Individual forecaster perceptions of the persistence of shocks to GDP

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.

September 9, 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 renewed interest in the nature of output fluctuations.[1] 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).

Krane (2011) and Bluedorn and Leigh (2018) investigate the beliefs or perceptions of professional forecasters regarding shocks to output, and the extent to which those shocks are expected 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.[2] 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 key innovation of our paper is to consider the heterogeneity in individual forecaster perceptions of the persistence of output growth. To this end we use the individual respondents to the US SPF.[3] Unless forecasters have identical perceptions of the persistence of output growth, the use of the consensus or aggregate may be misleading. Furthermore, 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

[1] Some of the seminal early contributions are Friedman's 'plucking model' (Friedman (1964, 1993)) and Beveridge and Nelson (1981).

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

[3] 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.

whether investigating forecasters' perceptions of the nature of output shocks using the consensus might also be misleading. In contrast, a finding that forecasters' perceptions of persistence are essentially the same would be the overarching rationale for considering the consensus forecasts.

We explore the heterogeneity in the directly-estimated correlations between forecast revisions at different horizons and in terms of decompositions of GDP into permanent and transitory components. For each forecaster in our data set, 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. We then use the relationships between short-horizon forecast revisions and medium-term and long-horizon forecast revisions (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. We do this two ways: by regressing the medium and long-term revisions on the short-term revision; and following Krane (2011) using the forecast revisions to estimate decompositions into temporary and permanent shocks.

Our first set of results suggest some mixed evidence. The means of the distribution of forecaster-specific regression coefficients as well as the regression coefficients implied by the variance decompositions are similar to those from the consensus estimate. However, we find notable variation in the distributions of these coefficients. A test for cross-sectional homogeneity in beliefs about persistence in a panel with time-specific fixed effects (see e.g., Canova and Ciccarelli (2009)) rejects the null for the medium-horizon revisions, but not for the long-horizon revisions. Furthermore, the mean of the forecaster-specific estimates of the variances of permanent and transitory shocks differ noticeably from those calculated using the consensus forecast.

To further explore whether the apparent differences in perceptions are real, or are simply a product of the difficulties that typically afflict studies of individual forecasters, we undertake a number of additional exercises. One difficulty is that the differences in the estimates obtained from the individual analysis may partly reflect the typically relatively small samples of forecast data available at the individual level. A second difficulty is that individual respondents are active participants in the survey at different times and the perceived effect of output shocks may depend on the state of the business cycle when the forecasts are made. In this case, the heterogeneity in the results may reflect time variation in whether the GDP shock is perceived to be a shock to labour productivity, a monetary policy shock, and so on.

Our methodological contribution is a solution to this problem. 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 responding to different sets of surveys.[4] We simulate a set of imaginary forecasters who match the actual forecasters in terms of when and how often they participate, but their long- or medium-horizon forecast revisions

[4] 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.

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 match the empirical distribution of the actual forecasters, we can infer that the observed spread of estimates of the actual forecasters is either random or is due to different forecasters participating during periods with different economic conditions.

We find evidence that this shuffling at each point in time does not produce different coefficient dispersion for the long-run forecasts, but does for the annual forecasts. This suggests the heterogeneity in the long-run persistence regressions may simply reflect small-sample variation, but the heterogeneity in the annual persistence estimates likely does not. Similarly, shuffling the long-term forecast revisions does not influence the variance decomposition results specifically, but shuffling the annual forecasts does change the heterogeneity observed in the perceived degree of decay of transitory shocks.

To examine the influence of different economic conditions on the observed heterogeneity, we see if the estimates of persistence parameters and shock decompositions are systematically correlated with the proportion of individual forecasters' projections that were made before the 2007-9 recession. We find no clear evidence of a correlation for the persistence regression parameters, but do find that forecasters' perceptions of the relative importance of permanent shocks compared to transitory shocks decreased after the Recession.

A closely-related paper is the study of perceived inflation persistence by Jain (2019).[5] Jain uses the US Survey of Professional Forecasters quarterly forecasts available from the current quarter up to one-year ahead to estimate perceived inflation persistence. Instead we choose to use 10-year ahead annual average forecast data for our study of output growth, which restricts the number of forecasts we are able to draw on (see section 2). 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, as well as the permanent-transitory shock decompositions.

The plan of the remainder of the paper is as follows. Section 2 describes the forecast data we use. Section 3 outlines two methods we use for measuring perceptions of persistence. Section 4 describes our empirical results. Section 5 offers some concluding remarks.

[5] Jain shows that, for a simple model of perceived inflation consisting of an unobserved persistent component and a (white noise) transitory term (see, e.g., Patton and Timmermann (2010, 2011)), 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. We relate our regression-approach persistence parameter estimates to the permanent-transitory decomposition of Krane (2011).

2 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.[6]

The US SPF has the advantage of providing individual-level forecast data. 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. This constitutes fewer medium to long-term horizons than e.g., the analysis of Krane (2011), and we explain below how the available forecast data is used.

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.[7]

Because we use individual-level forecast data, there will naturally be missing observations reflecting entry and exit to the survey, and non-response by otherwise regular participants in particular periods: respondents only provide forecasts to some surveys. Consequently we will need to assess the extent to which forecaster heterogeneity reflects forecasters being active during different economic conditions.

---

[6] 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.

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

3 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. We then explain the relationship between the two approaches.

3.1 Regression Approach

One approach to determining the perceived persistence of shocks to GDP is to consider directly 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.

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.[8] That is, $r_t\left[z_{t+h}\right] = z_{t+h|t} - z_{t+h|t-1}$, where $z_{t+h|t}$ is the forecast of z_{t+h} made at time t. The long-horizon regression dependent variable $r_t\left[\Delta y_{t,t+h}\right]$ is either $r_t\left[\Delta y_{t,10}\right]$, where $\Delta y_{t,10}$ is the 10-year annual average growth rate, or $r_t\left[\Delta y_{t,a}\right]$, where $\Delta y_{t,a}$ is the current-year annual growth rate. The SPF allows us to define the forecast revision $r_t\left[.\right]$ as the difference between the forecasts made in the first quarters of years $t-1$ and t. The right-hand-side 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 revision is again defined over the same two forecast origins, namely, 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 $\Delta y_{t,10}$.[9] The revision in the year t annual growth rate, $r_t\left[\Delta y_{t,a}\right]$, is naturally at an annual rate. 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. The availability of $r_t\left[\Delta y_{t,10}\right]$, $r_t\left[\Delta y_{t,a}\right]$ and $r_t\left[\Delta y_{t,cq}\right]$ means that we are able to estimate a regression of long-horizon forecast revisions (10-year) on the revision to current quarter forecasts, as well as a regression of medium-horizon forecast revisions (year-ahead) on current-quarter forecast revisions. We consider what the estimated perceived 'long-term' and 'medium-term' responses of output tell

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

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

us about the perceived persistence of shocks below, and the interpretation of the null that $\beta_h = 0$ in (1) in terms of permanent and transitory shocks.

We will estimate (1) for consensus forecast revisions, separately for each individual respondent with a minimum number of revisions, as well as in a pooled panel enabling a test of the assumption that $\beta_{h,i} = \beta_{h,j}$ for all i and j.

Notice that (1) bears some resemblance to a long-horizon regression, such as that used by Fama and French (1988) and many others to analyze whether continuously-compounded k-period stock returns are predictable from current period dividend yields. In (1), the long-horizon growth rate can be viewed as the cumulation of a sequence of annual (or quarterly) growth rates. Hodrick (1992) explores estimation and inference in such environments, and proposes a VAR analysis. However, we consider forecast *revisions* rather than actual data, or the forecasts themselves, and it seems unlikely that there will be much serial correlation in the revision series for individuals (e.g., between $r_{t,i}\left[\Delta y_{t,t+h}\right]$ and $r_{t-s,i}\left[\Delta y_{t-s,t-s+h}\right]$ for an individual i, for $s = 1, 2, \ldots$) or across individuals (e.g., between $r_{t,i}\left[\Delta y_{t,t+h}\right]$ and $r_{t-s,j}\left[\Delta y_{t-s,t-s+h}\right]$, for $s = 1, 2, \ldots$), suggesting a panel VAR analysis as in Canova and Ciccarelli (2013) is unlikely to be a fruitful approach. Under rational expectations, of course, the revisions will be uncorrelated. In section 4.2 we report some preliminary results consistent with this.

3.2 Permanent and Transitory Component Decomposition

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.

Our forecast data allows us to implement the approach of Krane (2011) as follows. Matching his equation (eqn. 6) we have:[10]

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

$$r_t\left[\Delta y_{t,a}\right] = 2.5w_t + \theta e_t + \rho u_t \tag{3}$$

$$r_t\left[\Delta y_{t,10}\right] = 3.85w_t + \frac{\theta e_t + \rho u_t}{10} \tag{4}$$

(2) - (4) are the only forecast revisions which can be constructed from the surveys held at times t, $Q1$ and $t-1$, $Q1$.[11] It is not possible to calculate both θ and ρ from this set of equations. So, as an identification restriction, we assume that $\theta = 1$, implying that the perceived permanent shock to the level of (log) GDP is fully absorbed at impact.

Under this identifying restriction, the shocks $\{w_t, e_t, u_t\}$, their variances, and ρ can be estimated as follows. We calculate the quantity $Z_{1t} = 2.5 \times r_t\left[\Delta y_{t,cq}\right] - r_t\left[\Delta y_{t,a}\right]$ (this cancels w_t) and regress this on $Z_{2t} = -2.4 \times r_t\left[\Delta y_{t,cq}\right] + 2.5 \times r_t\left[\Delta y_{t,a}\right] - r_t\left[\Delta y_{t,10}\right]$ (which does not depend on either w_t or e_t). Letting δ denote the regression coefficient defined by regressing Z_{1t} on Z_{2t}, straightforward calculations reveal that:

$$\hat{\rho} = \frac{2.4\hat{\delta} + 2.5}{2.4\hat{\delta} + 1}.$$

Substituting $\hat{\rho}$ into Z_2 gives $\hat{u}_t$, and from Z_{1t} we then calculate $\hat{e}_t$, and finally $\hat{w}_t$ from (2).

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

[10] Given eqn. (2) for quarterly growth, to obtain the expression for average annual GDP (eqn. (3)) we need to cumulate over the next three quarters:

$$\begin{aligned} r_t\left[\Delta y_{t,a}\right] &= \frac{1}{4}[w_t + e_t + u_t \\ &\quad +2w_t + \theta_1 e_t + \rho_1 u_t \\ &\quad +3w_t + \theta_2 e_t + \rho_2 u_t \\ &\quad +4w_t + \theta_3 e_t + \rho_3 u_t] \\ &= 2\frac{1}{2}w_t + \frac{1}{4}\left(1 + \theta_1 + \theta_2 + \theta_3\right)e_t + \frac{1}{4}\left(1 + \rho_1 + \rho_2 + \rho_3\right)u_t. \end{aligned}$$

Hence θ and ρ are the average of the first-year effects of e_t and u_t. The permanent shock on the growth rate has a coefficient of 1 in each period.

In eqn (4), the impact of w_t on the revision to average growth over 10 years is $\frac{1}{10}(9 \times 4)$, plus $\frac{1}{10} \times 2\frac{1}{2}$ (from (3)), which explains the 3.85. e_t and u_t are only assumed to affect the first-year.

[11] 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:

$$\beta_{10,IMP} = \frac{Cov\left(r_t\left[\Delta y_{t,10}\right], r_t\left[\Delta y_{t,cq}\right]\right)}{Var\left(r_t\left[\Delta y_{t,cq}\right]\right)} = \frac{3.85\sigma_w^2 + \frac{\sigma_e^2+\rho\sigma_u^2}{10}}{\sigma_w^2 + \sigma_e^2 + \sigma_u^2} \tag{5}$$

The implied population value of β_a is:

$$\beta_{a,IMP} = \frac{Cov\left(r_t\left[\Delta y_{t,a}\right], r_t\left[\Delta y_{t,cq}\right]\right)}{Var\left(r_t\left[\Delta y_{t,cq}\right]\right)} = \frac{2.5\sigma_w^2 + \sigma_e^2 + \rho\sigma_u^2}{\sigma_w^2 + \sigma_e^2 + \sigma_u^2}. \tag{6}$$

As an alternative to the OLS regression of (1), we can estimate β_{10} (β_a) by substituting the sample estimates of the population moments into (5) and (6). The perceived degrees of persistence will be smaller the smaller the variance of w_t, relative to that of e_t and u_t. Notice that (5) indicates that rejecting $\beta_{10} = 0$ does not indicate that a forecaster believes that there is a permanent effect of a shock to output on long-run growth, because $\beta_{10} \neq 0$ is consistent with $\sigma_w = 0$, because of the presence of the term $\frac{\sigma_e^2+\rho\sigma_u^2}{10}$ in the numerator. Hence failure to reject $\beta_{10} = 0$ does not indicate that the forecaster believes output fluctuates around a stable trend, because $\beta_{10} = 0$ does not imply $\sigma_w = \sigma_e = 0$. Consequently, we supplement the regression approach with an analysis of the perceived persistence of output using the decomposition of shocks into permanent and transitory components.

In equations (1) to (6) 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 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.[12]

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 Great Recession, but there are also peaks at the time of the 2001 recession, and in 1995, when the US economy appeared to falter. 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.[13]

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 0.051 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 0.643. The estimate for the ten-year forecasts suggests 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).

The remaining rows in table 1 report results for the individual forecasters, based on individual-specific regressions of (1) (with heteroscedastic and autocorrelation consistent standard errors). The table indicates that the means of the distributions of the forecaster-specific β_{10} and β_a estimates are close to the values we find for the revisions to the consensus forecasts. This is true for both the regression estimates ($\hat{\beta}_{10}$ and $\hat{\beta}_a$) and the implied estimates ($\beta_{10,IMP}$ and $\beta_{a,IMP}$). In this sense, the

[12] 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.

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

consensus gives accurate persistence estimates of the average forecaster. That said, the estimates for the individual respondents vary widely, from being negative (-0.09) to large and positive (0.19) for $\hat{\beta}_{10}$, and from 0.053 to 0.782 for $\hat{\beta}_a$. Just under a half of the $\hat{\beta}_{10}$ estimates are significantly different from zero (absolute value of the t-statistic exceeds 2), and of these all but one is positive. Hence for around a half of the individuals in the survey we do not reject the null that $\beta_{10} = 0$ for the long-horizon forecasts. At the medium-horizon there is less divergence in views. For 24 out of the 27 forecasters in the sample we reject the null that $\beta_a = 0$ in favour of $\beta_a > 0$, and the three for whom we do not reject are the forecasters who made the fewest survey responses. Although the marked divergence of perceptions primarily bears on the long-horizon (ten-year) response, in that the variation in β_a is more moderate than the variation in β_{10}, it remains to be seen whether the apparent differences across forecasters in both sets of estimates is 'real' or consistent with random variation.

A panel approach is applied to (1) to test the apparent heterogeneity in belief perceptions suggested by the dispersions of the $\hat{\beta}$ estimates in table 1. The test is in the spirit of Canova and Ciccarelli (2009) and Manzan (2011). In section 3 we argued that revisions are unlikely to be serially correlated. However, it might be that errors are clustered within time period t: a macro-shock at time t might be expected to affect all individuals forecasting at that time.[14] We include time fixed effects to account for possible within-year clustering.[15] In terms of (1), let $R_{h,i} = (r_{1,i}[\Delta y_{1,1+h}], \ldots, r_{T,i}[\Delta y_{T,T+h}])'$ and $R_{0,i} = (r_{1,i}[\Delta y_{1,1}], \ldots, r_{T,i}[\Delta y_{T,T}])'$ contain the long-horizon and current-quarter revisions for forecaster i, and $R_h = (R_{h,1}, \ldots, R_{h,N})'$ and $R_0 = (R_{0,1}, \ldots, R_{0,N})'$, then we estimate:

$$R_h = \mathbf{X}\boldsymbol{\beta}_h + v_h \tag{7}$$

where:

$$\mathbf{X} = \begin{bmatrix} R_{0,1} & 0 & & \cdots & 0 \\ & R_{0,2} & & & \\ \vdots & & R_{0,3} & & \\ & & & \ddots & \\ 0 & & \cdots & & R_{0,N} \end{bmatrix}$$

$\boldsymbol{\beta}_h = (\beta_1, \ldots, \beta_N)'$. We estimate (7) with time dummies, and the null of homogeneity is the set of $N-1$ restrictions that the β_i are all equal.

We find we do not reject the null of equal β's for the revisions to the 10-year growth rates

[14]Notice that the left and right-hand side revisions are made at time t, so it does not necessarily follow that the errors will be correlated over i for a given t.

[15]One may also wish to calculate a cluster-robust estimate of the variance matrix of the parameter estimates for the test of homogeneity. Clustering on time would allow for within-period correlation, while imposing no correlation across time. However, testing for homogeneity with clustered standard errors is not possible in our setup because the cluster-robust variance matrix is not full rank, as discussed by Cameron and Miller (2015).

($F\,(26,326)$ statistic of 0.99 with a p-value of 0.47), but we do reject for the year-ahead annual revisions ($F\,(26,338)$ statistic of 4.24 with a p-value of 0.00).

Finally, Hodrick (1992) proposes estimating the β-parameter from a VAR. However, we found little evidence of serial correlation in the forecast revisions, and did not pursue this approach.[16]

We undertake a number of additional exercises to better understand these findings. In section 4.4 we consider the effects of forecasting at different times. Then in section 4.5 we consider a bootstrap approach based on the individual regressions. Firstly, the results for the permanent-transitory decompositions are described in section 4.3.

4.3 Permanent-transitory decompositions

The right-hand-side columns of table 1 compare the individuals in terms of the variances of the components in the decomposition of the shocks à la Krane (2011) (under the identifying assumption that $\theta = 1$). In contrast to the regression approach, we do not find a close match between the estimates for the consensus forecasts and the average over the individual-specific estimates. Moreover, the estimates of the components (the variances σ_e^2, σ_w^2 and σ_u^2, and the parameter ρ) vary widely across individuals (see columns (7) to (10)), and the cross-sectional standard deviations are of a similar order of magnitude to the mean values. Table 2 shows the variation over forecasters in the perceived importance of the three shocks in explaining the current-quarter output growth revision. The cross-sectional mean suggests that transitory shocks explain around two thirds of the variance in current-quarter revisions, and permanent shocks to the level (e_t) explain the remaining third. These shares are about the same in the consensus forecast even though the shock variances are very different (see table 1). Permanent shocks to the growth rate (w_t) are of a smaller order of magnitude. Yet perceptions vary greatly over individuals, with the temporary shock accounting for as little as one third of the variation up to in excess of 90%. In section 4.5 we ask whether the inter-forecaster variation is statistically significant, given that the estimates of the component variances may not be precisely determined, and respondents will have been active at different times. In section 4.6 we consider whether there is any evidence that forecasters' perceptions about the relative importance of permanent and temporary shocks have been affected by experiencing the Great Recession.

The cross-sectional variability observed in the estimates of the components do not translate to the same extent to variability in the implied estimates $\beta_{10,IMP}$ and $\beta_{a,IMP}$ (see columns (11) and (12) of table 1). This suggests the estimates of the components may contain additional information on inter-forecaster perceptions relative to the regression-based estimates of persistence.

The cross-sectional characteristics of the implied β_a-estimate closely match those of the direct

[16] We ran a panel VAR-Granger causality Wald test in STATA 16.1 (STATA command pvargranger), having estimated a first-order Panel VAR with time fixed effects and robust standard errors (STATA command pvar). Neither lagged revision is statistically significant in the equation for the other revision.

regression estimates - the cross-sectional mean and standard deviation are virtually indistinguishable. (Recall that the implied estimates, the $\beta_{a,IMP,i}$ are calculated from substituting the sample estimates of the variances of the components, and the estimate of ρ, into (5), and the implied estimates of $\beta_{10,IMP,i}$ come from (6)). However the correspondence between the direct estimates and implied estimates for β_{10} is less close. The mean is similar, but the standard deviation of the implied estimates is less than a third of that of the direct estimates. That is, there is extraneous cross-sectional variation in the direct β_{10}-estimates relative to that implied by the permanent-transitory shock model, and this does not occur in the β_a-direct estimates. In section 4.5 we test whether the variation in the direct β_{10}-estimates is simply random, as well as testing whether the variation in the implied estimates is random.

Finally, in the notes to the table we indicate that the cross-sectional correlation between the direct estimates and the implied estimates is 0.95 for β_a, consistent with the close correspondence between the first two moments of the cross-sectional distributions we mentioned above, while that for β_{10} is only 0.35.

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 and β_a's, denoted $\hat{\beta}_{10,C_i}$ and $\hat{\beta}_{a,C_i}$, respectively: these are the β_{10} and β_a-estimates using the consensus forecasts for the surveys to which respondent i filed a forecast. If $\hat{\beta}_{10,i}$ and $\hat{\beta}_{10,C_i}$ (or $\hat{\beta}_{a,i}$ and $\hat{\beta}_{a,C_i}$) are highly correlated across respondents, one could attribute the cross-sectional variation in the β_{10}-estimates evident in table 1 to the 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 $\hat{\beta}_{10,C_i}$, along with the $\hat{\beta}_{10,i}$ estimates to aid comparison, as well as the estimates of $\hat{\beta}_{a,C_i}$ and $\hat{\beta}_{a,i}$. We also provide some summary statistics of the cross-section distributions. The cross-sectional standard deviation is more than halved for the $\hat{\beta}_{10,C_i}$, at 0.023, compared to 0.056 for the $\hat{\beta}_{10,i}$. 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. Similarly, the cross-sectional standard deviation is nearly halved for $\hat{\beta}_{a,C_i}$, compared to $\hat{\beta}_{a,i}$. The correlation between $\hat{\beta}_{10,i}$ and $\hat{\beta}_{C_i}$ 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. For the revisions of the annual forecasts, the correlation (between $\hat{\beta}_{a,C_i}$ and $\hat{\beta}_{a,i}$) is only 0.18, suggesting time of participation plays a relatively unimportant role in explaining cross-sectional differences.

4.5 Bootstrap test of the exchangeability of forecast revisions in medium-horizon and long-horizon revisions, and in permanent-transitory shock decompositions

The individual heterogeneity in the different β estimates in the individual time-series regressions, and in the relative importance attributed to permanent and transitory shocks, 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, say, are not significantly different from one another. We consider whether the simulated distributions of regression coefficients, and permanent-transitory shock decompositions, are consistent with the cross-sectional distributions of these quantities calculated using the actual forecast data.

4.5.1 Regression estimates

We illustrate the approach for the regression estimates, when the 10-year forecast revisions are randomly shuffled. Random assignment (with replacement) of the period t 10-year revisions is made to the active participants at time t. We consider whether the cross-sectional distribution of the actual $\hat{\beta}_i$ estimates is consistent with the bootstrap distribution obtained by random re-assignment.[17] 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 $r_{t,i}\left[\Delta y_{t,10}\right]$ denote the long-horizon forecast revision of respondent i to survey t, where

[17] 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.

$i = 1, \ldots, N$, 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 $\{r_{t,j}[\Delta y_{t,10}]\}_{j=1,\ldots,N}$ for each individual who recorded a forecast revision (at time $t = 1$). This creates $\left\{r^*_{t,j}[\Delta y_{t,10}]\right\}_{j=1,\ldots,N}$. Missing values in the actual data are replicated in the bootstrap sample ($r^*_{t,j}[\Delta y_{t,10}]$ is a missing if $r_{t,j}[\Delta y_{t,10}]$ is missing).

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

2. Given $\left\{r^*_{t,j}[\Delta y_{t,10}]\right\}_{j=1,\ldots,N;t=1,\ldots,T}$, we estimate individual regressions for each individual i using variation over t:

$$\hat{\beta}^*_i = \frac{\sum_{t=1}^{T}\left(r^*_{t,i}[\Delta y_{t,10}] - \overline{r^*_{t,i}[\Delta y_{t,10}]}\right)\left(r_{t,i}[\Delta y_{t,cq}] - \overline{r_{t,i}[\Delta y_{t,cq}]}\right)}{\sum_{t=1}^{T}\left(r_{t,i}[\Delta y_{t,cq}] - \overline{r_{t,i}[\Delta y_{t,cq}]}\right)^2},$$

where the $r_{t,i}[\Delta y_{t,cq}]$ are the actual current-quarter forecast revisions.

Missing values result in the corresponding rows of $\left[r^*_{t,i}[\Delta y_{t,10}] : r_{t,i}[\Delta y_{t,cq}]\right]$ 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 2 $R = 1000$ 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. 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 $\{r_{t,j}[\Delta y_{t,10}]\}_{j=1,\ldots,N;t=1,\ldots,T}$: 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.

For concreteness, we have described bootstrapping the long-horizon revisions, but we can adopt exactly the same process to bootstrap the medium-term revisions instead.

Table 4 records the regression results for random shuffling of the long-horizon revisions and of the medium-horizon revisions. 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. Consider first the 10-year revisions in panel A. 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 reports the results for the medium-term forecast revisions regressions. The individual β-estimate summary statistics are recorded in table 1, column (5): the mean and standard deviation of the cross-sectional distribution of actual estimates are 0.516 and 0.156, and the max and min values are 0.782 and 0.054. Both the mean and standard deviation are outside the 99% bootstrapped intervals, suggesting the actual distribution of the estimates is not consistent with that simulated under the null of exchangeable year-ahead forecasts.

Although the medium-horizon regression estimates are less dispersed than the long-horizon estimates, differences in perceptions at this horizon appear to be real.

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 ρ for each individual using (2) to (4), and we also estimate the implied parameter estimates from (5) and (6). The identification of the component 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. The year-ahead annual forecast revisions and the actual current-quarter forecast revisions are left unchanged.

We calculate the 99%, 95% and 90% confidence intervals 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 with the bootstrap confidence intervals (table 5 right panel), we find that all lie within the intervals. Hence all the parameters are consistent with the empirical estimates under random reshuffling 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.

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 the left panel of table 5. The actual cross-sectional mean of ρ is 0.27, which lies outside the 99% bootstrap interval [-0.093,0.120]. However, for the variances of the shocks, the actual cross-sectional mean is either within, or close to being within, the bootstrapped interval. One interpretation of these results is that interchangeability fails because of forecasters distinct perceptions about ρ, the perceived effect of temporary shocks on the annual forecasts.

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, for each forecaster we would consider whether the regression coefficients, and the components of the permanent-transitory shock decompositions, are constant before and after the Recession. Unfortunately there are too few forecasts available for most respondents to reliably detect time variation in a regression such as (1), or in the relative importance attributed to permanent and transitory shocks. 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.[18] It is evident that one forecaster was only active in the earlier period (the ratio for id 20 is 1) while some forecasters made as few as 10% in the period before the Recession.

To determine whether there is an association between perceptions of persistence and the extent to which a forecaster was active in one period rather than the other, we rank each forecaster in terms of *i*) a measure of persistence, such as their estimated β_{10} or β_a, or the proportion of the variance of the revision to current-quarter output growth due to transitory shocks, say, and *ii*) the proportion of their forecasts made pre-2008. We test for an association by testing whether Spearman's rank correlation coefficient is zero.[19] 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.[20]

Table 6 reports the results. The statistic r is positive for the direct regression estimates, and the implied regression estimates, indicating a tendency to report higher β-estimates of persistence before the Recession. But none of the rank correlations are statistically significant at the conventional 5% level in 2-sided tests. However, we do find clear evidence of time variation in the relative importance of the perceptions of permanent and temporary shocks. For both permanent shocks (to the level, e and the growth rate, w), there is a statistically significant positive relationship, and for the temporary shock a significantly negative relationship. That is, the proportion of $Var\left(r_t\left[\Delta y_{t,cq}\right]\right)$

[18] 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.

[19] Details of the rank correlation test are given in the notes to table 6.

[20] 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).

attributable to permanent shocks has fallen since 2008, and that attributable to temporary shocks has increased. This suggests that since the Great Recession the perception of the relative importance of permanent shocks to both the level and rate of growth has decreased.

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 from regressing the revisions to 10-year ahead GDP growth forecasts on revisions to current-quarter forecasts 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. There is less dispersion in the persistence estimates from the regression of revisions to current-year growth on current-quarter growth. A test for cross-sectional homogeneity in beliefs in a panel setting rejects the null for the year-ahead horizon, but not for the 10-year forecasts.

We investigate whether the cross-sectional differences in perceptions are real or reflect small forecast samples and forecasters being active at different times. We simulate a set of imaginary forecasters who match the actual forecasters in terms of when and how often they participate, but their 10-year or annual growth forecast revisions are randomly drawn from the set of actual revisions for that period. The distributions of persistence estimates across our imaginary forecasters match the empirical distribution of actual forecasters at the 10-year horizon, but not the annual, from which we infer the heterogeneity in the 10-year persistence regressions reflects small-sample variation, but the heterogeneity in the annual persistence estimates is real.

We interpret the meaning of the regression-based persistence estimates in terms of a model that decomposes output shocks into permanent and transitory components. The variance decomposition is calculated separately for each survey respondent, and we find that the actual cross-sectional distribution of the perceived degree of decay of transitory shocks differs from that obtained by shuffling the annual forecast revisions. Whereas the distributions of the component variance estimates across the real and imaginary forecasters broadly match.

We find some evidence that perceptions were affected by the 2007-9 recession. Forecasters' perceptions of the relative importance of permanent shocks compared to transitory shocks is systematically negatively correlated with the proportion of projections made before the 2007-9 recession, suggesting the relative importance of transitory shocks has increased in recent times.

We conclude that differences between forecasters in their perceptions of long-run (10-year) persistence may be illusory, whereas at medium horizons (such as one-year ahead) the differences in perceptions are real, and are attributable to different perceptions of the effects of transitory shocks.

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.

Beveridge, S., and Nelson, C. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics*, *7*(2), 151–174.

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.

Cameron, A. C., and Miller, D. L. (2015). A PractitionerâĂŹs Guide to Cluster-Robust Inference. *Journal of Human Resources*, **50**, 317–372.

Canova, F., and Ciccarelli, M. (2009). Estimating Multicountry VAR Models. *International Economic Review*, *50*(3), 929–959.

Canova, F., and Ciccarelli, M. (2013). Panel Vector Autoregressive Models: A Survey. Cepr discussion papers 9380, C.E.P.R. Discussion Papers.

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.

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.

Fama, E. F., and French, K. R. (1988). Dividend yields and expected stock returns. *Journal of*

Financial Economics, *22*(1), 3 – 25.

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.

Friedman, M. (1993). The 'plucking model' of business fluctuations revisited. *Economic Inquiry*, **31**, 171–177.

Friedman, M. (1964). Monetary Studies of the National Bureau. In *The National Bureau Enters Its 45th Year, 44th Annual Report*, pp. 7–25: NBER. reprinted in The Optimum Quantity of Money and Other Essays, by Milton Friedman, ch. 12, pp.261-84. Chicago: Aldine, 1969.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies*, **5**, 357–386. No. 3.

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.

Manzan, S. (2011). Differential interpretation in the Survey of Professional Forecasters. *Journal of Money, Credit and Banking*, **43**, 993–1017.

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.

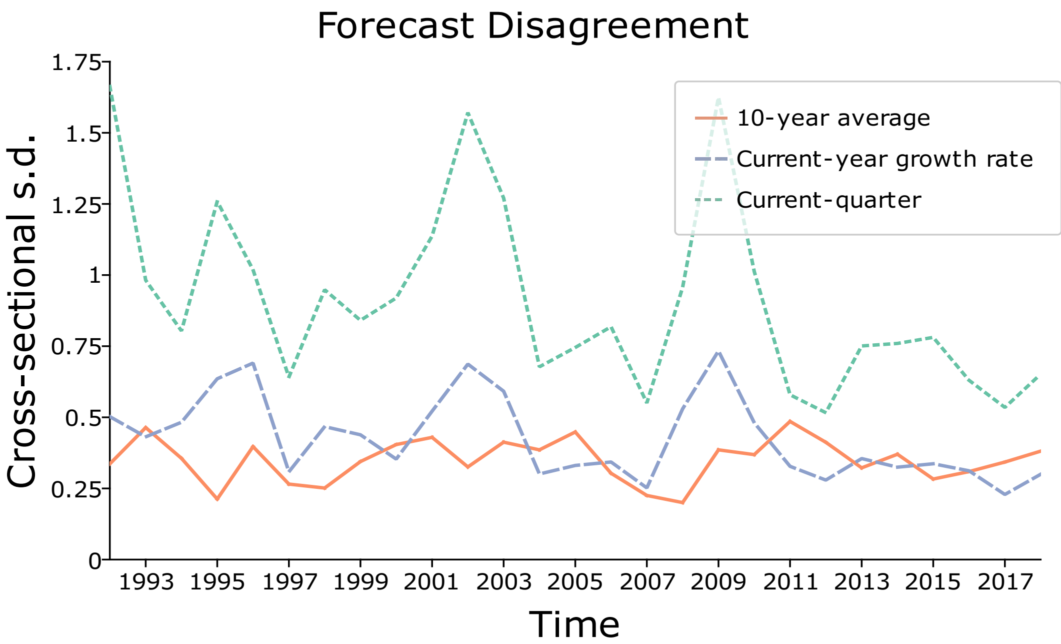

Figure 1: Forecaster Disagreement

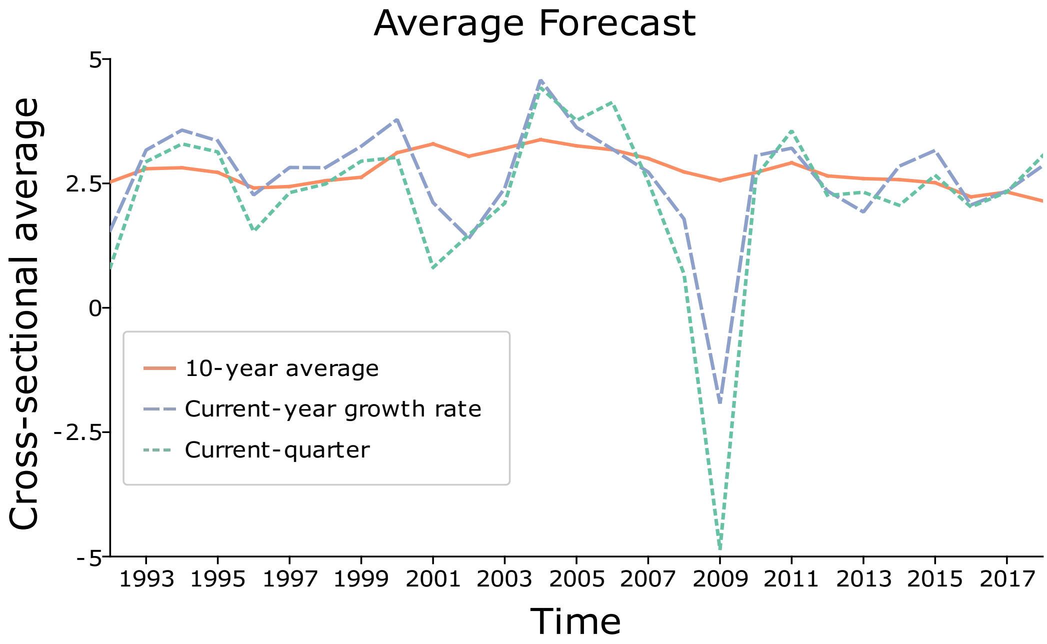

Figure 2: Consensus Forecasts

Table 1: Regression Estimates and Permanent and Transitory Component Parameters

		$r_t\,[\Delta y_{t,10}]$		$r_t\,[\Delta y_{t,a}]$		Permanent-Transitory Components						
id	No.	$\hat{\beta}_{10}$	t-stat	$\hat{\beta}_a$	t-stat	$\hat{\rho}$	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	0.051	3.355	0.643	15.815	0.458	1.023	0.003	1.963	0.068	0.646	.
421	22	0.009	0.591	0.441	12.824	0.268	0.881	0.004	2.377	0.051	0.469	0.591
428	21	0.063	1.770	0.588	5.918	0.099	1.212	0.006	1.009	0.068	0.595	0.714
426	21	0.014	0.617	0.570	7.746	0.390	1.603	0.008	3.283	0.065	0.593	0.619
433	19	0.153	3.456	0.670	4.898	-0.003	0.824	0.005	0.488	0.076	0.634	0.789
510	18	0.097	2.684	0.396	5.866	0.023	0.635	0.017	1.238	0.069	0.373	0.389
431	17	0.028	1.467	0.443	15.192	0.087	0.607	0.007	1.120	0.056	0.416	0.824
484	17	0.118	4.112	0.635	5.639	0.334	0.853	0.015	3.738	0.058	0.464	0.647
446	17	0.026	0.762	0.636	13.582	0.469	1.518	0.009	2.617	0.074	0.668	0.588
507	16	0.016	0.805	0.405	3.943	0.275	1.781	0.002	6.438	0.044	0.433	0.438
420	15	0.041	1.419	0.516	10.747	0.284	1.530	0.014	3.174	0.063	0.523	0.533
411	15	-0.023	-1.157	0.617	16.104	0.467	1.727	0.010	3.042	0.074	0.664	0.733
508	13	0.063	4.650	0.583	8.537	0.445	1.922	0.003	5.515	0.060	0.589	0.308
456	13	0.103	2.484	0.540	13.812	0.402	0.867	0.013	4.291	0.060	0.508	0.769
407	13	0.190	3.561	0.467	3.146	-0.225	0.622	0.009	0.659	0.063	0.384	0.615
463	13	0.049	0.920	0.640	15.939	0.446	2.415	0.023	3.717	0.080	0.671	0.769
548	12	0.021	2.316	0.475	6.749	0.441	1.214	0.004	10.488	0.051	0.499	0.083
512	12	0.051	2.265	0.617	20.506	0.502	1.783	0.005	4.925	0.066	0.636	0.583
518	12	0.040	2.505	0.665	10.404	0.531	2.414	0.004	5.201	0.070	0.681	0.250
504	12	-0.090	-3.019	0.250	2.604	0.225	1.072	0.019	6.073	0.044	0.347	0.250
483	11	0.039	0.739	0.626	12.734	0.483	4.190	0.031	8.387	0.075	0.660	0.545
516	10	0.076	3.992	0.782	11.981	0.530	3.913	0.001	3.404	0.078	0.781	0.600
557	10	0.029	0.652	0.368	1.565	0.020	1.302	0.003	2.279	0.041	0.378	0.100
555	10	0.044	4.059	0.347	2.230	0.228	2.988	0.006	18.266	0.035	0.337	0.100
527	10	-0.027	-1.572	0.053	0.629	0.052	0.878	0.004	12.961	0.012	0.113	0.200
535	10	0.070	2.400	0.599	19.554	0.536	1.391	0.008	6.496	0.066	0.620	0.100
524	10	0.021	0.258	0.642	14.782	0.334	3.238	0.043	3.288	0.091	0.677	0.400
20	10	0.003	0.029	0.357	1.480	-0.401	2.475	0.019	1.889	0.056	0.403	1.000
mean$_i$		0.045		0.516		0.268	1.698	0.011	4.680	0.061	0.523	
sd$_i$		0.056		0.156		0.242	0.981	0.010	4.006	0.016	0.150	
max$_i$		0.190		0.782		0.536	4.190	0.043	18.266	0.091	0.781	
min$_i$		-0.090		0.053		-0.401	0.607	0.001	0.488	0.012	0.113	
> 0			12		24							
< 0			1		0							

The first row 'Cons' is for the consensus forecasts. Subsequent rows refer to individual forecasters. The headers to columns (3)-(4), and (5)-(6), denote the dependent variable. In both cases the explanatory variable is $r_t\,[\Delta y_{t,cq}]$. The regression estimate t-statistics use heteroscedasticity and autocorrelation consistent standard errors (HACSEs). 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, $\beta_{10,IMP} = \frac{3.85\sigma_w^2 + \frac{\sigma_e^2 + \rho\sigma_u^2}{10}}{\sigma_w^2+\sigma_e^2+\sigma_u^2}$. The β_a in column (12) is the implied β_a calculated using sample estimates in place of the population moments, $\beta_{a,IMP} = \frac{2.5\sigma_w^2+\sigma_e^2+\rho\sigma_u^2}{\sigma_w^2+\sigma_e^2+\sigma_u^2}$. The correlation beween the individual β_{10} estimates in columns (3) and (11) is 0.35, 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: Proportion of Variance of Current-Quarter Revisions Due to Permanent and Transitory Shocks

	Permanent		Transitory
id	σ_e^2/V	σ_w^2/V	σ_u^2/V
421	0.270	0.001	0.729
428	0.544	0.002	0.453
426	0.328	0.002	0.671
433	0.626	0.004	0.370
510	0.336	0.009	0.655
431	0.350	0.004	0.646
484	0.185	0.003	0.812
446	0.366	0.002	0.632
507	0.217	0.000	0.783
420	0.324	0.003	0.673
411	0.361	0.002	0.637
508	0.258	0.000	0.741
456	0.168	0.002	0.830
407	0.482	0.007	0.511
463	0.392	0.004	0.604
548	0.104	0.000	0.896
512	0.266	0.001	0.734
518	0.317	0.001	0.683
504	0.150	0.003	0.848
483	0.332	0.002	0.665
516	0.535	0.000	0.465
557	0.363	0.001	0.636
555	0.141	0.000	0.859
527	0.063	0.000	0.936
535	0.176	0.001	0.823
524	0.493	0.007	0.501
20	0.565	0.004	0.431
mean_i	0.323	0.002	0.675
sd_i	0.149	0.002	0.150
max_i	0.626	0.009	0.936
min_i	0.063	0.000	0.370

Note than σ_e^2/V is shorthand for $\sigma_e^2/V\left[r_t\left[\Delta y_{t,cq}\right]\right]$, etc., that is, the variances of the shocks are relative to the variances of the revisions to the forecasts of current-quarter growth.

Table 3: The Effects of Participation

id	#	Ten-year				Annual			
		$\hat\beta_{10,i}$	t-stat	$\hat\beta_{10,C_i}$	t-stat	$\hat\beta_{a,i}$	t-stat	$\hat\beta_{a,C_i}$	t-stat
421	22	0.009	0.591	0.051	3.027	0.441	12.824	0.644	14.708
428	21	0.063	1.770	0.091	2.463	0.588	5.918	0.750	6.471
426	21	0.014	0.617	0.056	3.282	0.570	7.746	0.631	18.658
433	19	0.153	3.456	0.064	1.594	0.670	4.898	0.780	7.076
510	18	0.097	2.684	0.033	3.051	0.396	5.866	0.633	18.265
431	17	0.028	1.467	0.050	2.906	0.443	15.192	0.743	19.368
484	17	0.118	4.112	0.068	5.522	0.635	5.639	0.634	15.201
446	17	0.026	0.762	0.037	3.937	0.636	13.582	0.626	16.195
507	16	0.016	0.805	0.037	3.771	0.405	3.943	0.601	15.015
420	15	0.041	1.419	0.043	4.386	0.516	10.747	0.595	17.566
411	15	-0.023	-1.157	0.040	5.079	0.617	16.104	0.621	17.145
508	13	0.063	4.650	0.028	2.438	0.583	8.537	0.579	17.796
456	13	0.103	2.484	0.102	1.578	0.540	13.812	0.653	4.792
407	13	0.190	3.561	0.121	2.306	0.467	3.146	0.757	7.890
463	13	0.049	0.920	0.053	3.201	0.640	15.939	0.642	13.409
548	12	0.021	2.316	0.034	2.981	0.475	6.749	0.596	36.916
512	12	0.051	2.265	0.034	2.967	0.617	20.506	0.582	13.604
518	12	0.040	2.505	0.031	3.041	0.665	10.404	0.566	20.043
504	12	-0.090	-3.019	0.041	3.522	0.250	2.604	0.588	15.910
483	11	0.039	0.739	0.043	3.533	0.626	12.734	0.598	13.960
516	10	0.076	3.992	0.043	3.690	0.782	11.981	0.598	13.873
557	10	0.029	0.652	0.033	3.001	0.368	1.565	0.600	40.782
555	10	0.044	4.059	0.033	3.001	0.347	2.230	0.600	40.782
527	10	-0.027	-1.572	0.079	2.160	0.053	0.629	0.411	4.229
535	10	0.070	2.400	0.035	3.043	0.599	19.554	0.594	34.524
524	10	0.021	0.258	0.040	3.570	0.642	14.782	0.586	16.724
20	10	0.003	0.029	0.062	1.432	0.357	1.480	0.885	8.253
	mean_i	0.045		0.051		0.516		0.633	
	sd_i	0.056		0.023		0.156		0.088	
	> 0		12		24		24		27
	< 0		1		0		0		0

$\hat\beta_{10,C_i}$ and $\hat\beta_{a,C_i}$ denote the estimates using the consensus forecast revisions only from the surveys to which i responded.
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.
The correlation between $\hat\beta_{10,i}$ and $\hat\beta_{10,C_i}$ is 0.46, and the the correlation between $\hat\beta_{a,i}$ and $\hat\beta_{a,C_i}$ is 0.18.

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								
99%	0.012	0.054	0.030	0.078	0.080	0.338	-0.181	-0.010
95%	0.019	0.049	0.034	0.069	0.091	0.279	-0.133	-0.017
90%	0.022	0.047	0.037	0.065	0.097	0.253	-0.117	-0.022
Bootstrapping Not Conditioning on t								
99%	-0.025	0.026	0.029	0.072	0.040	0.228	-0.258	-0.041
95%	-0.019	0.019	0.034	0.067	0.050	0.194	-0.205	-0.049
90%	-0.015	0.016	0.035	0.064	0.056	0.172	-0.186	-0.055
Panel B. Annual forecast revisions on current-quarter forecast revisions								
Bootstrapping Conditioning on t								
99%	0.383	0.490	0.159	0.240	0.657	0.894	-0.174	0.092
95%	0.407	0.482	0.167	0.231	0.676	0.853	-0.133	0.065
90%	0.413	0.477	0.172	0.225	0.686	0.834	-0.103	0.055

The table presents the lower and upper values of a confidence interval at the specified level.

Table 5: Bootstrapping the error decomposition. Confidence intervals for the cross-sectional mean and variance of the components of the error decompostion

	Bootstrapping the annual forecasts				Bootstrapping the 10-year forecasts.			
Two-sided	Mean		Std. deviation		Mean		Std. deviation	
ρ								
99%	-0.093	0.120	0.266	0.784	0.241	0.284	0.216	0.277
95%	-0.047	0.103	0.296	0.573	0.247	0.280	0.221	0.267
90%	-0.031	0.092	0.311	0.533	0.251	0.278	0.225	0.263
σ_e^2								
99%	1.598	2.301	0.808	1.497	1.667	1.901	0.804	1.191
95%	1.683	2.212	0.889	1.438	1.689	1.870	0.854	1.146
90%	1.724	2.182	0.925	1.387	1.706	1.854	0.878	1.112
σ_w^2								
99%	0.011	0.012	0.010	0.011	0.008	0.013	0.003	0.009
95%	0.011	0.012	0.010	0.011	0.008	0.012	0.004	0.008
90%	0.011	0.012	0.010	0.011	0.009	0.012	0.004	0.008
σ_u^2								
99%	3.924	4.664	3.731	4.558	4.510	4.697	3.795	4.073
95%	4.059	4.589	3.844	4.474	4.536	4.676	3.850	4.053
90%	4.096	4.544	3.881	4.419	4.547	4.666	3.880	4.041
β_{10}								
99%	0.049	0.059	0.017	0.026	0.060	0.066	0.016	0.024
95%	0.051	0.058	0.019	0.025	0.060	0.065	0.016	0.023
90%	0.051	0.058	0.019	0.024	0.061	0.065	0.016	0.022
β_a								
99%	0.394	0.496	0.176	0.263	0.520	0.544	0.141	0.162
95%	0.409	0.487	0.184	0.251	0.523	0.541	0.144	0.159
90%	0.416	0.481	0.189	0.248	0.525	0.540	0.145	0.158

The table presents the lower and upper values of a confidence interval at the specified level.

Table 6: Forecaster Perceptions and the Great Recession: Rank Correlations

	$\hat{\beta}_{10}$	$\hat{\beta}_a$	$\beta_{10,IMP}$	$\beta_{a,IMP}$	ρ	σ_e^2/V	σ_w^2/V	σ_u^2/V
r	0.170	0.308	0.359	0.243	-0.176	0.596	0.520	-0.596
p-value	0.207	0.065	0.037	0.119	0.801	0.001	0.003	0.999

Note than σ_e^2/V is shorthand for $\sigma_e^2/V\left[r_t\left[\Delta y_{t,cq}\right]\right]$, etc., that is, the variances of the shocks are relative to the variances of the revisions to the forecasts of current-quarter growth.

We test whether there is a correlation between whether forecasters were primarily active before the Great Recession and their perceptions of persistence, measured by the regression aproach, and the permanent-transitory decomposition of shocks.

The Spearman rank correlation r lies between -1 and 1, and 0 indicates no relationship. For each test, there are two entries. The first row entry is the small-sample test statistic $r = 1 - \frac{6R}{N(N^2-1)}$, where R is the sum of squared differences between the ranks.

For large samples, the test statistic $\frac{6R-N\left(N^2-1\right)}{N(N+1)\sqrt{N-1}}$ is standard normal. We report the probability of obtaining a larger value (the probability in the right tail). In a two-sided test at the 5% level the null is rejected when this probability is less than 0.025 or greater than 0.975.