# Behind the scenes of expectations: interpreting survey forecasts

Federica Brenna*
KU Leuven

*Job Market Paper*
November 2022
Latest version

**Abstract**

Forecasts produced by experts can influence the expectations of the general public, and ultimately the real economy. In this paper I ask what type of structural drivers professional forecasters think will affect their projections? To what extent do they disagree about these drivers, and how uncertain are they about their magnitude? I model forecasts in a novel empirical macroeconomic setting, which allows me to decompose them into a model implied part and a judgement part, reflecting individual expectations of future shocks. The model takes into account multi-step ahead conditional forecasts, includes subjective uncertainty measured via different methods, and identifies shocks exploiting the time-varying volatility present in the forecasts. I find that throughout the sample, forecasters mostly disagree on the size of the shocks, while in periods of high volatility they give a larger weight to judgement and also disagree on the nature of shocks. My findings can inform policy makers by providing a deeper insight into the expectations formation process of forecasters from a structural perspective.

**Keywords:** Expectations Formation, Judgement, Identification via Stochastic Volatility, Subjective Uncertainty, Survey of Professional Forecasters
**JEL Codes:** C32, C33, C51, D84, E37

*This paper benefited greatly from comments and advice from my PhD advisor, Ferre De Graeve. I also thank, for useful comments and discussions at various stages of the project, Marta Bańbura, Žymantas Budrys, Michael Kobielarz, Christian Pröbsting, Francesco Ravazzolo, Jan David Schneider and Raf Wouters. The resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government.
Contact: federica.brenna@kuleuven.be, www.federicabrenna.com.

# 1 Introduction

Understanding which types of shocks drive main macroeconomic aggregates such as GDP and prices is an important, but difficult endeavor for policymakers. Particularly in times of higher uncertainty, knowing whether supply, demand or other factors are behind macro dynamics has direct relevance for policy decisions. A recent example is the high inflation period started during the COVID-19 pandemic, with other prominent cases in the Great Financial Crisis of 2008 or the high inflation of the '70s.

There is a specific group of agents whose opinion on these drivers is arguably very interesting for central banks and policymakers in general: professional forecasters. They are assumed to be "attentive" agents, who provide their forecasts after drawing from a vast array of data, news, quantitative models, as well as individual knowledge and judgement. Their responses can affect other agents' expectations, such as households and corporations (Carroll, 2003), and further feed back to investment and consumption decisions, ultimately affecting the real economy. What kind of structural shocks they assume are driving their forecasts? And how can their answers inform policy makers' decisions? This paper attempts to answer these questions by modelling professional forecasters' responses, accounting for the role of judgement, and identifying shocks from time variation in their volatility.

To do so, I assume each respondent produces their forecast by running a vector autoregression (VAR) model, to which they add a judgemental component. VARs are flexible and general specifications, able to encompass several data generating processes. Given that we do not know which model(s) are used by respondents to produce their forecast, any of them risks being misspecified. The objective of the paper, however, is not to guess forecasting methods used, rather to extract, from a general enough model of both forecasts and realizations, structural shocks, and analyze their variation, both over time and across agents. I collect individual responses from the Philadelphia Fed Survey of Professional Forecasters (SPF), for horizons between zero (i.e. nowcast) and four quarters ahead, for real GDP and the GDP deflator, and estimate a VAR on forecasts and data.

The estimation allows me to retrieve coefficients underlying respondents' forecasts, and, more importantly, the path of future shocks affecting them. This is possible thanks to the simple assumption that agents produce conditional forecasts, that is, they adjust their model forecasts by assuming a path for future shocks (or, in the words of Antolin-Diaz et al. (2021), a "structural scenario"). I identify structural shocks by exploiting the stochastic volatility present in the forecasts, and label them *supply* and *demand* shock, as defined traditionally: a positive demand shock increases both quantities and prices, and a positive supply shock increases quantities while decreasing prices. I then look at historical shock decompositions and at specific episodes in US history, both from the perspective of a representative (average) forecaster, and for the cross-section of individual forecasters.

I find that for the average respondent negative demand shocks caused the bulk of the Great Recession, although positive supply shocks were expected less than a year before the recession trough. As for the COVID pandemic, the strong rebound in GDP is attributed to a positive demand shock, with stronger supply expected in 2022, which did not fully materialize. As for inflation, supply shocks tend to dominate for the majority of the sample, with notable exceptions in the negative demand shocks during the Great Recession, and in the first half of the pandemic. The more recent surge in inflation is nearly completely attributed to a negative supply shock. Looking at cross-respondent variation, I find substantial disagreement both regarding the size and the nature of the shocks, which heightens in periods of high instability. While density forecasts provided by professional forecasters do contain additional information, they do not affect the model stochastic volatility in an evident manner.

This paper relates and contributes to three main strands of the literature: studies which model expectations structurally, papers on deviations of forecasters from the Full Information Rational Expectation (FIRE) paradigm, and empirical works analyzing surveys' uncertainty and disagreement.

The first related field is a relatively new, but growing, strand of the literature which looks at surveys from a structural perspective: among the most recent papers, Herbst and Winkler (2021) extract a supply and demand factor from the co-movements of individual survey responses across variables; Andre et al. (2022, 2021) provide empirical evidence of agents subjective models of the economy, finding substantial disagreement on economic drivers, both among experts and households. Less recently, Ricco et al. (2016) identify fiscal spending shocks from expectations revisions, and quantify the effects of fiscal communication (measured through an index of disagreement) on the propagation mechanism, while Krane (2011) characterizes forecasters' view on frequency and duration of shocks to GDP. I contribute to this literature by providing an empirical setting to extract structural shocks from survey forecasts, without resorting to any theoretical model. The use of identification via time-varying volatility in this context is also novel, and it allows me to identify shocks solely from the intrinsic forecasts' characteristics, without imposing ex ante restrictions.

The stylized facts found in empirical papers seem to suggest a departure from the "FIRE" paradigm of survey respondents, which is why several later works try to model the irrationality and incomplete information of expectation formation processes. These include papers on information rigidities (Coibion et al., 2018; Coibion and Gorodnichenko, 2012; Clements, 2022; Coibion and Gorodnichenko, 2015), imperfect or sticky information (Born et al., 2020; Farmer et al., 2021; Del Negro et al., 2022), over- or underreaction of respondents (Casey, 2021; Huang et al., 2022; Kohlhas and Walther, 2021; Bordalo et al., 2020), heterogeneous expectations (Dovern, 2015; Dovern and Hartmann, 2017) and inattentive forecasters (Giacomini et al., 2020; Andrade and Le Bihan, 2013). My paper departs from this literature by not taking any specific stance on the microfoundations behind processes of expectations formation, instead modelling expectations in a flexible, reduced form fashion. It then identifies structural drivers of forecasts, rather than looking at what type of irrational behaviors are consistent with specific types of revisions. In other words, while I acknowledge that forecasters may be irrational, and are certainly heterogeneous, be it in their priors or in their judgements, I do not try to assign their responses to a specific model, but I simply use them to extract implied structural shocks.

The last set of works includes, among others, papers on survey uncertainty (Binder et al., 2022; Abel et al., 2016; Boero et al., 2015; Clark et al., 2020) and its relationship with disagreement (Rich and Tracy, 2010, 2021).[1] In this paper, I further investigate the issue of measuring forecasters' uncertainty, by comparing the most common methods used to fit a continuous distribution to the discrete responses provided in surveys, and find differences across methods are larger in periods of higher economic turbulence. Besides, I map the observed, fixed event uncertainty to the model's, fixed horizon one, and use it to inform my estimation.

The rest of the paper is organized as follows: Section 2 describes the model and Section 3 the identification method. Section 4 introduces observed uncertainty, while Section 5 describes data and estimation. Section 6 presents results, and Section 7 concludes.

[1] A comprehensive review on survey of professionals can be found in Clements et al. (2022).

# 2 A model of conditional forecasts

I assume the average[2] SPF respondent runs a VAR model such as:

$$y_t = c + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + u_t \qquad\qquad u_t \sim N(0, \Lambda) \tag{1}$$

where $y_t$ is a vector of observables, $u_t$ is a vector of reduced form disturbances, $c$ is a vector of constants, $\phi$ is a matrix of lagged coefficients, $p$ is the number of lags, and $\Lambda$ is the variance-covariance matrix of $u_t$. In the following, I omit the constant and lags above one for easier calculations, without loss of generality.

When the respondent performs an unconditional forecast based on this model for horizon $h = 1$, they obtain:

$$\mathbb{E}_t[y_{t+1}] = \phi y_t \tag{2}$$

And for any horizon $h > 1$,

$$\mathbb{E}_t[y_{t+h}] = \phi^h y_t \tag{3}$$

Given this assumption, we define a VAR on forecasts as:

$$\mathbb{E}_t[y_{t+1}] = \phi \mathbb{E}_{t-1}[y_t] + \tilde{u}_t, \tag{4}$$

with $\tilde{u}_t = \phi u_t$, using $y_t = \phi y_{t-1} + u_t$ and $\phi y_{t-1} = \mathbb{E}_{t-1}[y_t]$. The coefficient $\phi$ is the same as in (1).

If each respondent were to run model 1, they would all obtain the same forecasts, given they all observe the same data. We know this is not the case both by looking at survey data (see Figure 1, displaying the time series of disagreement across forecasters) and thanks to several empirical analyses (e.g. Andrade et al., 2016; Glas, 2020; Mankiw et al., 2003).

[2]In the following, I describe the model for an average forecaster, to simplify notation. This model is analogous to the one run for each individual forecaster and for the pooled panel of forecasters.

Where do the differences between forecasts come from? In the literature, they have been attributed among others to imperfect information (Coibion and Gorodnichenko, 2015), asymmetric attention (Kohlhas and Walther, 2021), or heterogeneous priors (Giacomini et al., 2020; Patton and Timmermann, 2010). I propose they can be modeled in a general way as conditional forecasts: respondents adjust their (unconditional) model forecast by assuming future shocks are different than zero. Consider the conditional forecast for $y_{t+1}$ done at time $t$:

$$\mathbb{E}_t[y_{t+1}] = \phi y_t + u_{t+1|t} = \phi(\phi y_{t-1} + u_{t|t}) + u_{t+1|t} \tag{5}$$

Given $\phi y_{t-1} = \mathbb{E}_{t-1}[y_t] - u_{t|t-1}$, rewrite:

$$\mathbb{E}_t[y_{t+1}] = \phi\mathbb{E}_{t-1}[y_t] - \phi u_{t|t-1} + \phi u_{t|t} + u_{t+1|t}$$

The first two terms represent information available at time $t-1$ (the unconditional forecast and the shock assumed for period $t$ in the previous period), while the last two terms include information at time $t$, namely the current shock and the shock expected in $t+1$. The equation above can be further rearranged and be written in terms of a "restricted" VARMA model:

$$\mathbb{E}_t[y_{t+1}] = \phi\mathbb{E}_{t-1}[y_t] + \underbrace{\phi(u_{t|t} - u_{t|t-1})}_{\text{update due to new info}} + \underbrace{u_{t+1|t}}_{\text{judgement}},$$

where the autoregressive (AR) part is the first term, and the moving average (MA) part is the second term. The model is restricted in the sense that the coefficients for the AR and the MA parts are the same, unlike a traditional VARMA model. The MA term can be interpreted as the update due to new information which becomes available in period $t$, while the last term is an assessment of future shocks made by the respondent, which I call

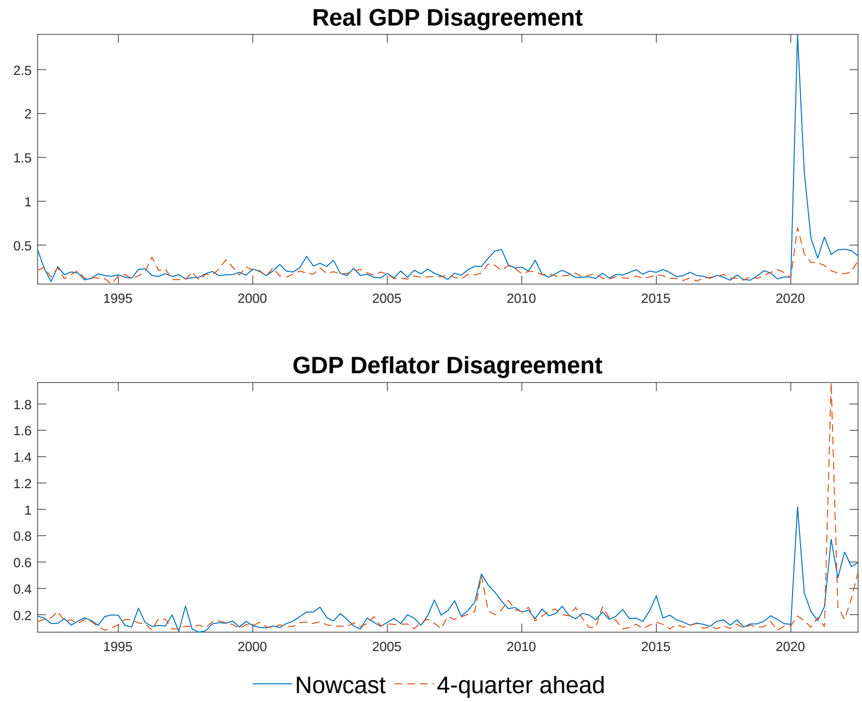


Figure 1: Disagreement for Real GDP and GDP Deflator q-o-q growth from Fed SPF forecasts, at h=0 and h=4. Disagreement is calculated as the standard deviation of point forecasts across individuals.

judgement. The generalized version for $h$-step ahead forecasts is[3]:

$$\begin{aligned}\mathbb{E}_t\left[y_{t+h}\right] &= \phi\mathbb{E}_{t-1}\left[y_{t+h-1}\right] + u_{t+h|t} \\ &+ \phi\left(u_{t+h-1|t} - u_{t+h-1|t-1}\right) + \phi^2\left(u_{t+h-2|t} - u_{t+h-2|t-1}\right) + ... + \phi^h\left(u_{t|t} - u_{t|t-1}\right) \quad (6)\end{aligned}$$

Equation (6) contains all conditions assumed by the respondent either at time $t$ or $t-1$, as well as the current shock ($\phi^h u_{t|t}$).

[3]Derivations are available in the Appendix.

## 2.1 State space representation

The model is cast in a state space form, both to allow for the time-varying nature of shock volatility, and to account for missing observations when using individual level data.

$$\mathbb{E}_t\left[y_{t+1}\right] = c + \phi \mathbb{E}_{t-1}\left[y_t\right] + \theta \sigma_{t+1} \mathbb{E}_t\left[\varepsilon_{t+1}\right] \qquad \varepsilon_t \sim N(0, I) \tag{7}$$

$$log(\sigma_t) = (1-\rho)\bar{\sigma} + \rho\, log(\sigma_{t-1}) + \psi_t \qquad \psi_t \sim N(0, \Sigma) \tag{8}$$

There are two transition equations: the first is the VAR described in (1) in its structural representation, with $\theta$ the matrix of contemporaneous coefficients. The second is the law of motion of the structural shocks' log standard deviation, which is an AR(1) process with long-term mean $\bar{\sigma}$ and autoregressive coefficient $\rho$.

As for the measurement equation, it is trivial:

$$\begin{bmatrix} \mathbb{E}_t\left[y_{t+h}^{OBS}\right] \\ \mathbb{E}_t\left[y_{t+h-1}^{OBS}\right] \\ \vdots \\ y_t^{OBS} \end{bmatrix} = Z \begin{bmatrix} \mathbb{E}_t\left[y_{t+h}\right] \\ \mathbb{E}_t\left[y_{t+h-1}\right] \\ \vdots \\ y_t \end{bmatrix} \tag{9}$$

where $Z = I_{N(h+1)}$, as there is a one-to-one relationship between model and observations, and the last row does not have the expectations operator, because it includes only data, known to the respondent at time $t$. Both $y_t^{OBS}...y_{t+h}^{OBS}$ and $y_t...y_{t+h}$ are expressed in log-differences. The Kalman filter is used to fill in missing observations for those quarters when a respondent did not provide a forecast.

# 3 Identification of structural shocks

The issue of identification in VARs is well-known and concerns any study aiming at providing a structural interpretation to model shocks, including the present one. I exploit the statistical properties of the data and use identification via stochastic volatility. This method is a special case of identification via time-varying volatility of Lewis (2021), himself building on the traditional identification through heteroskedasticity of Rigobon (2003). Other recent works using it are Bertsche and Braun (2022) and Chan et al. (2021). The common thread in these papers is the use of data characteristics as opposed to imposing assumptions, as in other popular techniques such as recursive identification or sign restrictions. The main advantage of this is avoiding the use of external information; the biggest drawback is the fact that while coefficients of the matrix of contemporaneous relationships $A_0$ are uniquely identified up to sign and column ordering, the shocks are not automatically "labelled", but need to be named ex-post, for example from looking at the impulse responses. In a bivariate system the labelling is relatively straightforward, as described in Section 6.

## 3.1 Identification via stochastic volatility

Identification via heteroskedasticity (and its variants) exploit changes in the variance of reduced form innovations across regimes or over time. While holding contemporaneous coefficients constant, a change in structural shock variances implies a change in the variance of innovations. The higher number of available moments allows to identify a higher number of coefficients, without any additional assumption. A condition for this identification to work, besides the presence of heteroskedasticity, is that structural variances do not change proportionately. Failure to satisfy this condition implies weak identification, as described in Lewis (2022).

I test for heteroskedasticity in the data by using the code provided by Lewis in his code for Lewis (2021), which computes the Cragg and Donald (1996) test for identification based on the first autocovariance of the outer product of residuals. The null hypothesis of zero autocovariance is rejected at the 10% level.

# 4 Model with uncertainty data

In macroeconomics is not common to have a direct measure of uncertainty surrounding expectations, despite increasing interest of policy makers and other agents in risk assessments. Some surveys, however, do collect density forecasts directly from respondents. While is not obvious how well this measure captures actual economic uncertainty, and several papers found it to be inaccurate or misspecified (Clements, 2018; Rossi and Sekhposyan, 2014; Glas and Hartmann, 2022; Bańbura et al., 2021), it is a wealth of information available in panel format for several variables, and gives a direct look into the forecast variance of survey respondents.

## 4.1 Measures of subjective uncertainty

The Philadelphia Fed SPF is among those surveys which require not only a point forecast, but also a probability assessment. More specifically, respondents are asked to assign probabilities between zero and one to one or more pre-determined bins, whose limits are the values the target variable may assume. The horizons for which they respond are fixed events, corresponding to the current and next calendar years. There are two issues to deal with before including observed probabilities into our model: the first is how to transform a discrete probability assessment into a continuous distribution, and the second how to connect the distribution to the forecast variances of the model, which are of the fixed horizon type ($h$-step ahead).

Several papers deal with the first issue, fitting a continuous distribution of some shape to a histogram whose within-bin distribution is not known. The most common approach is to assume the mass in each bin is concentrated in its mid-point (Rich and Tracy, 2010). Similarly, others assume a uniformly distributed mass (Abel et al., 2016). Fitting a normal (Giordani and Söderlind, 2003; Boero et al., 2015) or generalized beta (Engelberg et al., 2009) distribution are other widely used methods. For the normal or generalized beta method to work, it is necessary to have non-zero probabilities in at least three bins, as pointed out by Engelberg et al. (2009). Although I try all methods, for parsimony I select only two of them, uniform and generalized beta, the latter only calculated for quarters where at least three bins are assigned positive probability. Results for the four measures of uncertainty for the two variables of interest, real GDP and GDP deflator, are in Figure 2.[4]

The first thing to note is that the mass at midpoint and uniformly distributed probabilities method over-estimate uncertainty compared to the normal and generalized beta methods, as previously found in the literature (e.g. in Giordani and Söderlind, 2003). A second point regards the uncertainty of the GDP deflator: there is a structural break in the sample in 2014, where uncertainty drops markedly. This is due to an adjustment in the bin width, which goes from a percentage point to half a percentage point. While we could expect respondents to maintain their degree of uncertainty by assigning a positive probability to double the number of bins, what happened in fact was that most respondents kept using the same number of bins, resulting in a lower average uncertainty (as found also by Glas and Hartmann, 2022).

The literature deals less with the second issue of transforming fixed event in fixed horizon forecasts: exceptions include Ganics et al. (2020) for density, Dovern et al. (2012) and Knüppel and Vladu (2016) for point forecasts. I include in the model a second measurement equation, which makes the fixed event uncertainty comparable with the model's fixed

[4]The COVID period is excluded from this Figure for better readability, given its much larger scale. A Figure including COVID period is in the appendix.

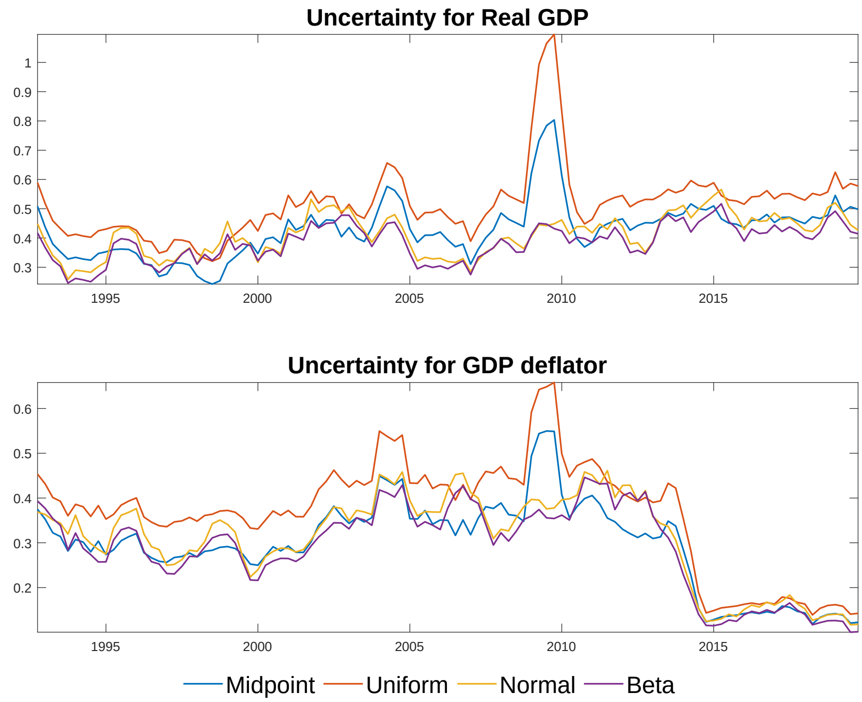


Figure 2: Aggregate uncertainty, given by the average of individual uncertainties, calculated from individual histograms and according to the methods in the legend: mass at midpoint, uniformly distributed mass, fitting a normal, fitting a generalized beta.

horizon forecast variance. Detailed results are available in appendix, but the intuition is that the forecast horizon varies depending on the quarter in which the survey is done. So for a survey done in the first quarter of the year, the corresponding horizon is four quarters, given that no information for the current year is available at that point in time. In the last quarter of the year, when asked for a current year assessment, the respondents only have one quarter to forecast, given that the previous three are known.

A final note on observed uncertainty: it is possible to derive two different measures of it from individual probability assessments: *aggregate* uncertainty and *average* uncertainty. The first is calculated by measuring the dispersion of the aggregate histogram of all respondents. The second comes from taking individual dispersions and averaging them over all respondents. Wallis (2005) proves that the first measure is equal to the sum of the second one and disagreement. For the purposes of this paper, the second measure is more precise as it accounts only for average individual uncertainty and is the selected one.

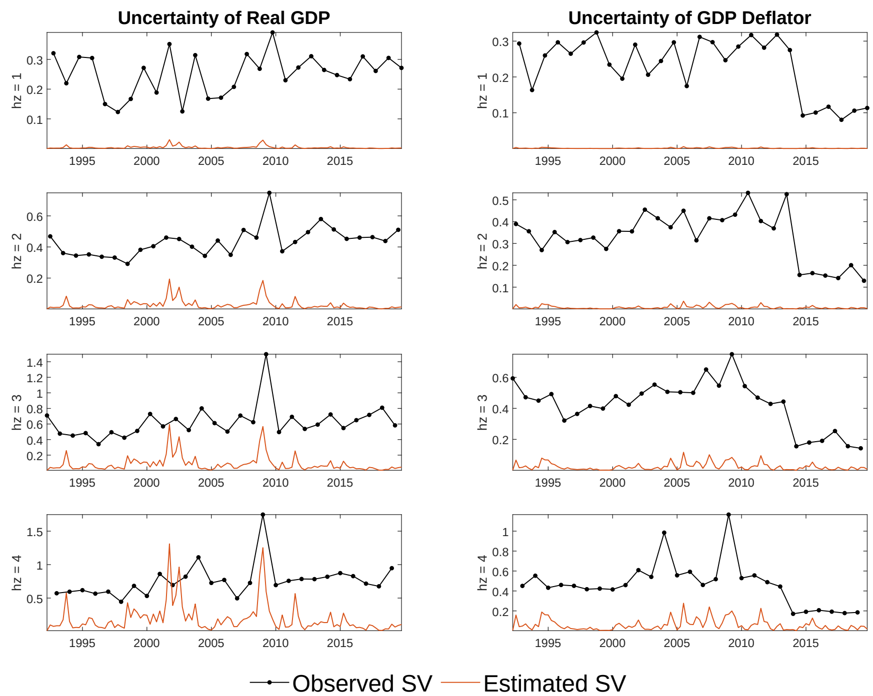


Figure 3: Aggregate observed uncertainty, calculated from assigning a uniform distribution to each bin assigned positive probability, and model implied uncertainty, for each of the four horizons.

## 4.2 Including subjective uncertainty in the baseline model

Figure 3 compares a measure of observed variance (the one obtained from assuming a uniform distribution within each bin's interval) with the variance implied by the model described in Section 2. Each column corresponds to one variable and each row to one forecast horizon. The black line has one observation per year, due to the fixed event nature of the survey question. I find that at shorter horizons uncertainty of SPF respondents is larger than the one estimated in the model, consistent with the findings of Clements and Galvão (2017), but most importantly that the two variances differ substantially. For these reasons, I run again the same model of Section 2, now including information on observed uncertainty. I do so by first, approximating the discrete survey responses with continuous distributions, and, second, mapping the observed uncertainty, which is a fixed event measure, to the fixed horizon forecasts of the model. The measurement equation allowing me to map each measure and model uncertainty is the following:

$$V_1^{OBS}(y_t) = d_{v,1} + Z_v \sigma_t + W_{v,1} u_t \tag{10}$$

$$V_2^{OBS}(y_t) = d_{v,2} + Z_v \sigma_t + W_{v,2} u_t \tag{11}$$

where $d_v$ is a level bias, $Z_v$ is a measurement matrix, $\sigma_t$ is the model time-varying variance, and $u_t$ is a measurement error with standard deviation $W_v$.

Figure 4 compares the uncertainty provided by respondents to the two model implied ones, for the two variables at each of the four horizons. The red line is the model uncertainty for the model including observed volatility, the yellow is the one for the model without observed volatility (the baseline model of Section 2). Both lines are adjusted for the estimated level bias $d_v$. While the two lines are fairly similar, there are specific periods during which the addition of observations has affected the volatility estimate: most notably, the early 2000s and Great Recession period for GDP, 2004 and the Great Recession for the deflator.

# 5 Data and Estimation

## 5.1 Data

The Survey of Professional Forecasters (SPF) is a quarterly survey of experts, conducted in the United States by the Philadelphia Fed around the middle of each quarter. At the time of filling in the survey, all respondents have access to the Bureau of Economic Analysis's

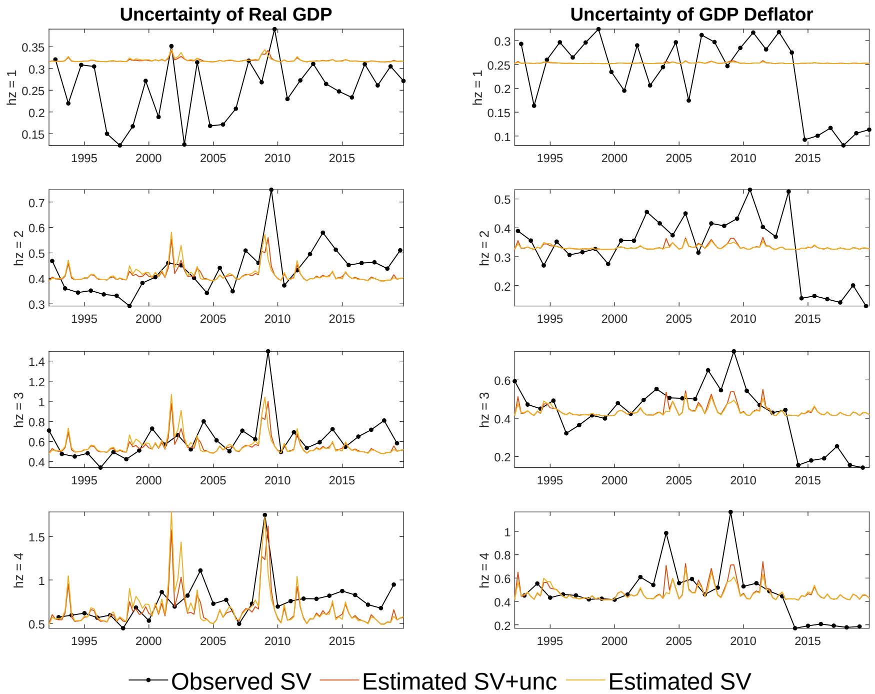


Figure 4: Aggregate observed uncertainty, calculated assuming a uniform distribution within each bin, fitted uncertainty implied by the model with observed volatility (red line), and fitted uncertainty implied by the model without observed volatility (yellow line).

(BEA) advance report of the national income and product accounts (NIPA), which contains the first estimate for GDP of the previous quarter. I use individual observations both of point forecasts (in levels) and probabilities (in annual growth rates) for real GDP and the GDP deflator, from the first quarter of 1992 to the first quarter of 2022. I exclude older data due to poorer availability and differences in target variables.

For each survey round $q$, I observe data from $q-1$ to $q+4$, for a total of six observations for each variable at each quarter.[5] The observation for quarter $q-1$ is effectively data: it represents the information available to all forecasters at the time of replying to the survey. In rare cases, some forecasters assume the figure will be revised and report a different number than everyone else in their responses, but here I assume this data point is the

[5]For part of the sample, higher horizons ($q+5$ to $q+8$) are available, but I do not use them here.

same for all forecasters, and that is the data point from which they start producing their forecasts. Given it does not include revisions, it is a real time observation. The observations from $q$ to $q+4$, instead, are "classic" nowcasts and forecasts.

I transform the point forecasts (and data) from level to quarter-on-quarter growth rates by taking first log-differences. When calculating growth rates for the data, I adjust for the level shifts caused by changes in the base year over the sample. To do so, I regress the data on dummies for the base years and on their growth rate available in the St. Louis FRED database, then subtract the dummy coefficients from the initial growth rate. Finally, I check for outliers in the individual data and assign missing values to observations larger than three times the standard deviation of each variable at each horizon. The reason for this is to avoid the possibility of misreporting, however, I do not adjust observations for particularly turbulent periods, since in those instances outliers might be simply due to a forecaster being much more optimistic (or pessimistic) than the average of respondents.

## 5.2 Estimation

**Particle filter and Rao-Blackwellization** I estimate the model with a frequentist approach. There are two considerations to take into account with regards to estimation: the non linear nature of the variance process, and the specific form of the model, the "restricted VARMA" specification described in Section 2.

Taken together, these two considerations call for the use of a particle filter: first, because the likelihood function of the variance conditional on all the other parameters is non-linear and cannot be derived analytically, and second, because the presence of stochastic volatility makes it harder to linearize the model (see, e.g. Zhang et al., 2015).

Given the computational intensity of the particle filter, I adjust it with a technique called "Rao-Blackwellization" (see, e.g. Hostettler and Särkkä, 2018), which allows to separate the problem into two parts, a linear and a non-linear one. This way, I can estimate the state equation for the point forecasts with the traditional Kalman filter, run it for each particle, then use the results to inform the estimation of the non-linear part.

**Maximization** Given the complexity of my objective function, in order to estimate optimal parameters I use the Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) optimization routine. The basic idea behind the routine is to approximate the inverse of the objective function's Hessian matrix. This avoids the use of gradients or second order derivatives and relies instead on simulation, as described by Andreasen (2010) in an application of the algorithm to DSGE models.

# 6 Results

## 6.1 Average of forecasters

The first set of results looks at average forecasts, representing the "consensus" at each point in time, and average uncertainty, calculated (as described in Section 4) as the average of individual standard deviations. This data set does not suffer from the issue of missing observations, and although it loses the dimension of heterogeneity across respondents, its analysis is of interest as it returns insights about the "representative" forecaster. As mentioned in Section 3, identification via stochastic volatility does not allow to name shocks ex ante. However, the sign of impulse responses can help (see Figure 5): for all the specifications, the response of inflation (the GDP deflator) to a shock increasing GDP by one unit is slightly positive, although close to zero, while the response of GDP to a shock raising inflation by one unit is negative. From these considerations, I label the first shock, moving GDP and inflation in the same direction, *demand*, and the second one *supply*. Finding a very small response of inflation to changes in GDP might reflect the characteristics of this variable over the sample considered. Estimating the model excluding the COVID period returns an even smaller coefficient for inflation, slightly negative.

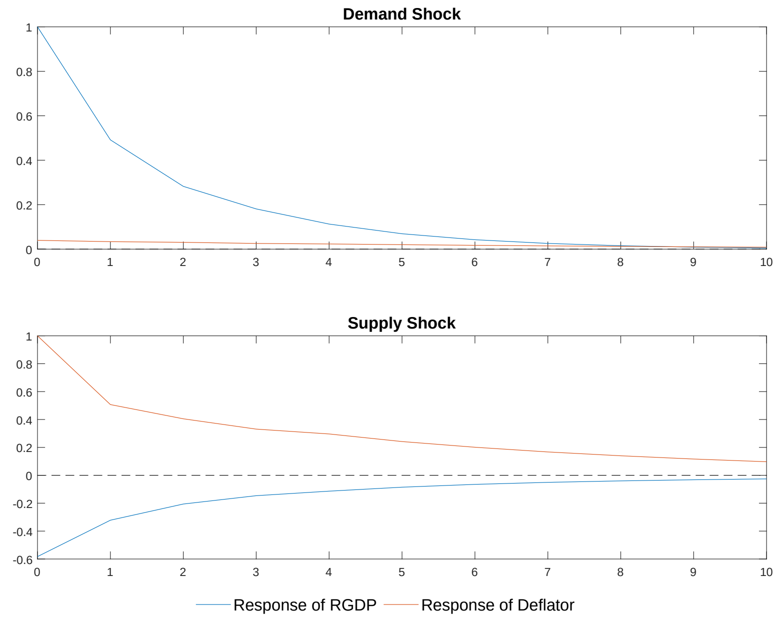


Figure 5: Impulse response functions for the average forecaster, response to a one unit shock.

#### 6.1.1 Historical shock decompositions

I first look at historical shock decompositions in the standard sense, namely I calculate to what extent each shock contributes to a variable's dynamics over the sample (Figure 6). I find negative demand shocks for GDP during the early 2000s, the Great Recession and at the beginning of the COVID period, with positive demand shocks counterbalanced by negative supply ones in the second half of the pandemic. For the GDP deflator, supply shocks dominate the sample, although demand contributes negatively to prices both during the Great Recession and in the first half of the pandemic.

Figure 7 shows a slightly more complex variance decomposition: the solid bars are similar to those of Figure 6, namely they represent the contribution of each shock at time $t$ to the dynamics of GDP and of the deflator respectively for the one quarter ahead and four quarter ahead forecasts. The shaded bars, on the other hand, represent the contribution of judgement to the forecasts, in other words the expected shock for that horizon and all the


Figure 6: Historical shock decomposition, y-o-y growth rates. The black line plots the real-time data.

previous expected shocks. In the setting of Section 2, the solid bars represent the current shocks $\varepsilon_{t|t}$, while the shaded bars are the sums of all expected future shocks, $\varepsilon_{t+1|t}$, ..., $\varepsilon_{t+4|t}$.

To highlight a couple of specific results, the first and second panels show how judgement regarding GDP growth at the beginning of the COVID pandemic changed from the four quarter ahead to the one quarter ahead forecast: at first, the average forecaster expected a positive demand shock to raise GDP, but later on, the expected demand shocks were negative. Expected shocks to the deflator have also affected forecasts considerably: the third panel shows that in the most recent period, as of the end of 2020, forecasters have been expecting a positive supply shock to bring inflation down, while actual supply shocks to inflation have consistently been positive.

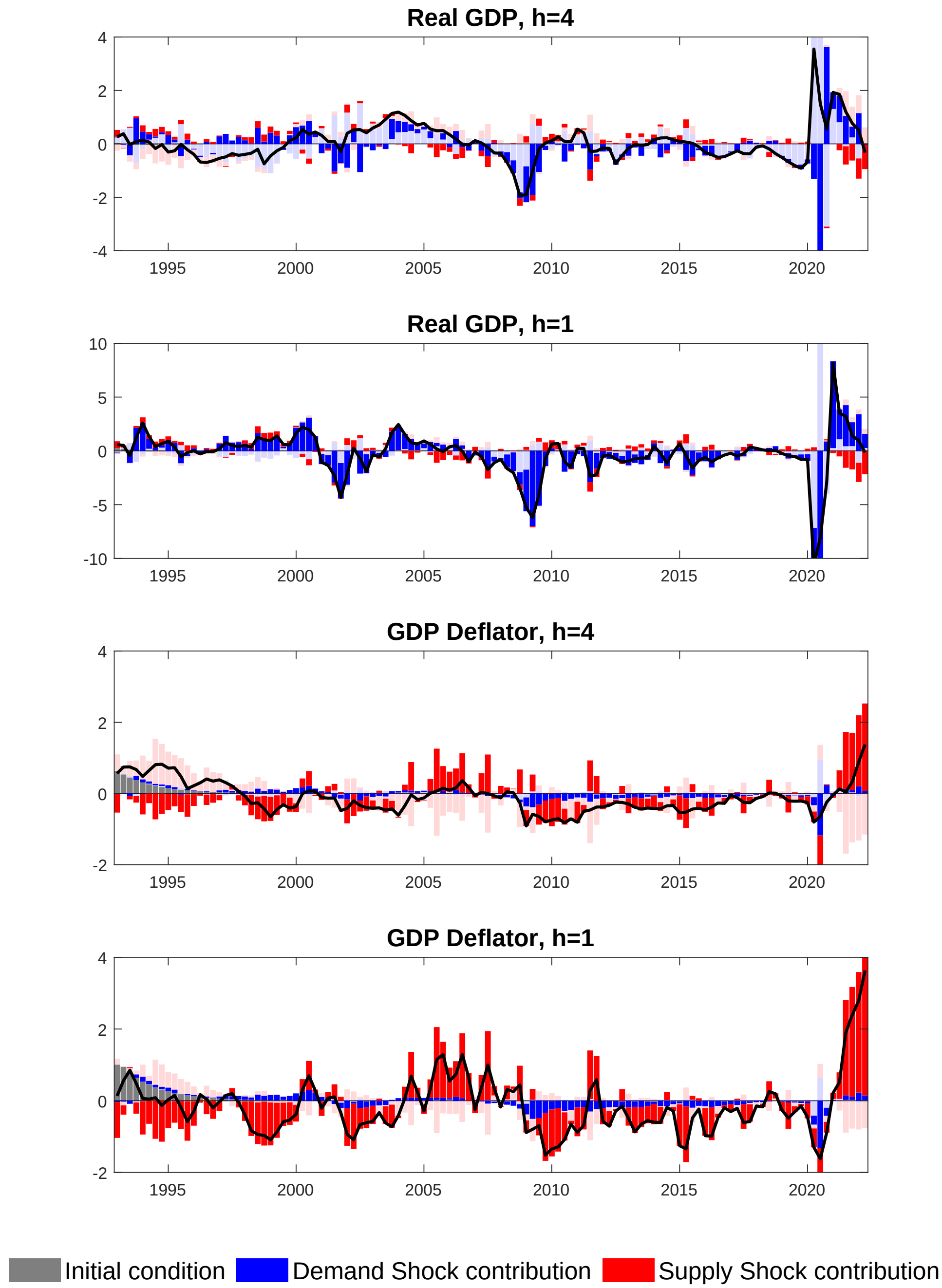


Figure 7: Historical shock decomposition of the forecast at each horizon, average forecaster. Solid bars represent current shocks, while shaded bars are the expected shocks ("judgement"). Horizon $h = 0$ represent the nowcast, $t - 1$ is only data.

### 6.1.2 Forecast error variance decomposition

Another way to look at results is by investigating the forecast error variance decomposition. This is the time-varying share of forecast variance due to each of the two shocks over the sample. Recall the model in state space form from Section 2:

$$\mathbb{E}_t\left[y_{t+1}\right] = c + \phi\mathbb{E}_{t-1}\left[y_t\right] + \theta\sigma_{t+1}\mathbb{E}_t\left[\varepsilon_{t+1}\right] \qquad \varepsilon_t \sim N(0, I) \tag{12}$$

$$log(\sigma_t) = (1-\rho)\bar{\sigma} + \rho\, log(\sigma_{t-1}) + \psi_t \qquad \psi_t \sim N(0, \Sigma) \tag{13}$$

The forecast variance at horizon $h = 1$ is:

$$\mathbb{E}_t\left[y_t y_t'\right] = \mathbb{E}_t\left[\theta\sigma_t\varepsilon_t\varepsilon_t'\sigma_t'\theta'\right] = \theta\mathbb{E}_t\left[\sigma_t\sigma_t'\right]\theta' = \theta\begin{bmatrix}\sigma^2_{1,t} & 0\\ 0 & \sigma^2_{2,t}\end{bmatrix}\theta' \tag{14}$$

where the last equality comes from the assumption $\mathbb{E}_t\left[\varepsilon_t\varepsilon_t'\right] = I$. To get the variance decomposition, I set one of the two $\sigma^2_t$ equal to zero throughout the sample, and re-simulate the path of the two unconditional variances without one of the two shocks. I then repeat for the second $\sigma^2_t$.

With this calculation I obtain Figure 8, where is clear the variance of the deflator is scarcely affected by demand shocks, with some notable exceptions: the early 2000s, 2014, and the COVID pandemic. The variance of GDP, on the other hand, has been influenced by supply shocks throughout the sample, to a larger extent in the mid '90s and early 2000s, then again in the latter half of the recent pandemic.

## 6.2 Pooled panel and individual forecasters

The second set of results accounts for individual differences by estimating a VAR in the form of a pooled panel, where, although coefficients are the same for the entire cross section, all individual information is used for the estimation, resulting in stochastic volatility and

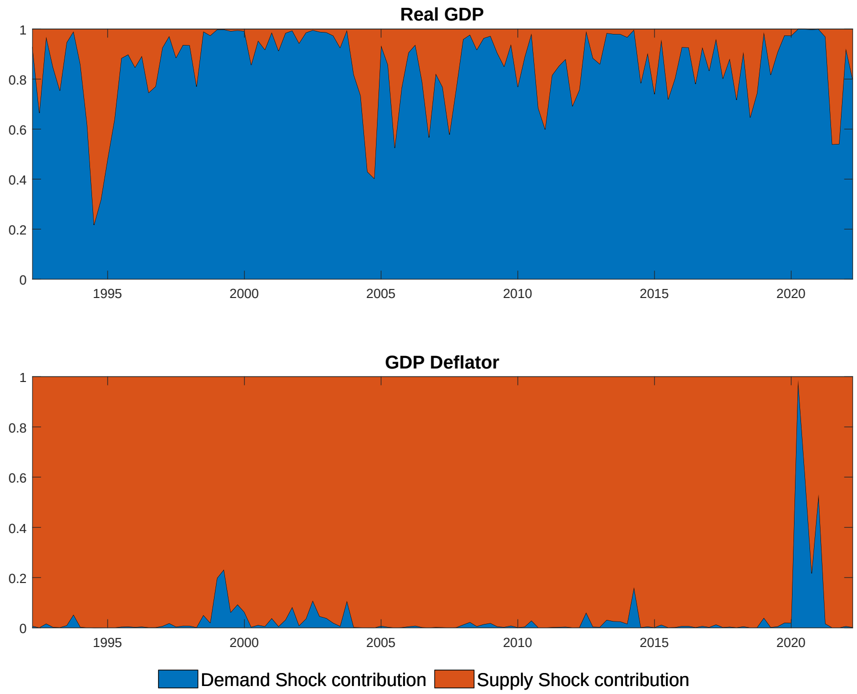


Figure 8: Forecast error variance decomposition over the sample.

shocks series for each respondent. This allows to analyze differences across respondents in their forecasts and judgement due to the differences in uncertainty (given the coefficients are constrained to be equal across respondents). Heterogeneity in uncertainty is clearly visible in Figure 9, which plot $16^{th}$, $50th$ and $84^{th}$ quantiles of respondents for the uncertainty surrounding both variables at each forecast horizon. The distribution of respondents is skewed to the left, with half of the respondents reporting a low uncertainty throughout the sample, and a much larger dispersion of uncertainty for the upper quantile.

# 7 Conclusions

In this paper, I characterize structurally professional forecasters responses by running a VAR model on their forecasts and the data they all observe. I assume respondents perform

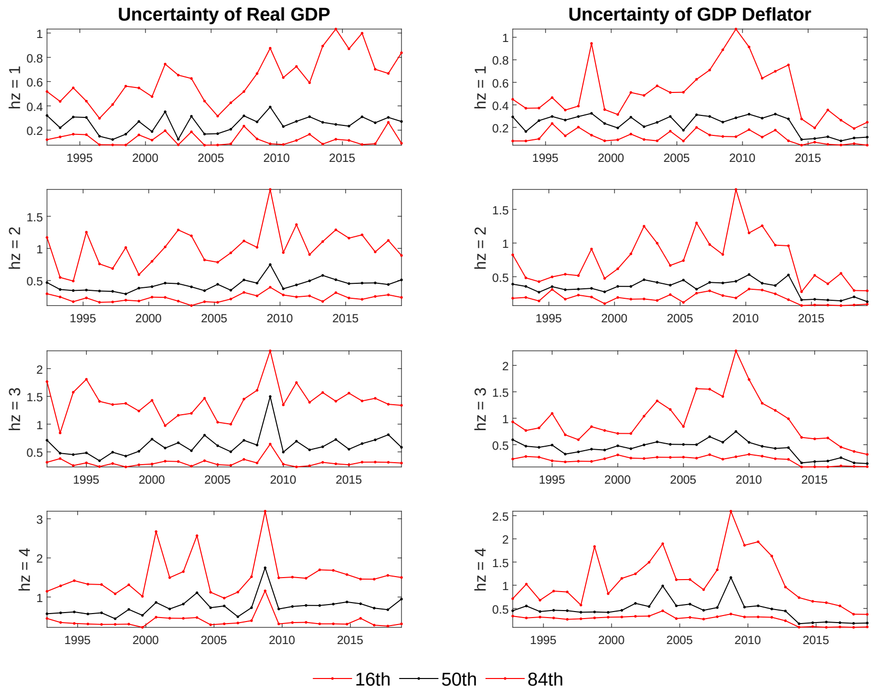


Figure 9: Observed uncertainty per variable and horizon, percentiles across respondents.

a conditional exercise and therefore expect different structural shocks to hit the economy over the forecast horizon. After estimating a VAR(MA) model for both the average and individual respondents, I identify shocks by exploiting variation in their volatility. Since I observe a discrepancy between the probability forecasts provided in the survey and the model implied forecast variance, I add the former to the model and re-estimate it, to check for any change in the coefficients or other results. While I find that for the most part adding observed probability forecasts to the model does not affect results, there are some specific periods in the sample where the model re-adjusts its volatility based on the new information.

I then observe forecast variances and historical shock decompositions, in order to analyze results in greater detail.

....

While I do not apply my findings to a structural model, it would be interesting to do so in order to inform more theoretical settings. Other possible extensions include applying this setup to different surveys and countries, adding more variables in order to identify more shocks, and shifting to a Bayesian setting to simplify estimation.

# References


Abel, J., Rich, R., Song, J., and Tracy, J. (2016). The Measurement and Behavior of Uncertainty: Evidence from the ECB Survey of Professional Forecasters. *Journal of Applied Econometrics*, 31(3):533–550.

Andrade, P., Crump, R. K., Eusepi, S., and Moench, E. (2016). Fundamental disagreement. *Journal of Monetary Economics*, 83:106–128.

Andrade, P. and Le Bihan, H. (2013). Inattentive professional forecasters. *Journal of Monetary Economics*, 60(8):967–982.

Andre, P., Haaland, I., Roth, C., and Wohlfart, J. (2021). Inflation narratives. *CEPR Discussion Paper No. DP16758.*

Andre, P., Pizzinelli, C., Roth, C., and Wohlfart, J. (2022). Subjective Models of the Macroeconomy: Evidence From Experts and Representative Samples. *The Review of Economic Studies.*

Andreasen, M. M. (2010). How to maximize the likelihood function for a DSGE model. *Computational Economics*, 35(2):127–154.

Antolin-Diaz, J., Petrella, I., and Rubio-Ramirez, J. F. (2021). Structural scenario analysis with SVARs. *Journal of Monetary Economics*, 117:798–815.

Aruoba, B. S. (2020). Term structures of inflation expectations and real interest rates. *Journal of Business & Economic Statistics*, 38(3):542–553.

Bańbura, M., Brenna, F., Paredes, J., and Ravazzolo, F. (2021). Combining Bayesian VARs with survey density forecasts: does it pay off? *ECB Working Paper*, 2021/2543.

Bertsche, D. and Braun, R. (2022). Identification of Structural Vector Autoregressions by Stochastic Volatility. *Journal of Business & Economic Statistics*, 40(1):328–341.

Binder, C., McElroy, T. S., and Sheng, X. S. (2022). The term structure of uncertainty: New evidence from survey expectations. *Journal of Money, Credit and Banking*, 54(1):39–71.

Boero, G., Smith, J., and Wallis, K. F. (2015). The Measurement and Characteristics of Professional Forecasters' Uncertainty. *Journal of Applied Econometrics*, 30(7):1029–1046.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2020). Overreaction in Macroeconomic Expectations. *American Economic Review*, 110(9):2748–2782.

Born, B., Dovern, J., and Enders, Z. (2020). Expectation Dispersion, Uncertainty, and the Reaction to News. *SSRN Electronic Journal*.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *the Quarterly Journal of economics*, 118(1):269–298.

Casey, E. (2021). Are professional forecasters overconfident? *International Journal of Forecasting*, 37(2):716–732.

Chan, J. C., Koop, G., and Yu, X. (2021). Large order-invariant Bayesian VARs with stochastic volatility. *arXiv preprint arXiv:2111.07225*.

Clark, T. E., McCracken, M. W., and Mertens, E. (2020). Modeling time-varying uncertainty of multiple-horizon forecast errors. *Review of Economics and Statistics*, 102(1):17–33.

Clements, M. (2018). Are macroeconomic density forecasts informative? *International Journal of Forecasting*, 34(2):181–198.

Clements, M. (2022). Forecaster efficiency, accuracy, and disagreement: Evidence using individual-level survey data. *Journal of Money, Credit and Banking*, 54(2-3):537–568.

Clements, M. and Galvão, A. B. (2017). Model and survey estimates of the term structure of us macroeconomic uncertainty. *International Journal of Forecasting*, 33(3):591–604.

Clements, M., Rich, R., and Tracy, J. (2022). Surveys of Professionals. *Federal Reserve Bank of Cleveland Working Papers*, 22(13).

Coibion, O. and Gorodnichenko, Y. (2012). What Can Survey Forecasts Tell Us about Information Rigidities? *Journal of Political Economy*, 120(1):116–159.

Coibion, O. and Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review*, 105(8):2644–2678.

Coibion, O., Gorodnichenko, Y., and Kamdar, R. (2018). The Formation of Expectations, Inflation, and the Phillips Curve. *Journal of Economic Literature*, 56(4):1447–1491.

Del Negro, M., Casarin, R., and Bassetti, F. (2022). A Bayesian Approach to Inference on Probabilistic Surveys. *SSRN Electronic Journal*.

Dovern, J. (2015). A multivariate analysis of forecast disagreement: Confronting models of disagreement with survey data. *European Economic Review*, 80:16–35.

Dovern, J., Fritsche, U., and Slacalek, J. (2012). Disagreement among forecasters in G7 countries. *Review of Economics and Statistics*, 94(4):1081–1096.

Dovern, J. and Hartmann, M. (2017). Forecast performance, disagreement, and heterogeneous signal-to-noise ratios. *Empirical Economics*, 53(1).

Engelberg, J., Manski, C. F., and Williams, J. (2009). Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters. *Journal of Business & Economic Statistics*, 27(1):30–41.

Farmer, L., Nakamura, E., and Steinsson, J. (2021). Learning about the long run. Technical report, National Bureau of Economic Research.

Ganics, G., Rossi, B., and Sekhposyan, T. (2020). From fixed-event to fixed-horizon density forecasts: Obtaining measures of multi-horizon uncertainty from survey density forecasts. *Banco de España Working Paper*.

Giacomini, R., Skreta, V., and Turen, J. (2020). Heterogeneity, Inattention, and Bayesian Updates. *American Economic Journal: Macroeconomics*, 12(1):282–309.

Giordani, P. and Söderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, 47(6):1037–1059.

Glas, A. (2020). Five dimensions of the uncertainty–disagreement linkage. *International Journal of Forecasting*, 36(2):607–627.

Glas, A. and Hartmann, M. (2022). Uncertainty measures from partially rounded probabilistic forecast surveys. *Quantitative Economics*, 13:979–1022.

Herbst, E. and Winkler, F. (2021). The Factor Structure of Disagreement. *Finance and Economics Discussion Series*, 2021(044):1–76.

Hostettler, R. and Särkkä, S. (2018). Rao–blackwellized gaussian smoothing. *IEEE Transactions on Automatic Control*, 64(1):305–312.

Huang, R., Pilbeam, K., and Pouliot, W. (2022). Are macroeconomic forecasters optimists or pessimists? A reassessment of survey based forecasts. *Journal of Economic Behavior & Organization*, 197:706–724.

Knüppel, M. and Vladu, A. (2016). Approximating fixed-horizon forecasts using fixed-event forecasts. *Bundesbank Discussion Paper*.

Kohlhas, A. N. and Walther, A. (2021). Asymmetric Attention. *American Economic Review*, 111(9):2879–2925.

Krane, S. D. (2011). Professional forecasters' view of permanent and transitory shocks to GDP. *American Economic Journal: Macroeconomics*, 3(1):184–211.

Lewis, D. J. (2021). Identifying Shocks via Time-Varying Volatility. *The Review of Economic Studies*, 88(6):3086–3124.

Lewis, D. J. (2022). Robust inference in models identified via heteroskedasticity. *Review of Economics and Statistics*, 104(3):510–524.

Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. *NBER macroeconomics annual*, 18:209–248.

Patton, A. J. and Timmermann, A. (2010). Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, 57(7):803–820.

Ricco, G., Callegari, G., and Cimadomo, J. (2016). Signals from the government: Policy disagreement and the transmission of fiscal shocks. *Journal of Monetary Economics*, 82:107–118.

Rich, R. and Tracy, J. (2010). The Relationships among Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts. *The Review of Economics and Statistics*, 92(1):200–207.

Rich, R. and Tracy, J. (2021). A Closer Look at the Behavior of Uncertainty and Disagreement: Micro Evidence from the Euro Area. *Journal of Money, Credit and Banking*, 53(1):233–253.

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics*, 85(4):777–792.

Rossi, B. and Sekhposyan, T. (2014). Evaluating predictive densities of us output growth and inflation in a large macroeconomic data set. *International Journal of Forecasting*, 30(3):662–682.

Wallis, K. F. (2005). Combining density and interval forecasts: a modest proposal. *Oxford Bulletin of Economics and Statistics*, 67:983–994.

Zhang, X.-P., Khwaja, A. S., Luo, J.-A., Housfater, A. S., and Anpalagan, A. (2015). Multiple Imputations Particle Filters: Convergence and Performance Analyses for Nonlinear State Estimation With Missing Data. *IEEE Journal of Selected Topics in Signal Processing*, 9(8):1536–1547.

# Appendix

# A A state-space model with stochastic volatility

Measurement and state equation for point forecasts:

$$y_t = Zx_t \tag{1}$$

$$x_t = c + Tx_{t-1} + R_t\sigma_t\varepsilon_t \qquad \varepsilon_t \sim N(0, I) \tag{2}$$

Where $Z = [I_N \ 0_N]$ is a matrix connecting observed forecasts to model forecasts and $x_t$ is an unobserved state variable following an $AR(1)$ process with drift $c$. $T$ is the time-varying matrix of coefficients and $R_tR_t'$ is the time varying variance covariance matrix of the standard normally distributed disturbances, $\varepsilon_t$.

Measurement and state equation for forecast variances:

$$V(y_t) = d_v + Z_v s_t + W_v u_t \qquad u_t \sim N(0, I) \tag{3}$$

$$log(\sigma_t) = (1-\rho)\bar{\sigma} + \rho\, log(\sigma_{t-1}) + \psi_t \qquad \psi_t \sim N(0, \Sigma) \tag{4}$$

Where $d_v$ is a bias, $W_vW_v'$ is the variance covariance matrix of the measurement error $u_t$ and $Z_v$ is the measurement matrix connecting observed variances and model variances. The unobserved state variable $s_t$ is the log standard deviation of $\varepsilon_t$ and it follows an $AR(1)$ process without drift, and has shocks with variance covariance matrix $R_vR_v'$, $R_v = \begin{bmatrix} \sigma_{u1} & 0 \\ 0 & \sigma_{u2} \end{bmatrix}$.

## A.1 Conditional forecasts at higher horizons

Recall the VAR defined in (5) in the main text, here expressed in structural form: **fix notation of next part**

$$x_t = c + Tx_{t-1} + A_0^{-1}\lambda_t^{1/2}\varepsilon_t \tag{5}$$

with $\lambda_{t+1} = (1-\rho)\bar{\lambda} + \rho\lambda_t + u_t$, $u_t \sim N(0, \Sigma_u)$. The conditional forecast at horizon $h = 1$ is equal to:

$$E_t[x_{t+1}] = c + Tx_t + E_t\left[A_0^{-1}\lambda_{t+1}^{1/2}\varepsilon_{t+1}\right] = c + Tx_t + A_0^{-1}\lambda_{t+1|t}^{1/2}\varepsilon_{t+1|t} \tag{6}$$

At horizon $h = 4$:

$$E_t\left[x_{t+4}\right] = (1+T+T^2+T^3)c + T^4 x_t + A_0^{-1}\lambda_{t+4|t}^{1/2}\varepsilon_{t+4|t} + TA_0^{-1}\lambda_{t+3|t}^{1/2}\varepsilon_{t+3|t} + T^2A_0^{-1}\lambda_{t+2|t}^{1/2}\varepsilon_{t+2|t} + T^3A_0^{-1}\lambda_{t+1|t}^{1/2}\varepsilon_{t+1|t} \quad (7)$$

$$E_{t-1}\left[x_{t+3}\right] = (1+T+T^2+T^3)c + T^4 x_{t-1} + A_0^{-1}\lambda_{t+3|t-1}^{1/2}\varepsilon_{t+3|t-1} + TA_0^{-1}\lambda_{t+2|t-1}^{1/2}\varepsilon_{t+2|t-1} + T^2A_0^{-1}\lambda_{t+1|t-1}^{1/2}\varepsilon_{t+1|t-1} + T^3A_0^{-1}\lambda_{t|t-1}^{1/2}\varepsilon_{t|t-1} \quad (8)$$

Using 6, rewrite 7:

$$E_t\left[x_{t+4}\right] = (1+T+T^2+T^3)c + \boldsymbol{T^4 c} + \boldsymbol{T^5} x_{t-1} + A_0^{-1}\lambda_{t+4|t}^{1/2}\varepsilon_{t+4|t} + ... + T^3A_0^{-1}\lambda_{t+1|t}^{1/2}\varepsilon_{t+1|t} + \boldsymbol{T^4 A_0^{-1}\lambda_t^{1/2}\varepsilon_t} \quad (9)$$

From 8, solving for $T^4 x_{t-1}$:

$$T^4 x_{t-1} = E_{t-1}\left[x_{t+3}\right] - (1+T+T^2+T^3)c - A_0^{-1}\lambda_{t+3|t-1}^{1/2}\varepsilon_{t+3|t-1} - ... - T^3A_0^{-1}\lambda_{t|t-1}^{1/2}\varepsilon_{t|t-1} \quad (10)$$

Substitute 10 in 9:

$$E_t\left[x_{t+4}\right] = (1+T+T^2+T^3+T^4)c + TE_{t-1}\left[x_{t+3}\right] - (T+T^2+T^3+T^4)c - TA_0^{-1}\lambda_{t+3|t-1}^{1/2}\varepsilon_{t+3|t-1} - T^2A_0^{-1}\lambda_{t+2|t-1}^{1/2}\varepsilon_{t+2|t-1} - T^3A_0^{-1}\lambda_{t+1|t-1}^{1/2}\varepsilon_{t+1|t-1} - T^4A_0^{-1}\lambda_{t|t-1}^{1/2}\varepsilon_{t|t-1} + A_0^{-1}\lambda_{t+4|t}^{1/2}\varepsilon_{t+4|t} + TA_0^{-1}\lambda_{t+3|t}^{1/2}\varepsilon_{t+3|t} + T^2A_0^{-1}\lambda_{t+2|t}^{1/2}\varepsilon_{t+2|t} + T^3A_0^{-1}\lambda_{t+1|t}^{1/2}\varepsilon_{t+1|t} + T^4A_0^{-1}\lambda_t^{1/2}\varepsilon_t \quad (11)$$

Rearrange terms:

$$E_t\left[x_{t+4}\right] = c + TE_{t-1}\left[x_{t+3}\right] + A_0^{-1}\lambda_{t+4|t}^{1/2}\varepsilon_{t+4|t} + TA_0^{-1}\left(\lambda_{t+3|t}^{1/2}\varepsilon_{t+3|t} - \lambda_{t+3|t-1}^{1/2}\varepsilon_{t+3|t-1}\right) + T^2A_0^{-1}\left(\lambda_{t+2|t}^{1/2}\varepsilon_{t+2|t} - \lambda_{t+2|t-1}^{1/2}\varepsilon_{t+2|t-1}\right) + T^3A_0^{-1}\left(\lambda_{t+1|t}^{1/2}\varepsilon_{t+1|t} - \lambda_{t+1|t-1}^{1/2}\varepsilon_{t+1|t-1}\right) + T^4A_0^{-1}\left(\lambda_t^{1/2}\varepsilon_t - \lambda_{t|t-1}^{1/2}\varepsilon_{t|t-1}\right) \quad (12)$$

# B The full model in companion form

For explanatory purposes, below is the full companion form of a "toy" model with two variables, two forecast horizons and one lag. The companion form allows to clearly see the autoregressive part (the first two terms), the moving average part (the second and third term) and the structural shock (obtained by multiplying the diagonal of the last matrix by the last vector).

$$\begin{bmatrix} \mathbb{E}_t[y_{t+2}] \\ \mathbb{E}_t[\pi_{t+2}] \\ \mathbb{E}_t[y_{t+1}] \\ \mathbb{E}_t[\pi_{t+1}] \\ y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} & 0 & 0 & 0 & 0 \\ \phi_{21} & \phi_{22} & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbb{E}_{t-1}[y_{t+1}] \\ \mathbb{E}_{t-1}[\pi_{t+1}] \\ \mathbb{E}_{t-1}[y_t] \\ \mathbb{E}_{t-1}[\pi_t] \\ y_{t-1} \\ \pi_{t-1} \end{bmatrix} \tag{13}$$

$$+ \begin{bmatrix} -\phi_{11} & -\phi_{12} & -\phi_{11}^2 & -\phi_{12}^2 & 0 & 0 \\ -\phi_{21} & -\phi_{22} & -\phi_{21}^2 & -\phi_{22}^2 & 0 & 0 \\ -1 & 0 & -\phi_{11} & -\phi_{12} & 0 & 0 \\ 0 & -1 & -\phi_{21} & -\phi_{22} & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon^y_{t+1|t-1} \\ \varepsilon^\pi_{t+1|t-1} \\ \varepsilon^y_{t|t-1} \\ \varepsilon^\pi_{t|t-1} \\ \varepsilon^y_{t-1|t-1} \\ \varepsilon^\pi_{t-1|t-1} \end{bmatrix} \tag{14}$$

$$+ \begin{bmatrix} 1 & 0 & \phi_{11} & \phi_{12} & \phi_{11}^2 & \phi_{12}^2 \\ 0 & 1 & \phi_{21} & \phi_{22} & \phi_{21}^2 & \phi_{22}^2 \\ 0 & 0 & 1 & 0 & \phi_{11} & \phi_{12} \\ 0 & 0 & 0 & 1 & \phi_{21} & \phi_{22} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon^y_{t+2|t} \\ \varepsilon^\pi_{t+2|t} \\ \varepsilon^y_{t+1|t} \\ \varepsilon^\pi_{t+1|t} \\ \varepsilon^y_{t|t} \\ \varepsilon^\pi_{t|t} \end{bmatrix} \tag{15}$$

$$\begin{aligned} \mathbb{E}_t\left[y_{t+2}\right] = c_1 + \phi_{11}\mathbb{E}_{t-1}\left[y_{t+1}\right] + \phi_{12}\mathbb{E}_{t-1}\left[\pi_{t+1}\right] + \varepsilon^y_{t+2|t} + \phi_{11}\left(\varepsilon^y_{t+1|t} - \varepsilon^y_{t+1|t-1}\right) \\ + \phi_{11}^2\left(\varepsilon^y_{t|t} - \varepsilon^y_{t|t-1}\right) + \phi_{12}\left(\varepsilon^\pi_{t+1|t} - \varepsilon^\pi_{t+1|t-1}\right) + \phi_{22}^2\left(\varepsilon^\pi_{t|t} - \varepsilon^\pi_{t|t-1}\right) \end{aligned} \tag{16}$$

$$\mathbb{E}_t\left[y_{t+1}\right] = \mathbb{E}_{t-1}\left[y_{t+1}\right] + \varepsilon^y_{t+1|t} - \varepsilon^y_{t+1|t-1} + \phi_{11}\left(\varepsilon^y_{t|t} - \varepsilon^y_{t|t-1}\right) + \phi_{12}\left(\varepsilon^\pi_{t|t} - \varepsilon^\pi_{t|t-1}\right) \tag{17}$$

$$y_t = \mathbb{E}_{t-1}\left[y_t\right] - \varepsilon^y_{t|t-1} + \varepsilon^y_{t|t} = \phi y_{t-1} + \varepsilon^y_{t|t} \tag{18}$$

# C Forecast variance

## C.1 Measurement equations for forecast variance

Respondents assign to each bin a probability between zero and one. The extremes of the bins represent the values taken by the annual average growth rate of RGDP and PGDP, defined as:

$$y_{OBS} = \left( \frac{\left(x_t^{(1)} + x_t^{(2)} + x_t^{(3)} + x_t^{(4)}\right)\frac{1}{4}}{\left(x_{t-1}^{(1)} + x_{t-1}^{(2)} + x_{t-1}^{(3)} + x_{t-1}^{(4)}\right)\frac{1}{4}} - 1 \right) \cdot 100, \tag{19}$$

where superscripts in parentheses are the quarter of reference in year $t$. After transforming and taking logs on both sides,

$$\ln\left(\frac{y_{OBS}}{100} + 1\right) = \ln\left( \frac{\left(x_t^{(1)} + x_t^{(2)} + x_t^{(3)} + x_t^{(4)}\right)\frac{1}{4}}{\left(x_{t-1}^{(1)} + x_{t-1}^{(2)} + x_{t-1}^{(3)} + x_{t-1}^{(4)}\right)\frac{1}{4}} \right)$$

Applying the approximation used by Aruoba (2020) to go from arithmetic mean to geometric mean, and well-known properties of logarithms:

$$\begin{aligned} &\ln\left( \frac{\left(x_t^{(1)} x_t^{(2)} x_t^{(3)} x_t^{(4)}\right)^{\frac{1}{4}}}{\left(x_{t-1}^{(1)} x_{t-1}^{(2)} x_{t-1}^{(3)} x_{t-1}^{(4)}\right)^{\frac{1}{4}}} \right) = \\ &\quad \frac{1}{4}\left(\ln x_t^{(1)} + \ln x_t^{(2)} + \ln x_t^{(3)} + \ln x_t^{(4)}\right) - \frac{1}{4}\left(\ln x_{t-1}^{(1)} + \ln x_{t-1}^{(2)} + \ln x_{t-1}^{(3)} + \ln x_{t-1}^{(4)}\right) \end{aligned} \tag{20}$$

The measurement equation for the model in log-differences needs to link the observed variance, expressed in average annual growth rates, to the model variance, expressed in log-difference between quarters. To do so, I start from equation 20 and add and subtract quantities, to finally obtain:

$$
\begin{gathered}
V\left[\ln\left(\frac{x}{100}+1\right)\right] \simeq \\
\frac{1}{16}V\left[4\Delta\ln x_t^{(1)}+3\Delta\ln x_t^{(2)}+2\Delta\ln x_t^{(3)}+\Delta\ln x_t^{(4)}+\Delta\ln x_{t-1}^{(2)}+2\Delta\ln x_{t-1}^{(3)}+3\Delta\ln x_{t-1}^{(4)}\right]
\end{gathered}
\tag{21}
$$

where $\Delta\ln x_t^{(1)} = \ln x_t^{(1)} - \ln x_{t-1}^{(4)}$, $\Delta\ln x_t^{(2)} = \ln x_t^{(2)} - \ln x_t^{(1)}$, ...

This implies higher weight assigned to the 4-q ahead forecast, which is the farthest and most uncertain, and decreasing weight the closer the horizon. The variance in each quarter is the sum of the forecasts variances for the periods that are still unknown, so in quarter one all of them, in quarter four only the last one.

$$V_t^{(1)} \simeq \frac{1}{16}V\left[4\Delta\ln x_t^{(1)}+3\Delta\ln x_t^{(2)}+2\Delta\ln x_t^{(3)}+\Delta\ln x_t^{(4)}\right] \tag{22}$$

$$V_t^{(2)} \simeq \frac{1}{16}V\left[3\Delta\ln x_t^{(2)}+2\Delta\ln x_t^{(3)}+\Delta\ln x_t^{(4)}\right] \tag{23}$$

$$V_t^{(3)} \simeq \frac{1}{16}V\left[2\Delta\ln x_t^{(3)}+\Delta\ln x_t^{(4)}\right] \tag{24}$$

$$V_t^{(4)} \simeq \frac{1}{16}V\left[\Delta\ln x_t^{(4)}\right] \tag{25}$$