BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM
DIVISION OF RESEARCH AND STATISTICS

Date: July 12, 2017

To: Distribution

From: Eric Engstrom and Manuel Gonzalez-Astudillo

Subject: Time variation in upside and downside risks to the staff baseline forecast

---

Summary

- This memo introduces time-varying (monthly) estimates for the magnitudes of upside and downside risks to the staff baseline forecasts for the unemployment rate, real GDP growth, and headline CPI inflation. (See figure 1. A full discussion of these estimates is deferred until section 3.)

- We document that indicators of economic and financial conditions that are available at the time that staff forecasts are constructed, including indexes of real activity, inflation, and financial market strain, are statistically significant indicators of risks to the staff forecasts. An index of macroeconomic uncertainty that has been developed in the academic literature is also a significant predictor of risks to staff forecasts.

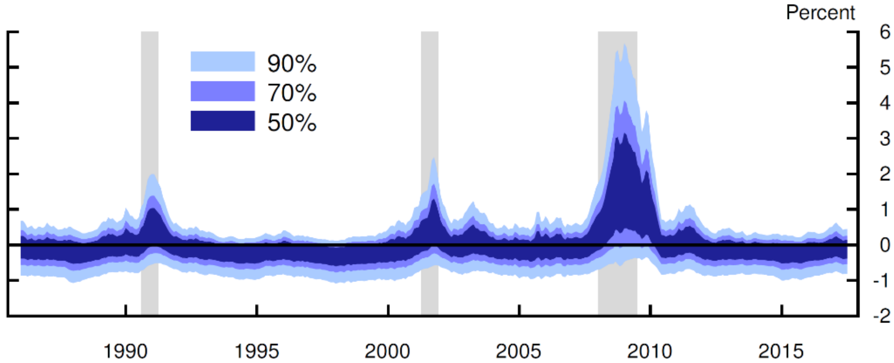
    - The degree of upside risk to the staff's forecasts of the unemployment rate varies substantially over time (top panel). Forecasts that are made during periods of relatively weak economic performance or when macroeconomic uncertainty is high are subject to larger-than-average upside risk. In contrast, the magnitude of downside risk to forecasts of the unemployment rate is relatively stable.

    - The degree of downside risk to the staff forecast of real GDP growth also varies substantially over time (middle panel), with greater downside risk during economic downturns and when macroeconomic uncertainty is elevated. The magnitude of upside risk to real GDP growth forecasts shows smaller, but still significant, variation.

    - The evidence for time-variation in the risks to the staff forecast for headline CPI inflation is a bit weaker in our sample, but still statistically significant (bottom panel). Indexes of inflation and macroeconomic uncertainty are the most useful forecasters of upside risk to the staff forecast for inflation.

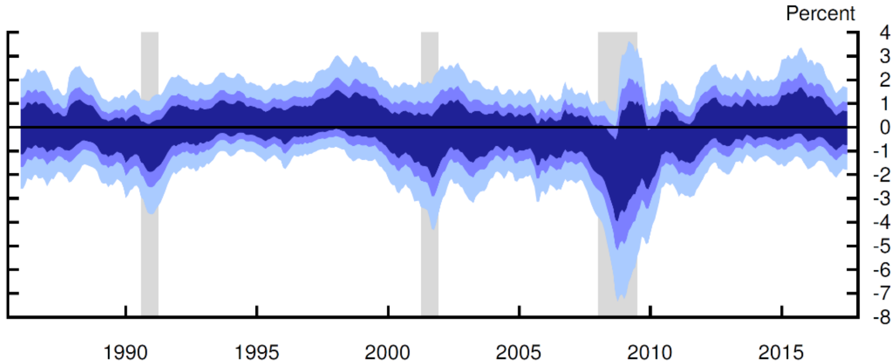
Figure 1: Predicted Distributions for Four-Quarter Ahead Forecast Errors

Unemployment Rate



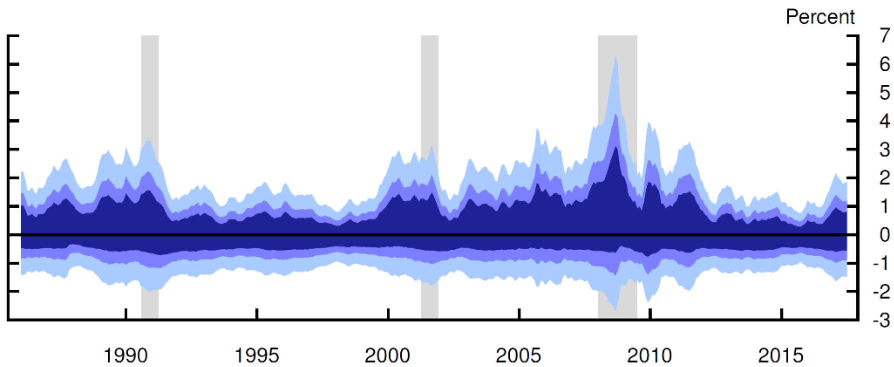
	July 2017
95th	0.4
85th	0.3
50th	-0.1
15th	-0.5
5th	-0.8

GDP Growth



	July 2017
95th	1.7
85th	1.0
50th	0.0
15th	-1.1
5th	-1.8

CPI Inflation



	July 2017
95th	1.9
85th	1.2
50th	0.1
15th	-0.9
5th	-1.5

Note: The exhibit shows estimates of quantiles of the predicted distribution of errors for four-quarter -ahead staff forecasts. The estimates are conditioned on indicators of real activity, inflation, financial market strain, and the volatility of high-frequency macroeconomic indicators. The tables show selected quantiles of the predictive distributions for the respective variables as of the current Tealbook.

Section 1: Introduction

One purpose of The Risk and Uncertainty section of the Tealbook is to provide policymakers with an assessment of the magnitude and balance of risks surrounding the staff's baseline macroeconomic forecast. To make this assessment, the staff relies on a broad range of indicators and analytic devices. This memo adds to that toolkit by providing direct estimates of the degree to which uncertainty and the skewness associated with staff forecast errors has varied systematically over time. We also document the degree to which these risks comove with indicators of cyclical position of the economy, financial market conditions, as well as a measure of uncertainty from the academic literature that is based macroeconomic data series that are published at a relatively high frequency.

Section 2: Data and Methodology

Staff forecast errors and revisions

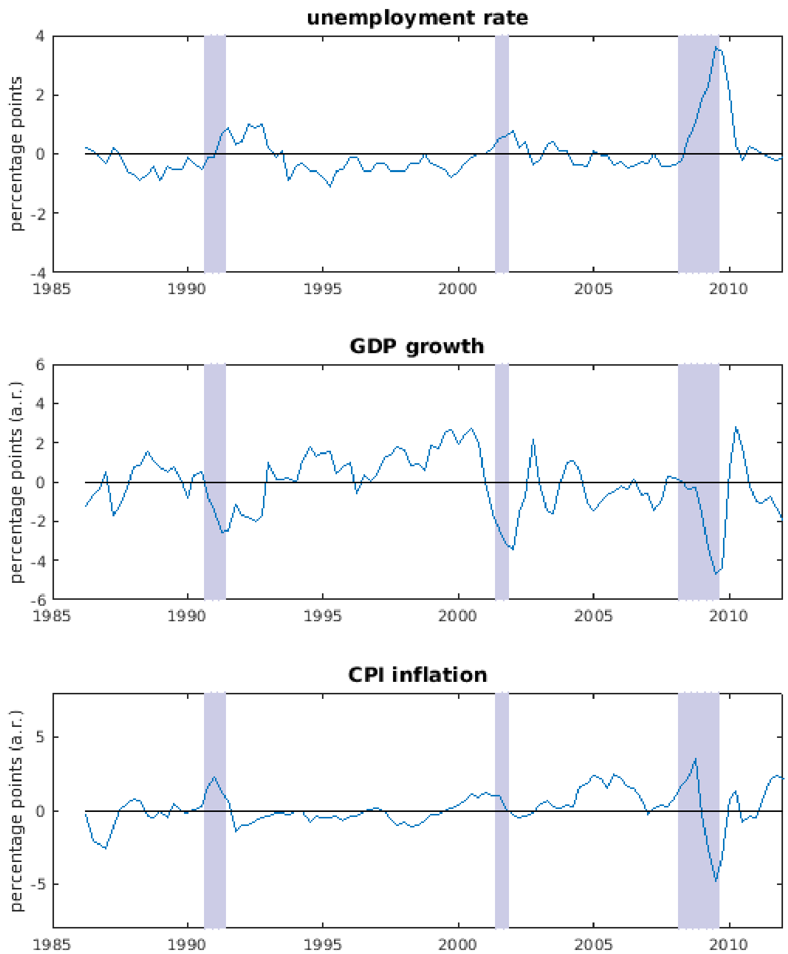
This memo investigates the ex-post errors or revisions to the staff forecast for GDP growth, the unemployment rate, and headline CPI inflation. For estimation, we use a quarterly dataset beginning in 1986:Q1[1] and extending through 2016:Q4. Most of our analysis focuses on revisions to four quarter-ahead forecasts. We define four-period forecast revisions for GDP growth as follows:

$$revgdp_t^4 = E_t\left[gdp_{t-3} + gdp_{t-2} + gdp_{t-1} + gdp_t\right] - E_{t-4}\left[gdp_{t-3} + gdp_{t-2} + gdp_{t-1} + gdp_t\right]$$

where the operator $E_t[\cdot]$ denotes the staff forecast that was constructed closest to the end of quarter *t*. Similarly, $E_{t-4}[\cdot]$ denotes the staff forecast that was made at the end of the year-earlier quarter, *t-4*. Note that the revision thus measures expectations taken four quarters apart, and that $E_t[gdp_{t-3} + gdp_{t-2} + gdp_{t-1} + gdp_t]$ denotes the current estimate of four-quarter growth for the period ending in the concurrent quarter. This estimate therefore reflects information in the "preliminary" or "final" estimates from the BEA for three of these four quarters. Revisions for CPI inflation and the unemployment rate are defined similarly.[2] Figure 2 plots time series for these revisions. Periods of elevated volatility are evident, generally around the times of recessions, with revisions to unemployment registering positive readings while revisions to GDP and CPI generally drop into negative territory. Our goal in this study is to determine to what degree these periods of higher-than-average upside or downside risk could have been predicted by information that was available at the time that the forecasts were made.[3]

[1] The start of our sample was chosen to avoid using data before the estimated start of the "Great Moderation" under the assumption that using data from the earlier periods might make our results less relevant for the current macroeconomic environment.
[2] For the unemployment rate, forecasts are for quarterly average levels in the final quarter of the forecast period.
[3] This memo does not investigate the predictability of the mode of forecast errors, which are assumed to be zero, and it does not address the informational efficiency of the mean or mode of the staff forecasts.

Figure 2: Four-quarter revisions to the staff forecast



Explanatory variables

Our aim is to investigate whether a small set of macroeconomic and financial market indicators can predict the upside or downside variance of staff forecast errors. The timing of the measurement of these instruments was set to ensure that they were available to forecasters at the time that the forecasts were made. Our main instruments include four indexes, which are plotted in figure 3:[4]

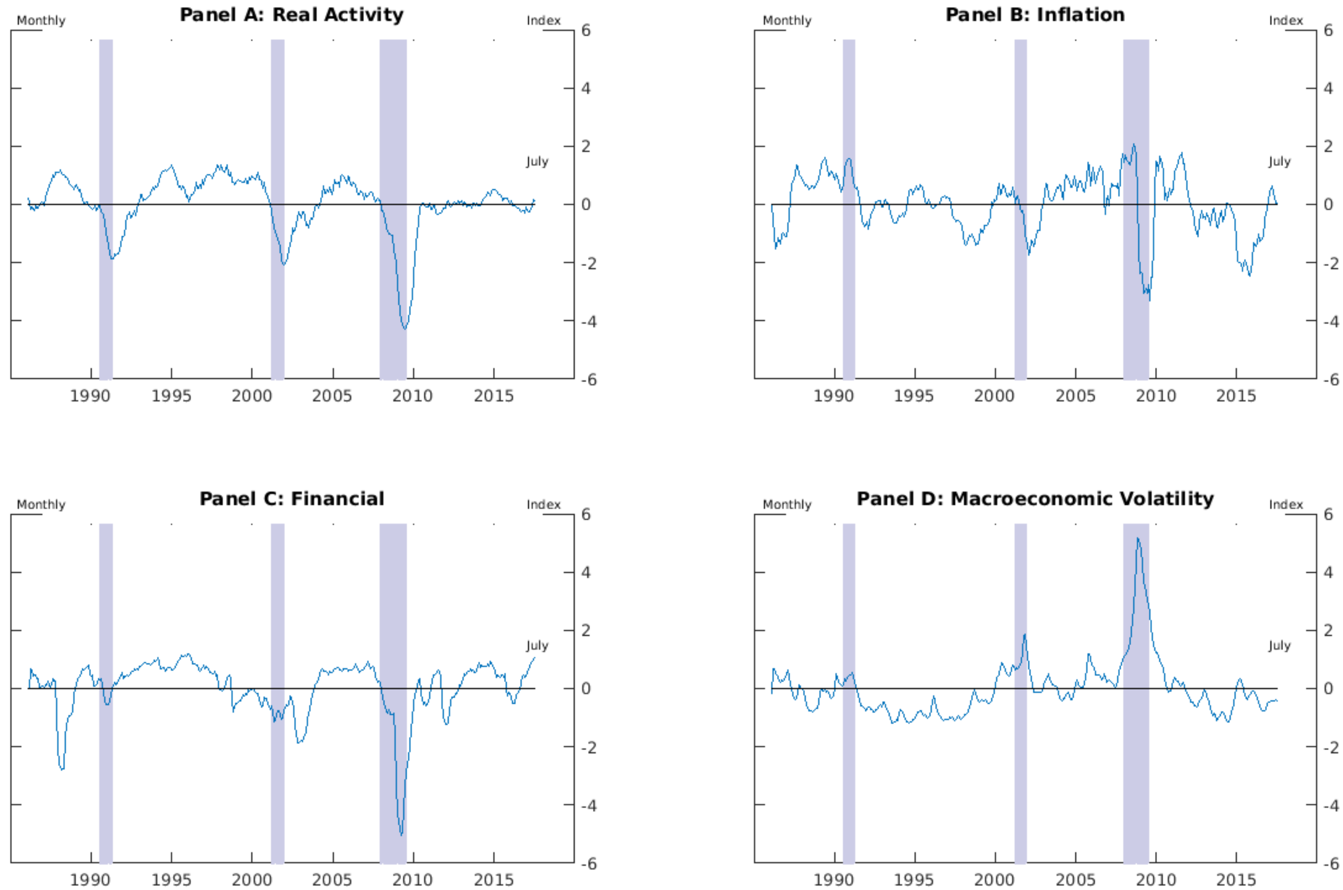
1) Real activity: A weighted average of the 12-month growth rates of nonfarm employment, industrial production, and an index of help wanted postings.[5]

[4] We tried a few other variables that did not produce significant results. These included a measure of U.S. economic policy uncertainty developed in the article, S. Baker, N. Bloom, and S. Davis, (2015), Measuring Economic Policy Uncertainty, No 21633, NBER Working Papers, National Bureau of Economic Research, Inc.

[5] The indexes for real activity and inflation were originally proposed in the article, A. Ang and M. Piazzesi(2003), “A No-Arbitrage Vector Regression of Term Structure Dynamics with Macroeconomic and Latent Variables,” *Journal of Monetary Economics*, vol.50, pp. 745-787.

2) Inflation: A weighted average of the 12-month growth rates of the consumer price index, producer price index, and an index of commodity prices from the Commodity Research Bureau.
3) Financial: A weighted average of option-implied and realized volatility for the S&P 500 index, corporate bond spreads, and two- and four-quarter equity returns. Each indicator is signed so that periods of stress drive the index lower.
4) JLN macroeconomic uncertainty: A weighted average of individual time-varying volatility estimates for a set of about 130 macroeconomic time series.[6]

Figure 3: Instruments



Empirical framework

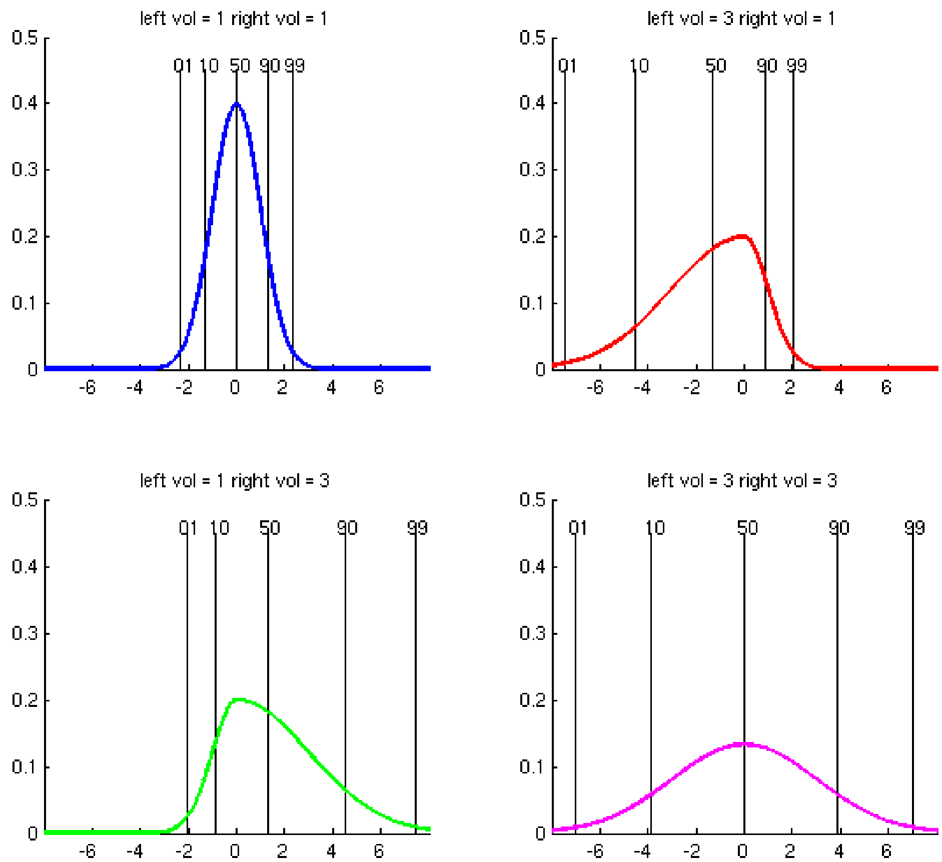
To investigate time-varying upside and downside volatility, we take a page from the Bank of England and the Sveriges Riksbank, which use a "double Gaussian" distribution to describe the potentially asymmetric uncertainties associated with their forecasts for real activity and inflation.[7] The double Gaussian distribution is essentially two "half Gaussian" distributions

---

[6] See K. Jurado, S. Ludvigson and S. Ng.(2015), "Measuring Uncertainty," *American Economic Review*, vol.105, pp. 1177-1216.
[7] See K. Wallis (2014), "The Two-Piece Normal, Binormal, or Double Gaussian Distribution: Its Origin and Rediscoveries," *Statistical Science*, vol. 29, pp. 106-112. For an alternative approach to measuring time-varying

pasted together at their (common) mode. The standard deviation of one distribution governs the downside tail, while the standard deviation of the other governs the upside tail. Figure 4 illustrates some examples of the double Gaussian distribution:

Figure 4: The Example of the Double-Gaussian Distribution



A rich range of potential shapes for the distribution is achieved by allowing the volatility on the left side of the mode to differ from that on the right side of the mode. Notice that the length of the left tail in the upper right panel, as measured as the distance from the 1st to the 50th percentiles of the distribution, is much greater than the corresponding length of the right tail. The opposite is true in the lower left panel. Formally, the double Gaussian density at time t for a random variable rev_{t+k} (for example, the. forecast error) that will be realized k periods ahead, is:

$$f\left(rev_{t+k};\sigma_t^{left},\sigma_t^{right}\right)=\begin{cases} A_t \exp\left[-rev_{t+k}^{\ 2}/2\sigma^2_{t,left}\right] & rev_{t+k}<0 \\ A_t \exp\left[-rev_{t+k}^{\ 2}/2\sigma^2_{t,right}\right] & rev_{t+k}\geq 0 \end{cases}$$

risks, see T. Adrian, N. Boyarchenko and D. Giannone, (2016), Vulnerable growth, No 794, Staff Reports, Federal Reserve Bank of New York.

The density for y_{t+k} on both sides of the mode (zero, in this case) is Gaussian, but with different standard deviations, $\sigma_{t,left}$ and $\sigma_{t,right}$. The parameter, A_t, is a normalization constant that depends only on those volatilities. To allow for time-variation in the distributions, the left and right volatilities are modeled as functions of the instruments that are available at time (t).

$$\sigma_{left,t} = g\left(X_t'\beta_{left}\right)$$
$$\sigma_{right,t} = g\left(X_t'\beta_{right}\right)$$

The vector X_t contains the explanatory variables, including a constant and some subset of the instruments depicted in Figure 3. The function, $g(x)$ ensures that the volatility measures are always positive.[8] The coefficients β_{left} and β_{right} are estimated by optimizing likelihood function:

$$\log\left(\ell\right)=\sum_{t=1}^{T}\log\left(f\left(rev_{t+k};\sigma_t^{left},\sigma_t^{right}\right)\right)$$

Asymptotic inference for parameters estimated by maximum likelihood has been well established under standard technical conditions. However, those conditions are not satisfied in our framework. First, our data sample is quite short, extending for only about 30 non-overlapping four-quarter periods, so standard asymptotic inference is potentially subject to small sample biases. Second, our estimations for four-quarter revisions use overlapping quarterly observations, creating an artificial serial dependence in the sequence of error realizations that could bias inference that relies on standard asymptotic results. For these reasons, we use bootstrapping techniques to calculate standard errors and test for the statistical significance of parameter estimates.

Choosing model specifications: univariate performance of instruments

Table 1 shows results from specifications in which σ_{left} and σ_{right} are allow to vary with one instrument only.[9] For the unemployment rate, three out of the four instruments significantly forecast upper tail risk. The coefficient in the real activity index is negative, suggesting that forecasts made during recessions are subject to greater right-side volatility (upside risk) for the unemployment rate forecast. The financial and macroeconomic uncertainty indexes show similarly that forecasts for the unemployment rate that are made during periods of financial strain are associated with greater upside risks to the unemployment rate. In contrast, little systematical variation is identified for downside risk to the unemployment rate forecast.

Regarding forecast errors for GDP growth, all four instruments demonstrate some explanatory power for the right- or left-side volatility. When the real activity index is low, as in recessions or periods of financial stress, downside volatility increases notably. Similar results obtain for the financial and macroeconomic uncertainty indexes. Somewhat puzzlingly, higher inflation is associated with lower upside risk to GDP growth, but greater downside risk.

---

[8] In particular, $g(x)$ is constructed piecewise, with $g(x) = x$ for $x \geq 1$ and $g(x) = exp(x\text{-}1)$ for $x<1$.

[9] These results are for four-quarter ahead forecast errors for the full estimation period, but results for one-quarter ahead forecast errors are broadly similar, as are results for an estimation period that ends in 2007, thereby excluding the Great Recession of 2008–2009.

Regarding inflation, the results are generally weaker. The coefficient on macroeconomic uncertainty is positive for the right tail, but most of the other coefficients are less significant.[10]

Table 1

Univariate models of upside and downside uncertainty for staff forecast errors

Four-quarter ahead forecast

forecast	uncertainty	explanatory factor			
		real	infl	fincl	JNL unc
unemp rate	upside	-1.3209 **	0.4030	-1.1874 **	0.8089 *
	downside	0.2263 **	0.0163	0.0288	-0.2496 **
real GDP growth	upside	-0.0683	-0.9865 ***	-0.3082 *	0.0687
	downside	-0.7156 **	0.6770 **	-0.9089 ***	0.8067 ***
CPI inflation	upside	-0.3088 *	0.4162 **	0.0126	0.9452 ***
	downside	-0.5077 *	0.2911	-0.2532	0.3054

Note. The symbols ***, **,* denote significance at the 10, 5, and 1 percent levels as determined by a block-bootstrapping procedure

Choosing model specifications: multivariate models

The univariate models suggest that all of the instruments show some promise as forecasters of right- or left-tail risk for at least one of the forecast error series. However, as is evident in figure 3, these series are all fairly highly correlated over time and they all show strong variation over the business cycle. We next proceed to determine whether there is evidence that using multiple explanatory variables simultaneously is useful. To make this determination, we use standard information criteria, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

Table 2 summarizes model selection tests for each of the four-quarter forecast error variables. A total of 16 specifications are investigated for each forecast error series: one “null” model, in which both right- and left-side volatility series are constant, four univariate models (the same ones for which parameters are reported in table 1), six possible bivariate models with two right hand side variables, four trivariate specifications, and one “kitchen sink” specification that uses all four instruments. Standard AIC and BIC tests are used.[11] A rank is calculated for each specification. The top eight specifications are highlighted in green and the bottom eight in red.

[10] Preliminary work suggests that the risks to forecast errors from survey-based estimates of macroeconomic variables, such as those from the Survey of Processional Forecasters, and even econometric models such as FRB/US demonstrate similar patterns similar to those documented in table 1.

[11] Note that we are examining the performance of models forecasting four-quarter ahead errors in a quarterly data set. While we use overlapping data to calculate parameter values in table 1 (being careful to appropriately bootstrap standard errors), we avoid using overlapping data for the model selection tests because the likelihood values (and thus the AIC and BIC scores) are hard to evaluate. We instead conduct tests separately for data from Q1, Q2, Q3, or Q4, and find the median rank for AIC and BIC across the four quarters.

Table 2

Model selection tests, four-quarter-ahead forecast errors

set of explanatory variables	unemp rate		real GDP growth		cpi inflation	
model	AIC	BIC	AIC	BIC	AIC	BIC
{null}	16	16	11	3	10	1
{real}	9	4	13	9	8	5
{infl}	14	14	3	2	3	3
{fincl}	15	15	7	5	16	10
{uncty}	5	2	1	1	2	2
{real,infl}	3	3	5	7	4	6
{real,fincl}	12	11	14	10	15	14
{real,uncty}	1	1	4	6	11	9
{infl,fincl}	11	10	16	12	5	7
{infl,uncty}	8	7	2	4	1	4
{fincl,uncty}	10	9	6	8	9	8
{real,infl,fincl}	6	8	12	15	12	13
{real,infl,uncty}	2	5	9	13	6	11
{real,fincl,uncty}	4	6	10	14	14	15
{infl,fincl,uncty}	13	13	8	11	7	12
{real,infl,fincl,uncty}	7	12	15	16	13	16

Note. In-sample period of from 1986-2010. Green denotes top-quartile model, yellow denotes second quartile, red denotes bottom half.

The AIC and BIC typically identify the more parsimonious models with one or two explanatory variables as optimal. Conversely, the best performaning model is rarely an element of the set of trivariate specifcations or the "kitchen sink" model. From this exercise, we conclude that it would be inadvisable to rely heavily on specifications with three or four explanatory variables because of the risk of overfitting the data, which would likely harm out-of-sample performance.

Choosing model specifications: Out-of-sample performance

While the previous two subsections investigated the in-sample performance of various specifications, in this subsection, we present results for the out-of-predictive power of the models that we consider. To test out-of-sample performance, we estimate each of the sixteen specifications in Table 2 over the period from 1986-2007. We then test the performance of various specifications and model combinations over the out-of-sample period from 2008–2016. Taking a cue from the forecasting literature, which suggests that better out-of-sample performance may be achieved by averaging across models, we evaluate the performance of both individual specifications and various model combination forecasts. In particular, the model forecasts that we evaluate are:

1) Null: constant upside and downside volatility estimated over the in-sample period
2) WTA(AIC): “winner-take-all,” the best performing model from the set of 16 as determined by performance in the in-sample period by the AIC criterion.
3) WTA(BIC): “winner-take-all,” the best of model from the set of 16 as determined by performance in the in-sample period by the BIC criterion.
4) EWT: An forecast calculated as an equal weighted average of the upside and downside volatility forecasts from all 16 models
5) RNK(AIC): A weighted average of the forecasts for upside and downside volatility from the sixteen models in which the weights are proportional to the inverse of the rank of the model as scored by AIC performance for the in-sample period.
6) RNK(BIC): A weighted average of the forecasts for upside and downside volatility from the 16 models. The weights are proportional to the inverse of the rank of the model as scored by BIC performance for the in-sample period.

The score for each forecast is calculated as the cumulative likelihood function for the out-of-sample period. Table 3 summarizes the results for this exercise

Table 3

Out-of-sample results for model combination schemes

	four-quarter		
	unemp rate	real GDP growth	cpi inflation
NULL	126.3	67.2	101.3
EWT	32.1	59.3	74.0
RNK(AIC)	28.8	59.2	73.3
RNK(BIC)	29.4	58.7	73.4
WTA(AIC)	27.9	61.0	82.2
WTA(BIC)	27.9	65.0	82.2

* Notes. Negative log likelihoods reported for the out-of-sample period
Pre-sample period: 86-07, postsample-period: 2008-2016
Highlighting denotes top three performing models.

According to table 3, the models with the best out-of-sample performance generally use some form of model combination (EWT or RNK). Moreover, the models that put higher weights on specifications that score the best according to the in-sample AIC or BIC are usually among the top performers. The null model under which the volatilities are constant generally performs very poorly, as do the winner-take-all specifications (except for the unemployment rate). Interestingly, the superior performance of simple model averaging schemes relative to those of singular specifications is consistent with the findings of researchers in several disparate forecasting contexts.[12] Based on these results, we choose as our main forecasting model the combination forecast RNK(AIC).

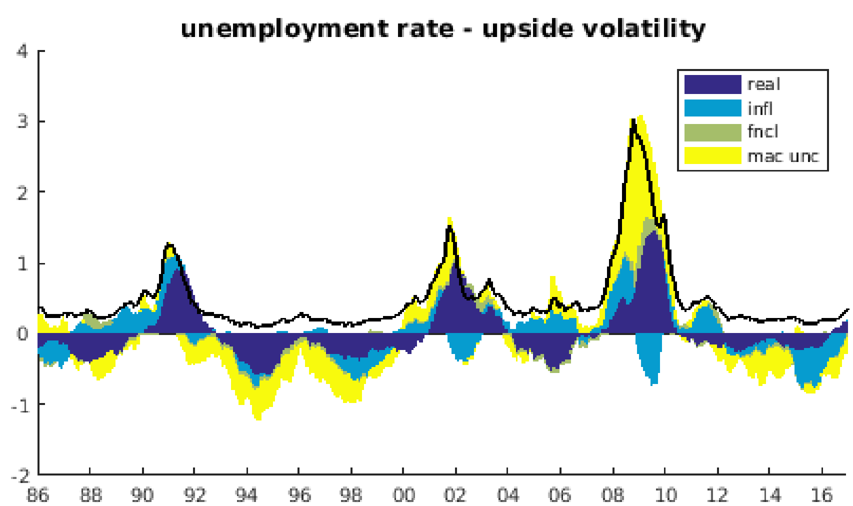
Section 3: Results

This section presents a fuller discussion of the results in figure 1 and provides a decomposition of the variation in various measures of upside and downside volatilities into components driven by the various instruments.

Unemployment rate

The forecasts for the distibution of unemployment rate forecast errors at the four-quarter ahead horizon are shown in the top panel in figure 1. There are two striking features of the predictive distributions. First, the downside volatility varies relatively little over time. That is, downside risk to the forecast is essentially constant, with the 5th percentile of the distribution hovering about -1 percentage point. Conversely, the upper tail of the distribution exhibits substantial volatility. Second, the variation of the upper tail is closely tied to the business cycle, rising from levels less than 1 percentage point to 3 percentage points or higher during periods of financial strain. Figure 5 shows how the various explanatory variables drive the forecasts of the upper and lower tail under the preferrred RNK(AIC) model.

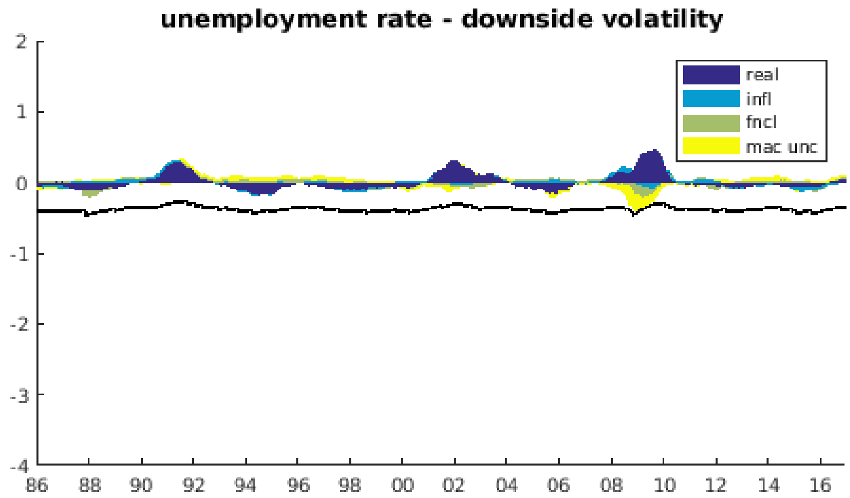
Figure 5: Components of the forecast for the distributiuon of unempoloyment rate forecast errors



---

[12] We did not investigate more complex Bayesian model averaging schemes. For a discussion of these more complex schemes and a discussion of the performance of complex schemes versus simpler ones, see G. Elliot and A. Timmerman (2016), *Economic Forecasting* (Princeton: Princeton University Press.)

Figure 5, continued:



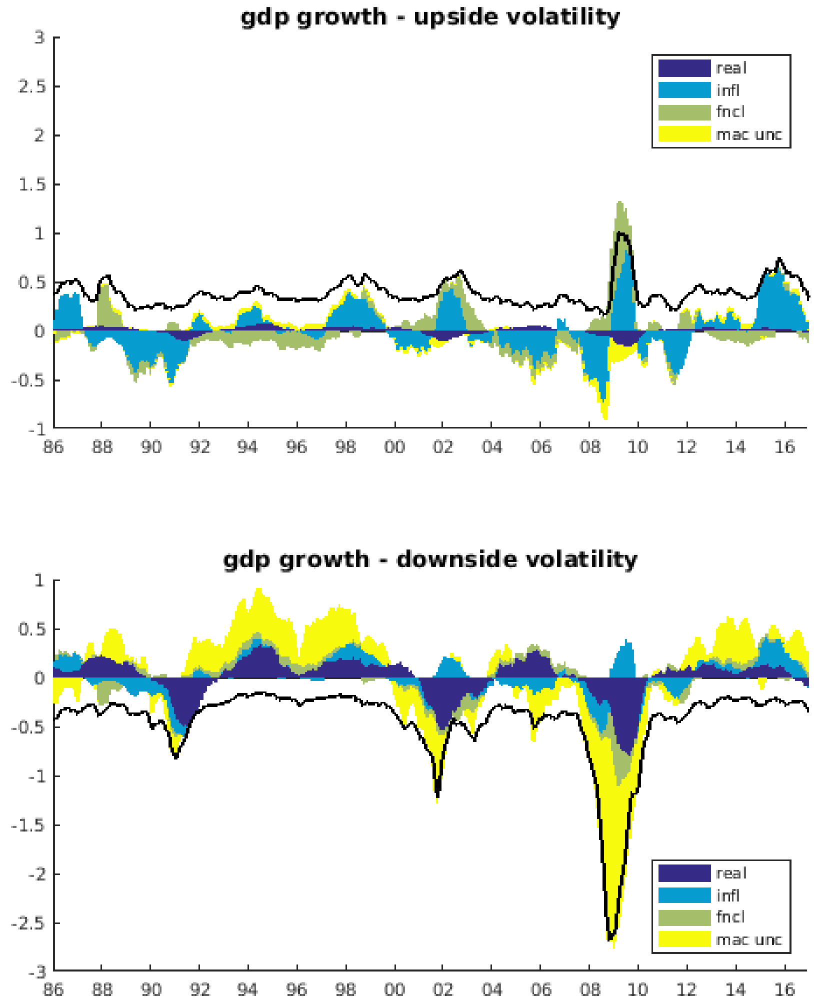
As can be seen the two explanatory variables that explain most of the variation in the forecast for the lower tail are the JLN measure of macroeconomic uncertainty and the real activity index. The inflation and financial market indexes contribute relatively little to the forecast errors.

GDP growth

The forecasts for the distibution of GDP growth at the four-quarter ahead horizon are shown in the middle panel in figure 1 on page 2. There are two striking features of the predictive distributions. First, the upside volatility varies relatively little over time. That is, upside risk to the forecast is essentially constant, wirth the 95th percentile of the distribution hovering about +2 percentage points. Conversely, the lower tail of the distribution exhibits substantial volatility. Second, the variation of the lower tail is closely tied to the business cycle. The lower tail of the distribution falls from a typical level of around -2 percentage points to -4 or -6 percentage points during business cycle downturns.

Figure 6 shows how the various explanatory variables drive the forecasts of the upper and lower tail under the preferrred model.

Figure 6: Components of the forecast for the distribution of GDP forecast errors

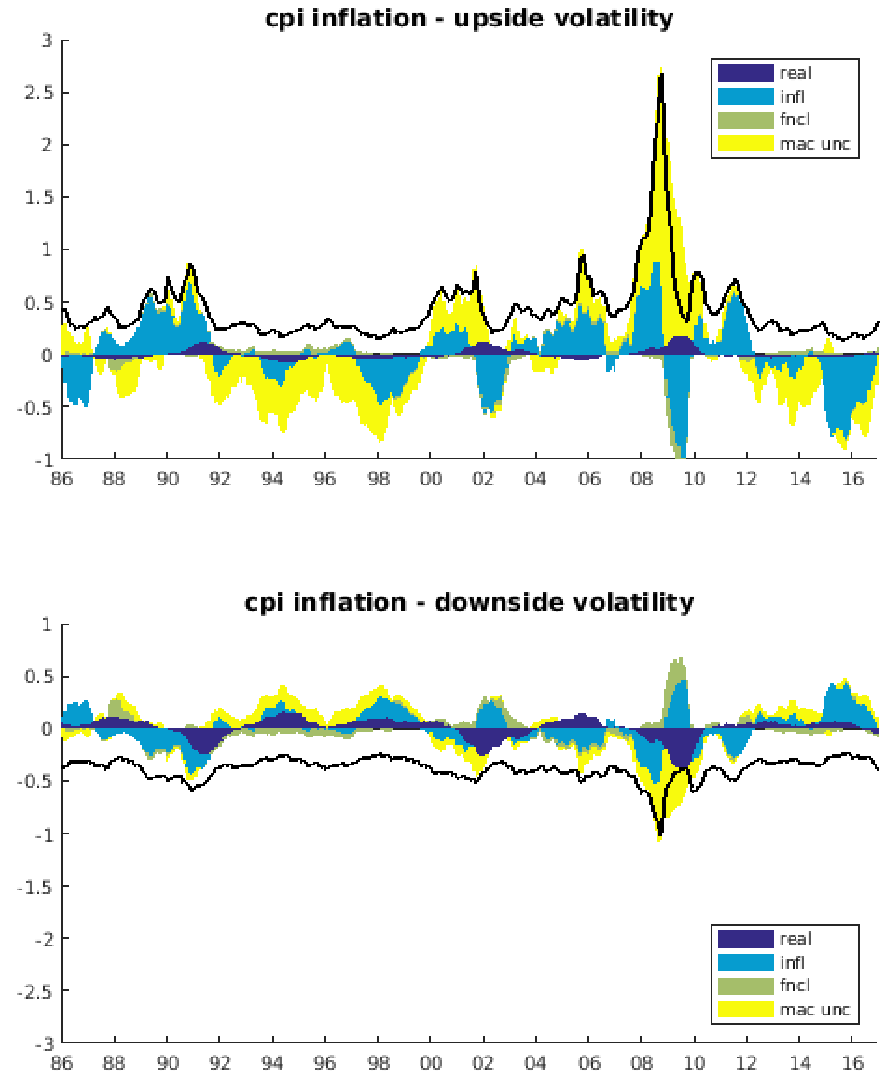


As can be seen the two explanatory variables that explain most of the variation in the forecast for the lower tail are the JLN measure of macroeconomic activity, and the real activity index. The inflation and financial market indexes contribute relatively little.

CPI inflation

The forecasts for the distibution of CPI inflation at the four-quarter ahead horizon are shown in the botton panel in figure 1. The downside volatility tends to vary relatively little. In contrast, upside volatility varies more substantially over time. Figure 7 shows how the various explanatory variabe drive the forecasts of the upper and lower tail under the preferrred model. The JLN index of macroeconomic uncertainty and the inflation index drive most of the variation in the upper tail of the distribution.

Figure 7: Components of the forecast for the distribution of CPI forecast errors



Section 4: Conclusion and recommendations

This memo provides evidence that the upside and downside risks to staff forecasts of real GDP growth, the unemployment rate, and headline CPI inflation vary substantially using instruments that are available in "real time." We intend to regularly consult the measures of risks presented in this memo and include them in an exhibit for the Risks and Uncertainty section of Tealbook A in order to provide additional information on the time-varying risks to the economic outlook, and their dependence on macroeconomic and financial market conditions.