ICMA Centre

Discussion Paper

Individual Forecaster Perceptions of the Persistence of Shocks to GDP

January 2020

Michael P. Clements
ICMA Centre, Henley Business School, University of Reading

Discussion Paper Number:
ICM-2020-02

The aim of this discussion paper series is to disseminate new research of academic distinction. Papers are preliminary drafts, circulated to stimulate discussion and critical comment. Henley Business School is triple accredited and home to over 100 academic faculty, who undertake research in a wide range of fields from ethics and finance to international business and marketing.

admin@icmacentre.ac.uk

www.icmacentre.ac.uk

Individual Forecaster Perceptions of the Persistence of Shocks to GDP.

Michael P. Clements*
ICMA Centre,
Henley Business School,
University of Reading,
Reading RG6 6BA
m.p.clements@reading.ac.uk.

January 23, 2020

Abstract

We analyze individual professional forecasters' beliefs concerning the persistence of GDP shocks. Despite substantial apparent heterogeneity in perceptions, with around one half of the sample of professional forecasters believing shocks do not have permanent effects, we show that these apparent differences may be largely due to short-samples and survey respondents being active at different times. When we control for these effects, using a bootstrap, we formally do not reject the null that individuals' long-horizon expectations are interchangeable at a given point in time. When we apply the same bootstrap approach to their medium-term expectations, we do reject the null. We explore this difference between long and medium-horizon forecasts by decomposing revisions in forecasts into permanent and transitory components.

JEL Cleassification codes: C53, C55, C83, E32.

Keywords. Long-horizon forecasts, output persistence, heterogeneous beliefs, bootstrap test.

*Helpful comments and suggestions from three referees are gratefully acknowledged.

1 Introduction

In the years following the severe 2007:Q4 to 2009:Q2 recession (in the US) there has been much interest in the nature of output fluctuations. Are the major economies expected to enjoy faster than normal economic growth to make up for the lost output during the recession, or is some of the loss in output permanent? If the economy follows a trend path, subject to transient fluctuations, then a period of faster growth might be expected to prevail. Alternatively, there might be long-term costs to recessions, and it might be the case that the 'cycle is the trend', as investigated by Aguiar and Gopinath (2007) for emerging-market economies, and by Bluedorn and Leigh (2018) for developed economies. That is, negative shocks today have a persistent effect on the future (so that the economy does not return to an immutable trend).

Following Krane (2011), Bluedorn and Leigh (2018) investigate the beliefs or perceptions of professional forecasters regarding shocks to output, and the extent to which those shocks are believed to have a temporary or permanent effect on output. If output is believed to fluctuate around a stable trend, then shocks would be expected to only have a transitory effect on output. This can be addressed by analyzing whether unexpected revisions to short horizon forecasts are associated with an expected long-run impact on the level of output in the future. Bluedorn and Leigh (2018) consider the long-term forecasts for 38 advanced and emerging economies, and Krane (2011) considers the US. Both consider the consensus forecasts.[1]

A key innovation of our paper is to consider the heterogeneity in individual forecaster perceptions of the persistence of output growth, and to this end we use the individual respondents to the US SPF.[2] Unless forecasters have identical perceptions of the persistence of output growth, the use of the consensus or aggregate may be misleading. Aggregate perceptions of persistence may change over time due to the changes in the composition of the panel through entry and exit: see, e.g., Engelberg, Manski and Williams (2011). The aggregate is known to be misleading for testing hypotheses about expectations formation, and in particular whether expectations are 'rational' (see, e.g., Figlewski and Wachtel (1981, 1983) and Keane and Runkle (1990, p.717)). Of interest is whether investigating forecasters' perceptions of the nature of output shocks using the consensus might also be misleading.

Our initial set of results based on regressions for each of the individual forecasters suggests the findings obtained for the consensus appear to be highly misleading, in the sense that the

[1] Krane (2011) the consensus of the Blue Chip Panel, and Bluedorn and Leigh (2018) the Consensus Economics forecasts.

[2] A number of authors have considered various other aspects of individual-level forecaster behaviour, including "inattentiveness" as an explanation of forecaster disagreement (see, e.g., Andrade and Le Bihan (2013)), as well as whether there are systematic differences between individuals over time (see, e.g., D'Agostino, McQuinn and Whelan (2012), Clements (2019)) and the accuracy of their perceptions of uncertainty (see, e.g., Clements (2014, 2018).)

A closely-related paper to our paper is the analysis of perceived inflation persistence by Jain (2019), which is discussed in the main text.

consensus results mask apparently very different views held by the individual forecasters. One possible explanation for the heterogeneity in the perceived permanence of GDP shocks is that the individual respondents are active participants at different times, and that the perceived effect of output shocks may be time dependent: it may depend on the state of the business cycle when the forecasts are made; or on whether the GDP shock is perceived to be a shock to labour productivity or to be neutral with regard to labour productivity, and so on. In addition, differences in the estimates obtained from the individual regressions may partly reflect the typically relatively small samples of forecast data available at the individual level. We devise a bootstrap test of whether the apparent differences in the estimates of persistence across individuals reflect real differences in perceptions, or a combination of small-sample effects and the responding to different sets of surveys. Our approach is similar in spirit to that used by D'Agostino *et al.* (2012) to determine whether apparent differences in forecasting ability across individuals reflect real differences in their ability to produce accurate forecasts.

We use the revisions between forecasts of the same target (i.e., the revision between fixed-event forecasts) to calculate real GDP growth expectations shocks. Short-horizon forecast revisions are used in conjunction with revisions to both medium-term and long-horizon forecasts - made at the same forecast origins - to estimate forecasters' perceptions of the medium and long-term responses of output to shocks to the current-level of output. The individual series of forecast revisions can be used to directly estimate the long-term perceived persistence of output growth, say, by regressing the long-term revision on the short-term revision, for each individual. In addition, following Krane (2011) we use the forecast revisions to estimate forecaster-specific decompositions into temporary and permanent shocks. This allows an exploration of the heterogeneity in the directly-estimated perceived responses in terms of widely-used decompositions of GDP into permanent and transitory components.

While our focus is on the individual respondents perceptions of persistence, and whether these are statistically significantly different one from another, Krane (2011) explains the reasons behind his focus on the consensus. These include forecast data availability; because the average is most likely to affect aggregate activity; because the individual level biases to optimism or pessimism might cancel; and because the average is a better predictor of future output. A finding that forecasters' perceptions of persistence are essentially the same would be the overarching rationale for considering the consensus forecasts.

The empirical contribution of our paper is the individual-level analysis of whether forecasters perceptions of output persistence are the same, or whether there is inter-forecaster heterogeneity in this respect. This contribution relies on a way of determining whether the apparent differences in perceptions are real, or are simply a product of the difficulties that typically afflict studies of individual forecasters: the relatively small samples of forecast observations due to limited participation in the survey, and the concomitant problem of forecasters being active during potentially quite different economic conditions. The methodological contribution is a solution to this general problem. Following D'Agostino *et al.* (2012), we simulate a set of imaginary forecasters who match the actual forecasters in terms of when and how often they participate, but their long-horizon forecast revisions are randomly drawn and allocated to them from the set of actual revisions for that period. If the distributions of persistence estimates across our imaginary forecasters - with randomly assigned long-horizon revisions - match the empirical distribution of the actual forecasters, we can infer that the observed spread of estimates of the actual forecasters is due to the small-sample nature of the persistence estimates and participating during different economic conditions (because we have built in exchangeability - no real differences between the imaginary forecasters).

A closely-related paper is the study of perceived inflation persistence by Jain (2019). Jain shows that, for a simple model of perceived inflation consisting of an unobserved persistent component and a (white noise) transitory term, a regression of the revision of the forecast of time $t+h$ between periods $t-1$ and t, on the revision of the forecast for period $t+h-1$, between the same forecast origins, estimates the persistence parameter of the (assumed first-order) permanent component. Jain exploits all the US SPF quarterly forecasts which are available up to one-year ahead (from the current survey quarter) to more efficiently forecast the persistence parameter using GLS. As explained in section 3, our decision to use 10-year ahead annual average forecast data for our study of output growth restricts the number of forecasts we are able to draw on. We leave for future research a comparison of the heterogeneity of perceived output persistence from a study using a large set of relatively short-horizon forecasts, as in Jain (2019), and the use of a restricted set of forecasts that comprises long-horizon forecasts, as here. Relative to Jain (2019), an innovation of the current paper is the bootstrap analysis of the inter-forecaster differences in perceptions.

The plan of the remainder of the paper is as follows. Section 2 outlines two methods of measuring perceptions of persistence. One uses a simple regression of the revision of long-horizon forecasts on the revision to short-horizon forecasts (or of the revision of medium-horizon forecasts on the revision to short-horizon forecasts), and the other decomposes the revisions to the short, medium and long-horizon forecasts into permanent and transitory components. Section 3 describes the forecast data we use. Section 4 describes our empirical results. The section begins with a description of the forecast disagreement at the relevant horizons, before describing the results of applying the approaches described in section 2. Because we find marked differences in the degree of perceived persistence across the forecasters in the sample, we bootstrap the regression equation approach for the long-horizon forecasts, to determine whether the apparent differences in perceptions generated by the regression approach are real. We carry out the same exercise for the regressions with the medium-horizon forecasts. We also use a bootstrap approach to explore the differences between the long and medium-horizons within the permanent and transitory shocks frameworks. Finally, we consider the evidence for changes in perceptions over time. Section 5 offers some concluding remarks.

2 Approaches to Gauging Forecaster Perceptions of the Permanence of GDP shocks

We begin by discussing two approaches to measuring the perceived persistence in real output. Krane (2011) suggests an approach to determining the relative importance of permanent and transitory components of GDP shocks using forecast revisions. He supposes that output (the log of real GDP) y_t can be decomposed into a permanent component, p_t, and a transitory component, c_t:

$$y_t = p_t + c_t$$

and shows that the shocks to these two components can be determined by considering the period $t-1$ to t revisions in the forecasts of y_{t+k} at different horizons, k. He supposes that the shock to the transitory component c_t, denoted u_t, will have no effect on the revision to y_{t+k} for sufficiently large k, but will have an effect at shorter horizons, and especially at $k=0$. There are two shocks to p_t: w_t affects the average trend rate of growth, and e_t the level. For sufficiently large k, we can assume the forecast revision to the *growth rate* (i.e., Δy_{t+k}) is equal to w_t, because c_t will have no affect on the forecast, given that it is transitory, and e_t will have been fully assimilated into y_{t+k-1} and y_{t+k}. For k between 0 and K, where K is large, Krane supposes that some proportion θ_k of e_t will affect the forecast revision (at that horizon), as will some proportion ρ_k of u_t. Krane (2011) has a rich enough set of forecast horizons to estimate the variances of the shock components, σ_u^2, σ_e^2, σ_w^2; as well as the impulse responses, θ_k and ρ_k. The US SPF has the advantage of providing individual level forecast data, but provides forecasts of fewer medium to long-term horizons, as well as fewer forecasts at each horizon, as most respondents only provide forecasts to some surveys. We explain how we adapt the approach given the available forecast data in section 3.

A more broadbrush approach is simply to consider the relationship between revisions to current output (growth) and long-horizon output (growth), without attempting to identify the relative variability of the shocks to the transitory and permanent components, or the (perceived) importance of the shocks at different horizons. If the shocks to the permanent component are relatively unimportant, then we would expect to also find little association between the revisions to short and long-horizon forecasts.

Following Bluedorn and Leigh (2018), we regress the revision in the forecast of the long-horizon average annual growth over the period t to $t+h$ on the revision in the forecast growth rate at t:

$$r_t\left(\Delta y_{t,t+h}\right) = \alpha + \beta_h . r_t\left(\Delta y_{t,t}\right) + v_t. \tag{1}$$

In (1), r_t is the revision in the forecasts made at t and $t-1$, and $\Delta y_{t,t+h}$ and $\Delta y_{t,t}$ are the long-horizon and current-quarter growth rates, respectively.[3] That is, $r_t\left(z_{t+h}\right) = z_{t+h|t} - z_{t+h|t-1}$, where

[3]Bluedorn and Leigh (2018) use the cumulative growth rate between t and $t+h$, whereas we use the annual average

$z_{t+h|t}$ is the forecast of z_{t+h} made at time t.

In (1), $\beta_h = 0$ corresponds to the perception that output fluctuates around a stable trend, for large h, such that a revised expectation for the short-term outlook is not associated with any change in expected long-term GDP growth.

3 Forecast Data: SPF Respondents' Forecasts

The US Survey of Professional Forecasters (SPF) is a quarterly survey of macroeconomic forecasters of the US economy that began in 1968, and is currently administered by the Philadelphia Fed (see Croushore (1993)). The SPF is made freely available by the Philadelphia Fed. It is a key database for academic research on survey expectations.[4]

Up until 1992, the survey was of short to medium-horizon forecasts. As of 1992:Q1, the survey asked for 10-year annual-average real GDP growth forecasts (SPF variable identifier RGDP10), although this information was only collected in response to first quarter of the year surveys. Hence we use the first-quarter surveys from 1992 to 2018, inclusive. In addition to the 10-year average forecasts, the survey provides forecasts of the current quarter, and each of the next four quarters, as well as forecasts of the current-year annual level of output, and of next year's annual level.

In terms of equation (1), the long-horizon regression LHS variable is given by $r_t\left[\Delta y_{t,10}\right]$, where $\Delta y_{t,10}$ is the 10-year annual average growth rate, and the forecast revision ($r_t\left[.\right]$) is between the forecasts made in the first quarters of years $t-1$ and t. The RHS variable is $r_t\left[\Delta y_{t,cq}\right]$. The target is the current-quarter (cq) growth rate, $\Delta y_{t,cq}$, and the two forecast origins are again the first quarter surveys of years $t-1$ and t. Hence the right-hand-side variable is the difference between a current-quarter forecast of the year t Q1 growth rate, and a forecast of the same target made in year $t-1$, Q1.

$\Delta y_{t,cq}$ is expressed as an annual growth rate to match $\Delta y_{t,10}$.[5] The SPF data also allow us to define the revision in the year t annual growth rate, denoted $r_t\left[\Delta y_{t,a}\right]$. This is the current-year annual growth forecast of year t, from the t,$Q1$ survey, minus the forecast of year t from the $t-1$, $Q1$ survey.[6] The availability of $r_t\left[\Delta y_{t,a}\right]$ means that we are able to consider a regression

growth rate over the period.

[4] An academic bibliography of research using the US SPF is maintained at: http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/academic-bibliography.cfm.

[5] If Y_t is the level of GDP in quarter t, $\Delta y_{t,cq}$ is calculated as $100\left[\left(\frac{Y_t}{Y_{t-1}}\right)^4-1\right]$.

[6] The forecasts are recorded as levels, except for RGDP10 (recorded as a percentage to two decimal places). To calculate the forecast of the annual growth rate in year t from the year $t-1$, Q1 survey, we use the forecasts of the current year and the next year. To calculate the year t, Q1 forecast of the annual growth rate in year t, we use the forecast of the level for year t, and the year t, Q1 vintage 'actual' values for year $t-1$ to construct the growth rate. (This is the end January vintage of data, which includes the advance or first estimate for the 4th quarter of year $t-1$. This will be available to the respondent when the forecasts are reported to the Survey). Hence we use only 'real time' vintage data available to the forecaster at each point in time. These data were taken from the Real Time Dataset for Macroeconomists maintained by the Philadelphia Fed:

of a revision to medium horizon forecasts on the revision to current quarter forecasts, as well as the regression using the revision to the long-horizon forecasts described in (1). The regression of $r_t\left[\Delta y_{t,a}\right]$ on $r_t\left[\Delta y_{t,cq}\right]$ is informative about the perceived 'medium-term' response of output, which we will denote as β_a, to distinguish it from the long-run response β_{10} defined by the regression of $r_t\left[\Delta y_{t,10}\right]$ on $r_t\left[\Delta y_{t,cq}\right]$.

Our forecast data supports allows us to implement the approach of Krane (2011) as follows. Matching his equation (eqn. 6) we have:

$$r_t\left[\Delta y_{t,cq}\right] = w_t + e_t + u_t \tag{2}$$

$$r_t\left[\Delta y_{t,a}\right] = w_t + \theta e_t + \rho u_t \tag{3}$$

$$r_t\left[\Delta y_{t,10}\right] = w_t \tag{4}$$

(2) - (4) are the only forecast revisions which can be constructed from the surveys held at times t, $Q1$ and $t-1$, $Q1$.[7] It is not possible to calculate both θ and ρ from this set of equations. We then consider two identification restrictions: Identification Condition I, $\rho = 0$, and Identification Condition II, $\theta = 1$. The first assumes that the transitory shock has no effect on the revision to the current-year annual growth rate forecast. Condition II supposes that $\theta = 1$, implying that the perceived permanent shock to the level of (log) GDP is fully absorbed at impact.

Under Identification Condition I, the shocks $\{w_t, e_t, u_t\}$, their variances, and θ can be estimated. Subtract (4) from both (2) and (3) gives, say $r_{t,S} = e_t + u_t$ and $r_{t,M} = \theta e_t$. We then estimate θ as the inverse of the slope in the regression of $r_{t,S}$ on $r_{t,M}$. Given $\widehat{\theta}$, calculate $e_t = r_{t,M}/\,\widehat{\theta}$, and $u_t = r_{t,S} - e_t$. Under Identification Condition II, the shocks and their variances along with ρ can be calculated as follows. As before, $r_{t,S} = e_t + u_t$, but now $r_{t,M} = e_t + \rho u_t$. Regressing $r_{t,S}$ on $r_{t,S} - r_{t,M}$ provides an estimate of $(1-\rho)^{-1}$, using which u_t and e_t can both be determined.

Notice that we can calculate the *implied* values of β_{10} and β_a from the shocks $\{e_t, u_t, w_t\}$ defined by (2) to (4), and these can be compared to the directly obtained estimates from the regressions of $r_t\left[\Delta y_{t,10}\right]$ (or $r_t\left[\Delta y_{t,a}\right]$) on $r_t\left[\Delta y_{t,cq}\right]$. The implied population value of β_{10} is given by the regression

---

https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/
and see Croushore and Stark (2001).

[7] Recall that we are limited to a consideration of the Q1 survey origins because the long-horizon 10-year forecasts are only reported to the Q1 surveys.

of (4) on (2), as:

$$\begin{aligned}\beta_{10,IMP} &= \frac{Cov\left(r_t\left[\Delta y_{t,10}\right], r_t\left[\Delta y_{t,cq}\right]\right)}{Var\left(r_t\left[\Delta y_{t,cq}\right]\right)} \\ &= \frac{\sigma_w^2}{\sigma_w^2+\sigma_e^2+\sigma_u^2}\end{aligned} \tag{5}$$

The implied population value of β_a will clearly depend on the Identification Condition. Under Identification Condition I:

$$\begin{aligned}\beta_{a,IMP,I} &= \frac{Cov\left(r_t\left[\Delta y_{t,a}\right], r_t\left[\Delta y_{t,cq}\right]\right)}{Var\left(r_t\left[\Delta y_{t,cq}\right]\right)} \\ &= \frac{\sigma_w^2+\theta\sigma_e^2}{\sigma_w^2+\sigma_e^2+\sigma_u^2}.\end{aligned} \tag{6}$$

Under Identification Condition II:

$$\beta_{a,IMP,II} = \frac{\sigma_w^2+\rho\sigma_u^2+\sigma_e^2}{\sigma_w^2+\sigma_e^2+\sigma_u^2}. \tag{7}$$

Hence as an alternative to the OLS regression of (1), with HACSE for inference, we can estimate β_{10} (β_a) by substituting the sample estimates of the population moments into (5) ((6) or (7)). Note that β_{10} does not depend on the identifying assumption at the population level. The perceived degrees of persistence will be smaller the larger the variances of u_t and e_t relative to that of w_t.[8]

In an Appendix we show how the implied estimates of β relate to the permanent-transitory decompositions of output used in the literature, and in particular by Jain (2019). The appendix also clarifies how our approach differs from that of Jain (2019) in terms of the forecasts we use.

In equations (1) to (7) we have omitted individual-forecaster scripts for notational convenience, but all terms in these equations are allowed to vary across forecasters. We estimate the regressions separately for each individual, and calculate the decompositions based on (2)-(4) separately for each individual.

4 Results

4.1 Disagreement regarding the long-term growth rate

As a preliminary exercise, we consider the term structure of disagreement across forecasters, for our sample of forecasts made from first-quarter surveys. A number of studies have considered the characteristics of forecaster disagreement at different forecast horizons. For example, Lahiri and Sheng (2008) and Patton and Timmermann (2010) consider the roles of differences in priors

[8]The form of (2) to (4) implies that $\beta_{10,IMP}$ in (5) is necessarily non-negative, and positive whenever $\sigma_w^2 \neq 0$.

about long-run growth rates and different models, versus differences in information signals (and their interpretation). The importance of information signals would be expected to diminish as the forecast horizon lengthens, assuming the variable being forecast is stationary (this is part of what it means for a variable to be stationary). As the horizon lengthens, the forecasts of stationary variables approach the long-run or unconditional expectation. If disagreement persists at long horizons, then one might infer that forecasters possess different priors about long-run means.[9]

Figure 1 shows short-horizon forecaster disagreement being higher than at longer horizons throughout the period, and being more responsive to business cycle conditions. The highest and most recent peak in the series occurs for the first quarter of 2009, during the Recession. Although the current-year growth forecast disagreement peaks at the same time as that of the current-quarter forecasts, the time-series movements are much less pronounced, and fluctuate around a lower level. These patterns are consistent with a diminished role for heterogeneous-signals at the longer horizon. But the series for the 10-year growth forecasts does not show a further marked decline in the level of disagreement, as might be expected. Figure 1 indicates a good deal of variability across respondents in their perceptions of the long-horizon outlook.[10]

Finally, figure 2 shows the dramatic effect on the average outlook for the short and medium (one-year ahead) term in 2009Q1. In 2009Q1 the average current quarter growth rate (annualized) was for a decline in GDP of 5%, with a slight dip in the 10-year average growth rate of less than a half a percentage point.

In section 4.2 we consider whether these summary statistics translate into different perceptions of the permanence of shocks to GDP by different survey respondents.

4.2 Regressions of long-horizon forecast revisions on short-horizon revisions

For the consensus, estimation of (1) results in a statistically significant estimate of β_{10} of 0.05 for the regression of the ten-year revision on the current-quarter revision: see table 1. The regression of the current-annual revision on the current-quarter gives an estimate of 0.64. The estimate for the ten-year forecasts suggests a shock to output (as measured by the short-horizon forecast revision) is perceived to have a permanent effect on output growth. A positive revision to the forecast of the current quarter GDP growth rate (annualized) of 1 percentage point is expected to raise the level of output by half a percentage point over the next 10 years (that, is by 0.05 percentage points for each of the next 10 years on average).

However, the estimates for the individual respondents vary widely, from being negative (-0.09) to large and positive (0.19). Just under a half of the estimates are significantly different from zero

[9] Both Lahiri and Sheng (2008) and Patton and Timmermann (2010) consider horizons up to two years ahead. Patton and Timmermann (2010) consider forecasts made every month of forecasts of real GDP growth and inflation for the current (calendar) year, and for next year. The forecasts analyzed by Lahiri and Sheng (2008) are also monthly, up to two years ahead, but for GDP growth for a number of industrialized countries.

[10] All the forecasts are at annual rates for comparability.

(absolute value of the t-statistic exceeds 2), and of these all but one is positive. Hence around a half of the individuals in the survey do not believe that shocks to GDP have a permanent effect.

Can we explain the widely divergent views across individuals? Firstly, note that there is much less heterogeneity in perceptions of the medium term GDP response. From the regression of the revision of the forecasts to the year-ahead growth rate on the current-quarter revision (columns (5) and (6)), it is apparent that most estimates are in the range of 0.4 to 0.7, and are statistically significant for 24 of the 27 individual forecasters. Clearly, the marked divergence of perceptions primarily bears on the long-horizon (ten-year) response.

4.3 Permanent-transitory decompositions

Tables 1 and 2 compare the individuals in terms of the variances of the components in the decomposition of the shocks à la Krane (2011). The right side of table 1 provides the estimates under the identifying assumption that $\rho = 0$ (Identification Condition I), and table 2 repeats for ease of comparison the direct estimates of $\hat{\beta}_{10,i}$ and $\hat{\beta}_{a,i}$ and provides estimates of the permanent and temporary components under the identifying assumption that $\theta = 1$ (Identification Condition II). Although the estimates of the specific components vary widely across individuals in both tables (see columns (7) to (10) of table 1, and columns (5) to (8) of table 2), and may not be precisely determined, the cross-sectional variability of the β_{10}- and β_a-estimates is reduced relative to the cross-sectional variability of the direct regression estimates of these parameters. (Recall that the implied estimates, $\beta_{10,IMP,i}$ and $\beta_{a,IMP,i}$ are calculated from substituting the sample estimates of the variances of the components, and the estimates of θ or ρ, into (5) and (6)/(7)). For example, the cross-sectional standard deviation of the the regression estimates $\hat{\beta}_{10}$ is 0.056 and that of the implied estimates is 0.041. The cross-sectional mean is also reduced, from 0.045 to 0.039.

A similar picture emerges for the medium-term estimates, β_a, in that the cross-sectional standard deviation of the implied estimates is reduced, from 0.156 for the regression estimates to 0.143/0.146, depending on the identification assumption. Generally, the identification assumption does make much difference to the cross-sectional distribution of either the implied long-run or implied medium-term persistence. That is, at least in terms of summary cross-sectional statistics of the estimates of perceived medium and long-horizon persistence, the identifying assumption (either $\rho = 0$ or $\theta = 1$ in (3)) is unimportant.

It remains the case that cross-sectional standard deviation of the β_{10} estimates is similar to the cross-sectional mean, but the standard deviation of the β_a estimates is only a third to a quarter of the cross-sectional mean.

Finally, in the notes to the table we indicate that the cross-sectional correlation between the direct estimates and the implied estimates is either 0.95 or 0.96 for β_a (depending on the identifying assumption), but only 0.64 for β_{10}.

Given that at the 10-year horizon there is much more diversity in beliefs (than at the medium or annual horizon), whether calculated directly by the regression approach, or indirectly from the trend-cycle model, in what follows we ask whether the apparent heterogeneity in the β_{10}-estimates is real, or whether it can be attributed to small-sample issues.

4.4 Different sample periods

We begin by directly confronting the possible effect of individual forecasters being active at different times. The consensus is based on the maximum sample of 26 observations (the first quarter surveys from 1992 to 2018, losing one observation to calculate the revision), while the number of observations for each respondent varies from our imposed minimum of 10 to a maximum of 22. It is possible that the range of estimates across individuals could be due to small sample issues, or to the individuals being active at different times, if we allow the possibility that the perceived relationship between short and long-horizon revisions is not constant over time, as discussed in the introduction.

We consider the effect of non-participation by calculating individual-specific consensus forecast β_{10}'s, denoted $\hat{\beta}_{C,t\in\mathcal{I}_i}$: this is the β_{10}-estimate using the consensus forecasts for the surveys to which respondent i filed a forecast. If $\hat{\beta}_{10,i}$ and $\hat{\beta}_{C,t\in\mathcal{I}_i}$ are highly correlated across respondents, we would conclude that the cross-sectional variation in the β_{10}-estimates evident in table 1 primarily reflects individuals being active at different times. A low correlation would instead suggest that time of participation is not important in explaining differences between forecasters' estimates. Table 3 presents the estimates of $\hat{\beta}_{C,t\in\mathcal{I}_i}$, along with the $\hat{\beta}_{10,i}$ estimates to aid comparison, as well as some summary statistics of the cross-section distributions. The cross-sectional standard deviation is more than halved for the $\hat{\beta}_{C,t\in\mathcal{I}_i}$, at 0.023, compared to 0.056 for the β_{10}. That the standard deviation is markedly lower is to be expected because all the consensus estimates draw on the same forecast observations and the β-estimates only differ by the sample period. The correlation between $\hat{\beta}_{10,i}$ and $\hat{\beta}_{C,t\in\mathcal{I}_i}$ is 0.46. Clearly different participation times explains *some* of the heterogeneity in perceived persistence, but a correlation coefficient of around a half does not clearly arbitrate between the different β_{10}-estimates primarily *i*) reflecting real differences in the perceptions of individual forecasters, or *ii*) small-sample variability in the estimates exacerbated by forecasters being active at different time periods.

4.5 Bootstrap test of the exchangeability of medium-horizon and long-horizon revisions

4.5.1 Regression estimates

The individual heterogeneity in the different β estimates obtained in the individual time-series regressions may reflect real differences in the perceptions of individual forecasters, or differences resulting from small-sample variability in the estimates exacerbated by forecasters being active at different time periods. We investigate this issue by testing the hypothesis that the individuals' forecast revisions are interchangeable conditional on the time period t. Under the null, individual forecaster perceptions of the warranted revisions to their 10 year forecasts are not significantly different from one another. We randomly assign the actual 10-year forecasts made at time t to the active participants at time t (with replacement), and we consider whether the cross-sectional distribution of the actual $\hat{\beta}_i$ estimates is consistent with this bootstrap distribution obtained by random re-assignment.[11] In so doing we condition on the actual short-horizon revisions in the SPF forecast data.

In detail, the bootstrap test is implemented as follows.

1. Let y_{it} denote the long-horizon forecast revision of respondent i to survey t, where $i = 1, \ldots, N$, and $N = 27$, and $t = 1, \ldots, T$, where $T = 26$ (denoting the first quarter surveys from 1993 to 2018), where some elements are missing values.

For $t = 1$, we randomly select with replacement from the non-missing set of values $\{y_{j,t=1}\}_{j=1,\ldots,N}$ for each $y_{i,t=1}$ for which $y_{i,t=1}$ is not missing, to create $\left\{y^*_{j,t=1}\right\}_{j=1,\ldots,N}$. Missing values in the actual data are replicated in the bootstrap sample ($y^*_{j,t=1}$ is a missing if $y_{j,t=1}$ is missing).

We repeat for $t = 2$, and so on up to $t = T$.

2. Given $\left\{y^*_{j,t}\right\}_{j=1,\ldots,N;t=1,\ldots,T}$, we estimate individual regressions for each individual i using variation over t:

$$\hat{\beta}^*_i = \frac{\sum_{t=1}^{T} \left(y^*_{it} - \overline{y}^*_i\right)\left(x_{it} - \overline{x}_i\right)}{\sum_{t=1}^{T} \left(x_{it} - \overline{x}_i\right)^2},$$

where the x_{it} are the actual current-quarter forecast revisions.

Missing values result in the corresponding rows of $[y^*_{it} : x_{it}]$ being deleted.

The mean, standard deviation, and extreme values of $\left\{\hat{\beta}^*_i\right\}$ over $i = 1, \ldots, N$ are saved.

3. We repeat steps 1 to 3 $R = 1000$ times, to calculate a bootstrap sample of R cross-section means, standard deviations, and extreme values of β-estimates. If, for example, the mean of the actual β-estimates lies within the 25th and 975th largest bootstrapped mean values, we conclude that the null of interchangeable long-horizon forecast revisions is not rejected at the 5% value.

Notice that the test accounts for the unbalanced nature of the panel, and the fact that some individuals respond less than half the time, because missing values in the forecast data are reproduced in the bootstrap samples. The small-sample estimation uncertainty that characterizes the empirical estimates will also feature in the bootstrap distributions of these estimates. As noted

[11] Our metric for comparing the actual $\hat{\beta}_i$ estimates and the bootstrapped estimates is in terms of the cross-sectional means and standard deviations, and the maximum and minimum values, as opposed to comparing the percentiles of the actual distributions of the $\hat{\beta}$ estimates and the percentiles of the bootstrap distributions. For example, D'Agostino *et al.* (2012) compare a percentile of the actual distribution of forecast accuracy against the estimate of this percentile under the null of equal accuracy, by calculating a confidence interval for this percentile (e.g., the median most accurate forecaster) from the bootstrap replications. In principle, we could do the same for the $\hat{\beta}$, but because the number of forecasters is relatively small in our context at 27, we consider just a few summary statistics: the mean and standard deviation, and the extreme values, rather than attempting a finer comparison.

in the Introduction, all that differs between the simulated data and the actual data is that the simulated data imposes interchangeability of revisions across respondents. If the estimates based on the actual data are consistent with the bootstrap estimates, then we can deduce that the actual forecasters' revisions are also interchangeable.

As a check on the bootstrap test based on a comparison of the cross-sectional moments, we implemented the above with a small but important change. At step 1, we randomly sampled from $\{y_{j,t}\}_{j=1,\ldots,N;t=1,\ldots,T}$: that is, we did not condition on t. Not conditioning on t supposes that there is no meaningful variation in the ten-year forecasts across time, and we would expect to reject this hypothesis. That we do so reassuringly suggests that there is predictability in the ten-year ahead forecasts.

Table 4 Panel A records the results when the y's are re-assigned randomly across forecasters (with replacement). It records the (two-sided) 1%, 5% and 10% critical values of the bootstrapped distributions of the cross-sectional mean, standard deviations, and extrema of the individual β-estimates. When we condition on t, the mean and standard deviations of the actual $\hat{\beta}_i$ estimates (0.045, and 0.056, respectively - see table 1) are consistent with the null. The same holds for the maximum and minimum values - these are 0.190, and -0.090, which lie well away from the tails of the bootstrapped values of these quantities.

When we do not condition on t, the bootstrap intervals for the mean are more or less symmetric about zero, and do not include the mean of the actual estimates. Hence we reject the null that there is no meaningful variation in the ten-year forecasts across time.

Panel B of table 4 indicates that the empirical cross-sectional distribution of persistence estimates is consistent with random re-assignment of the short-horizon forecast revisions (the x's) across respondents, conditional on t, and not just of the y's.

In summary: the long-horizon and short horizon forecast revisions are interchangeable across respondents, *at a given point in time*.

This raises the question of whether this is solely a long-horizon effect, or whether interchangeability is also a feature of medium term forecasts. We assess this by applying the bootstrap approach to regressions of the year-ahead growth rate forecasts on the current-quarter forecasts. The individual β-estimates are recorded in table 1, column (5). Although the cross-section standard deviation is only just over a quarter of the mean value, suggesting less diverse perceptions across individuals than at the 10-year horizon, nevertheless we find the apparent variability across individuals is real.

From table 1 the mean and standard deviation of the cross-sectional distribution of actual estimates (of column (5)) are 0.516 and 0.156, and the max and min values are 0.782 and 0.054. When we bootstrap the y's (annual forecast revisions), both the cross-sectional mean and standard deviation of the estimates from regressing the annual revisions on the current-quarter revisions are outside the 1% interval, suggesting the actual distribution of the estimates is not consistent with that simulated under the null of exchangeable year-ahead forecasts. See table 4 Panel B.

The results suggest that there are real differences between forecasters' perceptions of persistence at medium term horizons, such as one-year ahead forecasts, but that these do not hold at the ten-year horizon.

4.5.2 Shocks decompositions

In the previous section we tested whether the cross-sectional distributions of the regression estimates of the β's were affected by random shuffling across forecasters of their revisions to the 10-year ahead annual average forecasts, or their year-ahead annual average forecasts. In this section we carry out a similar exercise, but on each replication we estimate and record the variances of the shock components (and either θ or ρ) for each individual using (2) to (4), and we also estimate the implied parameter estimates from (5) and (6) or (7). The identification of the variances etc. requires the three forecasts - 10-year annual average, next-year annual growth, and current-quarter (annualized) growth.

Firstly, we randomly shuffle the 10-year forecast revisions. In terms of the bootstrap outlined in the previous section, y_{it} remains the long-horizon forecast revision of respondent i to survey t, but the x_{it} is now two-dimensional, comprising the revisions to the year-ahead annual forecasts, as well as the actual current-quarter forecast revisions.

We calculate the (two-sided) 1%, 5% and 10% critical values of the bootstrapped distributions of the cross-sectional means and standard deviations of the variances of the shocks. If we compare the means and standard deviations of the cross-sectional distributions of the 'actual estimates' (from table 1 for Identifying Condition I, and from table 2 for Identifying Condition II), with the bootstrap confidence intervals, we find that all lie within the intervals. Hence all the parameters are consistent with the empirical estimates under random re-shuffling of the 10-year forecast revisions. This is perhaps to be expected because the cross-sectional distribution of the 10-year revision regression estimates had been found to be unchanged. The results are given in table 6, in the Appendix.

More interesting is to shuffle the year-ahead annual growth revisions, to see whether the rejection of interchangeability in the regression estimates can be attributed to particular components. The results are reported in table 5. Firstly consider the right side of the table, calculated assuming $\theta = 1$ (Assumption II). The actual cross-sectional mean of ρ is 0.26, which lies outside the bootstrap interval [-0.125,0.126]. For both σ_e^2 and σ_u^2 the actual cross-sectional mean is either within, or close to being within, the bootstrapped interval. (Note that w_t and therefore σ_w^2 do not change across replications when only the year-ahead annual revisions are shuffled). One interpretation of these results is that interchangeability fails because of forecasters distinct perceptions about ρ, which measures the perceived effect of temporary shocks on the annual forecasts. Under Assumption I, the results do not support a simple interpretation. Both σ_e^2 and σ_u^2 lie outside their bootstrapped intervals, and the interval for the cross-sectional mean is large and inaccurately estimated for θ.

4.6 Were forecaster perceptions affected by the 2007-09 recession?

Of interest is whether the experience of the 2007:Q4 – 2009:Q2 Recession affected forecasters' perceptions of the persistence of output. Ideally we would consider the regression (1) for each forecaster, and whether the coefficients were constant over time, and especially whether there was evidence for a change after 2007. Unfortunately, there are too few forecasts available for most respondents to reliably detect time variation in a regression such as (1). A viable alternative is to consider whether respondents who were primarily active after the Recession have different perceptions relative to those who made a greater proportion of their forecasts before the Recession. The last column of table 1 records the proportion of pre-Recession forecasts made by each respondent relative to their total number of forecasts.[12] It is evident that one forecaster was only active in the earlier period (the ratio for id 20 is 1) while some forecasters only made 10% of their forecasts in the earlier period.

To determine whether there is an association between the estimated β, and the extent to which a forecaster was active in one period rather than another, we rank each forecaster in terms of their β, and the proportion of their forecasts made pre-2008. We test for an association by testing whether Spearman's rank correlation coefficient is zero. This allows us to test for an association without requiring linearity - the null will be rejected if there is a monotonic relationship between the two.[13]

Spearman's rank correlation r lies between -1 and 1, where 0 indicates no relationship, and a value of 1a perfect (monotonic) positive association. The coefficient is calculated as:

$$r = 1 - \frac{6R}{N\left(N^2 - 1\right)} \tag{8}$$

where R is the sum of squared differences between the two ranks. We follow the literature and calculate the Fisher transformation:

$$F\left(r\right) = \frac{1}{2}\ln\frac{1+r}{1-r}$$

such that $z = F\left(r\right).\sqrt{\frac{N-3}{1.06}} \sim N\left(0,1\right)$ under the null of statistical independence.

The test statistic z suggested no evidence of a correlation between β for the regression of the ten-year revision on the current-quarter revisions, and when the forecaster was active. (That is, there is no relationship between the ranks of columns (3) and (13) in table 1). When we considered instead the β's from the regression of the one-year ahead annual revisions on the current-quarter

[12] Our forecasts are made in the first quarters of the year, so we take the pre-Recession period to be 1993 to 2007, and the post Recession period to be 2008 to 2018.

[13] We do not make an allowance for the fact that the β-estimates are random variables with sampling uncertainty. This might be possible - see, e.g., Curran (2015).

revisions (the ranks of columns (5) and (13) in table 1), we obtained some weak evidence of a positive relationship: suggesting respondents who made more of their forecasts before 2008 perceive greater persistence between the year ahead annual output growth and current-quarter output growth.[14]

Using a bootstrap test, we concluded that long-horizon revisions were exchangeable amongst respondents at each point in time, but not across time. In this section we found no evidence of a change in perceptions about long-horizon persistence after the Recession. There was weak evidence that forecasters *en masse* might perceive lower persistence at the medium-term horizon after the Recession.

5 Conclusions

In this paper we contribute to a small but growing literature that seeks to better understand the behaviour of macro forecasters by considering the forecasts of individual survey respondents. Studies at the level of the individual forecaster are hampered by the relatively small samples of forecasts which are typically available. However, we show that such problems can be overcome. Our analysis of the perceptions of the persistence of GDP shocks by individual forecaster suggests considerable heterogeneity. Roughly a half of the respondents to the US panel of the Survey of Professional Forecasters do not expect any effect on output growth ten years down the line, while others expect a markedly higher effect than we obtain using the consensus forecasts (and consensus forecast revisions). According to the consensus view, a 1% point downward revision in the current quarter forecast (annualized) gives rise to an expected half a percentage point reduction in the level of output over the next ten years. (That is, a 0.05% point reduction for each of the next 10 years on average). The individual with the highest β estimate instead expects a reduction in the level of output of around 2%.

We show how one can determine whether these apparent differences in perceptions are real or not. We use a bootstrap to control for the relatively small samples of forecasts available for the individual respondents, and to control for the prevailing economic conditions. We compare features of the actual cross-sectional distribution of estimates of output persistence to those obtained assuming that individuals' forecast revisions at time t are interchangeable amongst those who were active at time t. The statistics of the distribution for the actual forecast data were consistent with the bootstrapped distributions under the null of interchangeability of the long-horizon forecast revisions.

However, there were found to be real differences between forecasters' perceptions of persistence at medium term horizons, such as one-year ahead forecasts.

A number of commentators have remarked on the persistence of the effects of the most recent

[14] The test statistic was positive, but the probability of obtaining a larger test statistic than we obtained was 0.07, so we would only reject at the 7% level in a one-sided test.

recession. For example, in the context of the UK, the comparison by Cribb and Johnson (2018) of the recoveries from recession since the 1920's, in terms of GDP per capita, makes this point forcibly. Is recent economic history reflected in the perceptions of professional forecasters? There are too few observations before and after the Recession for most individuals to be able to reliably determine whether there has been a change in persistence. Instead we exploit the interchangeability of individual respondents in the long-horizon forecast revisions regression to ask whether perceived persistence is correlated with whether the forecasts were made in the pre- or post-Recession periods.

In addition to the analysis based on estimates from the regressions involving long-term (and medium-term) revisions on short-horizon revisions, we conduct an analysis based on decomposing the forecast revisions into temporary and permanent components. At least under one of the identifying restrictions we consider, these decompositions are informative about the differences we find regarding the interchangeability of the long-horizon and medium-horizon forecast revisions. We find forecasters have different perceptions about the degree of persistence of temporary shocks.

Finally, the use of 10-year annual average growth expectations in the US SPF provide a direct measure of long-horizon forecasts, and this underpins our analysis. But the cost of using these forecasts is a reduced forecast sample (the 10-year forecasts are only available for first quarter of the year surveys). Future work might usefully examine this trade-off.

References

Aguiar, M., and Gopinath, G. (2007). Emerging Market Business Cycles: The Cycle Is the Trend. *Journal of Political Economy*, **115**, 69–102.

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

Bluedorn, J. C., and Leigh, D. (2018). Is the Cycle the Trend? Evidence From the Views of International Forecasters. Imf working papers 18/163, International Monetary Fund.

Clements, M. P. (2014). Forecast Uncertainty - Ex Ante and Ex Post: US Inflation and Output Growth. *Journal of Business & Economic Statistics*, **32(2)**, 206–216. DOI: 10.1080/07350015.2013.859618.

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

Clements, M. P. (2019). Forecaster efficiency, accuracy and disagreement: Evidence using individual-level survey data. Discussion paper, ICMA Centre, University of Reading.

Cribb, J., and Johnson, P. (2018). 10 years on - have we recovered from the financial crisis?. Discussion paper, Institute for Fiscal Studies https://www.ifs.org.uk/publications/13302.

Croushore, D. (1993). Introducing: The Survey of Professional Forecasters. *Federal Reserve Bank of Philadelphia Business Review, November*, 3–15.

Croushore, D., and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, **105(1)**, 111–130.

Curran, P. A. (2015). Monte Carlo error analyses of Spearman's rank test. mimeo, International Centre for Radio Astronomy Research, Curtin University, Australia.

D'Agostino, A., McQuinn, K., and Whelan, K. (2012). Are some forecasters really better than others?. *Journal of Money, Credit and Banking*, **44(4)**, 715–732.

Engelberg, J., Manski, C. F., and Williams, J. (2011). Assessing the temporal variation of macroeconomic forecasts by a panel of changing composition. *Journal of Applied Econometrics*, *26*(7), 1059–1078.

Figlewski, S., and Wachtel, P. (1981). The Formation of Inflationary Expectations. *The Review of Economics and Statistics*, *63*(1), 1–10.

Figlewski, S., and Wachtel, P. (1983). Rational Expectations, Informational Efficiency, and Tests Using Survey Data: A Reply. *The Review of Economics and Statistics*, *65*(3), 529–531.

Jain, M. (2019). Perceived Inflation Persistence. *Journal of Business & Economic Statistics*, *37*(1), 110–120.

Keane, M. P., and Runkle, D. E. (1990). Testing the rationality of price forecasts: new evidence from panel data. *American Economic Review*, **80(4)**, 714–735.

Krane, S. D. (2011). Professional Forecasters' View of Permanent and Transitory Shocks to GDP. *American Economic Journal: Macroeconomics*, *3*(1), 184–211.

Lahiri, K., and Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, **144(2)**, 325–340.

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.

Patton, A. J., and Timmermann, A. (2011). Predictability of output growth and inflation: A multi-horizon survey approach. *Journal of Business & Economic Statistics*, *29*(3), 397–410.

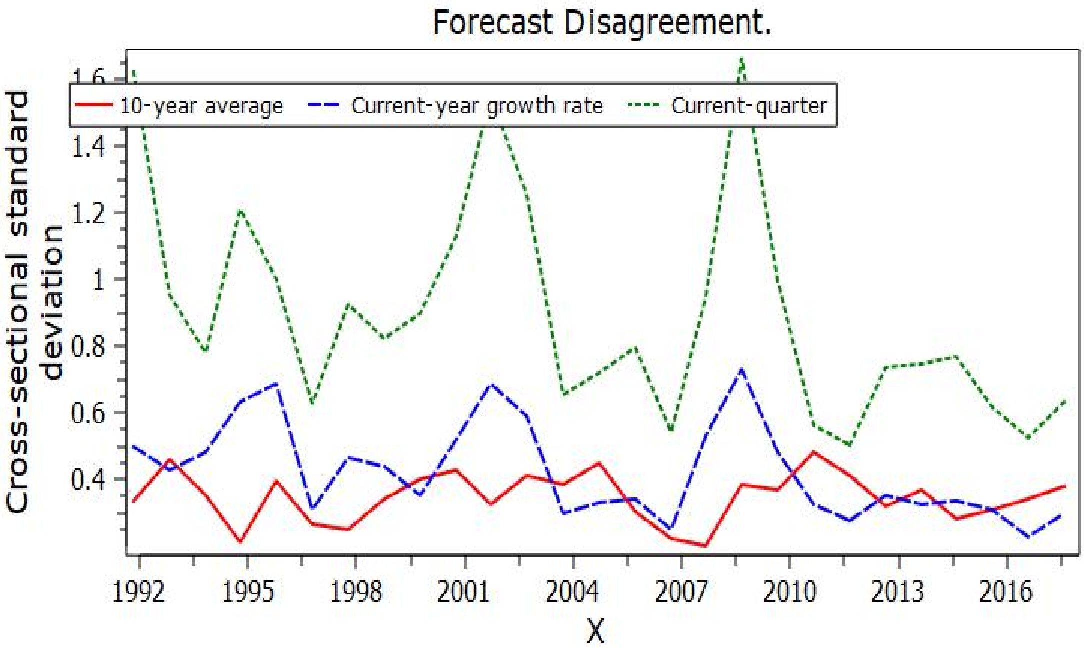

Figure 1: Forecaster Disagreement

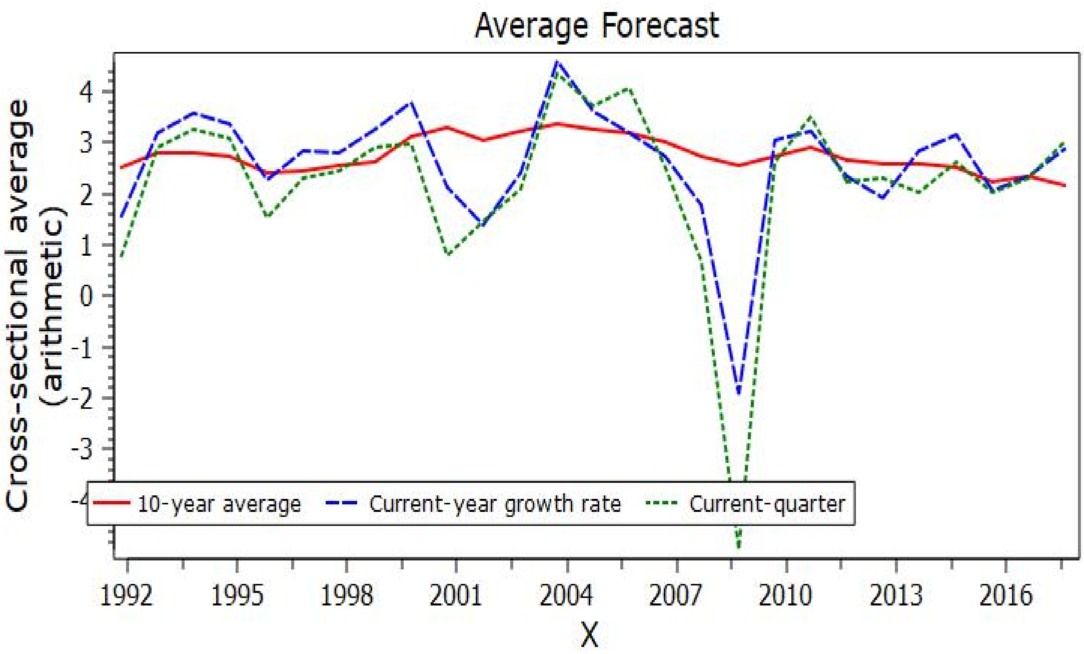

Figure 2: Consensus Forecasts

Table 1: Regression Estimates and Permanent and Transitory Component Parameters (Identification Restriction [I], $\rho = 1$)

		$r_t\left[\Delta y_{t,10}\right]$		$r_t\left[\Delta y_{t,a}\right]$								
id	No.	$\hat{\beta}_{10}$	t-stat	$\hat{\beta}_a$	t-stat	$\hat{\theta}$	σ_e^2	σ_w^2	σ_u^2	$\beta_{10,IMP}$	$\beta_{a,IMP}$	Pre-Crisis
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Cons	26	0.05	3.36	0.64	15.81	0.69	2.67	0.05	0.40	0.02	0.60	.
421	22	0.01	0.59	0.44	12.82	0.64	2.29	0.05	1.05	0.01	0.45	0.59
426	21	0.01	0.62	0.57	7.75	0.66	4.29	0.10	0.91	0.02	0.56	0.62
428	21	0.06	1.77	0.59	5.92	0.91	1.25	0.08	0.80	0.04	0.57	0.71
433	19	0.15	3.46	0.67	4.90	1.03	0.62	0.09	0.45	0.08	0.63	0.79
510	18	0.10	2.68	0.40	5.87	0.96	0.62	0.24	1.10	0.12	0.43	0.39
431	17	0.03	1.47	0.44	15.19	0.82	0.70	0.12	0.93	0.07	0.39	0.82
446	17	0.03	0.76	0.64	13.58	0.70	3.83	0.12	0.52	0.03	0.62	0.59
484	17	0.12	4.11	0.63	5.64	0.54	3.10	0.25	1.20	0.05	0.42	0.65
507	16	0.02	0.81	0.41	3.94	0.63	4.56	0.03	2.80	0.00	0.39	0.44
411	15	-0.02	-1.16	0.62	16.10	0.69	4.45	0.10	0.65	0.02	0.61	0.73
420	15	0.04	1.42	0.52	10.75	0.66	3.88	0.18	1.21	0.03	0.52	0.53
508	13	0.06	4.65	0.58	8.54	0.64	5.73	0.06	1.15	0.01	0.53	0.31
463	13	0.05	0.92	0.64	15.94	0.74	4.81	0.28	0.83	0.05	0.65	0.77
407	13	0.19	3.56	0.47	3.15	1.92	0.12	0.23	0.88	0.19	0.38	0.62
456	13	0.10	2.48	0.54	13.81	0.59	3.52	0.22	0.84	0.05	0.50	0.77
518	12	0.04	2.51	0.66	10.40	0.72	6.17	0.07	0.74	0.01	0.65	0.25
504	12	-0.09	-3.02	0.25	2.60	0.52	4.40	0.20	3.18	0.03	0.32	0.25
512	12	0.05	2.26	0.62	20.51	0.68	4.86	0.09	0.69	0.02	0.60	0.58
548	12	0.02	2.32	0.47	6.75	0.52	9.41	0.05	1.31	0.00	0.46	0.08
483	11	0.04	0.74	0.63	12.73	0.70	9.60	0.38	1.67	0.03	0.61	0.55
20	10	0.00	0.03	0.36	1.48	1.79	0.76	0.22	3.54	0.05	0.35	1.00
524	10	0.02	0.26	0.64	14.78	0.78	5.44	0.55	1.21	0.08	0.67	0.40
535	10	0.07	2.40	0.60	19.55	0.64	4.96	0.17	0.77	0.03	0.57	0.10
527	10	-0.03	-1.57	0.05	0.63	0.55	2.34	0.07	12.07	0.00	0.09	0.20
555	10	0.04	4.06	0.35	2.23	0.56	10.79	0.12	8.02	0.01	0.32	0.10
516	10	0.08	3.99	0.78	11.98	0.80	5.67	0.05	0.69	0.01	0.71	0.60
557	10	0.03	0.65	0.37	1.56	0.95	1.13	0.07	2.28	0.02	0.33	0.10
mean$_i$		0.045		0.516		0.79	4.05	0.16	1.91	0.039	0.494	
sd$_i$		0.056		0.156		0.34	2.80	0.12	2.54	0.041	0.143	
max$_i$		0.190		0.782		1.92	10.79	0.55	12.07	0.185	0.712	
min$_i$		-0.090		0.054		0.52	0.12	0.03	0.45	0.004	0.093	
>0			12		24							
<0			1		0							

The headers to columns (3)-(4), and (5)-(6), denote the dependent variable. In both cases the explanatory variable is $r_t\left[\Delta y_{t,cq}\right]$. The regression estimate t-statistics use heteroscedasticity and autocorrelation consistent standard errors. mean$_i$ and sd$_i$ are the cross-sectional means and standard deviations. max$_i$ and min$_i$ are the cross-sectional maximum and minimum. ' >0 ' and ' <0 ' are the number of regressions yielding statistically significant estimates at the 5% level. The $\beta_{10,IMP}$ in column (11) is the implied β_{10} calculated using sample estimates in place of the population moments, $\beta_{10,IMP} = \sigma_w^2 / \left(\sigma_w^2 + \sigma_e^2 + \sigma_u^2\right)$. The β_a in column (12) is the implied β_a calculated using sample estimates in place of the population moments, $\beta_{a,IMP} = \frac{\sigma_w^2 + \theta\sigma_e^2}{\sigma_w^2 + \sigma_e^2 + \sigma_u^2}$. The correlation beween the individual β_{10} estimates in columns (3) and (11) is 0.64, and between the estimates in columns (5) and the implied β_a in column (12) is 0.95. The last column (13) records the proportion of forecast observations made in response to the 'pre-Crisis' surveys, 1993 to 2007, inclusive.

Table 2: Regression Estimates and Permanent and Transitory Component Parameters (Identification Restriction [II], $\theta = 1$)

id	No.	$\hat{\beta}_{10}$	$\hat{\beta}_a$	$\hat{\rho}$	σ_e^2	σ_w^2	σ_u^2	$\beta_{10,IMP}$	$\beta_{a,IMP}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	26	0.05	0.64	0.43	0.83	0.05	1.88	0.02	0.61
421	22	0.01	0.44	0.27	0.77	0.05	2.39	0.01	0.45
426	21	0.01	0.57	0.38	1.33	0.10	3.38	0.02	0.57
428	21	0.06	0.59	0.11	1.02	0.08	1.01	0.04	0.57
433	19	0.15	0.67	-0.05	0.66	0.09	0.41	0.08	0.63
510	18	0.10	0.40	0.02	0.57	0.24	1.15	0.12	0.43
431	17	0.03	0.44	0.07	0.47	0.12	1.08	0.07	0.40
446	17	0.03	0.64	0.46	1.23	0.12	2.69	0.03	0.64
484	17	0.12	0.63	0.31	0.69	0.25	3.27	0.06	0.46
507	16	0.02	0.41	0.26	1.49	0.03	6.34	0.00	0.40
411	15	-0.02	0.62	0.46	1.42	0.10	3.36	0.02	0.63
420	15	0.04	0.52	0.28	1.33	0.18	3.04	0.04	0.52
508	13	0.06	0.58	0.42	1.56	0.06	5.00	0.01	0.56
463	13	0.05	0.64	0.46	2.01	0.28	3.78	0.05	0.66
407	13	0.19	0.47	-0.33	0.45	0.23	0.49	0.19	0.44
456	13	0.10	0.54	0.37	0.70	0.22	3.73	0.05	0.50
518	12	0.04	0.66	0.51	1.95	0.07	5.00	0.01	0.65
504	12	-0.09	0.25	0.23	0.91	0.20	7.01	0.02	0.34
512	12	0.05	0.62	0.50	1.45	0.09	4.56	0.01	0.62
548	12	0.02	0.47	0.42	0.99	0.05	10.07	0.00	0.47
483	11	0.04	0.63	0.48	3.46	0.38	8.31	0.03	0.64
20	10	0.00	0.36	-0.32	2.30	0.22	2.11	0.05	0.40
524	10	0.02	0.64	0.38	2.85	0.55	3.67	0.08	0.68
535	10	0.07	0.60	0.50	1.12	0.17	5.90	0.02	0.59
527	10	-0.03	0.05	0.06	0.72	0.07	13.46	0.00	0.11
555	10	0.04	0.35	0.21	2.49	0.12	17.12	0.01	0.31
516	10	0.08	0.78	0.50	3.17	0.05	3.09	0.01	0.76
557	10	0.03	0.37	0.02	1.02	0.07	2.38	0.02	0.33
mean$_i$		0.045	0.516	0.26	1.41	0.16	4.59	0.040	0.510
sd$_i$		0.056	0.156	0.24	0.83	0.12	3.89	0.042	0.145
max$_i$		0.190	0.782	0.51	3.46	0.55	17.12	0.194	0.758
min$_i$		-0.090	0.054	-0.33	0.45	0.03	0.41	0.004	0.108

Columns (1)-(4) repeat the information in table 1 for convenience. The headers to columns (3)-(4), and (5)-(6), denote the dependent variable. The $\beta_{10,IMP}$ in column (9) is the implied β_{10} calculated using sample estimates in place of the population moments, $\beta_{10,IMP} = \sigma_w^2 / \left(\sigma_w^2 + \sigma_e^2 + \sigma_u^2\right)$. The $\beta_{a,IMP}$ in column (10) is the implied β_a calculated using sample estimates in place of the population moments in $\beta_{a,IMP} = \frac{\sigma_w^2 + \rho\sigma_u^2 + \sigma_e^2}{\sigma_w^2 + \sigma_e^2 + \sigma_u^2}$. The correlation beween the individual β_{10} estimates in columns (3) and (9) is 0.64, and between the estimates in columns (4) and the implied β_a in column (10) is 0.96.

Table 3: The Effects of Participation

		$\hat{\beta}_{10,i}$		$\hat{\beta}_{10,C,t\in\mathcal{I}_i}$	
id	#	Estimate	t-stat	Estimate	t-stat
421	22	0.01	0.59	0.05	3.03
428	21	0.06	1.77	0.09	2.46
426	21	0.01	0.62	0.06	3.28
433	19	0.15	3.46	0.06	1.59
510	18	0.10	2.68	0.03	3.05
431	17	0.03	1.47	0.05	2.91
484	17	0.12	4.11	0.07	5.52
446	17	0.03	0.76	0.04	3.94
507	16	0.02	0.81	0.04	3.77
420	15	0.04	1.42	0.04	4.39
411	15	-0.02	-1.16	0.04	5.08
463	13	0.05	0.92	0.05	3.20
508	13	0.06	4.65	0.03	2.44
407	13	0.19	3.56	0.12	2.31
456	13	0.10	2.48	0.10	1.58
512	12	0.05	2.26	0.03	2.97
548	12	0.02	2.32	0.03	2.98
504	12	-0.09	-3.02	0.04	3.52
518	12	0.04	2.51	0.03	3.04
483	11	0.04	0.74	0.04	3.53
524	10	0.02	0.26	0.04	3.57
20	10	0.00	0.03	0.06	1.43
535	10	0.07	2.40	0.04	3.04
527	10	-0.03	-1.57	0.08	2.16
555	10	0.04	4.06	0.03	3.00
557	10	0.03	0.65	0.03	3.00
516	10	0.08	3.99	0.04	3.69
	mean_i	0.045		0.051	
	sd_i	0.056		0.023	
	> 0		12		24
	< 0		1		0

The mean and sd are the cross-sectional means and standard deviations. ‘ >0 ’ and ‘ <0 ’ are the number of regressions yielding statistically significant estimates at the 5% level.

Table 4: Bootstrap confidence intervals for various statistics of the cross-sectional distribution of persistence estimates from the regressions of the ten-year forecast revisions on the current-quarter forecast revisions, and of the annual forecast revisions on the current-quarter forecast revisions.

Two-sided	Mean		Standard deviation		Max		Min	
Panel A. 10-year forecast revisions on current-quarter forecast revisions								
Bootstrapping conditioning on t								
1%	0.014	0.052	0.029	0.076	0.083	0.314	-0.200	-0.004
5%	0.019	0.049	0.034	0.069	0.089	0.272	-0.135	-0.017
10%	0.021	0.047	0.036	0.065	0.097	0.255	-0.121	-0.022
Bootstrapping Not Conditioning on t								
1%	-0.022	0.023	0.030	0.070	0.040	0.216	-0.246	-0.038
5%	-0.018	0.017	0.032	0.065	0.048	0.187	-0.197	-0.048
10%	-0.015	0.015	0.034	0.062	0.055	0.168	-0.176	-0.056
Panel B. Annual forecast revisions on current-quarter forecast revisions								
Bootstrapping Conditioning on t								
1%	0.382	0.491	0.158	0.237	0.650	0.880	-0.195	0.092
5%	0.403	0.483	0.168	0.228	0.675	0.843	-0.123	0.065
10%	0.411	0.478	0.172	0.224	0.690	0.829	-0.099	0.051

The table presents the lower and upper critical values for two-sided tests at the given significance level.

Table 5: Bootstrapping the annual year-ahead forecasts. Confidence intervals for the cross-sectional mean and variance of the components of the error decompostions, under Assumptions I and II.

Two-sided	Mean		Std. deviation		Mean		Std. deviation	
	Assumption I, $\rho = 0$				Assumption II, $\theta = 1$			
θ or ρ								
1%	-28.115	35.277	0.596	252.691	-0.125	0.126	0.272	0.922
5%	-5.925	10.003	0.889	100.283	-0.045	0.106	0.295	0.580
10%	-1.527	4.436	1.009	39.003	-0.023	0.095	0.307	0.529
σ_e^2								
1%	2.778	3.672	2.097	3.120	1.357	1.977	0.713	1.345
5%	2.932	3.567	2.229	2.992	1.442	1.909	0.771	1.275
10%	2.985	3.514	2.312	2.902	1.474	1.866	0.798	1.229
σ_w^2								
1%	0.155	0.155	0.118	0.118	0.155	0.155	0.118	0.118
5%	0.155	0.155	0.118	0.118	0.155	0.155	0.118	0.118
10%	0.155	0.155	0.118	0.118	0.155	0.155	0.118	0.118
σ_u^2								
1%	2.296	3.154	2.187	3.946	3.957	4.573	3.666	4.341
5%	2.395	3.052	2.374	3.647	4.013	4.498	3.729	4.253
10%	2.442	2.974	2.467	3.511	4.049	4.456	3.762	4.205
β_{10}								
1%	0.038	0.039	0.038	0.040	0.038	0.040	0.038	0.041
5%	0.038	0.039	0.038	0.039	0.038	0.039	0.038	0.040
10%	0.038	0.039	0.038	0.039	0.038	0.039	0.038	0.040
β_a								
1%	0.384	0.480	0.158	0.236	0.392	0.486	0.158	0.240
5%	0.394	0.471	0.165	0.228	0.402	0.478	0.170	0.231
10%	0.401	0.467	0.171	0.222	0.409	0.473	0.174	0.225

The table presents the lower and upper critical values for two-sided tests at the given significance level.

6 Appendix

6.1 Model of output growth with one persistent and one transitory component

Patton and Timmermann (2010, 2011) model output growth as comprising a persistent component (x_t) and a temporary component (u_t):

$$\Delta y_t = x_t + u_t.$$

$$x_t = \phi x_{t-1} + \varepsilon_t, \;\; |\phi| < 1, \tag{9}$$

where u_t and ε_t are iid, $u_t \sim \left(0, \sigma_u^2\right)$, $\varepsilon_t \sim \left(0, \sigma_\varepsilon^2\right)$, and are uncorrelated contemporaneously and at all leads and lags. Here, ϕ measures the persistence of the *growth rate* of output. Jain (2019) attributes to each forecaster a model of this sort, and x_t, u_t and ε_t (and σ_u^2 and σ_ε^2) carry a forecaster subscript. We omit that script for convenience. Relative to the model in the main text, there is a single shock to the permanent component x_t, given by ε_t, and the transitory shock is u_t. Jain (2019) uses quarterly forecasts of quarterly quantities (CPI inflation), and letting $r_{t,h} = E_t\left(\Delta y_{t+h}\right) - E_{t-1}\left(\Delta y_{t+h}\right)$, derives:

$$r_{t,h} = \phi^h \varepsilon_t, \text{ for } h > 1 \tag{10}$$

$$r_{t,0} = \varepsilon_t + u_t \tag{11}$$

(Recognizing that the expectations operators should also have a forecaster script.) Then it follows immediately that the regressions of $r_{t,h}$ on $r_{t,h-1}$ for $h = 1, \ldots, 4$ (determined by the availability of the SPF survey data) each estimate the parameter ϕ.

As explained in the main text, we use the $Q1$ survey forecasts, so t indexes years, and different horizon forecasts - the current quarter, the current year, and the 10-year average. Nevertheless, our data can be approximately interpreted within Jain's approach as follows. In our case, $r_{t,0} = u_t + \varepsilon_t$ is the current quarter revision (at an annual rate) between the year $t-1$, $Q1$ and year t, $Q1$ forecasts (of the year t, $Q1$ value). The revision to the annual growth rate forecast (between year $t-1$, $Q1$ and year t, $Q1$) is $r_t\left(\Delta y_{a,t}\right) = r_{t,1}$, where $r_{t,1} = \phi \varepsilon_t$ from (11). If we approximate the 10-year average growth rate $\Delta y_{10,t}$ by $\frac{1}{10}\sum_{s=1}^{10} \Delta y_{t+1}$, where Δy_t is the annual growth rate between years t and $t-1$, we can write the revision as

$$\begin{aligned} r_t\left(\Delta y_{10,t}\right) &= \frac{1}{10}\sum_{s=1}^{10} r_t\left(\Delta y_{t+s}\right) = \frac{1}{10}\sum_{s=1}^{10} r_{t,s} \\ &= \frac{1}{10}\left(\frac{1-\phi^{11}}{1-\phi}\right)\varepsilon_t. \end{aligned}$$

From the expressions for the current-quarter revisions, $r_t\left(\Delta y_{cq,t}\right) = r_{t,0}$, and the ten-year revisions, we find:

$$\beta_{10} = \frac{Cov\left(r_t\left(\Delta y_{10,t}\right), r_t\left(\Delta y_{cq,t}\right)\right)}{Var\left(r_t\left(\Delta y_{cq,t}\right)\right)} = \frac{\frac{1}{10}\left(\sigma_\varepsilon^2 \frac{1-\phi^{11}}{1-\phi}\right)}{\sigma_u^2 + \sigma_\varepsilon^2} \tag{12}$$

and from the current and annual:

$$\beta_a = \frac{Cov\left(r_t\left(\Delta y_{a,t}\right), r_t\left(\Delta y_{cq,t}\right)\right)}{Var\left(r_t\left(\Delta y_{cq,t}\right)\right)} = \frac{\phi\sigma_\varepsilon^2}{\sigma_u^2 + \sigma_\varepsilon^2} \tag{13}$$

These two equations clarify that neither β_{10} nor β_a directly estimates the parameter ϕ, often interpreted as measuring persistence in the permanent/transitory shock framework. These expressions are based on a different framework from the decomposition in the main text, and differ from equations (5) to (7). But unsurprisingly the perceived long- and medium-term responses given by (12) and (13) are increasing in the variance of permanent shocks (σ_ε^2), 'persistence' (ϕ) and decreasing in the variance of transitory shocks (σ_u^2).

6.2 Bootstrapped error decompositions - Bootstrapping the 10-year forecasts

Table 6: Bootstrapping the 10-year forecasts. Confidence intervals for the cross-sectional mean and variance of the components of the error decompostions, under Assumptions I and II.

Two-sided	Mean		Std. deviation		Mean		Std. deviation	
	Assumption I, $\rho = 0$				Assumption II, $\theta = 1$			
θ or ρ								
1%	0.740	0.926	0.185	1.016	0.233	0.288	0.194	0.307
5%	0.753	0.858	0.201	0.628	0.246	0.284	0.200	0.266
10%	0.756	0.838	0.210	0.528	0.251	0.282	0.204	0.255
σ_e^2								
1%	3.871	4.367	2.539	3.391	1.373	1.594	0.647	1.008
5%	3.924	4.304	2.646	3.267	1.399	1.574	0.695	0.946
10%	3.957	4.280	2.702	3.214	1.411	1.560	0.711	0.917
σ_w^2								
1%	0.107	0.183	0.044	0.122	0.107	0.183	0.044	0.122
5%	0.113	0.169	0.051	0.108	0.113	0.169	0.051	0.108
10%	0.117	0.165	0.055	0.103	0.117	0.165	0.055	0.103
σ_u^2								
1%	1.876	2.001	2.434	2.709	4.380	4.753	3.466	4.195
5%	1.886	1.986	2.465	2.688	4.429	4.708	3.585	4.112
10%	1.893	1.976	2.481	2.634	4.453	4.680	3.648	4.062
β_{10}								
1%	0.027	0.046	0.020	0.057	0.028	0.047	0.020	0.057
5%	0.029	0.043	0.023	0.050	0.029	0.043	0.024	0.050
10%	0.030	0.041	0.024	0.047	0.030	0.042	0.025	0.048
β_a								
1%	0.489	0.517	0.134	0.161	0.503	0.531	0.137	0.160
5%	0.492	0.514	0.137	0.158	0.507	0.528	0.139	0.159
10%	0.494	0.512	0.139	0.156	0.509	0.526	0.141	0.157

The table presents the lower and upper critical values for two-sided tests at the given significance level.