefficients on ev_6 is actually *positive*. High levels of realized volatility forecast low employment in the future, but high levels of expected volatility actually forecast high

[13]Christiano, Motto, and Rostagno (2014) study a model with a data-generating process that gives a formal justification for the ordering we use. Suppose realized volatility, rv_t follows the process

$$rv_t = \phi rv_{t-1} + \varepsilon_{0,t} + \varepsilon_{1,t-1} \tag{16}$$

where $\varepsilon_{1,t}$ is observable to agents on date t. Investors' expectation of volatility at date $t+1$ is thus $E_t rv_{t+1} = \phi rv_t + \varepsilon_{1,t}$. In period t, the innovation in rv_t is $\varepsilon_{0,t}$, while the innovation in $E_t rv_{t+1}$ is $\phi\varepsilon_{0,t} + \varepsilon_{1,t}$. If those innovations are rotated with a Cholesky factorization in which rv is ordered first, the identified shock to rv is $\varepsilon_{0,t}$ and the identified shock to $E_t rv_{t+1}$ is $\varepsilon_{1,t+1}$. That is how we intuitively understand the method we use.

employment. As is often found with VARs, the coefficients are not statistically significant, and their difference only has a p-value of 0.14. Nevertheless, this basic result will appear consistently through our results and will drive the other results we report below (which will themselves have more statistical power).

The second row of table 3 replaces log employment with log industrial production and finds similar, though statistically weaker results. The third and fourth rows report the coefficients on employment and industrial production from our main VAR that includes those two variables, rv, ev_6, and the Fed funds rate. Finally, the right-hand panel of table 3 reports results from the 1988–2006 subsample that eliminates the two biggest jumps in realized volatility. In all cases, we find similar results.

5.3.2 Impulse responses

We now examine impulse response functions, which describe the full dynamic response of the variables in the economy to each of the identified shocks. We begin by estimating a version of the VAR that has been used in the past literature that does not control for realized volatility. Here the VAR includes one-month implied volatility, ev_1 (here the VXO, which is closely related to the VIX).[14] All IRFs that we report are in response to unit standard deviation shocks. In figure 4, the shock to volatility expectations has a half-life of approximately 8 months and reduces employment and industrial production by statistically and economically significant amounts. The peak reduction in both following a volatility shock is approximately 0.35 percent. The magnitude of these responses is in line with those obtained by Basu and Bundick (2015), for example, though slightly larger. As noted above, by only looking at one volatility indicator (in this case VXO), it is impossible to distinguish whether the macroeconomic effects are caused by, or related to, the component of realized volatility with which VXO is correlated or the expectations component.

Figure 5 presents our main VAR results. The figure has four columns for the responses of rv, ev_6, employment, and industrial production to the shocks. We measure responses to unit standard deviation shocks to both realized and expected volatility.

The first row of figure 5 shows that a shock to realized volatility is highly transitory: the IRF falls by half within two months, and by three fourths within 5 months (compared to the eight-month half-life of shocks to the VXO). Consistent with the results in section 4, then, realized volatility appears to have a highly transitory component. Naturally, the shock to realized volatility also affects volatility expectations – the dynamics of 6-month volatility expectations appear to line up reasonably well with what is implied by the IRF for realized volatility itself.

As to the real economy, a unit standard deviation increase in realized volatility is associated with statistically and economically significant declines in both employment and industrial production, reducing them both by 0.3 percent. In terms of the model, this is consistent with the view that

[14] We use the VXO for this analysis, which is the equivalent of the VIX for the S&P 100 rather than the S&P 500, since it has a longer time series than the VIX; this is why previous literature has focused on the VXO rather than the VIX. In practice, the same results hold with VIX and VXO.

jumps in the stock market (which are typically negative) are associated with declines in output. That is, $k_{Y,J} > 0$. And, interestingly, the effect of realized volatility on the economy actually seems to be more strongly negative than that of the VXO (though that difference is not statistically significant).

The effects of a shock to volatility expectations are much different. First, as we would expect, a shock to volatility expectations forecasts high realized volatility in the future at a high level of statistical significance. That result alone is important: it says that shocks to ev_6, after controlling for the current shock to rv, are able to significantly predict future volatility. That is, option prices contain information about future volatility even after controlling for current and past realized volatility. Surprisingly, though, increases in volatility expectations are associated with no significant change in employment or industrial production. Employment declines only trivially, and industrial production appears to actually increase.

The bottom row of panels in figure 5 reports the difference in the IRFs for the rv and ev_6 shocks. The first two panels show that the two shocks have highly similar impacts on the future paths of realized and expected volatility (though on impact they obviously have different effects on rv simply due to the ordering of the identification). The cumulative effect of the two shocks on future rv (ignoring the impact period) is 0.76 for the rv shock and 0.87 for the ev_6 shock. In other words, the two shocks have highly similar effects on cumulative future volatility. Similarly, the rv shock raises future volatility expectations (i.e. the sum of the second panel in the top row) by 0.80, while the ev_6 shock raises them by 1.13.

		Cumulative responses	
		rv	ev_6
Shocks:	rv	0.76	0.80
	ev_6	0.87	1.13

The two shocks thus have approximately equal effects on future uncertainty – if anything, the ev shock has bigger effects on future uncertainty. But the figure shows that the two shocks in fact have significantly different effects on employment and output. Again, rv has much more negative effects, and that difference is statistically significant.

So the rv and ev_6 shocks have similar effects on future uncertainty but markedly different effects on the economy. What can explain that difference? In terms of the VAR, the only qualitative difference between the two shocks is that the rv shock has a large effect on realized volatility *on impact*. It is this initial impact effect that seems to drive output down. The results in figure 5 therefore show that periods of high realized volatility are associated with declines in activity, but shocks to volatility expectations have no significant effect on the economy.

5.3.3 Forecast error variance decompositions

To further understand the importance of the rv and ev shocks, figure 6 reports forecast error variance decompositions. As in figure 5, we report the effect of the rv shock, the ev_6 shock, and

their difference. The realized variance shock explains 10 percent of the variance of employment and 5 percent of the variance of industrial production at most horizons, while the point estimates for the fraction of the variance accounted for by expected volatility are close to zero. The upper end of the 95-percent confidence interval for the *ev* shock is below 5 percent for the first 10 months. The upper end of the 95-percent confidence interval for the *rv* shock, though, reaches as high as 25 percent for employment and 20 percent for industrial production 10 months ahead, indicating that RV can potentially be an important driver of the real economy (though this is not a causal statement – our model says that this can be simply due to common shocks to RV and output).

The bottom row of figure 5 shows that the fraction of the forecast error of volatility expectations accounted for by the two shocks is not significantly different, again emphasizing that they have similar effects on volatility expectations – the main difference between the two shocks is that one is affecting realized variance on impact, and the other is not. We see, though, that the difference in the fraction of the forecast error of employment and industrial production explained by the two shocks is meaningfully different economically, and marginally significantly at shorter horizons.

So in terms of the model, the estimated IRFs imply $k_{Y,\sigma}$ is zero or potentially even positive. The variance decompositions furthermore show that shocks to expected volatility have little or no importance for economic outcomes. Shocks to volatility expectations, after controlling for realized volatility and ARCH effects, seem to be associated with little change in the state of the real economy.

5.3.4 Quarterly data

In order to examine the effects of our two shocks on a wider range of variables, we also estimate a VAR using quarterly data similar to that of Basu and Bundick (2015) that includes, in addition to the two volatility series, GDP, consumption, investment, hours, the GDP deflator, the M2 money supply, and the Fed Funds rate (using the Wu and Xia (2014) shadow rate when the zero lower bound binds). Figure A.6 shows that following an increase in realized volatility, we obtain the same comovement emphasized by Basu and Bundick (2015): output, consumption, investment, and hours worked all decline, all statistically significantly. As we would expect, investment is most sensitive to the shock to realized volatility, with a peak response four times larger than that of output and six times larger than that of consumption. For expected volatility we again find no statistically or economically significant effects, with small initial declines and subsequent rebounds. Furthermore, the magnitude of the point estimates for the declines following expected volatility shocks is again not only statistically insignificant but also far smaller than the declines in response to realized volatility.

To summarize, then, we confirm the usual result that increases in one-month stock market volatility expectations are contractionary when they are included alone. But when the VAR includes both realized and expected volatility, we find that it is the increase in realized volatility that is associated with contractions, while increases in expected volatility are weakly associated with expansions.

5.4 Alternative identification schemes

A drawback of the results in the previous section is that they rely on the strong VAR structure and the Cholesky ordering of the shocks. Furthermore, the identification of the volatility expectations shock is potentially affected by time-varying risk premia. This section therefore shows that we obtain qualitatively similar results using two alternative estimation and identification schemes.

5.4.1 Estimation through local projections

In order to ensure that our results are not driven by the particular structure of the VAR that we examine, we now estimate impulse responses using the nonparametric method of Jordà (2005). The Jordà method uses local projections that estimate impulse responses essentially as partial derivatives. They rely much less strongly on the internal propagation of a VAR. Specifically, the Jordà projection estimates the partial derivative of the expected value of a variable, $E_t Y_{t+j}$, with respect to the current values of the state variables, X_t. That is, it yields $\partial E_t Y_{t+j}/\partial X_t$, which is exactly the definition of an impulse response function in a VAR, but the Jordà method is far less reliant on the strict VAR structure. In particular, the IRF at horizon j is obtained through direct regressions of Y_{t+j} on X_t and its lags. The difference between the Jordà projection and the VAR IRF is therefore equivalent to the difference between a direct and an iterated forecast (see Marcellino, Stock, and Watson (2006)).

Figure 7 reports estimates and confidence intervals for IRFs using the Jordà (2005) projection method. The red dotted lines indicate the IRFs from figure 5 estimated with OLS. Across all nine panels in the figure, the local projections, even though they have none of the internal propagation of the VAR, yield IRFs that are highly similar to what we obtained through OLS estimation of the VAR. Increases in rv are associated with increases employment and industrial production, while increases in ev_6 are not, and the difference in the IRFs for the two variables is again statistically significant. The results in figure 7 thus imply that there is nothing about the VAR itself – i.e. some sort of strange feedback between the endogenous variables – that is driving our results.

5.4.2 Identifying expected volatility shocks as news shocks

Since our goal is to understand shocks to expectations of future volatility, our analysis is closely related to the literature on news shocks, which often examines the effects of news about future productivity on the economy.[15] The main obstacle to estimating the effects of shocks to expectations of future productivity is that productivity expectations are not directly observable. Our data on market expectations of future stock market volatility distinguish us from the news shock literature because we are able to obtain a direct measure of news. That said, we can still use the econometric methods from the news shock literature to estimate the effects of volatility news. Moreover, those methods can be potentially helpful when the volatility risk premium varies over time.

[15] See Beaudry and Portier (2006), Beaudry, Nam, and Wang (2011), Arias, Rubio-Ramirez, and Waggoner (2014), and Barsky and Sims (2011), among others.

We identify news shocks using the method proposed by Barsky and Sims (B–S; 2011). The B–S identification yields two structural shocks (leaving the remainder of the structural shocks unidentified; in that sense it is set identified). The first shock is the pure surprise to volatility, which is precisely our realized volatility shock in the Cholesky factorization (this link is analytically exact and holds in all samples). The second shock is constructed to be the linear combination of the remaining shocks (up to a normalization) that accounts for the largest fraction of the forecast error variance of realized volatility over some set horizon, which we choose to be 12 months. In other words, the identified news shock under the B–S method is chosen so as explain as much of the future variation in realized volatility as possible. The B–S identification is valid under our model because the shock to volatility expectations that is orthogonal to rv_t, $\varepsilon_{\sigma,t}$ is strictly orthogonal to the other shocks in the model.

If our market-based measure of volatility expectations represents the statistical expectation of future volatility, then on average we would expect the B–S method to identify it as the news shock. In any sample, obviously, the other variables in the VAR may potentially also appear to forecast future volatility. On the other hand, if the volatility expectation measure is contaminated by risk premia, then any variable that is correlated with risk premia will also help forecast future volatility, and the B–S method will include them as part of the news shock. The B–S identification therefore yields a measure of volatility news that is robust to the presence of time-varying risk premia. For the B–S method to have power, we do not need to assume that market-based expectations are equal to statistical expectations; rather, we simply need it to be the case that market-based expectations contain *some* information about the future path of realized volatility.

Figures 8 and 9 report results using two versions of the B–S identification. First, in figure 8 we allow the news shock to depend on all the variables in the VAR. Figure 8 shows that the identification in this case is extremely weak in the sense that the confidence intervals are so wide as to be uninformative. The reason for this result is that the set of coefficients that determines the combination of reduced-form shocks that represent the new shock is very poorly identified.

In order to help strengthen the identification, figure 9 uses the B–S identification, but not allowing the shocks to the macro variables (the Fed funds rate, employment, and industrial production) to be included as part of the news shock. We also include in this VAR a measure of the slope of the term structure of volatility expectations to give more information about future volatility. So we construct a news shock in figure 9 as the combination of the reduced-form innovations to 6-month volatility expectations and the slope of the volatility term structure. Duffee (2011) shows that except for knife-edge cases, financial term structures (e.g. the volatility term structure we examine) contain all information about physical expectations available to investors, even if there are time-varying risk premia. The identification in figure 9 thus can consistently identify the shock to volatility expectations.

Figure 9 shows that we then obtain similar results to the main VAR in figure 5. We estimate a weakly negative response of employment to the news shock that is not statistically significant, while the initial response of industrial production is actually positive. Again, the bottom row of

figure 9 shows that the responses of employment and industrial production to the news and surprise shocks are statistically significantly different (though in this case only marginally so).

5.5 Robustness tests

We examine a range of perturbations of our main specification from figure 5. First, we consider alternative orderings of the variables in the VAR. The effects of the ordering depend ultimately on the correlation matrix of the innovations, which we report below:

	rv	ev_6	Fed Funds	Empl.	IP
rv	1				
ev_6	0.66	1			
Fed Funds	-0.02	-0.07	1		
Empl.	0.04	0.10	0.03	1	
IP	-0.01	0.10	0.03	0.55	1

Consistent with standard models of equity volatility (e.g. ARCH and GARCH models), the shocks to realized and expected volatility are strongly correlated (though far from collinear). Their ordering in the VAR is therefore obviously relevant for the results. However, their innovations are almost completely uncorrelated with those to the other variables, implying that the relative ordering of the financial and macro variables is unlikely to affect the results. We confirm that intuition in the appendix.

Figure A.5 reports results from a monthly VAR analogous to that of figure 5 where we reverse the ordering of realized and expected volatility. As predicted by the model, we then find that shocks to expected volatility appear to have negative effects on the economy. This shock essentially combines the positive shock to expected volatility in figure 5 (orthogonalized to realized volatility) and adds some of the realized volatility shock, thus mixing zero and negative effects.

Importantly, though, figure A.5 shows that even after we orthogonalize the realized volatility shock to expected volatility – i.e. if we just look at a change in realized volatility that has no direct effect on expectations, therefore a purely transitory shock to RV – we continue to find a negative effect on the real economy. The point estimate of the effect is of a similar (though slightly smaller) magnitude as the one we obtain in the main specification in figure 5. The negative relationship between the real economy and realized volatility thus appears to be a robust feature of the data, while the effects of expected future volatility depend on how the innovations are rotated.

The result that when expected volatility is ordered first, both it and realized volatility have negative effects on the economy is again consistent with the theoretical model. The part of realized volatility that is orthogonal to the increase in expected volatility still affects output through the coefficient $k_{Y,J}$. Conversely, the full shock to expectations in this ordering now includes both effects through $k_{Y,\sigma}$ and also through ARCH effects – part of the increase in volatility expectations comes from the jumps, so both $k_{Y,\sigma}$ and $k_{Y,J}$ in this ordering affect the estimated response of output to volatility expectations.

Figures A.7 and A.8 in the appendix report a range of additional robustness tests. Figure A.7 shows the response of log employment to the two volatility shocks, and figure A.8 reports the response of log industrial production. In each figure, each row corresponds to a different specification of the model. The left panels report responses to unit standard deviation shocks to rv, while the middle panels report the responses to unit standard deviation shocks to volatility expectations, ev_6. The right panel reports the difference between the two impulse-response functions.

The first row in the two figures shows the results obtained without detrending the macroeconomic series, which appears to have little effect on the results from a qualitative and quantitative standpoint. The second row reports a version of the main VAR using quarterly data instead of monthly. The results are again very similar. The third row re-estimates the VAR using only the period 1988 to 2006, i.e. excluding the two large volatility spikes in 1987 and 2008. The results are again similar to our baseline estimates, though the responses of industrial production are both now shifted upwards (rv has zero average effect and ev has a significantly positive effect). The fourth row of the figures orders rv and ev_6 last in the VAR, and obtains similar results (the decline in employment following a shock to ev is now marginally significantly negative, but still only half the size of the decline following a shock to rv).

5.6 Confounding effects of time-varying risk premia

Our main results measure volatility expectations with ev_6, which is a market-based measure of volatility expectations. ev_6 therefore depends not only on the statistical expectation of volatility, but also on a risk premium. A natural question, then, is whether variation in the risk premium embedded in ev_6 would be expected to affect our results.

Our first response to this concern is the Barsky–Sims (2011) identification that shows that when we identify volatility shocks as news shocks our results are unaffected. That is, our results are nearly unchanged if we explicitly account for the possibility that risk premia vary over time. However, even without the Barsky–Sims identification, there are still arguments for using ev_6 directly in the VAR.

Risk premia are generally viewed as varying countercyclically (e.g. Campbell and Cochrane (1999), among many others). Since volatility is high in bad times, it earns a negative risk premium (Coval and Shumway (2001)). A countercyclical risk premium would thus imply that when output is low, ev_6 should be *higher* than average (i.e. investors are particularly risk averse, so they are willing to pay more for the protection that a market-based expectation like ev_6 provides against high-volatility states).

So countercyclical variation in risk premia should induce a *negative* relationship between volatility expectations and the state of the economy, exactly the opposite of what we see in the data. In order for our results on the effects of volatility expectations to be due to fluctuations in the risk premium, it would have to be the case that investors are relatively *less* risk averse when output is low or falling.

As to realized volatility, risk premia should not have any direct effects. Realized volatility is measured simply as the sum of squared daily returns on the S&P 500. A day is sufficiently small that any realistic variation in the conditional mean return on the market would have a trivial effect on realized volatility.

5.7 Alternative measures of volatility

While stock market volatility has been widely studied and is our preferred measure, there are a number of other measures that have been used recently. We focus on the index of policy uncertainty constructed by Baker, Bloom, and Davis (BBD; 2015). That index potentially measures information that economic agents have that is independent of realized volatility.

Figure 10 report the results of the VARs. The top panel reports the response of the economy to a BBD shock in a VAR with only BBD to capture uncertainty. We find that innovations in the BBD index are associated with declines in employment and IP. The bottom panel (second and third rows) reports results from a *separate* VAR where realized volatility is ordered first and the BBD index second. We see that when we control for rv, the BBD index becomes only very weakly contractionary. Moreover, increases in the BBD index actually predict *declines* in realized volatility (though not statistically significantly). The estimated effects of an rv shock on the economy appear unchanged whether we include the BBD index or ev_6 in the VAR.

Jurado, Ludvigson, and Ng (2015) provide an important measure of uncertainty that depends on forecast uncertainty for a wide range of macro and financial variables. In VARs, their measure drives out both the VIX and realized stock market volatility as a predictor of future output growth. Their time series for uncertainty is measured purely based on the history of realized volatility (squared forecast errors). We therefore view the results of Jurado, Ludvigson, and Ng (2015) as consistent with ours, in the sense that they both say that realized volatility (or some transformation of it) is an important driver of fluctuations in the real economy. Our results are also consistent with more recent work by Ludvigson, Ma, and Ng (2015), who find that a large fraction of the variation in uncertainty in the economy is actually an endogenous response to shocks, rather than an exogenous driving force.

6 Contractionary volatility or volatile contractions?

The analysis above shows that fluctuations in expected future stock market volatility are associated with, if anything, expansions in output, while shocks to realized volatility are robustly associated with contractions. Those results immediately suggest that contractions are simply volatile, rather than news about volatility being contractionary. In this section, though, we discuss general theoretical implications of the two hypotheses about the relationship between volatility and the real economy. Those implications then motivate us to enrich our results above.

While there is broad interest in the effects of shocks to volatility, there is no canonical model

of volatility fluctuations, either causing or caused by business cycles. We therefore attempt to draw testable implications from a more basic description of contractionary volatility and volatile contractions.

6.1 Contractionary volatility

When we refer to contractionary volatility, we mean the hypothesis that an exogenous increase in volatility in the economy induces a recession (reducing output, employment, consumption, and investment). In the language of the model in section 2, the hypothesis is that $k_{Y,\sigma} < 0$. In general, such a hypothesis assumes that the uncertainty is about the *future*, implying that the impact of the shock occurs before the realization of the volatility. That is, the assumption is that on date t, agents receive information that the variance of shocks on date $t+1$ or later will be high ($\varepsilon_{\sigma,t}$). So the news arrives before the volatility itself does.

Fernandez-Villaverde et al. (2013), Basu and Bundick (2015), Leduc and Liu (2015) are leading recent examples of such models. We claim that there are at least three basic implications that can be drawn from the contractionary volatility hypothesis. As a specific example of this type of model, we have calibrated a model similar to that of Basu and Bundick (2015) and show that the implications claimed below hold in that case.

6.1.1 Implications:

1. *Unexpected increases in expected volatility, i.e. news about future volatility, cause low or falling economic activity.* This is really the basic assumption of the contractionary volatility hypothesis. In the most basic form of the hypothesis, then, there should be a negative relationship between economic activity and shocks to expected volatility.

2. *News about high future volatility should earn a negative risk premium.* The hypothesis is that news about high future volatility is associated with reductions in output and consumption. If periods of low output are also periods with high marginal utility, then people will be willing to pay to hedge shocks to volatility expectations, in the sense that assets whose returns covary positively with expected volatility will earn negative returns. For example, under power utility, if consumption is low when expected volatility is high, then the Arrow–Debreu price of such states is relatively high, and assets that insure investors against high volatility should earn negative risk premia.

3. *Stock returns and innovations to output need not be skewed.* The key assumption of the contractionary volatility hypothesis is that the key shock is one to expected *future* uncertainty. That is, in period t there is news that risk on dates $t+1$ or later is higher. Stock returns and output growth will thus be unexpectedly negative on date t and volatile on date $t+1$. But skewness is a measure of the correlation between the level and volatility (i.e. $E\left[\varepsilon^3\right] = E\left[\varepsilon \cdot \varepsilon^2\right]$). Since negative shocks are not contemporaneously correlated with volatile shocks

in this model, there is no force generating skewness. That is not to say that a contractionary volatility could not have skewness in it. Rather, skewness is not a fundamental characteristic that must be common to contractionary volatility models.

6.2 Volatile contractions

The volatile contractions hypothesis says that recessions are periods of high volatility. It is fundamentally a story about skewness: negative shocks are volatile shocks, so denoting shocks as ε, $E\left[\varepsilon^3\right] < 0$. In terms of the model in section 2, this simply says that $k_{Y,j} > 0$. There is no claim about causation or timing. Rather the view is that bad news is volatile news – jumps *are* high realized volatility and tend to be negative. This type of model may be illustrated with a simple real business cycle model that features negatively skewed innovations to the level of technology. There is a long literature studying mechanisms that can generate skewness endogenously. Models with frictions that bind mainly in recessions (e.g. financial frictions) tend to generate skewness (see, for example, Kocherlakota (2000) and Ordonez (2013)). Other mechanisms, such as learning (van Nieuwerburgh and Veldkamp (2006)) can also generate asymmetries, or the fundamental shocks might simply be skewed to the left.

6.2.1 Implications:

The subjects of the implications here parallel those for contractionary volatility, but apply to realized rather than expected volatility.

1. *Increases in realized volatility are correlated with low or falling economic activity.* Under the volatile contractions hypothesis, there is no causal relationship between expected volatility and activity ($k_{Y,\sigma} = 0$). There will be, however, a clear correlation between realized volatility and low output ($k_{Y,J} > 0$). Again, this is really just the basic assumption of the volatile contractions hypothesis.

2. *Surprises in realized volatility should earn a negative risk premium.* The key implication of the volatile contractions hypothesis is that high realized volatility occurs in bad states of the world. We would thus expect investors to be willing to pay for insurance against realized volatility.

3. *Stock returns and innovations to output will be negatively skewed.* Again, the key assumption in this case is that negative shocks are volatile shocks. This is fundamentally a statement that $E\left[\varepsilon^3\right] < 0$. In terms of our toy model, this follows from the skewness of $\varepsilon_{J,t}$, which is a Poisson process.

6.3 Distinguishing the hypotheses

The results we obtain above showing that realized volatility is contractionary while expected volatility is not provide direct evidence on the first implication of the two models. However, there are two

additional implications of the contractionary volatility and volatile contractions models that allow them to be distinguished. We now show that the data on risk premia and skewness also support the view that contractions are volatile, rather than that volatility is contractionary.

6.4 Risk premia

In the main analysis above, we used claims on stock return variance as state variables in a VAR. But since those assets are financial assets, they themselves have returns. Assets that are risky, in the sense that their performance is positively correlated with the state of the economy (more formally, negatively correlated with state prices), should earn positive risk premia, while assets that are hedges, in the sense that their performance is *negatively* correlated with the state of the economy, should earn *negative* risk premia.

To better understand the information content in the term structure of risk premia, it is useful to take a simple transformation of the market-expectations of volatility, constructing *variance forwards*. While $EV_{n,t}$ represents an expectation to realized volatility between times $t+1$ and $t+n$, variance forwards $F_{n,t}$ represent expectations of realized variance *only during period* $t+n$. Variance forwards can be constructed in a similar way as $EV_{n,t}$ (note that these forwards represent claims to future variance, i.e. the square of RV, which is more convenient for the risk premia interpretation below).

Consider now the return corresponding to holding a variance forward for a single month. A forward with a maturity of one month has a price of $F_{1,t}$ at time t. After one month, it expires paying off realized variance RV^2_{t+1}. This claim therefore represents a direct hedge for the unpredictable component of realized volatility, and its expected return tells us about the risk premium associated with unpredictable shocks to realized volatility.

The second forward, with maturity 2, has a price of $F_{2,t}$ at time t. After one period, that asset becomes a 1-month forward. The one-month holding period return, therefore, will be $(F_{1,t+1} - F_{2,t})/F_{2,t}$. Similarly, for any forward with maturity $n > 1$, the one-month holding period return is $(F_{n-1,t+1} - F_{n,t})/F_{n,t}$. As is clearly visible from these formulas, for all $n > 1$, the one-month return of an n-month forward depends on the *change in the expectation* of future variance (since the forward price F is just a market expectation of variance n periods ahead). Therefore, all forwards except the very first one (F_1) are contracts that hedge *news* about future variance. The expected returns of these forwards reflect the risk premium associated with news about future variance.

Under the contractionary volatility hypothesis, then, we would expect that the 6-month variance forward earns a negative risk premium, while there is no particular prediction about the premium for realized volatility. Conversely, the volatile contractions hypothesis says that news about future volatility need not earn any premium. But volatile contractions imply that the 1-month forward, since it is exposed to realized volatility, which is high in downturns, should earn a negative risk premium.

The question, then, is whether risk premia are higher for short-term (1-month) or longer-term

(6-month) variance forwards. Dew-Becker et al. (2016) study that question in detail. We report here a simple summary of the risk premia and also extend their sample back to the 1980's.

Figure 11 (left panel) plots the average annual Sharpe ratios earned by forward variance claims between 1996 and 2015. We focus on this sample because it allows us to construct the returns from data on variance swaps, which have the highest quality price information (they are direct claims on variance, whereas the model free implied volatilities we study require numerically integrating across many option prices). The solid line plots the sample mean Sharpe ratio, while the dotted lines bound the 95-percent confidence interval. The results in the figure are stark: only the one- and two-month variance forwards earn returns that are statistically significantly negative – all other point estimates are positive, and some are even statistically significantly so.

As an economic matter, the annual Sharpe ratios that we observe on the one-month variance swap are extremely large, at -1.3. The return earned by an investor who sells one-month variance claims has been three to five times larger than that earned by an investor in the aggregate stock market. In other words, investors have paid enormous premia for protection against periods of high realized volatility, but the premia associated with shocks to expected future volatility are not distinguishable from zero (and are statistically significantly smaller than that on realized volatility).

The Sharpe ratios directly reveal what states of the world investors have paid for protection against. Specifically, under power utility, one may show that for any asset return x, the Sharpe ratio is

$$\frac{ER\left[x\right]}{SD\left[x\right]} = -RRA \times std\left(\Delta c\right) \times corr\left(x, \Delta c\right) \tag{17}$$

where RRA is the coefficient of relative risk aversion. Assets with larger Sharpe ratios therefore have larger correlations with consumption growth, Δc. More generally, assets that have a higher correlation with marginal utility should earn more negative Sharpe ratios. The fact that the largest Sharpe ratios are earned by one-month variance claims suggests that investors have historically viewed realized volatility as being more correlated with their marginal utility than news about future volatility.[16]

In order to be able to use our longer sample running back to the 1980's, figure 11 (right panel) reports the average shape of the term structure of variance forwards constructed using the CME options data (which is estimated more precisely than the month-to-month returns). The term structure reported here is still extremely informative about risk premia. The average return on an n-month variance claim is:

$$E\left[\frac{F_{n-1,t} - F_{n,t-1}}{F_{n,t}}\right] \approx \frac{E\left[F_{n-1}\right] - E\left[F_n\right]}{E\left[F_n\right]} \tag{18}$$

$$= E\left[R_{n,t}\right] - 1 \tag{19}$$

The slope of the average term structure thus indicates the average risk premia. If the term structure

[16]Both here and in Dew-Becker et al. (2016) we emphasize that our asset pricing data simply measures historical returns: we make no claim about what risk premia will be in the future.

is upward sloping, then the prices of the variance claims fall on average as their maturities approach, indicating that they have negative average returns. If it slopes down, then average returns are positive.

Figure 11 plots the average term structure of variance forward prices for the period 1983–2013. The term structure is strongly upward sloping for the first two months, again indicating that investors have paid large premia for assets that are exposed to realized variance and expected variance one month in the future. But the curve quickly flattens, indicating that the risk premia for exposure to fluctuations in expected variance farther in the future have been much smaller.

The asset return data says that investors appear to have been highly averse to news about high realized volatility, while shocks to realized volatility have not seemed related to marginal utility. The confidence intervals that we obtain are sufficiently wide that we cannot claim that shocks to expected future volatility do not earn an economically meaningfully negative risk premium. What we can say, though, is that investors seem to have cared over our sample *much more* about surprises in realized volatility than in expected volatility.

6.5 Skewness

The volatile contractions model is fundamentally about skewness in shocks. There are large literatures studying skewness in both aggregate stock returns and economic growth. We therefore provide just a brief overview of the literature and the basic evidence.

Table 4 reports the skewness of monthly and quarterly changes in a range of measures of economic activity. Nearly all the variables that we examine are negatively skewed, at both the monthly and quarterly levels. One major exception is monthly growth in industrial production, but that result appears to be due to some large fluctuations in the 1950's. When the sample is cut off at 1960, the results for industrial production are consistent with those for other variables.

In addition to real variables, table 4 also reports realized and option-implied skewness for S&P 500 returns.[17] The implied and realized skewness of monthly stock returns is substantially negative, and in fact surprisingly similar to the skewness of capacity utilization. The realized skewness of stock returns is less negative than option-implied skewness, which is consistent with investors demanding a risk premium on assets that have negative returns in periods when realized skewness is especially negative (i.e. that covary positively with skewness).

In addition to the basic evidence reported here, there is a large literature providing much more sophisticated analyses of asymmetries in the distributions of output and stock returns. Morley and Piger (2012) provide an extensive analysis of asymmetries in the business cycle and review the large literature. They estimate a wide range of models, including symmetrical ARMA specifications, regime-switching models, and frameworks that allow nonlinearity. The models that fit aggregate output best have explicit non-linearity and negative skewness. Even after averaging across models using a measure of posterior probability, which puts substantial weight on purely symmetrical

[17] We obtain option-implied skewnesss from the CBOE's time series of its SKEW index, which is defined as $SKEW = 100 - 10 \times Skew\,(R)$. We thus report $10 - SKEW/10$.

models, Morley and Piger find that their measure of the business cycle is substantially skewed to the left, consistent with the results reported in table 4.

Finally, the finance literature has long recognized that there is skewness in aggregate equity returns and in option-implied return distributions (see Campbell and Hentschel (1992), Ait-Sahalia and Lo (1998), and Bakshi, Kapadia, and Madan (2003), for recent analyses and reviews). The skewness that we measure here appears to be pervasive and has existed in returns reaching back even to the 19th century (Campbell and Hentschel (1992)).

Taken as a whole, then, across a range of data sources and estimation methods, there is a substantial body of evidence that fluctuations in the economy are negatively skewed. In a world of negative skewness, it is not surprising that measures of realized volatility are correlated with declines in activity, simply because skewness is related to the third moment: $E\left[\varepsilon^3\right] = E\left[\varepsilon \cdot \varepsilon^2\right]$.

7 Conclusion

The goal of this paper is to understand whether shocks to uncertainty have negative effects on the economy. Our contribution is to estimate a range of models that include measures of both volatility expectations and also realized volatility. We find that shocks to expected volatility in the future, after controlling for current realized volatility, do not have meaningful negative effects on the economy. The evidence we present favors the view that bad times are volatile times, not that volatility causes bad times. A leading hypothesized explanation for the slow recovery from the 2008 financial crisis has been that uncertainty since then has been high. Our evidence suggests that uncertainty may not have been the driving force, and that economists should search elsewhere for an explanation to the slow recovery puzzle.

References

Adrian, Tobias and Joshua Rosenberg, "Stock Returns and Volatility: Pricing the Short-Run and Long-run Components of Market Risk," *The Journal of Finance*, 2008, *63(6)*, 2997–3030.

Ait-Sahalia, Yacine and Andrew W Lo, "Nonparametric estimation of state-price densities implicit in financial asset prices," *The Journal of Finance*, 1998, *53* (2), 499–547.

____ , Mustafa Karaman, and Loriano Mancini, "The Term Structure of Variance Swaps, Risk Premia and the Expectations Hypothesis," 2015. Working paper.

Alexopoulos, Michelle, "Read All about It!! What Happens Following a Technology Shock?," *American Economic Review*, 2011, *101* (4), 1144–1179.

____ and Jon Cohen, "Uncertain Times, Uncertain Measures," 2009. Working paper.

Andersen, Torben G. and Oleg Bondarenko, "Construction and Interpretation of Model-Free Implied Volatility," Working Paper, National Bureau of Economic Research September 2007.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, "The Cross-Section of Volatility and Expected Returns," *The Journal of Finance*, 2006, *61(1)*, 259–299.

Arias, Jonas, Juan Francisco Rubio-Ramirez, and Daniel F Waggoner, "Inference based on SVARs identified with sign and zero restrictions: Theory and applications," 2014. 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.

Barsky, Robert B and Eric R Sims, "News shocks and business cycles," *Journal of Monetary Economics*, 2011, *58* (3), 273–289.

Basu, Susanto and Brent Bundick, "Uncertainty Shocks in a Model of Effective Demand," Working Paper 18420, National Bureau of Economic Research 2015.

___ and ___ , "Uncertainty Shocks in a Model of Effective Demand," 2015. Working paper.

Beaudry, Paul and Franck Portier, "Stock Prices, News, and Economic Fluctuations," *The American Economic Review*, 2006, *96* (4), 1293–1307.

___ , Deokwoo Nam, and Jian Wang, "Do mood swings drive business cycles and is it rational?," Technical Report 2011. Working paper.

Black, Fischer 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.

___ , Scott Baker, and Steven J. Davis, "Measuring Economic Policy Uncertainty," 2015. Working paper.

Bollerslev, Tim, "Generalized autoregressive conditional heteroskedasticity," *Journal of econometrics*, 1986, *31* (3), 307–327.

___ , George Tauchen, and Hao Zhou, "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 2009, *22(11)*, 4463–4492.

Broadie, Mark and Jerome Detemple, "American option valuation: new bounds, approximations, and a comparison of existing methods," *Review of Financial Studies*, 1996, *9* (4), 1211–1250.

Campbell, John Y. and John H. Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 1999, *107* (2), 205–251.

___ and Ludger Hentschel, "No news is good news: An asymmetric model of changing volatility in stock returns," *Journal of Financial Economics*, 1992, *31* (3), 281–318.

___ , Martin Lettau, Burton G. Malkiel, and Yexiao Xu, "Have Individual Stocks Become More Volatile? an Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 2001, *56* (1), 1–43.

Carr, Peter and Liuren Wu, "Variance Risk Premiums," *Review of Financial Studies*, 2009, *22(3)*, 1311–1341.

Coval, Joshua D. and Tyler Shumway, "Expected Option Returns," *The Journal of Finance*, 2001, *56(3)*, 983–1009.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, "The Price of Variance Risk," 2016. Working paper.

Duffee, Gregory R., "Information in (and not in) the Term Structure," *Review of Financial Studies*, 2011, *24(9)*, 2895–2934.

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.

Eisfeldt, Andrea L. and Adriano A. Rampini, "Capital Reallocation and Liquidity," *Journal of Monetary Economics*, 2006, *53* (3), 369–399.

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.

Feldman, Ron, Kenneth Heinecke, Narayana Kocherlakota, Samuel Schulhofer-Wohl, and Thomas Tallarini, "Market-Based Expectations as a Tool for Policymakers." Working paper.

Fernandez-Villaverde, Jesus, Pablo Guerron-Quintana, Keith Kuester, , and Juan Rubio-Ramirez, "Fiscal Volatility Shocks and Economic Activity," 2013. Working paper.

Gatheral, Jim and Antoine Jacquier, "Convergence of Heston to SVI," *Quantitative Finance*, 2011, *11* (8), 1129–1132.

___ and ___ , "Arbitrage-free SVI volatility surfaces," *Quantitative Finance*, 2014, *14* (1), 59–71.

Gourio, Francois, "Disaster Risk and Business Cycles," *American Economic Review*, 2012, *102(6)*, 2734–2766.

___ , "Credit Risk and Disaster Risk," *American Economic Journal: Macroeconomics*, 2013, *5(3)*, 1–34. Working paper.

Guvenen, Fatih, Serdar Ozkan, and Jae Song, "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 2014, *122* (3), 621–660.

Homescu, Cristian, "Implied Volatility Surface: Construction Methodologies and Characteristics," 2011. Working paper.

Jiang, George J. and Yisong S. Tian, "The Model-Free Implied Volatility and its Information Content," *Review of Financial Studies*, 2005, *18* (4), 1305–1342.

___ **and** ___ , "Extracting Model-Free Volatility from Option Prices: An Examination of the VIX Index," *The Journal of Derivatives*, 2007, *14* (3), 35–60.

Jordà, Òscar, "Estimation and inference of impulse responses by local projections," *American economic review*, 2005, pp. 161–182.

Jurado, Kyle, Sydney Ludvigson, and Serena Ng, "Measuring Uncertainty," *American Economic Review*, 2015, *105* (3), 1177–1216.

Kocherlakota, Narayana, "Creating business cycles through credit constraints," *Federal Reserve Bank of Minneapolis Quarterly Review*, 2000, *24* (3), 2–10.

Leduc, Sylvain and Zheng Liu, "Uncertainty Shocks are Aggregate Demand Shocks," 2015. Working paper.

Lee, Roger W., "The Moment Formula for Implied Volatility at Extreme Strikes," *Mathematical Finance*, 2004, *14* (3), 469–480.

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.

Marcellino, Massimiliano, James H Stock, and Mark W Watson, "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series," *Journal of Econometrics*, 2006, *135* (1), 499–526.

Martin, Ian, "Simple Variance Swaps," 2013. Working paper.

Morley, James and Jeremy Piger, "The asymmetric business cycle," *Review of Economics and Statistics*, 2012, *94* (1), 208–221.

Nieuwerburgh, Stijn Van and Laura Veldkamp, "Learning asymmetries in real business cycles," *Journal of Monetary Economics*, 2006, *53* (4), 753–772.

Ordonez, Guillermo, "The Asymmetric Effects of Financial Frictions," *Journal of Political Economy*, 2013, *121* (5), 844–895.

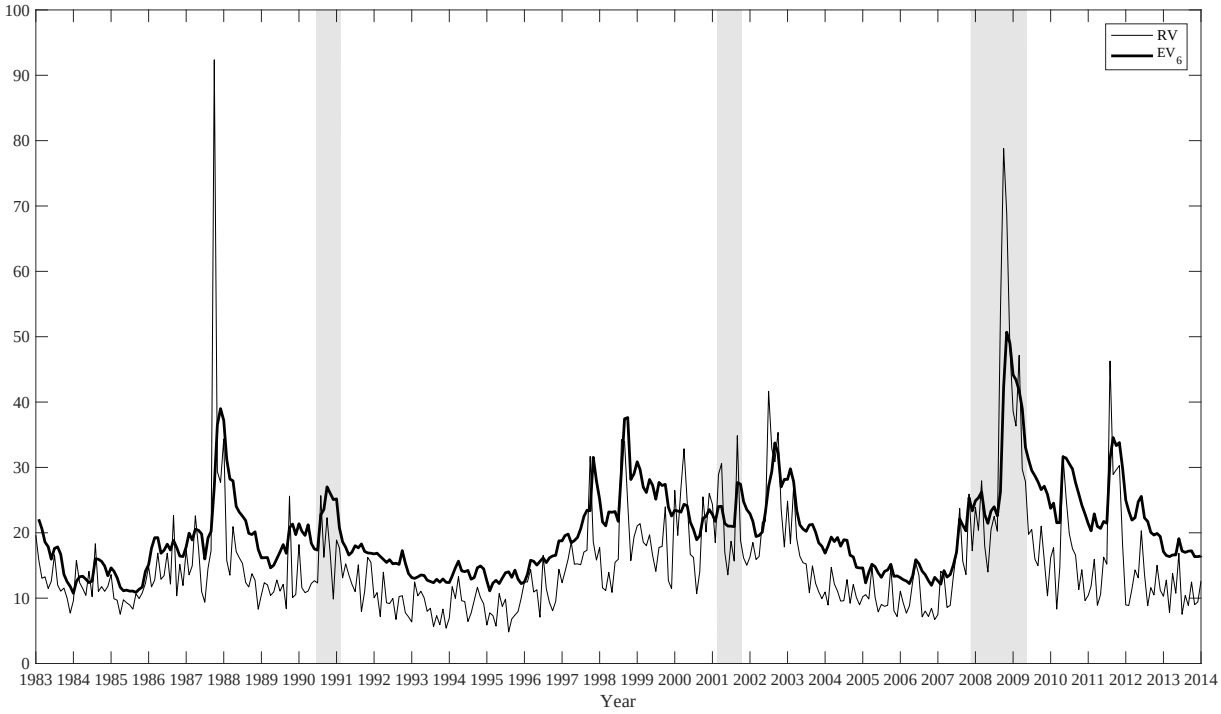
Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron, "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy*, 2004, *112* (3), 695–717.

Systems, Zeliade, "Quasi-Explicit Calibration of Gatheral's SVI model," Technical Report, Zeliade Systems 2009.

Taylor, Stephen J., Pradeep K. Yadav, and Yuanyuan Zhang, "The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks," *Journal of Banking & Finance*, 2010, *34* (4), 871–881.

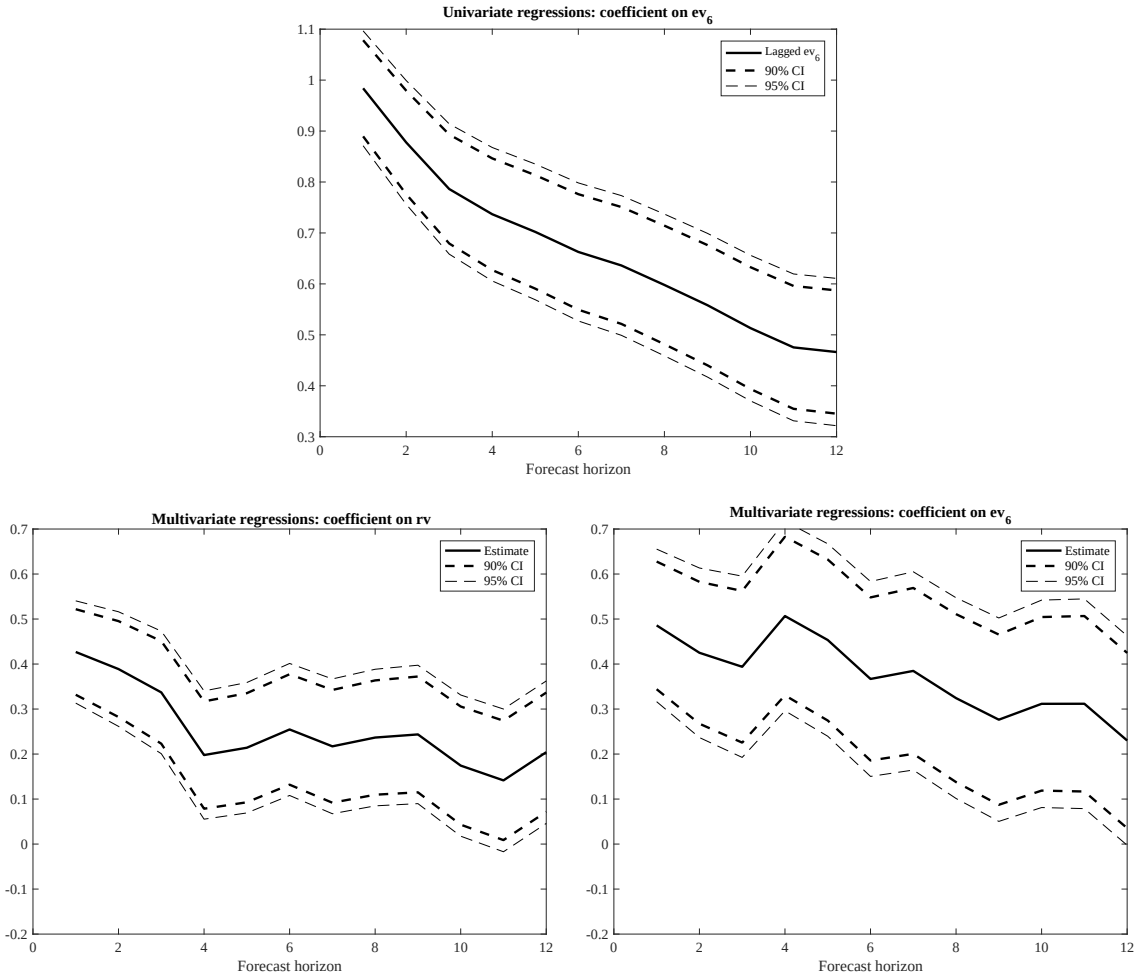
Wu, Jing Cynthia and Fan Dora Xia, "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound," 2014. Working paper.

Figure 1: Time series of realized variance and expectations



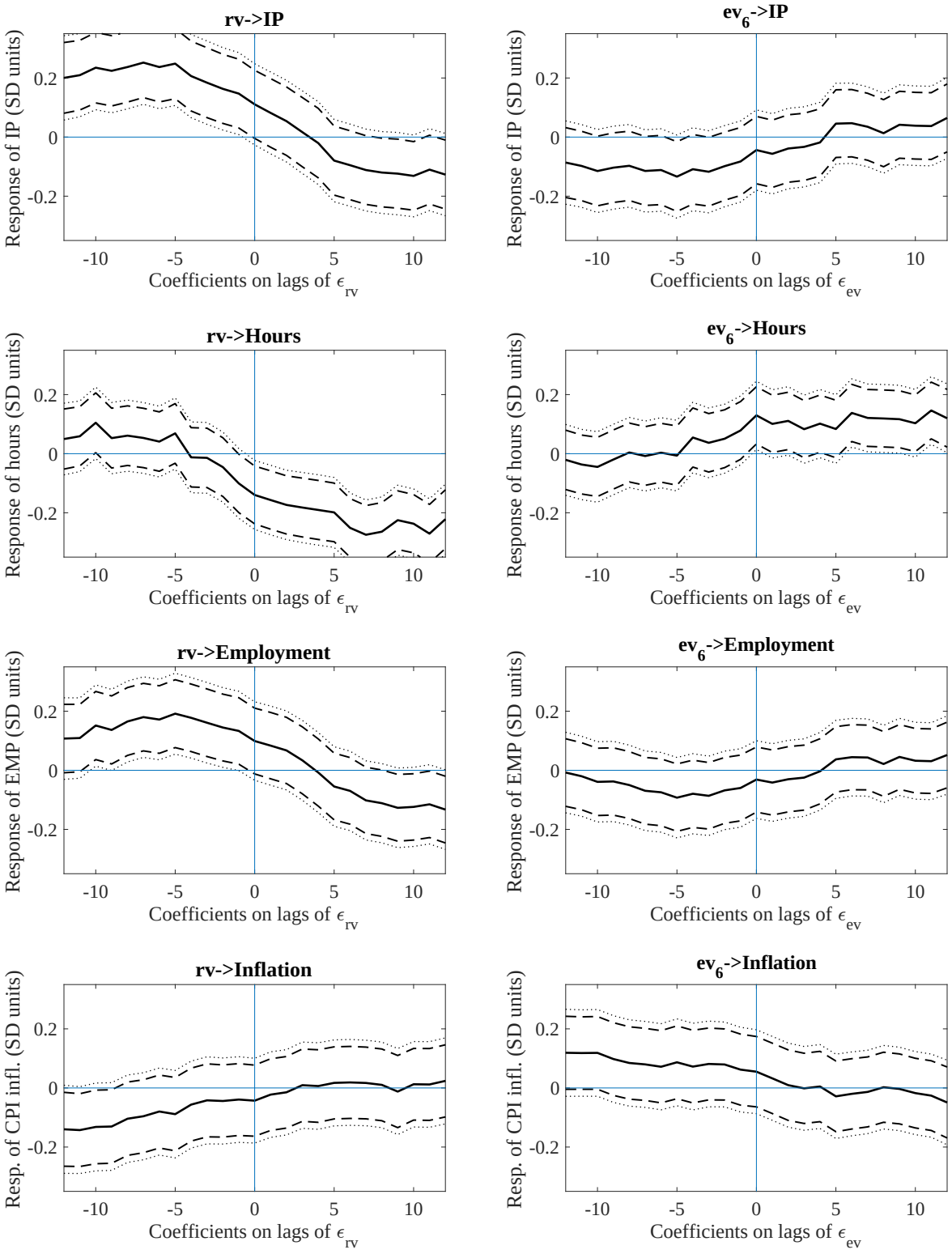
Note: Time series of realized variance (RV), and 6-month expectations (EV_6), in annualized units. Grey bars indicate NBER recessions.

Figure 2: Predictive regressions



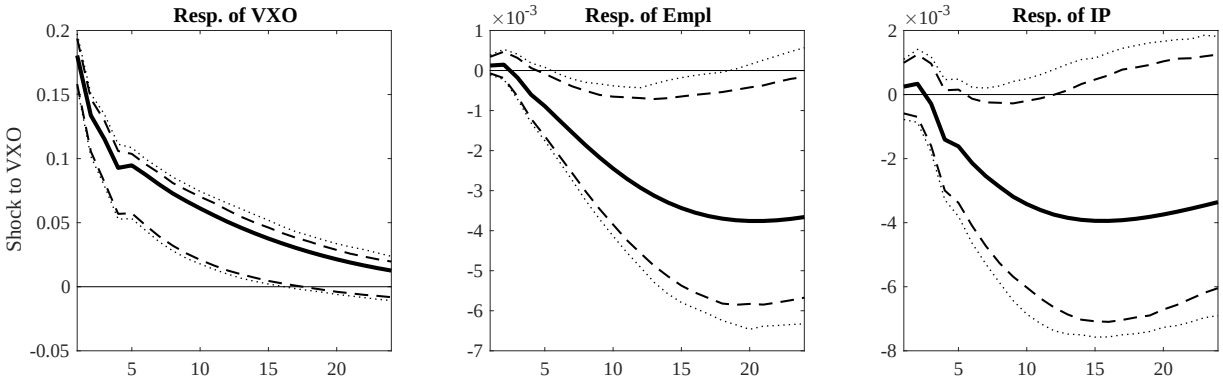
Note: Coefficients of regressions of rv on lagged rv and expected volatility, ev_6, at different lags (X axis). Top row reports coefficient of univariate regressions of rv on lagged ev_6, for different lags. Bottom row reports coefficients of multivariate regressions of rv onto lagged rv and lagged ev_6, for different lags (respectively, left and right panel).

Figure 3: Correlation of macro variables to rv and expected volatility shocks



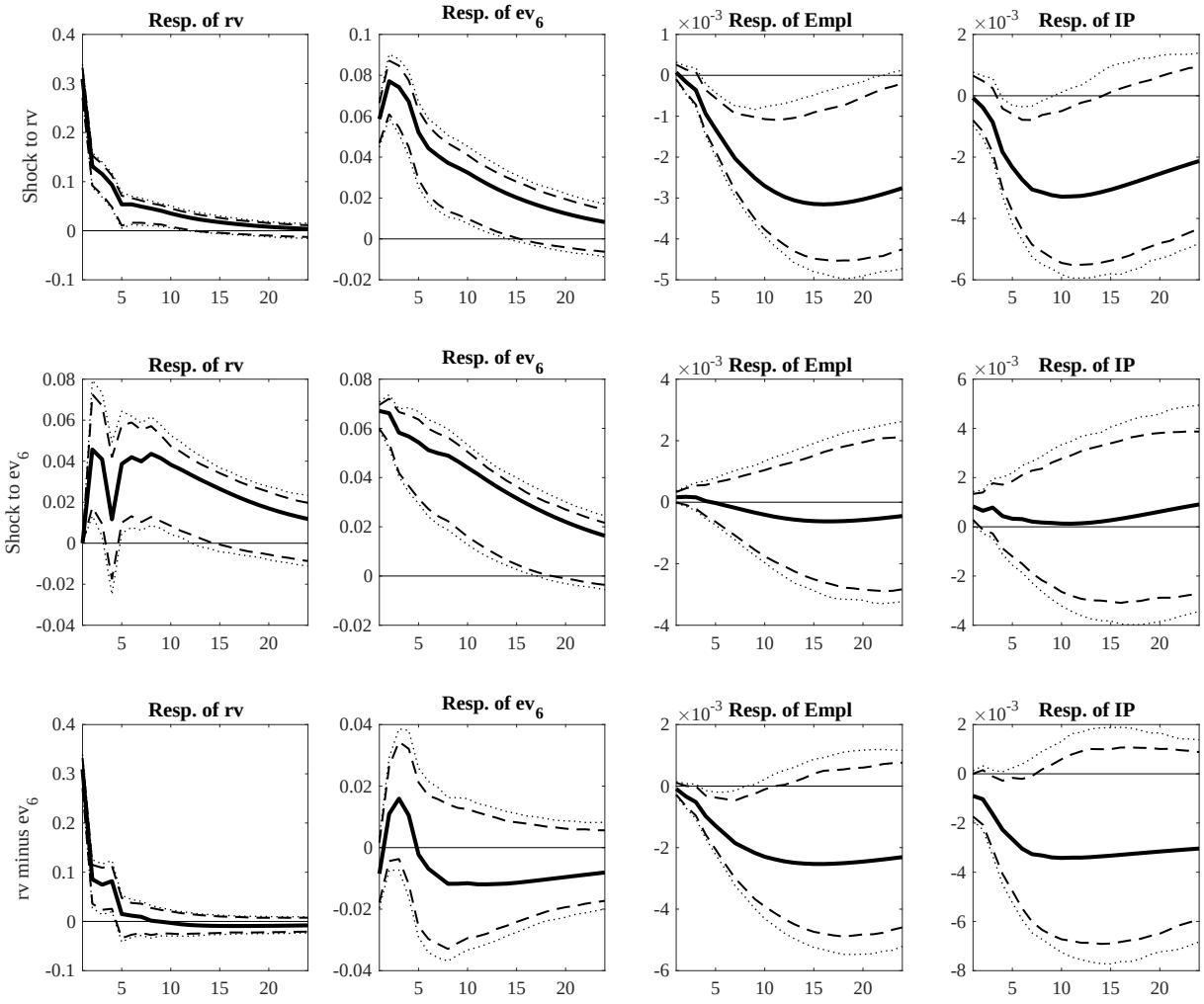
Note: The figure shows the response of macroeconomic variables to shocks to rv (left) and expected volatility, ev_6 (right), before and after the shock (lag 0 on the X axis). In particular, the solid line reports the coefficient of regressions of macroeconomic variables (Industrial Production, first row, Hours Worked, second row, Employment, third row, and Inflation, fourth row) on leads and lags of shocks to rv (left column) and expected volatility ev_6 (right column). Dashed lines indicate 90% conficence intervals, dotted lines 90% confidence intervals. The sample covers 1983 to 2014.

Figure 4: Impulse response functions from VAR (with only VXO)



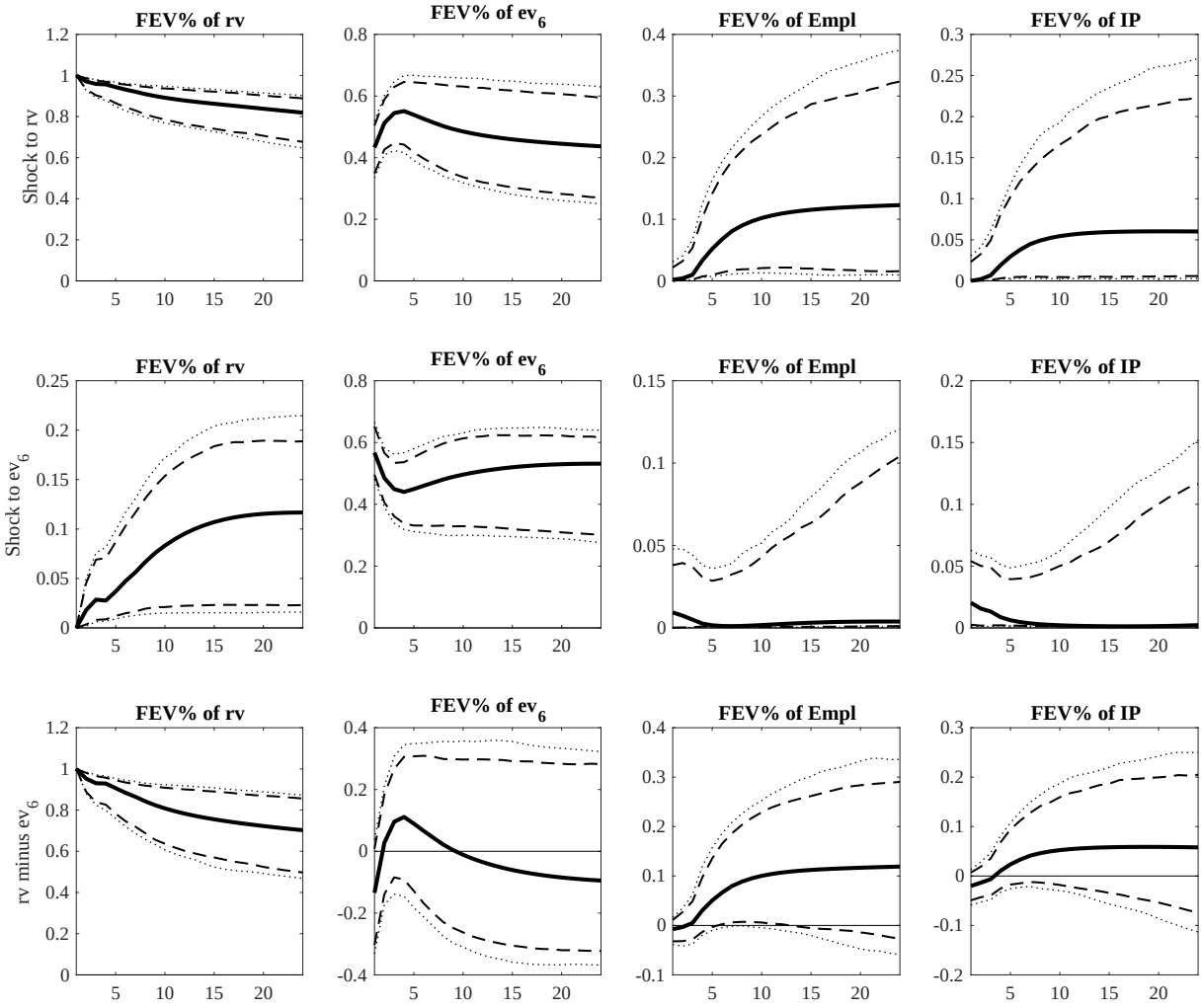
Note: The figure shows impulse response functions (with 90% and 95% CI) of volatility (measured by VXO), employment and industrial production to a shock to VXO, in a VAR with VXO, federal funds rate, log employment and log industrial production. Impulse response functions reported correspond to structural shocks with Choleski decomposition (in the order reported above). Sample covers 1986-2014. All macroeconomic series were detrended with a one-sided HP filter.

Figure 5: Impulse response functions from VAR (ordering rv first and ev_6 second)



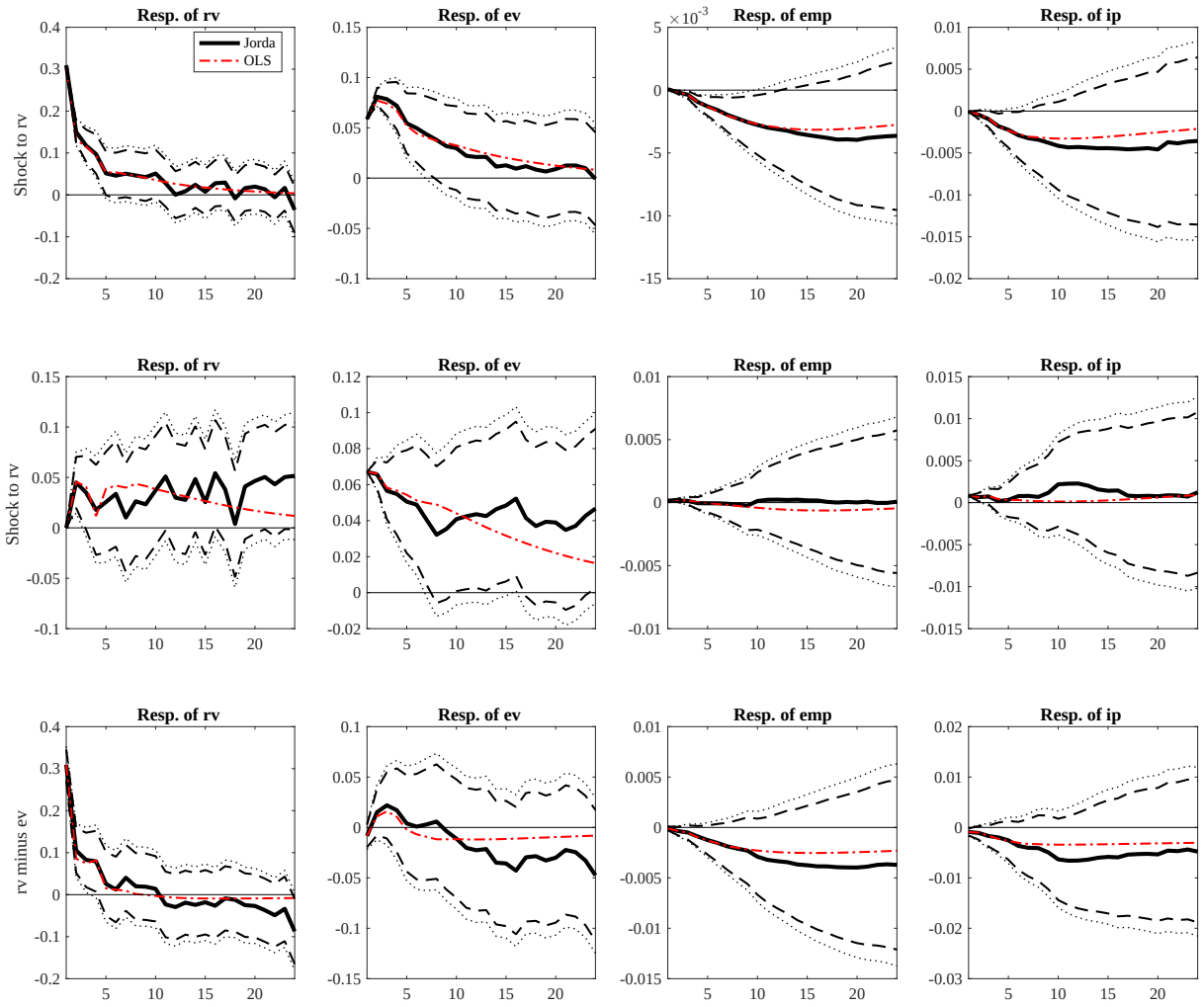
Note: The figure shows impulse response functions (with 90% and 95% CI) of rv, volatility expectations (ev_6), employment and industrial production to shocks to rv and ev_6, in a VAR with rv, ev_6, federal funds rate, log employment and log industrial production. Impulse response functions reported correspond to structural shocks with Choleski decomposition (in the order reported above). First row shows the response to a one-standard-deviation shock to rv. Second row shows the response to a one-standard-deviation shock to ev. Third row shows the difference in the response to the two shocks. Sample covers 1983-2014. All macroeconomic series were detrended with a one-sided HP filter.

Figure 6: Forecast error variance decomposition



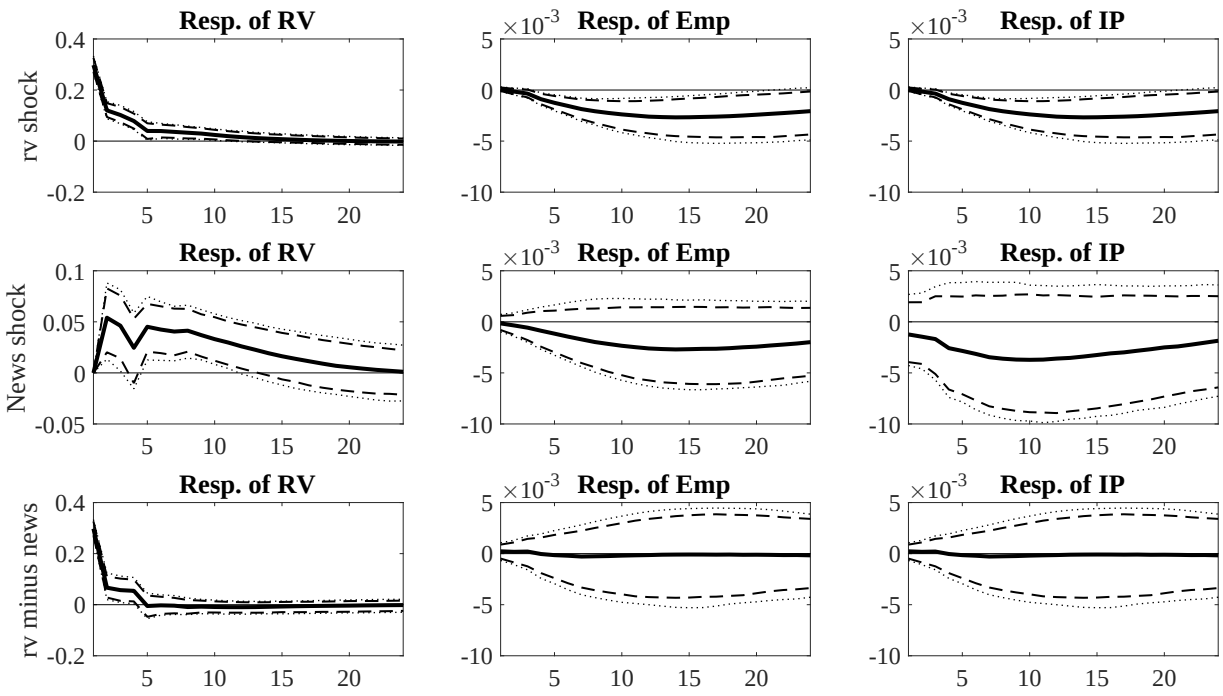
Note: The figure shows the fraction of the forecast error variance (FEV) of rv, volatility expectations (ev_6), employment and industrial production to shocks to rv and ev_6, in the VAR of figure 5. The figure also reports 90% and 95% confidence intervals. First row shows the fraction of the FEV due to the rv shock. Second row shows the fraction of FEV due to the ev shock. Third row shows the difference in the fractions of FEV due to the rv and the ev shock.

Figure 7: Jorda local projections



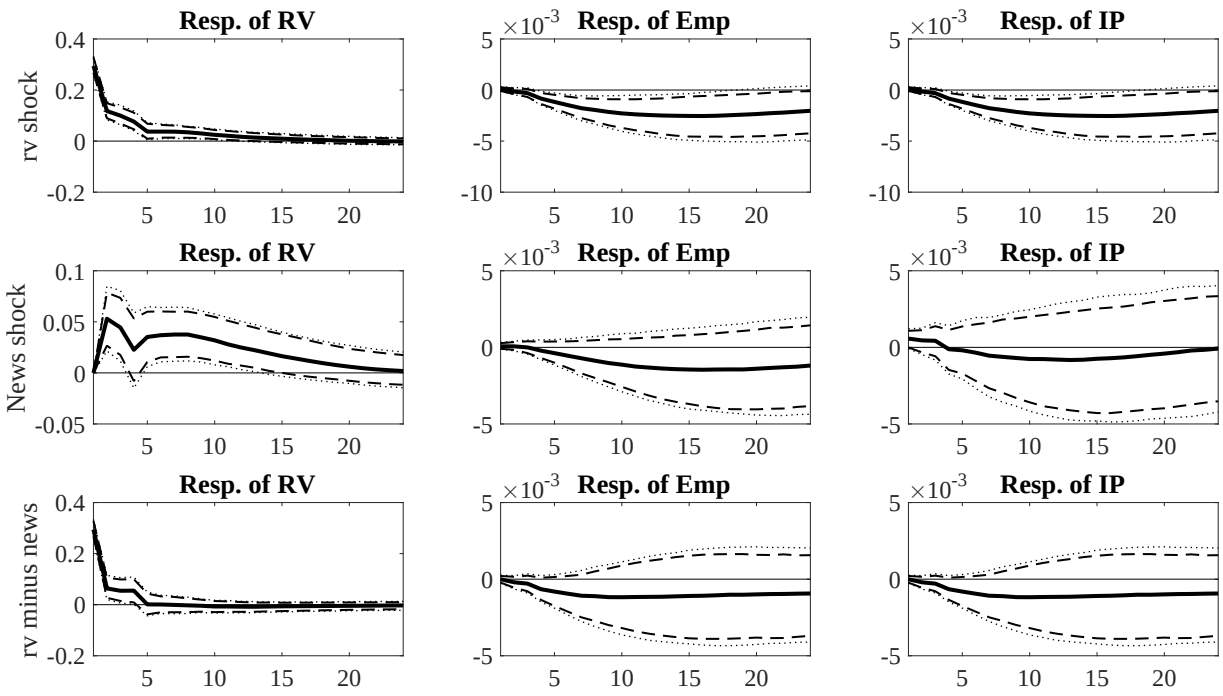
Note: The figure shows impulse response functions from the local projections (Jorda 2005), with 90% and 95% CI, of rv, volatility expectations (ev_6), employment and industrial production to shocks to rv and ev_6, in a VAR with rv, ev_6, federal funds rate, log employment and log industrial production. The responses reported correspond to structural shocks with Choleski decomposition (in the order reported above). First row shows the response to a one-standard-deviation shock to rv. Second row shows the response to a one-standard-deviation shock to ev. Third row shows the difference in the response to the two shocks. The figure also reports the VAR-based impulse-response function for comparison. Sample covers 1983-2014. All macroeconomic series were detrended with a one-sided HP filter.

Figure 8: Barsky and Sims procedure (unrestricted)



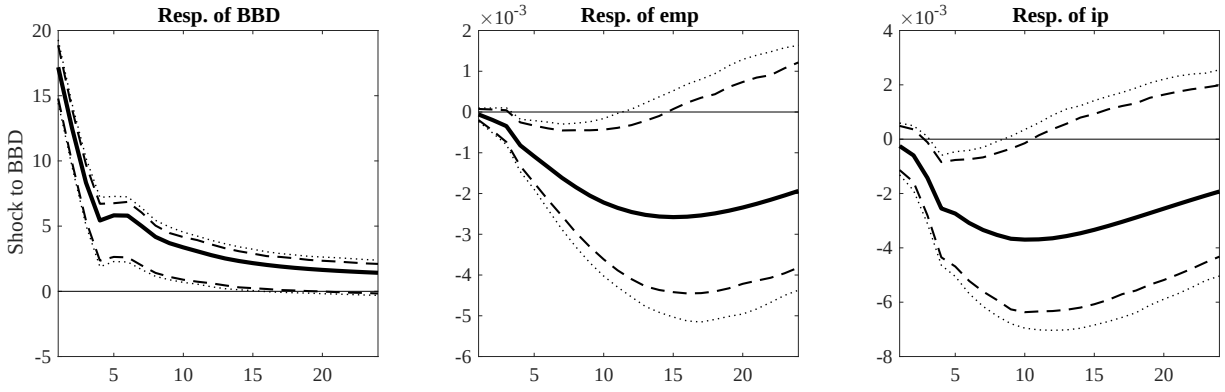
Note: The figure shows the impulse response function to shocks to realized volatility (first row) and to news about future volatility (second row). News and volatility surprises are estimated using the VAR of figure 5 using the technique of Barsky and Sims (2012). The third row reports the difference in the impulse response functions to the two shocks.

Figure 9: Barsky and Sims procedure (restricted)

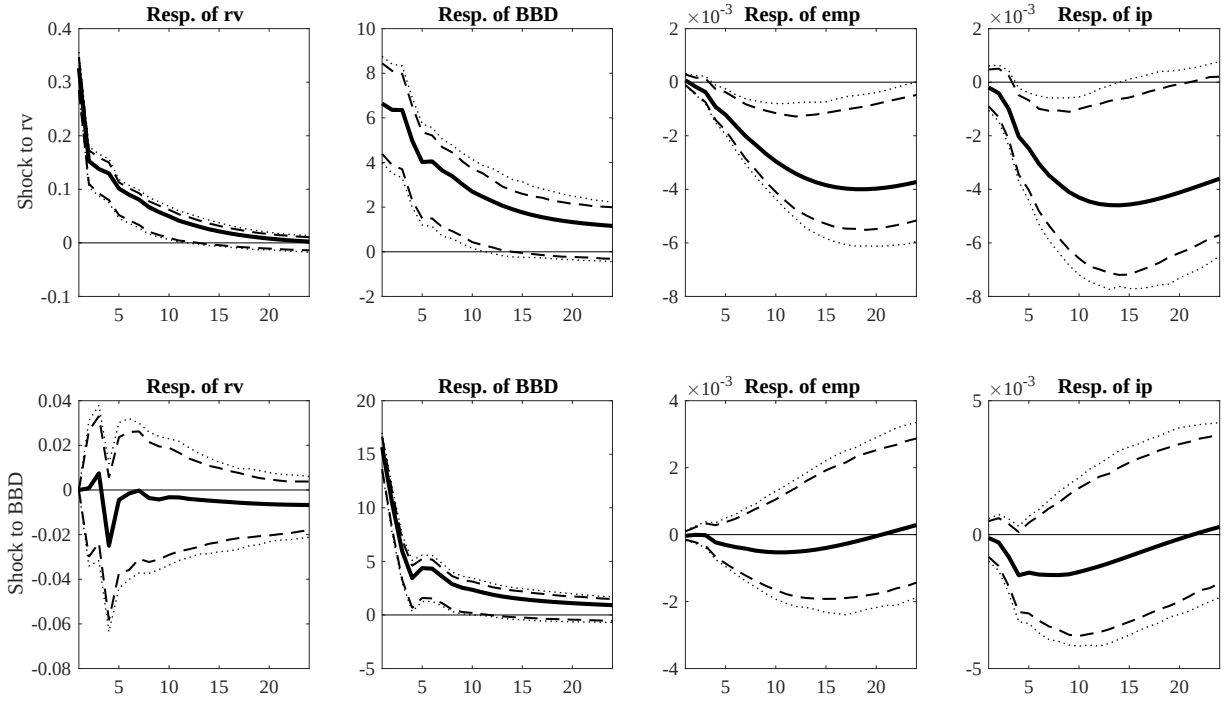


Note: The figure shows the impulse response function to shocks to realized volatility (first row) and to news about future volatility (second row). News and volatility surprises are estimated using the VAR of figure 5 using the technique of Barsky and Sims (2012), where the volatility expectation is restricted to load only on the slope and the level of the term structure of VIX claims. The third row reports the difference in the impulse response functions to the two shocks.

Figure 10: VAR with Baker, Bloom and Davis measure (BBD; 2015)



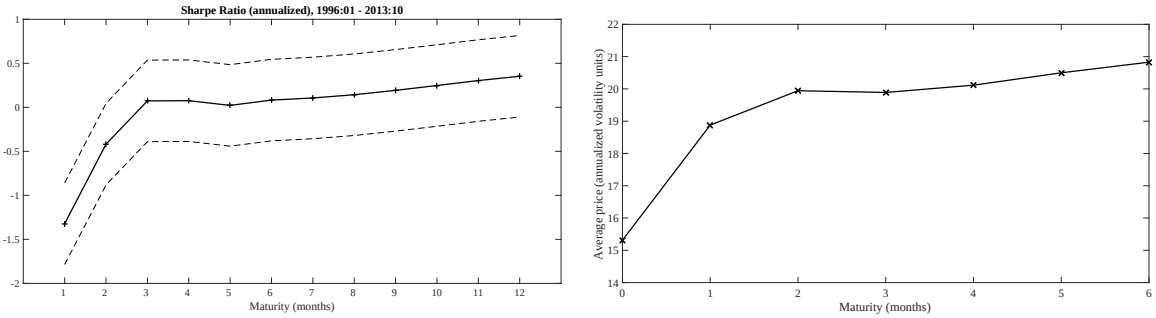
(a) VAR with only BBD



(b) VAR with both rv and BBD

Note: The figure shows impulse response functions (with 90% and 95% CI) of uncertainty (measured as in Baker, Bloom and Davis (BBD, 2015)), employment and industrial production to a shock to the volatility measures, in a VAR with volatility measures, federal funds rate, log employment and log industrial production. The top panels shows a VAR using only the BBD measure as a volatility measure, the bottom panel instead uses with rv and BBD. Impulse response functions reported correspond to structural shocks with Choleski decomposition (in the order reported above). Sample covers 1985-2014. All macroeconomic series were detrended with a one-sided HP filter.

Figure 11: Forward variance claims: returns and prices



(a) Sharpe ratios

(b) Average prices

Note: Panel A shows the annualized Sharpe ratio for the forward variance claims, constructed using Variance Swaps. The returns are calculated assuming that the investment in an n-month variance claim is rolled over each month. Dotted lines represent 95% confidence intervals. All tests for the difference in Sharpe ratio between the 1-month variance swap and any other maturity confirm that they are statistically different with a p-value of 0.03 (for the second month) and < 0.01 (for all other maturities). The sample used is 1996-2013. For more information on the data sources, see Dew-Becker et al. (2015). Panel B shows the average prices of forward variance claims of different maturity, constructed from option prices, for the period 1983–2014. All prices are reported in annualized volatility terms. Maturity zero corresponds to average realized volatility.

Table 1: Descriptive statistics

Panel A: Descriptive statistics	Mean	Std.	Skewness
RV	15.31	9.34	3.62
EV_6	19.99	6.60	1.36

Panel B: Correlations	rv	ev_6	Unempl.	Cap. Ut.	S&P 500 ret.
rv	1.00	0.78	0.17	-0.33	-0.20
ev_6	0.78	1.00	0.06	-0.38	-0.14
Unemployment	0.17	0.06	1.00	-0.70	0.10
Capacity Utilization	-0.33	-0.38	-0.70	1.00	-0.06
S&P 500 return	-0.20	-0.14	0.10	-0.06	1.00

Panel C: Decomposition of $V(rv)$	$V(rv_{t+1} - E_t^1)$	$V(E_{t+5}^1 - E_t^6)$	$V(E_t^6)$
	48%	31%	21%

Note: The table reports various statistics on realized volatility, forwards and their relationship. Panel A reports the mean, standard deviation and skewness of realized volatility and expected volatility ev_6 (the market expectation of realized variance up to 6 months in the future). Panel B reports the correlations between those variables and with macroeconomic and financial variables: unemployment, capacity utilization, and the S&P 500 return. Panel C computes a variance decomposition for the variance of rv, $V(rv)$, into the surprise component $V(rv_{t+1} - E_t^1)$, the volatility news 1 to 5 months ahead, $V(E_{t+5}^1 - E_t^6)$, and the volatility news 6 months ahead, $V(E_t^6)$. The decomposition is obtained using the 1-month and 6-month market expectation, ev_1 and ev_6, to predict volatility at the 1-month and 6-month horizon. Sample period is 1983-2014.

Table 2: Predictability of rv

Predictors	(1)	(2)	(3)
ev_6	0.82***	0.45***	0.46***
	(0.11)	(0.11)	(0.12)
rv		0.32***	0.32***
		(0.08)	(0.09)
FFR			0.01
			(0.02)
$Empl.$			-0.41
			(1.43)
IP			0.36
			(1.00)
N	377	377	377
Adj. R2	0.38	0.43	0.42

Note: The figure reports the results of linear predictive regressions of 6-month rv on various macroeconomic variables, with Hansen-Hodrick standard errors (6 lags).

Table 3: Cumulative coefficients on rv and ev lags

Panel A: Full Sample		VAR specification	rv	ev	diff.	p-value
Effect on:	Empl.	Empl., rv, ev	-0.11	0.07	-0.18	0.14
	IP	IP, rv, ev	-0.20	0.14	-0.34	0.39
	Empl.	IP, Empl., rv, ev	-0.10	0.06	-0.16	0.20
	IP	IP, Empl., rv, ev	-0.21	0.15	-0.36	0.41

Panel B: 1988-2006		VAR specification	rv	ev	diff.	p-value
Effect on:	Empl.	Empl., rv, ev	-0.15	0.13	-0.29	0.16
	IP	IP, rv, ev	-0.25	0.31	-0.56	0.37
	Empl.	IP, Empl., rv, ev	-0.16	0.03	-0.19	0.38
	IP	IP, Empl., rv, ev	-0.11	0.60	-0.71	0.29

Note: The figure reports the sum of the coefficients of IP and Employment on the lags of rv and ev in VAR specifications that include rv, ev, as well as IP or Employment (or both). Columns 1 and 2 report the sums of the coefficients on the lags of rv and ev, respectively; column 3 reports the difference in these sums, and column 4 reports p-values for tests of this difference. Panel A performs the analysis on the full sample (1983-2014), while Panel B focuses on the sample 1988-2006.

Table 4: Skewness

Panel A: real economic activity	Monthly	Quarterly	Start of sample (year)
Employment	-0.41	-0.41	1948
Capacity Utilization	-1.02	-1.30	1967
IP	0.17	-0.16	1948
IP, starting 1960	-0.93	-1.28	1960
Y		-0.11	1947
C		-0.28	1947
I		-0.03	1947

Panel B: skewness of S&P 500 monthly returns	
Implied (since 1990)	-1.81
Realized (since 1926)	0.36
Realized (since 1948)	-0.42
Realized (since 1990)	-0.61

Note: Panel A reports the skewness of changes of employment, capacity utilization, industrial production (beginning both in 1948 and in 1960), GDP, consumption and investments. The first column reports the skewness of monthly changes, the second column the skewness of quarterly changes. Panel B reports the realized skewness of S&P 500 monthly returns in different periods, as well as the implied skewness computed by the CBOE using option prices.

A.1 Construction of model-free implied volatility, EV_n

In this section we describe the details of the procedure we use to construct model-free implied volatility (market expectations of future volatility) at different horizons, starting from our dataset of end-of-day prices for American options on S&P 500 futures from the CME.

A.1.1 Main steps of construction of EV

A first step in constructing the model-free implied volatility is to obtain implied volatilities corresponding to the observed option prices. We do so using a binomial model.[1] For the most recent years, CME itself provides the implied volatility together with the option price. For this part of the sample, the IV we estimate with the binomial model and the CME's IV have a correlation of 99%, which provides an external validation on our implementation of the binomial model.

Once we have estimated these implied volatilities, we could in theory simply invert them to yield implied prices of European options on forwards. These can then be used to compute EV directly as described in equation (10).

In practice, however, an extra step is required before inverting for the European option prices and integrating to obtain the model-free implied volatility. The model-free implied volatility defined in equation (10) depends on the integral of option prices over *all strikes*, but option prices are only observed at discrete strikes. We are therefore forced to interpolate option prices between available strikes and also extrapolate beyond the bounds of observed strikes.[2] Following the literature, we fit a parametric model to the Black–Scholes implied volatilities of the options and use the model to then interpolate and extrapolate across all strikes (see, for example, Jiang and Tian (2007), Carr and Wu (2009), Taylor, Yadav, and Zhang (2010), and references therein). Only after this extra interpolation-extrapolation step, the fitted implied volatilities are then inverted to yield option prices and compute EV according to equation (10). To interpolate and extrapolate the implied volatility curve, we use the SVI (stochastic volatility inspired) model of Gatheral and Jacquier (2014).

In the next sections, we describe in more detail the interpolation-extrapolation step of the procedure (SVI fitting) as well as our construction of EV after fitting the SVI curve. Finally, we report a description of the data we use and some examples and diagnostics on the SVI fitting method.

A.1.2 SVI interpolation: theory

There are numerous methods for fitting implied volatilities across strikes. Homescu (2011) provides a thorough review. We obtained the most success using Gatheral's SVI model (see Gatheral and Jacquier 2014). SVI is widely used in financial institutions because it is parsimonious but also known

[1] See for example Broadie and Detemple (1996) and Bakshi, Kapadia, and Madan (2003), among others.

[2] See Jiang and Tian (2007) for a discussion of biases arising from the failure to interpolate and extrapolate.

to approximate well the behavior of implied volatility in fully specified option pricing models (e.g. Gatheral and Jacquier (2011)); SVI also satisfies the limiting results for implied volatilities at very high and low strikes in Lee (2004), and, importantly, ensures that no-arbitrage conditions are not violated.

The SVI model simply assumes a hyperbolic relationship between implied variance (the square of the Black–Scholes implied volatility) and the log moneyness of the option, k (log strike/forward price).

$$\sigma_{BS}^2(k) = a + b\left(\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}\right)$$

where $\sigma_{BS}^2(k)$ is the implied variance under the Black–Scholes model at log moneyness k. SVI has five parameters: a, b, ρ, m, and σ. The parameter ρ controls asymmetry in the variances across strikes. Because the behavior of options at high strikes has minimal impact on the calculation of model-free implied volatilities, and because we generally observe few strikes far above the spot, we set $\rho = 0$ (in simulations with calculating the VIX for the S&P 500 – for which we observe a wide range of options – we have found that including or excluding ρ has minimal impact on the result).

We fit the parameters of SVI by minimizing the sum of squared fitting errors for the observed implied volatilities. Because the fitted values are non-linear in the parameters, the optimization must be performed numerically. We follow the methodology in Zeliade (2009) to analytically concentrate a and b out of the optimization. We then only need to optimize numerically over σ and m (as mentioned above, we set $\rho = 0$). We optimize with a grid search over $\sigma \times m = [0.001, 10] \times [-1, 1]$ followed by the simplex algorithm.

For many date/firm/maturity triplets, we do not have a sufficient number of contract observations to fit the implied volatility curve (i.e. sometimes fewer than four). We therefore include strike/implied volatility data from the two neighboring maturities and dates in the estimation. The parameters of SVI are obtained by minimizing squared fitting errors. We reweight the observations from the neighboring dates and maturities so that they carry the same amount of weight as the observations from the date and maturity of interest. Adding data in this way encourages smoothness in the estimates over time and across maturities but it does not induce a systematic upward or downward bias. We drop all date/firm/maturity triplets for which we have fewer than four total options with $k < 0$ or fewer than two options at the actual date/firm/maturity (i.e. ignoring the data from the neighboring dates and maturities).

When we estimate the parameters of the SVI model, we impose conditions that guarantee the absence of arbitrage. In particular, we assume that $b \leq \frac{4}{(1+|\rho|)T}$, which when we assume $\rho = 0$, simplifies to $b \leq \frac{4}{T}$. We also assume that $\sigma > 0.0001$ in order to ensure that the estimation is well defined. Those conditions do not necessarily guarantee, though, that the integral determining the model-free implied volatility is convergent (the absence of arbitrage implies that a risk-neutral probability density exists – it does not guarantee that it has a finite variance). We therefore eliminate observations where the integral determining the model-free implied volatility fails to converge numerically. Specifically, we eliminate observations where the argument of the integral

does not approach zero as the log strike rises above two standard deviations from the spot or falls more than five standard deviations below the strike (measured based on the at-the-money implied volatility).

A.1.3 Construction of EV from the SVI fitted curve

After fitting the SVI curve for each date and maturity, we compute the integral in equation (10) numerically, over a range of strikes from -5 to +2 standard deviations away from the spot price.[3] We then have EV for every firm/date/maturity observation. The model-free implied volatilities are then interpolated (but not extrapolated) to construct EV at maturities from 1–6 months for each firm/date pair.

A.1.4 Data description and diagnostics of SVI fitting

Our dataset consists of 2.3 million end-of-day prices for all American options on S&P 500 futures from the CME.

When more than one option (e.g. a call and a put) is available at any strike, we compute IV at that strike as the average of the observed IVs. We keep only IVs greater than zero, at maturities higher than 9 days and lower than 2 years, for a total of 1.9 million IVs. The number of available options has increased over time, as demonstrated by Figure A.2 (top panel), which plots the number of options available for EV estimation in each year.

The maturity structure of observed options has also expanded over time, with options being introduced at higher maturities and for more intermediate maturities. Figure A.1 (top panel) reports the cross-sectional distribution of available maturities in each year to estimate the term structure of the model-free implied volatility. The average maturity of available options over our sample was 4 months, and was relatively stable. The maximum maturity observed ranged from 9 to 24 months and varied substantially over time.

Crucial to compute the model-free implied volatility is the availability of IVs at low strikes, since options with low strikes receive a high weight in the construction of EV. The bottom panel of Figure A.1 reports the minimum observed strike year by year, in standard deviations below the spot price. In particular, for each day we computed the minimum available strike price, and the figure plots the average of these minimum strike price across all days in each year; this ensures that the number reported does not simply reflect outlier strikes that only appear for small parts of each year.

Figure A.1 shows that in the early part of our sample, we can typically observe options with strikes around 2 standard deviations below the spot price; this number increases to around 2.5 towards the end of the sample.

[3] In general this range of strikes is sufficient to calculate EV. However, the model-free implied volatility technically involves an integral over the entire positive real line. Our calculation is thus literally a calculation of Andersen and Bondarenko's (2007) corridor implied volatility. We use this fact also when calculating realized volatility.

These figures show that while the number of options was significantly smaller at the beginning of the sample (1983), the maturities observed and the strikes observed did not change dramatically over time.

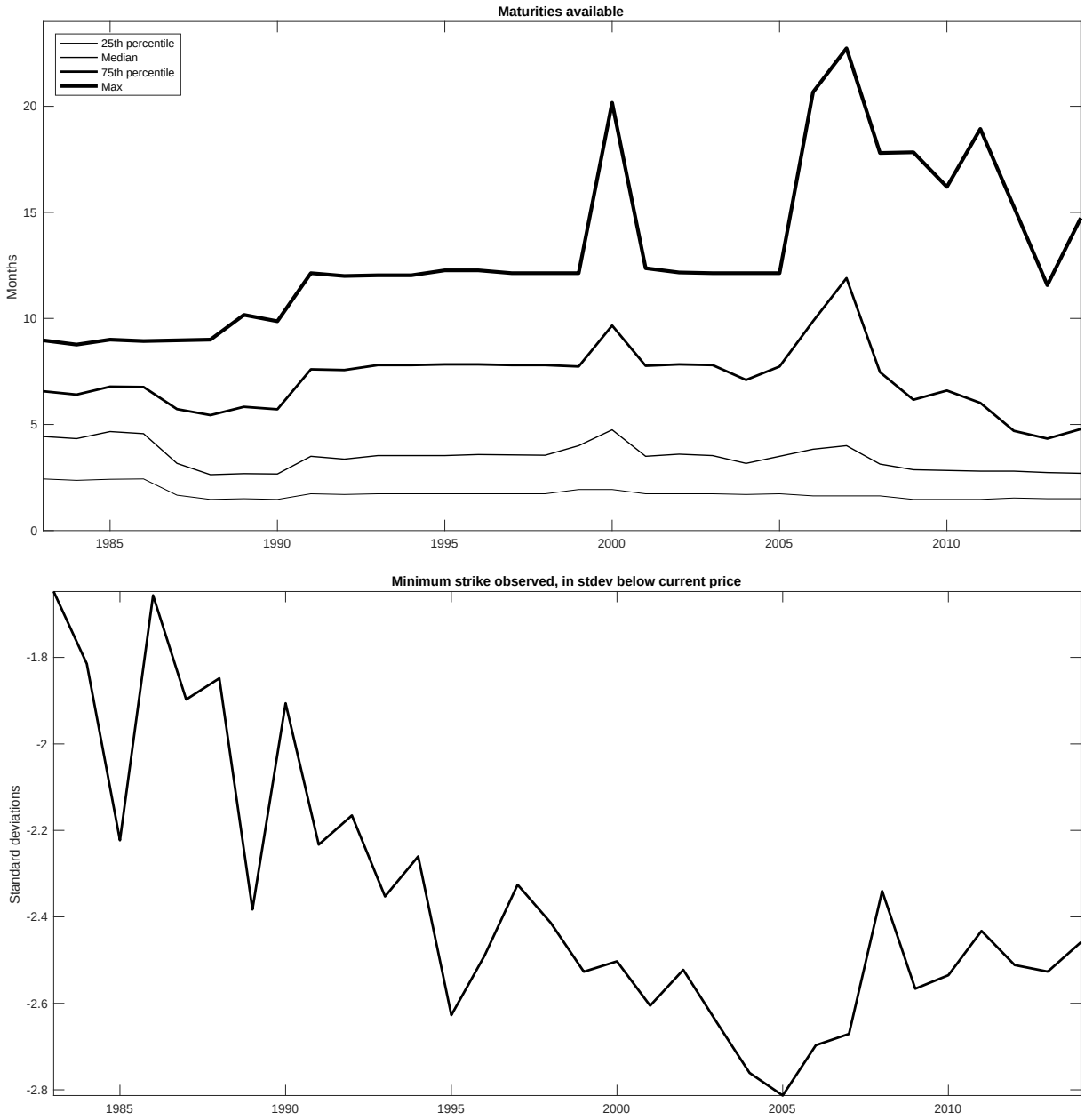
Figure A.3 shows an example of the SVI fitting procedure for a specific day in the early part of our sample (November 7th 1985). Each panel in the figure corresponds to a different maturity. On that day, we observe options at three different maturities, of approximately 1, 4, and 8 months. In each panel, the x's represent observed IVs at different values of log moneyness k. The line is the fitted SVI curve, that shows both the interpolation and the extrapolation obtained from the model.

Figure A.4 repeats the exercise in the later part of our sample (Nov. 1st 2006), where many more maturities and strikes are available.

Both figures show that the SVI model fits the observed variances extremely well. The bottom panel of Figure A.2 shows the average relative pricing error for the SVI model in absolute value. The graph shows that the typical pricing error for most of the sample is around 0.02, meaning that the SVI deviates from the observed IV by around 2% on average. Only in the very first years (up to 1985) pricing errors are larger, but still only around 10% of the observed IV.

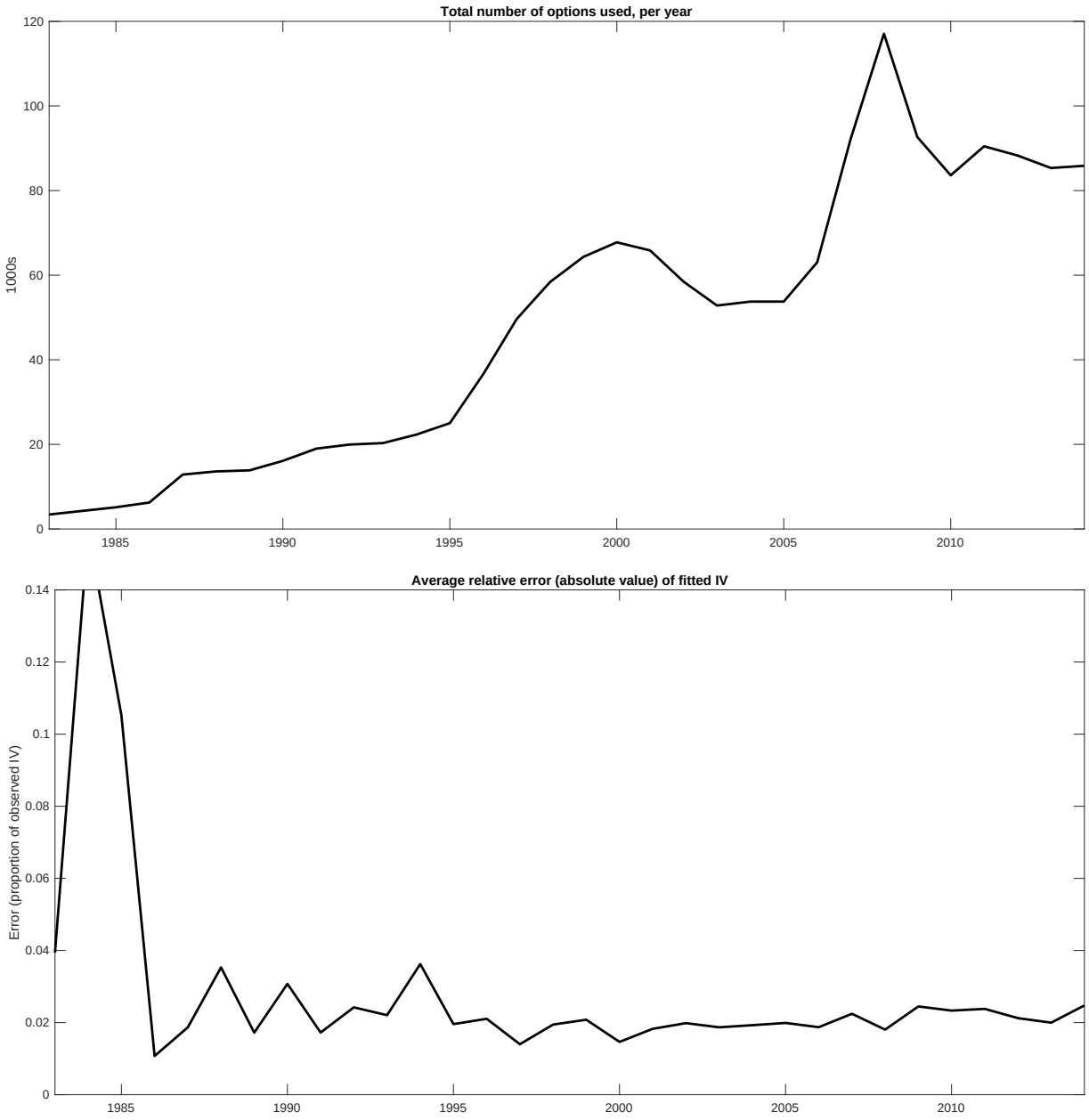
Overall, the evidence in this section shows that our observed option sample since 1983 has been relatively stable along the main dimensions that matter for our analysis – maturity structure, strikes observed, and goodness of fit of the SVI model.

Figure A.1: Maturities and strikes in the CME dataset



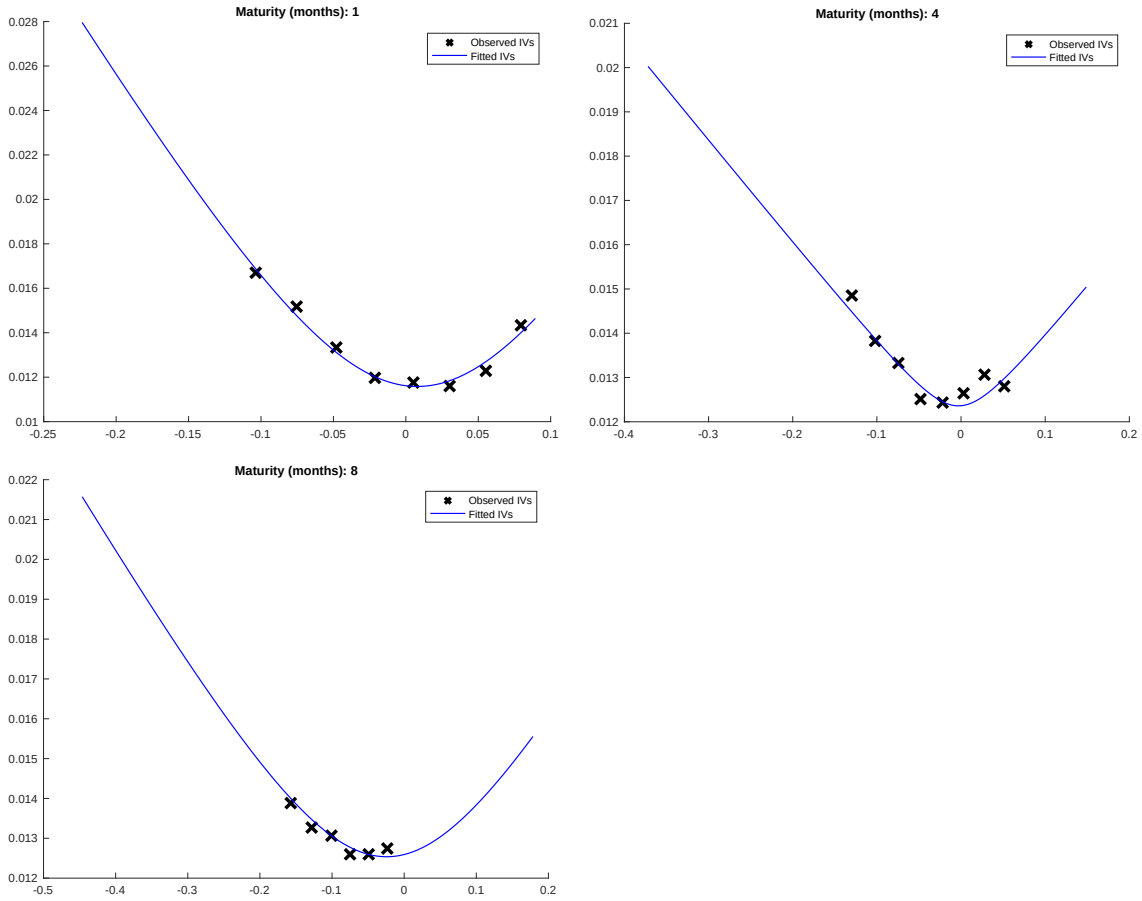
Note: Top panel reports the distribution of maturities of options used to compute the VIX in each year, in months. Bottom panel reports the average minimum strike in each year, in standard deviations below the forward price. The number is obtained by computing the minimum observed strike in each date and at each maturity (in standard deviations below the forward price), and then averaging it within each year to minimize the effect of outliers.

Figure A.2: Number of options to construct the VIX and pricing errors



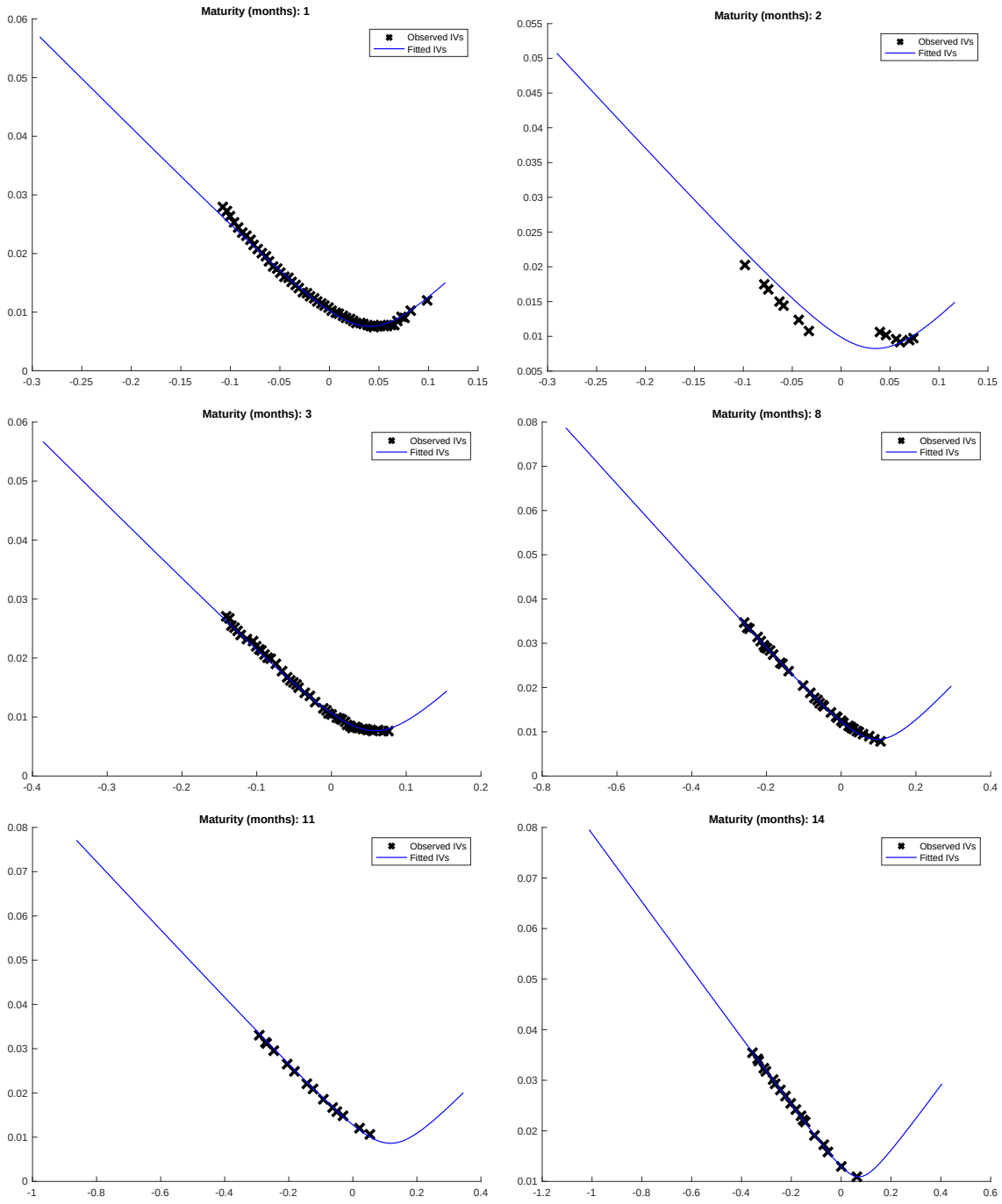
Note: Top panel reports the number of options used to compute the VIX in each year, in thousands. Bottom panel reports the average absolute value of the pricing error of the SVI fitted line relative to the observed implied variances, in proportional terms (i.e. 0.02 means absolute value of the pricing error is 2% of the observed implied variance).

Figure A.3: SVI fit: 11/7/1985



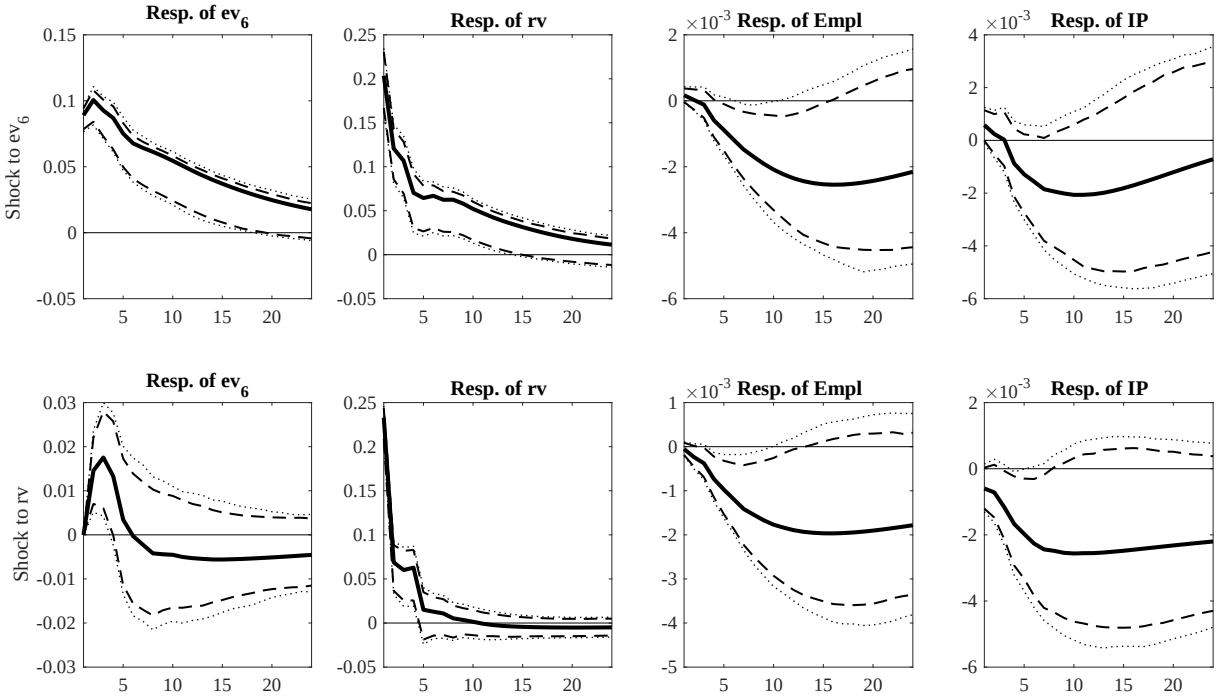
Note: Fitted implied variance curve on 11/7/1987, for the three available maturities. X axis is the difference in log strike and log forward price. x's correspond to the observed implied variances, and the line is the fitted SVI curve.

Figure A.4: SVI fit: 11/1/2006



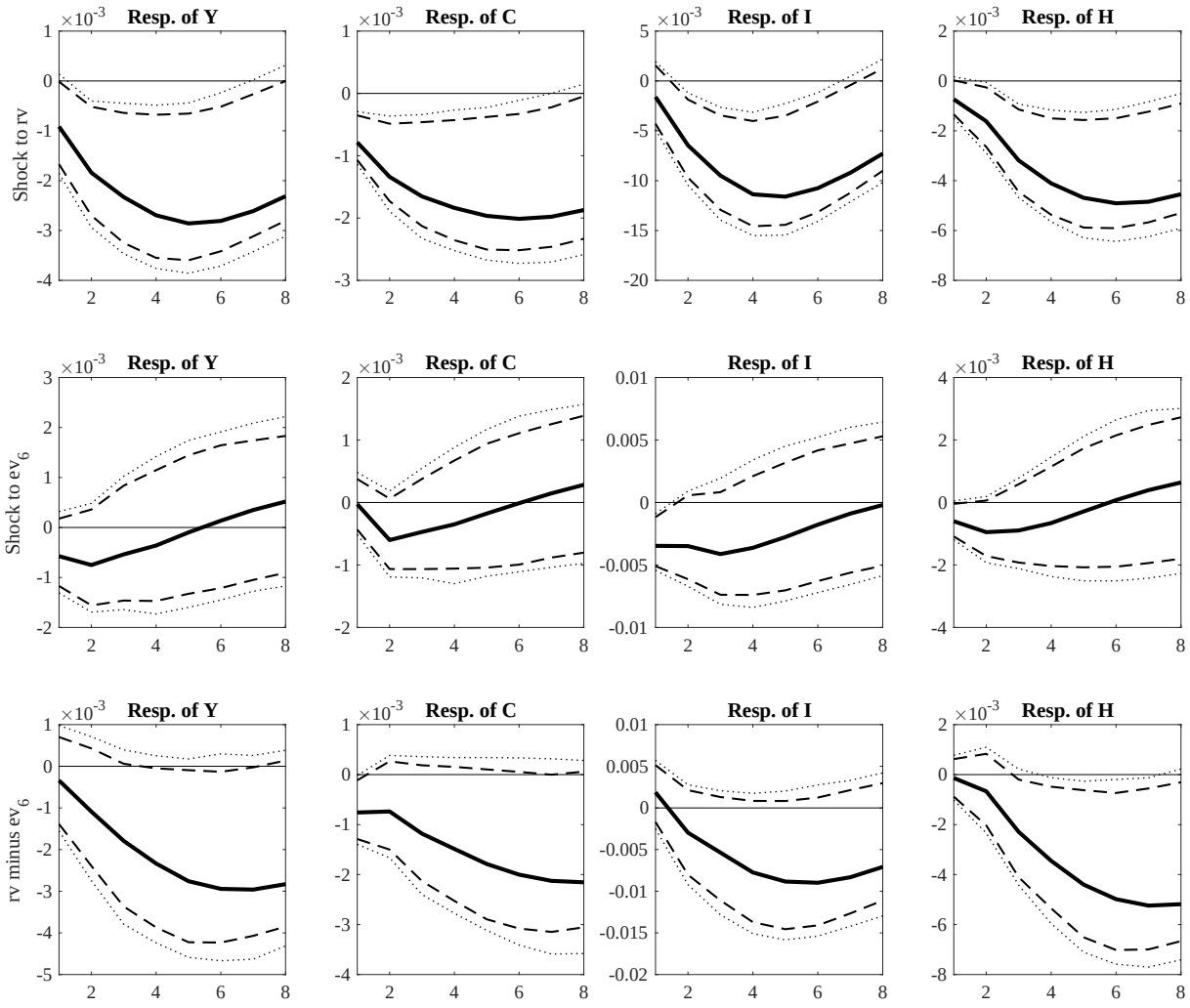
Note: Fitted implied variance curve on 11/1/2006, for the three available maturities. X axis is the difference in log strike and log forward price. x's correspond to the observed implied variances, and the line is the fitted SVI curve. On 11/1/2006 also a maturity of 5 months was available (not plotted for reasons of space).

Figure A.5: Impulse response functions from VAR (ordering *ev* first and *rv* second)



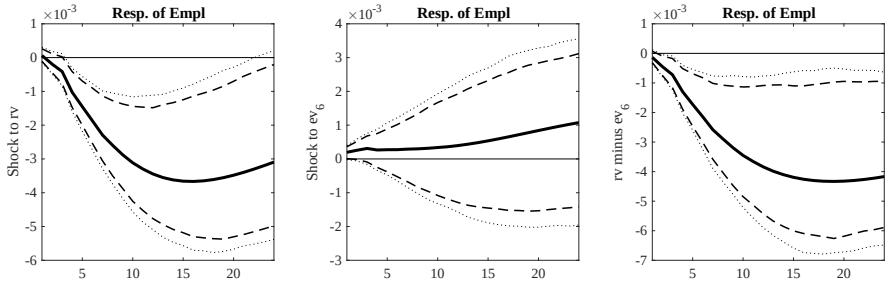
Note: The figure shows impulse response functions (with 90% and 95% CI) of *rv*, volatility expectations *ev*, employment and industrial production to shocks to *ev* and *rv*, in a VAR with *rv*, *ev*, federal funds rate, log employment and log industrial production. Impulse response functions reported correspond to structural shocks with Choleski decomposition (in the order reported above). Sample covers 1983-2014. All macroeconomic series were detrended with a one-sided HP filter.

Figure A.6: Quarterly VAR specification

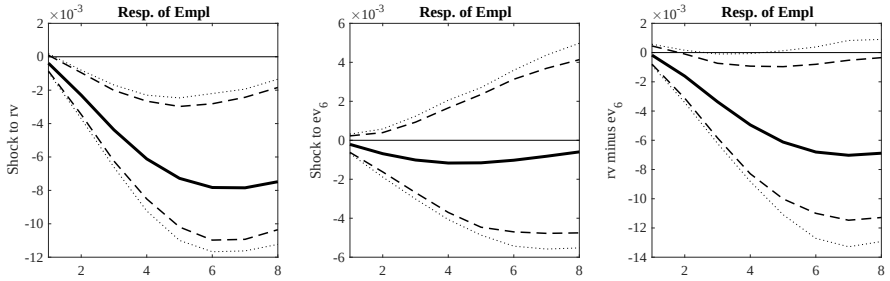


Note: The figure shows impulse response functions (with 90% and 95% CI) from a VAR with rv, volatility expectations ev, and the macroeconomic series from Basu and Bundick (2015): GDP (Y), consumption (C), investment (I), hours (H), the GDP deflator (DEF), M2 and the FFR. The figure reports IRF of Y, C, I, H to shocks to rv and ev. Impulse response functions reported correspond to structural shocks with Choleski decomposition (in the order reported above). First row shows the response to a one-standard-deviation shock to rv. Second row shows the response to a one-standard-deviation shock to ev. Third row shows the difference in the response to the two shocks. Sample covers 1983-2014.

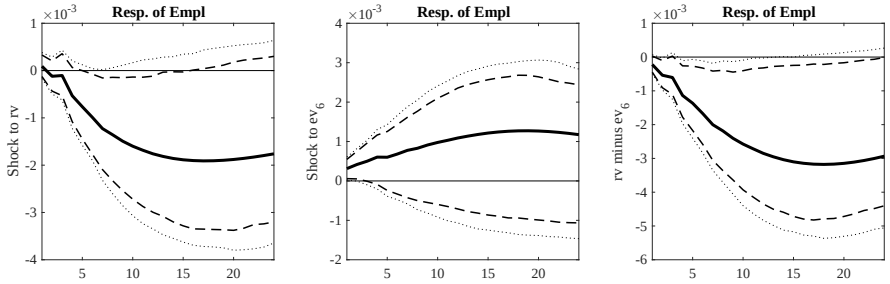
Figure A.7: Robustness (I): response of Employment to RV and expectations shocks across specifications



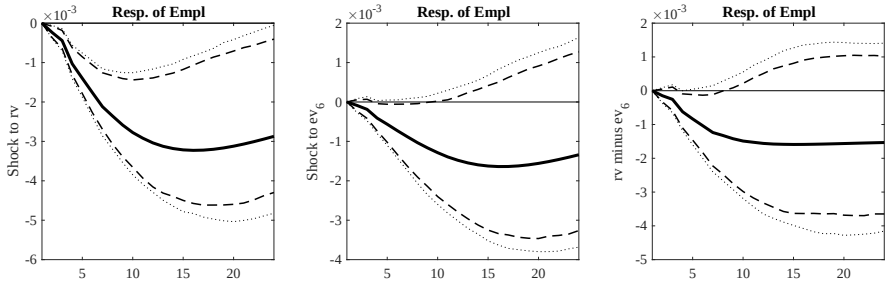
(a) Without detrending the macroeconomic time series



(b) VAR estimated at the quarterly frequency



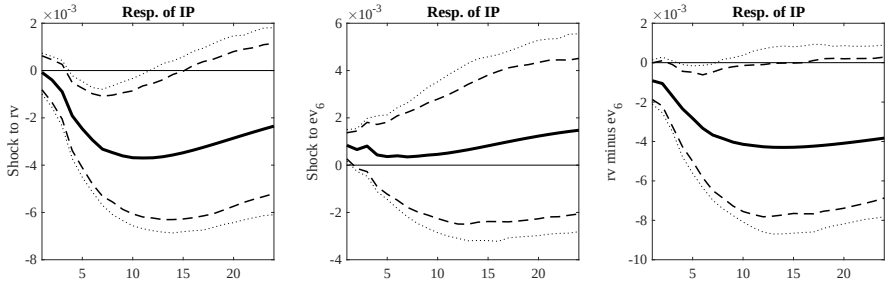
(c) Subperiod 1988-2006 (excluding 1987 crash and financial crisis)



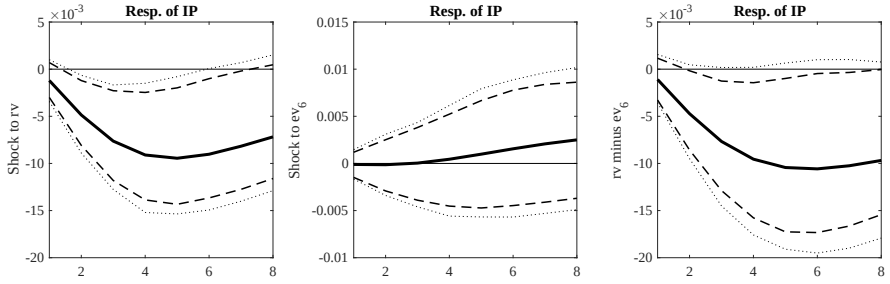
(d) Ordering RV and expectations last in the VAR

Note: The Figure reports the response of employment to RV shocks (left panels) and volatility expectations (middle panels) in different specification, with the difference in the right panel. Each row of the figure corresponds to a different model specification. Row (a) does not detrend the macroeconomic time series. Row (b) repeats the exercise using quarterly, rather than monthly, data. Row (c) estimates the VAR in the subsample 1988-2006, which excludes both RV peaks (1987 crash and financial crisis). Row (d) orders the RV shock second to last and the expectation shock last in the VAR.

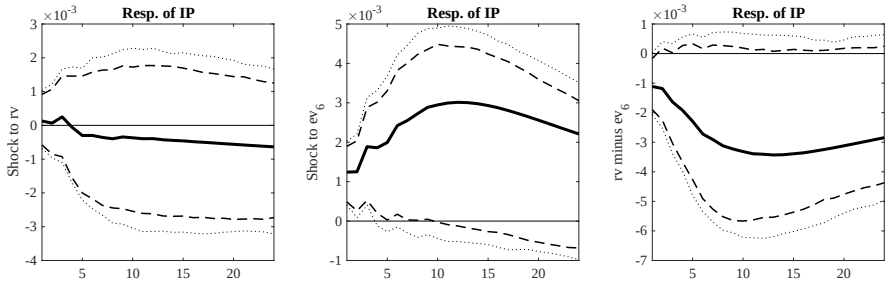
Figure A.8: Robustness (II): response of IP to RV and expectations shocks across specifications



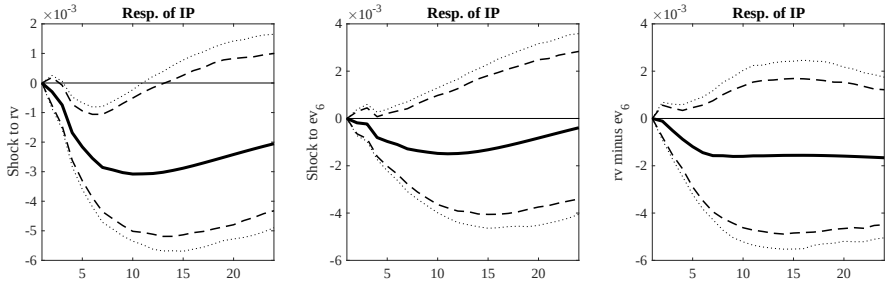
(a) Without detrending the macroeconomic time series



(b) VAR estimated at the quarterly frequency



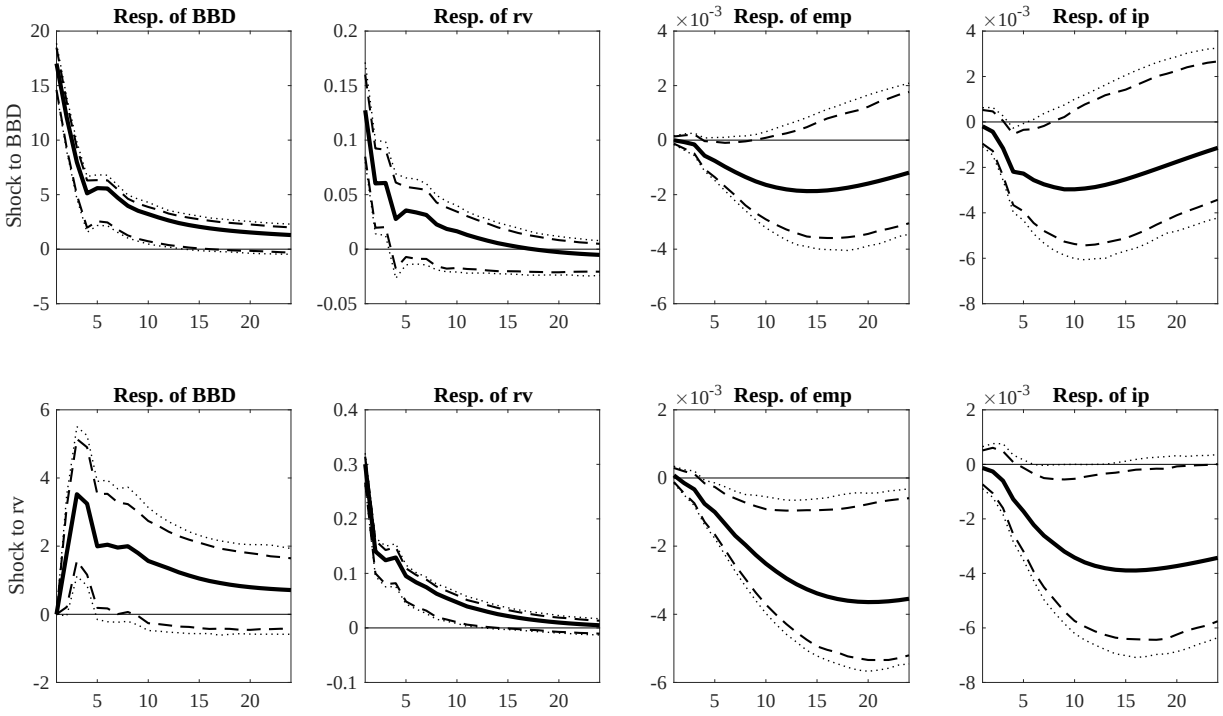
(c) Subperiod 1988-2006 (excluding 1987 crash and financial crisis)



(d) Ordering RV and expectations last in the VAR

Note: The Figure reports the response of IP to RV shocks (left panels) and volatility expectations (middle panels) in different specification, with the difference in the right panel. Each row of the figure corresponds to a different model specification. Row (a) does not detrend the macroeconomic time series. Row (b) repeats the exercise using quarterly, rather than monthly, data. Row (c) estimates the VAR in the subsample 1988-2006, which excludes both RV peaks (1987 crash and financial crisis). Row (d) orders the RV shock second to last and the expectation shock last in the VAR.

Figure A.9: VAR with both rv and Baker, Bloom and Davis measure (BBD; 2015), BBD ordered first



Note: The figure shows impulse response functions (with 90% and 95% CI) of rv, uncertainty (measured as in Baker, Bloom and Davis (BBD, 2015)), employment and industrial production to shocks to rv and BBD, in a VAR with BBD, rv, federal funds rate, log employment and log industrial production. Impulse response functions reported correspond to structural shocks with Choleski decomposition (in the order reported above). Sample covers 1985-2014. All macroeconomic series were detrended with a one-sided HP filter.