Do Professional Forecasters Pay Attention to Data Releases?

Michael P. Clements

No 956

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Do Professional Forecasters Pay Attention to Data Releases?

Michael P. Clements*

Department of Economics

University of Warwick

January 26, 2011

Abstract

We present a novel approach to assessing the attentiveness of professional forecasters to news about the macroeconomy. We find evidence that professional forecasters, taken as a group, do not always update their estimates of the current state of the economy to reflect the latest releases of revised estimates of key data.

Key words: Professional forecasters, data revisions, inattention.

JEL classification: C53

*Department of Economics, University of Warwick, Coventry, CV4 7AL, UK. m.p.clements@warwick.ac.uk

1 Introduction

Recently Mankiw and Reis (2002) have proposed a 'sticky-information' model (SIM) of expectations formation that supposes that agents update their expectations periodically, due to the costs of acquiring and processing information, rather than continuously, as implied by the rational expectations hypothesis (REH). Their SIM is applied by Mankiw, Reis and Wolfers (2003) to the analysis of disagreement amongst forecasters, and more recently by Coibion and Gorodnichenko (2008), while Carroll (2005) provides an epidemiological foundation for the model - the acquisition of the latest information is akin to the spread of a disease through a population. First and foremost, the sticky-information model explains the prevalence of disagreement amongst forecasters: the REH assumes that all agents have access to all relevant information and know the structure of the economy, and so have identical expectations. Mankiw *et al.* (2003) regard the SIM as a halfway house between the REH, which assumes too much of agents, and the oft-used alternative of adaptive expectations (AE), which assumes too little (because agents are expected to draw on information other than lags of the variable of interest when they formulate their expectations). Whereas the 'island model' of Lucas (Lucas (1973)) also generates disagreement among agents, the level of disagreement is determined independently of the state of the macroeconomy. By way of contrast, the SIM is capable of explaining the empirical finding that the level of disagreement is endogenously determined (see, e.g., Mankiw *et al.* (2003)).

However, there are of course a number of other explanations of disagreement, so that it would be wrong to suppose that the existence of disagreement lends support to the SIM. For example, Lahiri and Sheng (2008) allow disagreement about i) forecasters' initial beliefs, ii) the weights attached to these prior beliefs, and iii) the interpretation of public information, while Patton and Timmermann (2008) stress prior beliefs and individuals receiving different signals. Both of these studies follow Carroll (2003, 2005) and others in supposing that professional forecasters pay attention to news and base their forecasts on the latest information available to them.

By and large, the literature views the SIM as possibly being relevant for laymen but not for professional forecasters. This is most clearly evident in Carroll (2003) who supposes that the forecasts of professionals embody the latest news and information, and that information is gradually diffused to laymen (e.g., respondents to consumer surveys) according to the sticky information model. However this view has not gone unchallenged, and Coibion and Gorodnichenko (2008) find no evidence that consumers and professional forecasters update their forecasts at different rates.

In this paper we present a novel way of assessing the validity of the SIM for professional forecasters. Specifically, the question we ask is whether professional forecasters are attentive to the latest news about the macroeconomy when they form their expectations. The answer is that professional forecasters, taken as a group, do not always update their estimates of the current state of the economy to reflect the latest releases of revised estimates of key data. As ever when testing a theory of expectations formation a number of auxiliary assumptions need to be made. Our assessment of the SIM is not different in this respect, although the assumptions we make are arguably relatively weak, and at least in part testable. We are careful to consider other possible explanations for the empirical findings that are taken as support for the SIM. By and large the SIM appears to be the most plausible explanation.

The basis of our assessment of the SIM is very simple. Suppose in the third quarter of the year, a forecast is made of the third and fourth quarter values of a variable y. The current quarter - the third quarter - will need to be forecast, as data for the third quarter will not be available until the next quarter. But the forecaster will have access to estimates of the first and second quarter values of y. The forecaster reports their quarterly forecasts y_3 and y_4, and the value for the second quarter, y_2. They also report a forecast of the annual value, x. Under some assumptions, if the forecaster is attentive to the latest data released by the Bureau of Economic Analysis, then the sum of the reported values, $\sum_{i=2}^{4} y_i$, plus the latest estimate of y_1 available at the time the forecasts are made, y_1^3 (i.e., the third-quarter vintage value of the variable in quarter one) should equal the forecast of the annual value. Suppose instead the annual forecast is based on the second-quarter estimate of y_1, namely, y_1^2. If we now (mistakenly) assume the forecaster uses y_1^3, we would find a discrepancy between the annual and quarterly 'forecasts', i.e., $\delta = x - (y_1^3 + \sum_{i=2}^{4} y_i)$, which is negatively correlated with the data revision, $y_1^3 - y_1^2$, the difference between the vintage estimates of the first quarter available in the third quarter and in the second quarter. If the forecaster were 'attentive', the correlation would be zero, and δ would on average be zero. It would only differ from zero because of idiosyncratic errors: reporting errors, rounding errors, etc.

The key assumption we make is that the forecasts satisfy the national accounts identity: the sum of the forecasts of the future quarters and the estimates of the past quarters sum to the forecast of the year. We discuss why this condition might not hold, and present evidence that suggests it does hold for our sample. We also consider the possibility that the forecaster *assumes* that y_1^3 will be revised, and reports an annual forecast x which is consistent with *their* estimate of the first quarter value.

The majority of the literature on the SIM focuses on inflation, but we supplement this by drawing on evidence for a broad range of macroeconomic variables. These include real GDP and its major components drawn from the US National Income and Product Accounts (NIPA). There is no reason to believe a priori that inflation expectations are generated differently from expectations of other key variables, so we draw on a range of variables.

The plan of the remainder of the paper is as follows. Section 2 details our approach to testing forecaster attentiveness to new data releases. Section 3 allows that forecasters might attempt to predict the revised values of past data, and considers how this will affect the interpretation of our tests for attentiveness. Section 4 describes the forecast dataset. Section 5 reports the results of the empirical tests, and section 6 offers some concluding remarks.

2 Testing forecaster attentiveness to data releases

The key assumption we make is that the forecast of the annual value is equal to the average of the forecasts of the corresponding quarters.[1] Although this might appear to be a mild assumption,[2] it need not hold, as it requires that the forecasters report as their predictions their conditional expectations. This rules out asymmetric loss: see e.g., Elliott, Komunjer and Timmermann (2008). We find that to the extent that we can assess the validity of this assumption it appears to hold by and large. As we explain below, 'adding up' can be tested for the responses to the Q1 and Q2 surveys, but is not testable for the Q3 and Q4 surveys. But there seems no reason why forecasters should behave differently in this regard depending on the survey quarter.

First we need to introduce some notation. Let $x_{i,t,k}$ be the annual forecast of the year t value of x made by individual i at time k, where k is one of the four quarters of year t ($k \in \{1, 2, 3, 4\}$). Let $y_{i,t,j,k}$ be the forecast of the value of the variable in quarter j of year t, made at the time

[1] The forecasts SPF forecasts of GDP and its components are reported at annual rates, so the annual values are averages rather than sums. We drop this in the exposition for simplicity.

[2] Patton and Timmermann (2008) make the same assumption in their analysis of the term-structure of forecast disagreement.

of quarter k (again by individual i). Note that $y_{i,t,j,k}$ with $j = k$ will denote a 'current-quarter' forecast, as the first estimate of $y_{t,j}$ will not be available until $j + 1$. Similarly, $x_{i,t,4}$ is also a forecast. The quantities x_t and $y_{t,j}$ without the individual and forecast origin subscripts denote the actual values. Superscripts denote the data vintage, when necessary, so that $x_t^{\tau,q}$ and $y_{t,j}^{\tau,q}$ are the estimates of x_t and $y_{t,j}$ in the data vintage of year τ, quarter q. The annual and quarterly data satisfy $x_t^{\tau,q} = \sum_{j=1}^{4} y_{t,j}^{\tau,q}$ by construction. We usually economize on notation for 'same year' vintages by writing say, $y_{t,1}^{3}$ to denote the third-quarter vintage estimate of the value of y in the first quarter of the year (rather than the more cumbersome $y_{t,1}^{t,3}$). We make use of the Real-time data sets of Croushore and Stark (2001), because these match up with the SPF in the sense that the latest data on y available at the time a respondent files a return to the kth quarter of year t survey is given in the corresponding real-time dataset, consisting of $y_{t,k-1}^{k}$, $y_{t,k-2}^{k}$,... (where e.g., $k - 1 < 0$ or $k - 2 < 0$ indexes a quarter from an earlier year, etc.). Hence these datasets provide the actual data that would have been available to the survey members each time they responded to a particular survey.

The nature of the Bureau of Economic Analysis (BEA) data revisions is such that $y_{t,j}^{j+2}$ is always revised relative to $y_{t,j}^{j+1}$. $y_{t,j}^{j+1}$ is the BEA advance estimate, and $y_{t,j}^{j+2}$ is known as the BEA 'final' estimate. The data are then unrevised except for three annual revisions in the July of each year: see, e.g., Fixler and Grimm (2005, 2008) and Landefeld, Seskin and Fraumeni (2008). Hence $y_{t,j}^{j+3} = y_{t,j}^{j+2}$ except when $j + 3$ is a third quarter (so when the data relate to the fourth quarter of the year, i.e., $j \in Q4$). Exceptions to this rule may occur when there are benchmark revisions. Hence the nature of the data revisions process creates the possibility that inattentive forecasters will base their predictions on out-dated estimates of some of the quarterly values.

Our approach rests on the nature of the SPF questionnaire and how the information it provides on current-year quantities implicitly changes depending on the quarter of the year in which the survey is held (see Zarnowitz (1969), Zarnowitz and Braun (1993) and Croushore (1993) for details of the SPF). Specifically, respondents are asked to provide forecasts of the previous quarter, the current quarter, and each of the next four quarters (so up to and including a forecast of the same quarter in the following year) as well as forecasts for the current and following year. Respondents are not told what to report. They are provided with an estimate of the previous quarter's value.

Bearing this in mind, consider the surveys held in the first quarter of the year ($k = 1$). The econometrician will observe $y_{i,t,j,1}$, $j = 1,2,3,4$ and $x_{i,t,1}$ for all respondents i. Hence we can calculate $\delta_{i,t,1} = x_{i,t,1} - \sum_{j=1}^{4} y_{i,t,j,1}$ for all i. Because for the $Q1$ surveys the forecasts of all the quarters are reported, we can see whether 'adding up' holds, or more generally whether the discrepancies are zero-mean and unsystematic, once we allow for reporting and rounding errors. Whether the forecasters are fully attentive or not is irrelevant to the calculation of $\delta_{i,t,1}$. A similar situation prevails for Q2 surveys ($k = 2$), $\delta_{i,t,2} = x_{i,t,2} - \sum_{j=1}^{4} y_{i,t,j,2}$, as respondents report values for the four quarters of the year. Forecasts from these two origins are not informative (within our approach) about the SIM, but are used to assess the underlying assumption that the quarters sum to the annual totals.

Consider now the forecasts made from the third-quarter survey, as discussed in the introduction. Let $\hat{y}_{i,t,1,3}$ denote i's value for the first quarter, which is unreported and hence unknown to the econometrician, whereas the forecasts of the other three quarters are reported. We calculate $\delta_{i,t,3}$ *assuming* that $\hat{y}_{i,t,1,3} = y^3_{t,1}$, and denote this by $\delta_{i,t,3}\left(y^3_{t,1}\right)$, so that:

$$\delta_{i,t,3}\left(y^3_{t,1}\right) \equiv x_{i,t,3} - \left(y^3_{t,1} + \sum_{j=2}^{4} y_{i,t,j,3} \right).$$

Under the null that the forecaster pays attention to data releases and conditions their forecasts on $y^3_{t,1}$, we would expect adding up to hold apart perhaps from an idiosyncratic error ($\varepsilon_{i,t,3}$, which reflects reporting/rounding/computational errors), so that:

$$H_0 : E\left[\delta_{i,t,3}\left(y^3_{t,1}\right) \mid Z_{t,3}\right] = 0 \tag{1}$$

given that $\delta_{i,t,3}\left(y^3_{t,1}\right) = \varepsilon_{i,t,3}$. $Z_{t,3}$ contains variables known at the time of the survey. Under the null, the discrepancy δ should not be systematically related to these variables. The obvious candidate variables are simply the public releases of estimates of $y_{t,1}$, i.e., $\left\{y^3_{t,1}, y^2_{t,1}\right\}$.

Under the alternative, the forecaster uses an earlier estimate of $y_{t,1}$, say $\hat{y}_{i,t,1,3} = y^2_{t,1}$. Suppose $\delta_{i,t,3}\left(y^2_{t,1}\right) = x_{i,t,3} - \left(y^2_{t,1} + \sum_{j=2}^{4} y_{i,t,j,3}\right) = \varepsilon_{i,t,3}$, so that adding up holds using $y^2_{t,1}$ for this inattentive respondent. Under this alternative:

$$H_1 : E\left[\delta_{i,t,3}\left(y^3_{t,1}\right) \mid \left(y^3_{t,1} - y^2_{t,1}\right)\right] < 0 \tag{2}$$

because $\delta_{i,t,3}\left(y_{t,1}^{3}\right)=\varepsilon_{i,t,3}-\left(y_{t,1}^{3}-y_{t,1}^{2}\right)$,[3] and where we have used $Z_{t,3}=y_{t,1}^{3}-y_{t,1}^{2}$. This is the obvious choice of $Z_{t,3}$ when the inattentiveness takes the form of the forecaster using $y_{t,1}^{2}$ in place of $y_{t,1}^{3}$. For this choice of Z_t we have an unambiguous prediction of the sign of the correlation being negative.

Forecasts from fourth-quarter surveys also allow an assessment of whether full use is made of the latest information. Respondents report values for the third and fourth quarters. We calculate the discrepancies assuming the estimates for the first two quarters are drawn from the data vintage available at the time the forecasts are filed, namely, $y_{t,1}^{4}$ and $y_{t,2}^{4}$:

$$\delta_{i,t,4}\left(y_{t,1}^{4},y_{t,2}^{4}\right)=x_{i,t,4}-\left(y_{t,1}^{4}+y_{t,2}^{4}+\sum_{j=3}^{4}y_{i,t,j,4}\right).$$

Suppose the forecaster used out-dated information, say $y_{t,1}^{3}$ and $y_{t,2}^{3}$, in place of $y_{t,1}^{4}$ and $y_{t,2}^{4}$. As explained, the nature of the BEA revisions process is such that $y_{t,2}^{4}-y_{t,2}^{3}$ is non-zero, but $y_{t,1}^{4}-y_{t,1}^{3}=0$. Under the null of attentiveness we have:

$$H_0:E\left[\delta_{i,t,4}\left(y_{t,1}^{4},y_{t,2}^{4}\right)\mid Z_{t,4}\right]=0 \tag{3}$$

whereas for a 'one-quarter' inattentive forecaster with $\delta_{i,t,4}(y_{t,1}^{3},y_{t,2}^{3})=\varepsilon_{i,t,4}$, we have:

$$H_1:E\left[\delta_{i,t,4}\left(y_{t,1}^{4},y_{t,2}^{4}\right)\mid\left(y_{t,2}^{4}-y_{t,2}^{3}\right)\right]<0. \tag{4}$$

We have assumed in this section that the inattentive forecaster uses $y_{t,1}^{2}$ in place of $y_{t,1}^{3}$ when responding to a third-quarter survey, and $y_{t,2}^{3}$ instead of $y_{t,2}^{4}$ for fourth-quarter surveys. We show in the following section that this is simply for expositional purposes, and that the results are unchanged if instead the forecaster provides an estimate of $y_{t,1}$ or $y_{t,2}$ based on out-dated data releases. In the next section we consider a potentially more serious objection, which is that respondents may estimate future vintage values (of say, $y_{t,1}^{3}$, in the case of the third-quarter surveys) having observed the latest data release.

[3] To obtain this expression, simply subtract $\left(y_{t,1}^{3}-y_{t,1}^{2}\right)$ from both sides of $\delta_{i,t,3}\left(y_{t,1}^{2}\right)=\varepsilon_{i,t,3}$.

3 Respondents forecast revisions to past data

We distinguish between respondents who provide their own estimates of future releases of past data based on the latest available data estimates (attentive forecasters) and those who provide their own estimates in ignorance of the latest official statistics.

3.1 Attentive forecasters

A possible objection to our approach is that the Q3-survey forecaster may not take $y_{t,1}^{3}$ at face value. If it were the case that statistical agencies (such as the BEA) produced data estimates which were not efficient, in the sense that subsequent revisions to these data were predictable, then some respondents might report forecasts (annual and quarterly) that are consistent on the basis of their estimates of the 'post-revision' values of past data, rather than the latest publically available values at the time of the survey return. In this section we consider the implications of this behaviour for our assessment of attentiveness, and whether there is any evidence to indicate that the respondents do attempt to forecast data revisions.

As discussed by Croushore (2011), a government statistical agency would produce 'noisy estimates', whereby the revision between vintages is correlated with the earlier data release, if it simply reports its sample information. Producing an optimal estimate that adds 'news' requires the use of judgment and 'subjective procedures' that the agency may shy away from. The empirical evidence on the nature of revisions is mixed (again see Croushore (2011)), with Mankiw and Shapiro (1986) finding that GDP revisions add news but more recently Aruoba (2008) concluding that such revisions are predictable.

We illustrate in the context of the Q3 survey quarter. Suppose the respondent's estimate of $y_{t,1}$ is $\hat{y}_{t,1,3}$ (dropping the i subscript for convenience). If attentive, then $\hat{y}_{t,1,3} = y_{t,1}^{3}$, and $\delta_{t,3}\left(y_{t,1}^{3}\right) = \varepsilon_{t,3}$. For a general $\hat{y}_{t,1,3}$, the econometrician calculates $\delta_{t,3}\left(y_{t,1}^{3}\right) = \varepsilon_{t,3} - \left(y_{t,1}^{3} - \hat{y}_{t,1,3}\right)$. Of concern is that the 'attentive' forecaster who uses the estimate $\hat{y}_{t,1,3}$ in place of $y_{t,1}^{3}$ might be mistaken for an inattentive forecaster, as would occur in our approach if using an estimate induces a correlation between $\delta_{t,3}\left(y_{t,1}^{3}\right)$ and $Z_{t,3}$.

Suppose the investigator uses $Z_{t,3} = y_{t,1}^{3} - y_{t,1}^{2}$, then we have:

$$Cov\left(\delta_{t,3}, y_{t,1}^{3} - y_{t,1}^{2}\right) = Cov\left(\varepsilon_{t,3}, y_{t,1}^{3} - y_{t,1}^{2}\right) - Cov\left(y_{t,1}^{3} - \hat{y}_{t,1,3}, y_{t,1}^{3} - y_{t,1}^{2}\right)$$

where $Cov\left(\varepsilon_{t,3}, y^3_{t,1} - y^2_{t,1}\right) = 0$ by the assumption that $\varepsilon_{t,3}$ is an idiosyncratic error. To evaluate the second term, write $y^3_{t,1} - \hat{y}_{t,1,3} = (y_{t,1} - \hat{y}_{t,1,3}) + \left(y^3_{t,1} - y_{t,1}\right)$. The first term is the error in the respondent's own estimate of the post-revision value ($y_{t,1}$), and the second term is (minus) the error in the latest-data release value as a prediction of the post-revision value. It follows that $y_{t,1} - \hat{y}_{t,1,3}$ is orthogonal to $\left\{y^3_{t,1}, y^2_{t,1}\right\}$ if $\hat{y}_{t,1,3}$ is an efficient forecast of the 'post-revision' value $y_{t,1}$ using an information set that includes $\left\{y^3_{t,1}, y^2_{t,1}\right\}$. Assuming $\hat{y}_{t,1,3}$ is an efficient forecast,

$$Cov\left(\delta_{t,3}, y^3_{t,1} - y^2_{t,1}\right) = Cov\left(y_{t,1} - y^3_{t,1}, y^3_{t,1} - y^2_{t,1}\right)$$

so that the correlation between $\delta_{t,3}\left(y^3_{t,1}\right)$ and $Z_{t,3}$ will depend on the nature of the BEA data releases and does not depend on the respondent's forecast. Under noise, the data revisions $y_{t,1} - y^3_{t,1}$ and $y^3_{t,1} - y^2_{t,1}$ will typically be correlated.

In summary, if data revisions are noise, then revised estimates of earlier-vintage estimates are in principle predictable. If a forecaster bases their survey response of the current-year quarterly and annual values on their own estimate of the future vintage value of past quarterly value(s), then we would find a non-zero correlation.[4] This suggests care in the interpretation of a finding of a non-zero correlation. However, because the respondents to each survey are asked to provide estimates for the value of the previous quarter, we can assess whether respondents tend to provide their own estimates or simply report the official figure. A casual examination of the data suggests that forecasters tend to report the official figure as their estimate of the previous quarter, and this is borne out by formal testing in section 5. This may suggest they are not interested in reporting the 'final' value, or that they believe the first estimate to be an unbiased forecast of the final value. This does not rule out the possibility that respondents to a Q3 survey base their forecast return on their own estimate of the Q1 value, but there is no obvious reason why they should take the official Q2 figure at face value yet forecast the Q1 value.

3.2 Inattentive forecasters

We show that inattentive forecasters who use estimates of past data will (correctly) show up as inattentive irrespective of the properties of data revisions. We again illustrate with the Q3 survey

[4] Note that it is irrelevant whether the *reported* annual and quarterly forecasts are of the first-release values, or whether they target a later release. Hence for the Q1 and Q2 surveys, for which all the quarterly values are reported, it is immaterial what vintage is being targetted in terms of assessing whether the forecasts 'add up'.

quarter. Suppose the respondent's estimate of $y_{t,1}$ is $\widetilde{y}_{t,1,3}$, where their inattentiveness is manifest in the conditioning of the estimate on out-dated data, say, $\widetilde{y}_{t,1,3} = E\left(y_{t,1} \mid y^2_{t,1}\right)$. Then $\delta_{t,3}\left(y^3_{t,1}\right) = \varepsilon_{t,3} - \left(y^3_{t,1} - \widetilde{y}_{t,1,3}\right)$. With $Z_{t,3} = y^3_{t,1} - y^2_{t,1}$, we have:

$$Cov\left(\delta_{t,3}, y^3_{t,1} - y^2_{t,1}\right) = Cov\left(\varepsilon_{t,3}, y^3_{t,1} - y^2_{t,1}\right) - Cov\left(y^3_{t,1} - \widetilde{y}_{t,1,3}, y^3_{t,1} - y^2_{t,1}\right) \tag{5}$$

Firstly, suppose data revisions are news. Then the best an inattentive but 'efficient' forecaster can do is $\widetilde{y}_{t,1,3} = y^2_{t,1}$, and so:

$$Cov\left(\delta_{t,3}, y^3_{t,1} - y^2_{t,1}\right) = -Var\left(y^3_{t,1} - y^2_{t,1}\right).$$

Secondly, suppose that revisions are noise, and therefore predictable to some degree.

Then we can write $\widetilde{y}_{t,1,3} = y^2_{t,1} + \nu_t$, and substituting into (5) results in:

$$Cov\left(\delta_{t,3}, y^3_{t,1} - y^2_{t,1}\right) = -Var\left(y^3_{t,1} - y^2_{t,1}\right) + Cov\left(\nu_t, y^3_{t,1} - y^2_{t,1}\right) \tag{6}$$

In the extreme case that $\nu_t = y^3_{t,1} - y^2_{t,1}$ then $Cov\left(\delta_{t,3}, y^3_{t,1} - y^2_{t,1}\right) = 0$. The interpretation is that when revisions are perfectly predictable, $\widetilde{y}_{t,1,3} = y^3_{t,1}$, it is obviously irrelevant whether forecasters are 'inattentive' or not. Given that revisions are not perfectly predictable, our approach will signal inattentive forecasters by a negative correlation whether they use an earlier-vintage value, or an estimate of their own based on an earlier-vintage data release.

4 Data

We consider variables forecast by the SPF respondents to the 1981:3 survey through to the 2008:4 survey. Although the SPF[5] began in 1968, prior to the 1981:3 survey, forecasts were not reported for real GDP. Because it is a survey of professional forecasters, authors such as Keane and Runkle (1990) have argued that one can reasonably assume that the reported forecasts reflect the forecasters' expectations, which might not be true when ordinary individuals and firms are surveyed. We

[5] The survey was originally the ASA-NBER Survey of Forecasts by Economic Statisticians, administered by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER). Since June 1990 it has been run by the Philadelphia Fed, renamed as the SPF - Survey of Professional Forecasters: see Zarnowitz (1969), Zarnowitz and Braun (1993) and Croushore (1993).

consider only the forecasts made by regular respondents[6]. Table 1 lists the variables we study, their SPF mnemonics, and also their codes in the Real-Time Data Sets for Macroeconomists (RTDSM) (see Croushore and Stark (2001)). The variables are nominal and real GDP, and the GDP deflator, as well as real consumption and four measures of capital expenditure.

The SPF forecast data are provided as levels (rather than growth rates), and we analyse these directly. As well as the regular data revisions, which are key to our test of attentiveness, there are also benchmark revisions, reflecting methodological changes in measurement or collection procedures, and which may also incorporate base-year changes. All the benchmark revisions bar two occur in the first quarters of the year. The revisions that take place in a first quarter will have no effect on our assessments of attentiveness for any of the survey quarters of these years, as all the forecasts are made on the new base-year data, and all the relevant data releases by the BEA for our purposes are in terms of the new base. In principle, the benchmark revisions in the fourth quarters of 1993 and 1994 should not be problematic either, in the sense that (1) and (3) should hold for the third and fourth quarters for attentive forecasters. In 1993:Q3, for example, adding up holds for the attentive forecaster using the $y_{1993:Q1}^{1993:Q3}$ value for the first quarter; but equally in 1993:Q4 adding up holds for the attentive forecaster using the $y_{1993:Q1}^{1993:Q4}$ and $y_{1993:Q2}^{1993:Q4}$ values for the first and second quarters (albeit that these are now measured on the new basis). In practice, we find that the size of the benchmark revisions do appear to affect our tests based on the fourth quarter survey returns, as described in section 5.1.

5 Results

We firstly consider the evidence that bears on whether the quarterly forecasts sum to the annual forecasts. Recall that for the surveys in the first two quarters of each year, we have reported quarterly values for all the four quarters of the year, as well as the annual (average). This is because respondents report the quarterly levels for the previous quarter and the current quarter, as well as future quarters. Table 2 reports selected percentiles of the distribution of the discrepancies $\delta_{i,t}$ across individuals and years for surveys in each quarter of the year. More precisely, we report the distribution of the percentage discrepancy relative to a measure of the level of the variable, to

[6] Defined as those who filed returns on at least 12 occasions.

take account of changes in the levels of the variables over time from base-year changes.[7]

For the Q1 and Q2 surveys there is little evidence that the quarterly forecasts do not add up to the annual totals. For example, for most of the variables we find that 80% of the discrepancies are around or less than 0.1 of a percent of the level in absolute value (that is, the 10 and 90% percentiles are ± 0.1 of a percent). There is also some evidence that the discrepancies are larger for the more volatile components such as investment, compared to consumption.

For the third quarter and fourth quarter surveys we calculate the discrepancies assuming that respondents use the latest vintage estimates for quarters for which their survey returns do not provide a value. That is, assuming the forecasters are attentive. So, for the third quarter surveys we assume their forecasts are based upon the Q3 vintage estimate of the first quarter, $y_{t,1}^{3}$, and for the fourth quarter surveys, we assume their implicit values for the first two quarters are $y_{t,1}^{4}$ and $y_{t,2}^{4}$. From table 2 it is apparent that the discrepancies are of a similar magnitude to those for the first two quarters. However, it would be wrong to take this as evidence that all forecasters are attentive to the latest vintage estimates, and incorporate these into their forecasts, as we have given no indication of the expected discrepancies that we would expect to find if respondents used earlier-vintage estimates.

A formal test of the null of attentiveness, which will have power to reject the null if respondents do not incorporate latest-vintage estimates into their forecasts, is based on testing (1) and (3): whether the discrepancies are systematically correlated with data revisions with the predicted sign. Table 3 reports the OLS estimates of β_0 and β_1 in the separate first-quarter and second-quarter survey regressions:

$$\delta_{i,t,3} = \beta_0 + \beta_1 \left(y_{t,1}^{3} - y_{t,1}^{2} \right) + \varepsilon_{i,t,3} \tag{7}$$

and:

$$\delta_{i,t,4} = \beta_0 + \beta_1 \left(y_{t,2}^{4} - y_{t,2}^{3} \right) + \varepsilon_{i,t,4} \tag{8}$$

as well as the p-values of the individual null hypotheses $\beta_0 = 0$ and $\beta_1 = 0$ (against two-sided alternatives). These results are shown in the columns headed 'No scaling'. The standard p-values from estimating these regressions by OLS do not make any allowance for the fact that the $\varepsilon_{i,t,.}$ from

[7] For the first and second quarter surveys, we scale by the first estimate of the variable for the previous quarter (and multiply by one hundred), and for the third and fourth quarter surveys we take the first estimate of the first and second quarters, respectively. The reason for the latter choice is so that the scaling of the discrepancy is by the same variable as we use to scale the data revisions in the regressions that follow.

a given survey, t, may be correlated across individuals because of common macroeconomic shocks or other factors. To control for this we also estimate pooled regressions based on the approach of Keane and Runkle (1990) and Bonham and Cohen (2001), see the appendix. The coefficient estimates are unchanged relative, and thus are not repeated, but the corrected p-values are recorded in the table. The findings are not especially sensitive to this, and for the third-quarter forecasts, we find that for 4 of the 8 variables the estimate of β_1 is negative and the null that $\beta_1 = 0$ is rejected at conventional significance levels. The rejection of the null in favour of a negative correlation between the discrepancy and the data revision is consistent with inattentiveness. Note that this occurs for the 'headline' macroeconomic variable, real GDP. As a check on whether the base-year changes over the sample period might have affected the results, by causing heteroscedasticity in the disturbances, we also estimated (7) and (8) with the discrepancies and revisions scaled by the level of the variable (and multiplied by one hundred).[8] The results are little affected.

The results for the fourth quarter surveys (lower panel of table 3) offer little evidence against attentiveness. The slope is significantly different from zero for 3 of the variables using the corrected standard errors (column (3)), but for only one of these variables is the slope negative. In section 5.1 we investigate the time-series properties of the two sets of revisions, $\left\{y_{t,1}^3 - y_{t,1}^2\right\}$, and $\left\{y_{t,2}^4 - y_{t,2}^3\right\}$, to see whether these may account for the starkly different results between the third and fourth quarter surveys reported in table 3.

5.1 Time series properties of the revisions series relevant for the tests based on the third and fourth quarter surveys

Table 4 reports the means and standard deviations of the two series of revisions $\left\{y_{t,1}^3 - y_{t,1}^2\right\}$ and $\left\{y_{t,2}^4 - y_{t,2}^3\right\}$ which underpin the tests of (1) and (3). The summary statistics are based on the 28 revisions over the period 1981 to 2008 in each case. The revisions relevant for the fourth quarter surveys $(y_{t,2}^4 - y_{t,2}^3)$ are larger and more variable than those for the third quarter. However, as is apparent from table 5 which provides the time series of the revisions, the fourth-quarter revisions series contains two very large observations for each variable which correspond to the 1993 and

[8] For the Q3 surveys, the LHS and RHS variables are divided by $y_{t,1}^2$, so that the regressor becomes $100 \times \left(y_{t,1}^3 - y_{t,1}^2\right)/y_{t,1}^2$, the percentage revision in the value of the estimate for the first quarter. For the Q4 surveys, we divide by $y_{t,2}^3$, so the regressor again has the interpretation of being the percentage revision. As the same scaling is used for the dependent variables, this accounts for our choice of denominator for the Q3 and Q4 survey discrepancies in table 2.

1999 benchmark revisions.[9] It seems extremely unlikely that the respondents would not have been aware of 'one-off' revisions on the scale of those made in 1993 and 1994. At the same time, if we exclude these two observations, the table shows that although the average size of the fourth-quarter revisions was often on a par with the third-quarter revisions (see e.g., nominal and real output, and consumption) the *variability* was only around a fifth as large (as measured by the standard deviation). Hence the two large benchmark revisions coupled with the relatively little variability in the revisions faced by the fourth-quarter respondents offers an explanation for the very different results found for the third and fourth quarters in table 3 - there was insufficient variability in the latter case for inattentive forecasters to be detected.

5.2 Do survey respondents aim to produce forecasts which add up?

A key assumption underlying our analysis is that the forecast of the annual value is equal to the average of the forecasts of the corresponding quarters. Table 6 reports the results of regressing $\delta_{i,t,1}$ and $\delta_{i,t,2}$ on a constant, as well as of carrying out this regression having first scaled the variables. We obtain similar results whether or not we scale the variables to account for the level shifts due to re-basings. These regressions constitute a direct test of adding up for the responses to the first and second quarter surveys. For the scaled regressions using corrected standard errors there is no evidence against adding up.

5.3 Do survey respondents attempt to second guess the official national accounts data?

We argued in section 3 that even attentive forecasters might generate non-zero correlations between δ and data revisions if they attempt to predict future revisions to the official data releases. We do not observe whether respondents to third and fourth quarter surveys provide their own estimates of data pertaining to the first and second quarters. However, for each survey, we do observe the respondents' estimates of the value for the previous quarter, and these can be compared to the latest official data. Table 7 reports the results of testing whether the differences between the latest official estimates and the survey responses are significantly different. We report the estimates of

[9] For real GDP, for example, the 1999 percentage fourth-quarter revision was 12%, compared to an average revision over the period when 1993 and 1999 are excluded of just a half of one percentage point.

the mean differences, and p-values of whether this is significantly different from zero (based on corrected standard errors). In no cases are the means significant at a conventional 5% level, which suggests that there is no evidence against the hypothesis that respondents take the official statistics at face value.

Of course, these tests would only reject the null if respondents forecasted future revisions and revisions were non-zero mean. The evidence in Aruoba (2008) suggests that there is strong evidence against the null that revisions to US macro-series are zero-mean. Hence our tests should have power to reject if respondents attempted to forecast revisions.

6 Conclusions

We have shown that the responses of professional forecasters of the US SPF are consistent with the sticky information model of expectations formation. Our approach requires that individual respondents seek to provide sets of forecasts which are consistent in the sense that the forecasts of the quarters are consistent with the annual level. This holds by construction in the national accounts data, yet need not be true of the forecasts if the forecasts are optimal for an asymmetric loss function, for example. To the extent that 'adding up' is testable, we find little evidence that it does not hold. Our approach also requires that agents do not attempt to forecast revisions to the official data releases. Again, to the extent that this is testable, there is no evidence to suggest that agents do not take the official data at face value. Under these assumptions we can deduce that agents are using outdated estimates of past data if there is a negative correlation between the discrepancy (between the annual forecast and the quarterly values) and data revisions to past quarterly data. We find significantly negative correlations for half the macro-variables we consider based on returns to third-quarter surveys, but much less evidence in favour of the sticky-information hypothesis for the returns to the fourth-quarter surveys. However, our explanation for the findings for the fourth-quarter surveys is not inconsistent with the hypothesis. It is that the nature of the revisions series relevant for these survey returns is such that our tests are likely to have low power.

Our approach to testing the sticky-information hypothesis has the virtue of being simple. It does not require knowledge of, or that we make assumptions about, the models or forecasting methods that agents use to generate their expectations. Some may dismiss this hypothesis out of hand on the grounds that it is unreasonable to assume that professional forecasters would not pay attention to the latest releases of data by the Bureau of Economic Analysis. Be that as it may, in the absence of an alternative explanation for the negative correlations between the discrepancies and revisions, we propose the sticky-information hypothesis as a possible explanation of this phenomenon.

References

Aruoba, S. B. (2008). Data revisions are not well-behaved. *Journal of Money, Credit and Banking*, **40**, 319–340.

Bonham, C., and Cohen, R. (2001). To aggregate, pool, or neither: Testing the rational expectations hypothesis using survey data. *Journal of Business and Economic Statistics*, **190**, 278–291.

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

Carroll, C. (2005). The epidemiology of macroeconomic expectations. In Blume, L. E., and Durlauf, S. N. (eds.), *The Economy as an Evolving Complex System, III.* Oxford: Oxford University Press.

Coibion, O., and Gorodnichenko, Y. (2008). What can survey forecasts tell us about informational rigidities?. Mimeo, College of William and Mary.

Croushore, D. (1993). Introducing: The Survey of Professional Forecasters. *Federal Reserve Bank of Philadelphia Business Review*, **November/December**, 3–13.

Croushore, D. (2011). Forecasting with real-time data vintages. In Clements, M. P., and Hendry, D. F. (eds.), *Forecasting Handbook*: Oxford University Press. Forthcoming.

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

Elliott, G., Komunjer, I., and Timmermann, A. (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss. *Journal of the European Economic Association*, **6**, 122–157.

Fixler, D. J., and Grimm, B. T. (2005). Reliability of the NIPA estimates of U.S. economic activity. *Survey of Current Business*, **85**, 9–19.

Fixler, D. J., and Grimm, B. T. (2008). The reliability of the GDP and GDI estimates. *Survey of Current Business*, **88**, 16–32.

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

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

Landefeld, J. S., Seskin, E. P., and Fraumeni, B. M. (2008). Taking the pulse of the economy. *Journal of Economic Perspectives*, **22**, 193–216.

Lucas, R. (1973). Some international evidence on inflation-output tradeoffs. *American Economic Review*, **63**, 326–334.

Mankiw, N. G., and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics*, **117**, 1295–1328.

Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. mimeo, National Bureau of Economic Research, Cambridge MA.

Mankiw, N. G., and Shapiro, M. D. (1986). News or noise: An analysis of GNP revisions. *Survey of Current Business (May 1986), US Department of Commerce, Bureau of Economic Analysis*, 20–25.

Patton, A. J., and Timmermann, A. (2008). Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. Mimeo, Department of Economics, University of Oxford.

Zarnowitz, V. (1969). The new ASA-NBER Survey of Forecasts by Economic Statisticians. *The American Statistician*, **23, No. 1**, 12–16.

Zarnowitz, V., and Braun, P. (1993). Twenty-two years of the NBER-ASA quarterly economic outlook surveys: aspects and comparisons of forecasting performance. In Stock, J., and Watson, M. (eds.), *Business Cycles, Indicators, and Forecasting*, pp. 11–84: Chicago: University of Chicago Press and NBER.

Table 1: Macroeconomic Variables in the SPF

Variable	SPF code	RTDSM code
Nominal GDP (GNP)	NGDP	NOUTPUT
GDP price index (implicit deflator, GNP deflator)	PGDP	P
Real GDP (GNP)	RGDP	ROUTPUT
Real personal consumption	RCONSUM	RCON
Real nonresidential fixed investment	RNRESIN	RINVBF
Real residential fixed investment	RRESINV	RINVRESID
Real federal government expenditure	RFEDGOV	RGF
Real state and local government	RSLGOV	RGSL

The SPF data were downloaded from http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/. See Croushore (1993). The Real-time Date Sets for Macroeconomists (RTDSM) were taken from http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/. See Croushore and Stark (2001). Both the survey data and the real-time data were downloaded in February 2009.

Table 2: Distributions of scaled discrepancies by survey quarter

		Percentiles						
SPF code	#	5	10	25	50	75	90	95
First quarter surveys								
NGDP	650	-0.049	-0.022	-0.005	0.000	0.005	0.043	0.207
PGDP	640	-0.070	-0.041	-0.012	0.000	0.023	0.077	0.221
RGDP	648	-0.083	-0.028	-0.001	0.000	0.003	0.029	0.071
RCONSUM	621	-0.055	-0.029	-0.004	0.000	0.001	0.031	0.052
RNRESIN	616	-0.197	-0.052	-0.004	0.000	0.007	0.069	0.157
RRESINV	618	-0.231	-0.081	-0.011	0.000	0.013	0.127	0.382
RFEDGOV	594	-0.214	-0.071	-0.007	0.000	0.012	0.095	0.214
RSLGOV	596	-0.137	-0.045	-0.005	0.000	0.008	0.080	0.163
Second quarter surveys								
NGDP	731	-0.083	-0.025	-0.004	0.000	0.004	0.033	0.107
PGDP	729	-0.073	-0.040	-0.015	0.000	0.022	0.052	0.215
RGDP	737	-0.046	-0.027	-0.005	0.000	0.001	0.022	0.047
RCONSUM	713	-0.062	-0.026	-0.001	0.000	0.002	0.027	0.052
RNRESIN	702	-0.185	-0.051	-0.004	0.000	0.010	0.103	0.240
RRESINV	700	-0.278	-0.080	-0.013	0.000	0.017	0.151	0.419
RFEDGOV	675	-0.213	-0.080	-0.011	0.000	0.011	0.080	0.423
RSLGOV	675	-0.120	-0.047	-0.004	0.000	0.008	0.080	0.213
Third quarter surveys								
NGDP	736	-0.096	-0.026	-0.004	0.000	0.004	0.044	0.180
PGDP	738	-0.101	-0.045	-0.018	0.000	0.024	0.053	0.151
RGDP	736	-0.043	-0.022	-0.001	0.000	0.005	0.032	0.087
RCONSUM	715	-0.050	-0.025	-0.002	0.000	0.007	0.044	0.113
RNRESIN	700	-0.336	-0.108	-0.010	0.000	0.015	0.135	0.352
RRESINV	707	-0.493	-0.132	-0.016	0.000	0.025	0.180	0.718
RFEDGOV	678	-0.262	-0.108	-0.015	0.000	0.015	0.128	0.340
RSLGOV	682	-0.183	-0.073	-0.008	0.000	0.011	0.071	0.214
Fourth quarter surveys								
NGDP	796	-0.051	-0.023	-0.003	0.000	0.004	0.027	0.057
PGDP	796	-0.106	-0.046	-0.021	0.000	0.017	0.044	0.132
RGDP	803	-0.046	-0.025	-0.003	0.000	0.003	0.018	0.035
RCONSUM	771	-0.046	-0.028	-0.002	0.000	0.003	0.031	0.048
RNRESIN	758	-0.253	-0.082	-0.009	0.000	0.008	0.069	0.193
RRESINV	755	-0.364	-0.119	-0.014	0.000	0.015	0.124	0.412
RFEDGOV	728	-0.212	-0.096	-0.015	0.000	0.011	0.095	0.252
RSLGOV	734	-0.105	-0.041	-0.007	0.000	0.009	0.055	0.113

The first-quarter discrepancies are scaled as $100 \times \delta_{i,t,1}/y^{t,1}_{t-1,4}$; the second-quarter as $100 \times \delta_{i,t,2}/y^{t,2}_{t,1}$; the third-quarter as $100 \times \delta_{i,t,3}/y^{t,2}_{t,1}$, and the fourth-quarter as $100 \times \delta_{i,t,4}/y^{t,3}_{t,2}$.

Table 3: Tests of the null of attentiveness for the third and fourth quarter survey forecasts

	No scaling			Scaled	
	β_0 β_1	$p_{\beta_0=0}$ $p_{\beta_1=0}$	$p^*_{\beta_0=0}$ $p^*_{\beta_1=0}$	β_0 β_1	$p^*_{\beta_0=0}$ $p^*_{\beta_1=0}$
Third quarter surveys					
NGDP	0.798	0.677	0.651	0.018	0.451
	0.052	0.201	0.163	0.065	0.131
PGDP	0.009	0.489	0.529	0.011	0.269
	0.015	0.528	0.565	-0.023	0.298
RGDP	0.110	0.725	0.778	0.004	0.502
	-0.016	0.002	0.017	-0.024	0.004
RCONSUM	0.081	0.779	0.795	0.005	0.407
	-0.020	0.015	0.027	-0.021	0.012
RNRESIN	0.014	0.942	0.942	0.032	0.278
	-0.007	0.235	0.234	0.004	0.706
RRESINV	-0.011	0.897	0.896	0.010	0.756
	-0.030	0.029	0.026	-0.031	0.033
RFEDGOV	0.166	0.075	0.086	0.050	0.065
	0.007	0.630	0.645	-0.002	0.919
RSLGOV	0.062	0.648	0.510	0.000	0.979
	-0.019	0.128	0.017	-0.018	0.175
Fourth quarter surveys					
NGDP	-0.925	0.457	0.690	0.004	0.835
	-0.021	0.348	0.615	-0.011	0.679
PGDP	0.009	0.473	0.359	0.007	0.379
	0.002	0.696	0.599	0.002	0.636
RGDP	-0.590	0.058	0.072	-0.009	0.098
	0.002	0.161	0.189	0.002	0.277
RCONSUM	-0.476	0.140	0.092	-0.014	0.119
	0.006	0.026	0.011	0.007	0.110
RNRESIN	-0.422	0.042	0.091	-0.035	0.229
	0.013	0.038	0.088	0.011	0.198
RRESINV	-0.063	0.478	0.473	-0.005	0.938
	0.006	0.590	0.586	0.016	0.521
RFEDGOV	-0.026	0.742	0.779	0.008	0.660
	0.001	0.797	0.827	0.000	0.934
RSLGOV	0.061	0.618	0.621	0.010	0.437
	-0.001	0.816	0.817	-0.002	0.707

The entries are based on regression equations (7) and (8), respectively, for the third and fourth-quarter surveys. We report the estimates and individual p-values that the corresponding coefficients are zero, along with the p-values based on corrected standard errors. The scaled regressions are of $100 \times \delta_{i,t,3}/y^{t,2}_{t,1}$ on $100 \times \left(y^{t,3}_{t,1} - y^{t,2}_{t,1}\right)/y^{t,2}_{t,1}$ and of $100 \times \delta_{i,t,4}/y^{t,3}_{t,2}$ on $100 \times \left(y^{t,4}_{t,2} - y^{t,3}_{t,2}\right)/y^{t,3}_{t,2}$ for the third and fourth quarter surveys, respectively. For the regressions in the scaled variables, we report the estimates and corrected p-values. [$p^*_{\beta_0=0}$ denotes a corrected p-value.]

Table 4: Summary statistics of data revision series

	Third-quarter revisions		Fourth-quarter revisions			
			All fourth-quarter revisions		Excluding 1993 and 1999 revisions	
	Mean	Std dev.	Mean	Std dev.	Mean	Std dev.
NGDP	0.053	0.625	0.232	0.640	0.066	0.151
PGDP	0.089	0.420	-0.290	1.641	0.008	0.113
RGDP	-0.055	0.726	0.558	2.362	0.057	0.133
RCONSUM	-0.052	0.783	0.451	2.024	0.034	0.115
RNRESIN	-0.458	2.715	0.905	3.027	0.234	0.510
RRESINV	-0.237	2.425	0.532	2.475	-0.024	1.034
RFEDGOV	-0.303	1.201	0.782	3.353	0.114	0.710
RSLGOV	0.077	1.120	0.618	2.647	0.027	0.211

The revisions are all scaled - divided by the earlier estimate and multiplied by 100. The third-quarter revisions are $\left\{y_{t,1}^{3}-y_{t,1}^{2}\right\}$ and the fourth-quarter revisions are $\left\{y_{t,2}^{4}-y_{t,2}^{3}\right\}$.

Table 5: Time series of fourth-quarter revisions, $\left\{y^4_{t,2} - y^3_{t,2}\right\}$

	NGDP	PGDP	RGDP	RCONSUM	RNRESIN	RRESINV	RFEDGOV	RSLGOV
1981	0.167	0.080	0.086	-0.052	1.003	-0.830	-0.549	-0.055
1982	-0.072	-0.180	0.108	-0.136	-0.892	2.296	0.823	0.000
1983	-0.052	-0.294	0.243	0.010	0.742	2.714	-0.508	0.287
1984	-0.047	0.039	-0.085	0.235	0.148	-1.935	-0.404	0.112
1985	-0.010	-0.046	0.036	-0.091	0.182	-1.774	-0.689	0.437
1986	-0.160	-0.043	-0.117	0.058	0.440	-0.207	1.169	0.218
1987	-0.083	-0.054	-0.029	-0.048	0.852	-0.606	1.931	-0.413
1988	0.352	0.379	-0.028	0.186	0.204	-0.629	1.937	0.066
1989	0.131	-0.077	0.209	0.208	0.235	0.371	0.674	0.086
1990	-0.152	0.042	-0.195	0.135	0.375	0.661	0.261	-0.189
1991	-0.144	0.086	-0.230	-0.249	0.931	-0.522	0.113	-0.208
1992	0.146	0.107	0.039	0.058	0.586	0.897	-0.772	-0.089
1993	1.945	0.294	1.646	1.018	3.434	4.141	1.439	2.793
1994	0.094	0.002	0.092	0.020	-0.182	0.000	-0.830	0.206
1995	0.260	0.068	0.192	0.213	-0.118	0.136	0.062	0.134
1996	0.093	0.005	0.109	-0.094	0.806	0.249	0.212	0.087
1997	0.369	0.088	0.279	0.032	-0.101	0.432	-0.411	-0.037
1998	0.106	0.004	0.101	0.073	0.305	0.395	0.066	-0.036
1999	2.844	-8.642	12.494	10.717	15.831	11.379	17.484	13.819
2000	0.085	-0.028	0.108	0.016	-0.961	-0.640	-0.054	-0.399
2001	-0.147	-0.037	-0.106	0.092	-0.272	-0.342	0.054	-0.215
2002	0.068	0.018	0.048	-0.028	-0.211	-0.541	0.016	-0.146
2003	0.236	0.009	0.222	0.117	0.109	0.122	0.090	0.327
2004	0.070	0.007	0.062	0.131	0.818	0.231	0.000	-0.033
2005	0.015	0.039	-0.025	0.015	-0.039	0.234	0.259	0.048
2006	0.026	0.000	0.025	0.025	0.408	-1.315	-0.284	0.264
2007	0.094	-0.013	0.106	0.035	0.668	-0.708	-0.173	0.032
2008	0.267	0.015	0.229	-0.074	0.035	0.681	-0.025	0.212

Table 6: Tests of 'adding up' for the first and second quarter survey forecasts

	No scaling			Scaled	
	β_0	$p_{\beta_0=0}$	$p^*_{\beta_0=0}$	β_0	$p^*_{\beta_0=0}$
	First quarter surveys				
NGDP	3.171	0.009	0.026	0.031	0.884
PGDP	0.024	0.154	0.225	-0.012	0.123
RGDP	1.227	0.403	0.419	0.214	0.333
RCONSUM	-0.024	0.948	0.950	-0.654	0.207
RNRESIN	0.070	0.756	0.804	-0.155	0.074
RRESINV	0.069	0.544	0.512	0.004	0.902
RFEDGOV	0.228	0.463	0.472	-0.816	0.253
RSLGOV	-0.002	0.984	0.985	-0.252	0.178
	Second quarter surveys				
NGDP	1.289	0.373	0.369	0.030	0.920
PGDP	0.037	0.012	0.020	-0.001	0.846
RGDP	-0.494	0.115	0.082	-0.074	0.710
RCONSUM	-0.384	0.144	0.074	-0.230	0.378
RNRESIN	-0.117	0.592	0.657	-0.088	0.273
RRESINV	-0.221	0.144	0.157	-0.077	0.222
RFEDGOV	0.119	0.197	0.324	-0.257	0.113
RSLGOV	0.122	0.190	0.175	-0.227	0.314

The entries are based on regressions of $\delta_{i,t,1}$ and $\delta_{i,t,2}$ on an intercept (for the first and second surveys, respectively). We report the estimates of the intercept and the p-value that the population value is zero, as well as p-value based on corrected standard errors. The results for the scaled regressions are of $100 \times \delta_{i,t,1}/y^{t,1}_{t-1,4}$ and of $100 \times \delta_{i,t,2}/y^{t,2}_{t,1}$ on an intercept, for the first and second quarter surveys, respectively. Reported are estimates and corrected p-values. [$p^*_{\beta_0=0}$ denotes a corrected p-value.]

Table 7: Official figures for the previous quarter and respondents' estimates

	β_0	$p_{\beta_0=0}$
First quarter surveys		
NGDP	0.031	0.884
PGDP	-0.012	0.123
RGDP	0.214	0.333
RCONSUM	-0.654	0.207
RNRESIN	-0.155	0.074
RRESINV	0.004	0.902
RFEDGOV	-0.816	0.253
RSLGOV	-0.252	0.178
Second quarter surveys		
NGDP	0.030	0.920
PGDP	-0.001	0.846
RGDP	-0.074	0.710
RCONSUM	-0.230	0.378
RNRESIN	-0.088	0.273
RRESINV	-0.077	0.222
RFEDGOV	-0.257	0.113
RSLGOV	-0.227	0.314
Third quarter surveys		
NGDP	0.076	0.796
PGDP	-0.002	0.850
RGDP	0.000	1.000
RCONSUM	-0.053	0.592
RNRESIN	0.092	0.420
RRESINV	-0.093	0.311
RFEDGOV	-0.333	0.059
RSLGOV	-0.289	0.224
Fourth quarter surveys		
NGDP	0.297	0.591
PGDP	0.011	0.253
RGDP	-0.098	0.502
RCONSUM	-0.038	0.569
RNRESIN	-0.115	0.382
RRESINV	0.022	0.438
RFEDGOV	-0.031	0.752
RSLGOV	0.027	0.648

The entries for survey quarter q are based on a regression of $y^{t,q}_{t,q-1} - y_{i,t,q-1,q}$ on an intercept, to test whether the official estimate for the quarter before the survey quarter $(q-1)$ diverges systematically from the respondents' estimates of those quarters. The p-values are based on 'corrected' standard errors: see the main text.

7 Appendix: The estimation of the pooled regression.

To get the 'correct' standard errors for the regression (7) or (8) we adapt the approach of Keane and Runkle (1990) and Bonham and Cohen (2001) as follows. We assume that for an individual i:

$$E\left[\varepsilon_{i,t}^2\right] = \sigma_0^2$$

and that for any pair of individuals i, j:

$$E\left[\varepsilon_{it}\varepsilon_{jt}\right] = \delta_0^2.$$

We follow Keane and Runkle (1990) and estimate σ_0^2 and δ_0^2 from the residuals of the pooled OLS regression (which imposes microhomogeneity: the same intercepts and slope parameters over all individuals), whereas Bonham and Cohen (2001) use the residuals from separate regressions for each individual. Hence:

$$\hat{\sigma}_0^2 = \frac{1}{\overline{\overline{T}}}\sum_{i=1}^{N}\sum_{t_i}\hat{\varepsilon}_{i,t_i}^2$$

where t_i runs over all the surveys to which i responded, T_i is the number of forecasts made by i, $\overline{T} = \sum_{i=1}^{N} T_i$. Similarly:

$$\hat{\delta}_0^2 = \frac{1}{\overline{\overline{T}}}\sum_{i=1}^{N}\sum_{\substack{j=1,\\ j\neq i}}^{N}\sum_{t_{ij}}\hat{\varepsilon}_{i,t_{ij}}\hat{\varepsilon}_{j,t_{ij}}$$

where t_{ij} runs over all the surveys to which i and j responded, T_{ij} is the number of such forecasts, and $\overline{T} = \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} T_{ij}$.

We can then construct the estimator $\hat{\Sigma}$ of $\Sigma = E\left(\varepsilon\varepsilon'\right)$, where $\varepsilon = [\varepsilon_{1,1}\ \varepsilon_{1,2}\ \ldots\varepsilon_{1,T};\ldots;\varepsilon_{N,1}\ \varepsilon_{N,2}\ \ldots\varepsilon_{N,T}]'$, using $\hat{\sigma}_0^2$ and $\hat{\delta}_0^2$. Write the model as:

$$Y = X\gamma + \varepsilon$$

where Y and X are ordered conformably with ε (all the time observations on individual 1, then on individual 2 etc.) and where X has two columns, the first being the intercept, and $\gamma = (\alpha_0\ \alpha_1)'$. $\hat{\gamma}$ is obtained by deleting the rows of Y and X corresponding to missing observations (as in the calculation of the $\hat{\varepsilon}_{i,t}$ residuals). The covariance matrix for $\hat{\gamma}$ is given by the usual formula $\left(X'X\right)^{-1} X'\hat{\Sigma}X\left(X'X\right)^{-1}$ where X is again compressed to eliminate missing values, and the corresponding rows (and equivalent columns) are deleted from $\hat{\Sigma}$.