Nowcasting is not just Contemporaneous Forecasting

Jennifer L. Castle, Nicholas W.P. Fawcett and David F. Hendry[*]
Department of Economics, Oxford University.

Abstract

We consider the reasons for nowcasting, the timing of information and sources thereof, especially contemporaneous data, which introduce different aspects compared to forecasting. We allow for the impact of location shifts inducing nowcast failure and nowcasting during breaks, probably with measurement errors. We also apply a variant of the nowcasting strategy proposed in Castle and Hendry (2009) to nowcast Euro-area GDP growth. Models of disaggregate monthly indicators are built by automatic methods, forecasting all variables that are released with a publication lag each period, then testing for shifts in available measures including survey data, switching to robust forecasts of missing series when breaks are detected.

Keywords: Nowcasting; contemporaneous information; *Autometrics*; location shifts; impulse-indicator saturation; robust nowcasts; Euro-area GDP growth.

JEL Classification: C51; C52.

1 Introduction

On 24 July 2009 the Office for National Statistics (ONS) published GDP growth figures for the second quarter of 2009: The 0.8% decline in GDP induced severe forecast failure–the official data was worse than any forecast.[1] Statistical agencies such as the ONS must base their flash estimates on incomplete and uncertain data, hence there is a need for nowcasts that combine both actual data for components that are known with forecasts of components that are unknown. Often surveys are used to 'infill' the unknown components to arrive at the flash estimate (see e.g., Ashley, Driver, Hayes, and Jeffery, 2005), a reasonable strategy when the surveys are good leading indicators, but such a methodology can prove disastrous in times of structural change. The *Financial Times* (25 July 2009) reported that 'the ONS said that its estimates were even less reliable than normal because the economy's unpredictability meant its models had "broken down".'[2] A similar problem has beset the National Institute's nowcasting approach (see Mitchell, 2009). Any nowcasting strategy must be robust to structural breaks, requiring a methodology that can both detect structural breaks and rapidly adapt when such breaks occur. It is precisely at the point at which accurate nowcasts are needed (turning points) that the current nowcasting methods may be breaking down. This paper considers the reasons for nowcasting and the implications for how to produce robust nowcasts.

---

[1]The estimate of GDP growth has since been revised to a 0.7% decline in the second release for 2009Q2 on 28 August 2009, see www.statistics.gov.uk/cci/nugget.asp?id=192.

[2]www.ft.com/cms/s/0/c5529f00-7886-11de-bb06-00144feabdc0.html

There are four practical reasons why nowcasting–'forecasting' the current state–is needed for economic aggregates, which play a key role in economic policy decisions. Unfortunately, each of the reasons is matched by a corresponding problem, as we now explain.

First and foremost, and the central basis for 'nowcasting', is that not all disaggregated contemporaneous data are available at the time their information is required to construct the relevant aggregate. Statistical agencies decide when to publish preliminary estimates based on their view of the optimal trade-off between timeliness and accuracy. For preliminary estimates of GDP, for example, the information is only available up to 24 days after the end of the quarter for which a 'nowcast' is required for the UK,(see Reed, 2000, 2002), whereas the Euro-area preliminary estimates are published 45 days after the end of the quarter. This 'missing data problem' is ever present, so that aggregates cannot be constructed just by adding up the relevant observed disaggregates.

Secondly, many economic time series are themselves only preliminary, or 'flash', estimates, which are subject to potentially substantial revisions in due course as more information accrues, so are not necessarily a reliable and accurate guide to current conditions. Figure 1 (from Walton, 2006) demonstrates that initial estimates of GDP growth are a poor guide to actual GDP growth out-turns (defined as the latest available vintage in Autumn 2005) with revisions of more than 2% for some quarters and magnitudes, directions and turning points differing substantially between the two measures. Faust, Rogers and Wright (2007) report that revisions to GDP announcements are quite large in all G7 countries, and are fairly predictable in the UK, mainly because of reversion to the mean, which they interpret as due to removing measurement 'noise'. Garratt and Vahey (2006) consider UK macroeconomic indicators and find that most preliminary measurements are biased predictors, with some revision series affected by multiple structural breaks, and Mitchell (2004) finds that poor construction of nowcasts can explain much of the revision to UK GDP growth. Thus, 'measurement error problems' are also pervasive, and consequently, the 'observed data' cannot be taken at face value, and may be no more accurate than nowcast values.
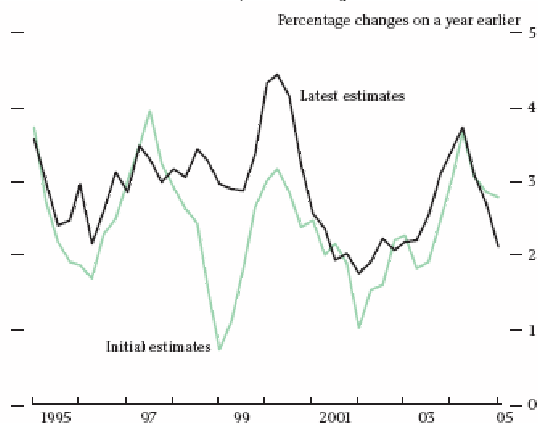

Figure 1: GDP outturns: initial and latest estimates

Thirdly, different components of the data entering the aggregates for which nowcasts are required are unavailable in different periods (see e.g., Clements and Hendry, 2003), and so are missing on a non-systematic basis. Consequently, a consistent subset of information is rarely available, inducing a 'changing database problem' which is bound to affect any system for nowcasting.

Fourthly, even when actual values of the relevant variables are available, nowcasts should still be produced. These would now allow for immediate evaluation, to act as an 'early warning signal' when nowcasts departed too radically from the measured series. That could enable rapid action to be taken following any deterioration in the performance of a nowcasting model or method, or warn of measurement

problems in the 'actual' series. Thus, a 'break problem' always threatens to disrupt the accuracy of any nowcasts.

The interaction of these four problems poses major uncertainties for nowcasters. At first sight, the difficulties seems similar to those facing forecasters. However, a key difference is the presence of contemporaneous observations–which logically could never occur in a forecasting context–and there are many possible sources of such evidence, including rapidly and frequently observed data on disaggregates such as retail sales; correlates like road traffic and air passenger numbers or energy consumption; surveys; and more recent innovations including Google (see Choi and Varian, 2009) and prediction markets. Castle and Hendry (2009) propose a framework for nowcasting that seeks to address three of these issues, namely the 'missing data', 'changing database' and 'break' problems. Here we also consider 'measurement errors', and the use of contemporaneous data to try and ascertain the extent of any anomalies being due to those, as against breaks. We also apply our approach, building on the example in Ferrara, Guegan and Rakotomarolahy (2008).

Given the need for nowcasts, section 2 considers the range of information available. Then section 3 discusses the potential role for automatic model selection to handle all the available information while allowing for multiple past breaks at unknown dates, as well as contemporaneous location shifts, which together entail that nowcasting models always involve more variables than observations. This aspect builds on the approach in Castle, Doornik and Hendry (2009a), based on *Autometrics* with impulse-indicator saturation for removing breaks, outliers and past data contamination: see Doornik (2009a), Hendry, Johansen and Santos (2008), and Johansen and Nielsen (2009). Section 4 discusses problems affecting economic nowcasting due to the trade-off between timeliness and data quality, with section 5 discussing the conflict between breaks and measurement errors, leading to an application of the proposed nowcasting strategy to Euro-area GDP growth in section 6. Section 7 concludes.

2   Range of information available

There are many possible sources of contemporaneous data, including up-to-date values for some of the relevant disaggregates, rapidly and frequently observed outcomes for variables such as retail sales; correlated variables like road traffic and air passenger numbers or energy consumption; qualitative surveys of consumers and businesses about their plans and expectations; and more recent innovations including Google (see Choi and Varian, 2009) and prediction markets. Models that exploit related series, possibly in combination with univariate time-series models, can help improve the accuracy with which any missing data are estimated, by adding the proxy as an explanatory variable in a model for the variable to be forecast. The advantage of automatic *Gets* is that all covariate information can be included in the general model at the outset and the data characteristics will determine whether the explanatory variables are relevant in explaining the individual disaggregate series. The model selection for the individual forecast models could be undertaken for every new release of data, which is feasible for automatic model selection. Hence, if a break is detected, the model will be updated as soon as information on the break is available. The robustness of the model specification and parameter constancy can be tested alongside the break detection tests, but the distinction between internal and external breaks discussed in Castle, Fawcett and Hendry (2009b) points to different approaches in even those two cases.

2.1   Disaggregates

Castle and Hendry (2009) consider the first of these, and propose how to utilize data on the subset of available disaggregates to construct the desired aggregate in a framework that addresses the 'missing

data', 'changing database', and 'break' problems, using *Autometrics* with impulse-indicator saturation to remove past breaks, outliers and data contamination, and robust forecasting devices to offset current location shifts (see Hendry, 2006). However, they do not include additional variables.

Impulse-indicator saturation (IIS) adds an impulse indicator for every observation to the candidate regressor set: Hendry *et al.* (2008) establish one feasible algorithm, and derive the null distribution for an IID process, and Johansen and Nielsen (2009) generalize their findings to general dynamic regression models (possibly with unit roots), and show that there is a small efficiency loss under the null of no breaks at a nominal significance level α, despite investigating the potential relevance of T additional variables for a sample of size T when αT is small.[3]

2.2 Covariate information

Variables like retail sales are observed more frequently (monthly) and rapidly than aggregates like GDP (quarterly), and while they include some of the information needed for expenditure measures of GDP, they are also potentially correlated with other variables that are only available with greater latency. For example, retail sales are published approximately 32 days after the end of the reference period, whereas consumption data, for which retail sales is a reasonable proxy, is only released approximately 64 days after the end of the reference period along with the first full GDP release (excluding the flash estimate). Other proxies, such as new passenger car registrations, construction output and industrial new orders are also available more rapidly than components of the expenditure or income based measures of GDP.

2.3 Google query data

Choi and Varian (2009) show that *Google Trends* data can help improve their forecasts of the current level of activity for a number of different US economic time series, including automobile, home, and retail sales, as well as travel behaviour. The associated Google query variable is added to simple linear seasonal-AR models to measure the additional 'predictive' power, and although the resulting models are used for forecasting, their focus is on 'predicting the present'. Other correlated contemporaneous variables include road traffic and air passenger numbers or energy consumption and surveys of consumers and businesses about their plans and expectations. In an *Autometrics* approach, such additional data series would all be added to the information set for the relevant aggregate (Doornik, 2009b, illustrates this approach), or could be used in the Castle and Hendry (2009) framework to nowcast missing disaggregates.

The advantage of an automatic *Gets* framework that allows for more variables than observations is that it will enable very general models initially given the magnitude of databases available, with impulse-indicator saturation to detect outliers and location shifts. The selected model specification could be maintained to update within quarters, but the models can be re-selected intermittently to update if breaks have occurred. Furthermore, the quantity of data available will enable measures of variance via the realized volatility of inter-day activity, and possibly even complete distributions via a non-parametric approach. The optimal level of aggregation for such data is an empirical question that again can be tested within the *Gets* framework. Further developments to handle mixed frequencies of data could draw on a MIDAS-type of approach (see, *inter alia*, Ghysels, Santa-Clara and Valkanov, 2004, Ghysels, Sinko and Valkanov, 2006, and Clements and Galvão, 2008).

---

[3]Because many other uses of the words impulse and indicator exist, we have used the more precise designation of impulse-indicator saturation.

2.4 Prediction market data

Prediction markets for the relevant measures could be another valuable source of ancillary information: see Wolfers and Zitzewitz (2004), Gil and Levitt (2006) and Croxson and Reade (2008) *inter alia*. These information markets, like Iowa Electronic Markets for political outcomes and Betfair for sporting competitions, are claimed to have more accurate predictions than polls, surveys, or expert judgement, and to have predicted high profile events well, such as the size of President Obama's victory. Although, it is unclear how well they forecast relative to robust econometric equations, as location shifts occur intermittently (see e.g., Hendry and Reade, 2008a), the probabilities do adjust very rapidly to new information. As participants are risking real money with their views, prediction market real-time measures could be a valuable addition to measuring and nowcasting the present state of the economy. For example, the July 2009 Iowa Electronic Market is for bets on the Federal Reserve's Monetary Policy decisions, which thereby could indicate implicit knowledge about the state of the US economy.

2.5 Surveys

Survey information could be used directly to modify estimates of the forecast origin values, or as possible additional regressors, or as part of a signal extraction approach to estimating missing data on the disaggregates, or as one of the devices to be pooled. We doubt their likely efficacy as leading indicators following the critiques in Emerson and Hendry (1996) and Diebold and Rudebusch (1991), since their *ex post* performance is usually superior to the *ex ante*, reflected in regular revisions of the indicator components of indexes. However, we do test whether surveys are able to provide timely detection of structural breaks in the empirical example in section 6.

3 Selecting nowcasting models

Castle and Hendry (2009) outline the use of automatic model selection for selecting nowcasting models, allowing for all the available information, multiple past breaks and contemporaneous location shifts, building on Castle *et al.* (2009a). The selection algorithm is automated in *Autometrics* with impulse-indicator saturation (see Doornik, 2009a). The *Gets* methodology commences from a general model specification. In a nowcasting context with multiple vintages of data, this enables inclusion of highly correlated vintages. *Gets* can handle perfectly collinear variables, and so will have power to select between preliminary releases and revised releases of data.

Nowcasting models tend to rely on bridge equations that can only handle a limited set of covariates, but are useful in handling mixed frequency data. The bridge equations are specified at a higher frequency than the variable to be nowcast, enabling a bridge between, say, monthly and quarterly data. Salazar and Weale (1999) use interpolated estimates of GDP, correcting for the impact of measurement error, to obtain monthly data and use a VAR approach on the higher frequency data. They find that the monthly model is useful for nowcasting quarterly GDP once one or two months' hard data becomes available. Methods proposed in the literature to get around the bridging problem include combining predictions from many small models (see Diron, 2006, and Kitchen and Monaco, 2003), common factor models (see Giannone, Reichlin and Small, 2008), mixed frequency interpolation in a regression context or state space form (see Chow and Lin, 1971 for the former and Harvey and Pierse, 1984 for the latter). Mitchell, Smith, Weale, Wright and Salazar (2005) use an interpolation approach to nowcast monthly GDP using monthly indicators and then aggregate up to obtain early estimates of quarterly GDP. This is the approach used by the National Institute for Economic and Social Research for nowcasting monthly GDP growth.

A *Gets* approach allows for large sets of covariates to be included in the initial model specification, circumventing the need for bridging models.

The proposed nowcasting approach relies on well-specified econometric models to both forecast the disaggregate covariates and nowcast the aggregate. This requires all in-sample breaks to be modelled. We use two methods to ensure breaks are captured. First, impulse-indicator saturation (IIS) is applied to all selected models. Castle *et al.* (2009a) examine the ability of IIS to detect multiple breaks, and show it can find up to 20 breaks in 100 observations. Secondly, all models are selected recursively to enable breaks to be rapidly updated in the model specification. This procedure is feasible in a real-time context due to automation of the selection algorithm, which requires that it is fast and easy to implement. Castle *et al.* (2009b) show that recursive updating is beneficial when there is an external break (rolling windows could also be used, but can be costly if the size of the rolling window is small relative to the full sample).

4   Problems in economic nowcasting

Nowcasting is characterized by a trade-off between timeliness and data quality. More timely data are subject to larger revisions due to initial measurement errors, but allow breaks at the forecast origin to be detected more rapidly, enabling a switch to robust forecasting. Hence, the 'optimal' publication lag is an empirical question, and the views of statistical agencies differ from country to country. Furthermore, data are released with different frequencies (e.g., GDP at the quarterly frequency and retail sales at the monthly frequency). Again, there is a trade-off as to the 'optimal' frequency to model the data at. Although time disaggregation is beneficial for detecting breaks sooner, Castle and Hendry (2008) show that the impact of location shifts is not mitigated by higher frequency data: when the objective is to forecast (say) a quarterly outcome one quarter ahead, quarterly and monthly models are equally affected by forecast-origin breaks. However, in nowcasting, some contemporaneous data are already available for the period to be nowcast, so higher-frequency observations could help.

There is a signal extraction problem at T as the final observation will contain measurement error, so an 'extreme observation' could be an outlier that will be revised or a more permanent location shift. The two explanations are observationally equivalent with one data point. Measurement errors in dynamic models lead to autocorrelated residuals, but after the data are revised, there are only a few observations at the end of the sample to detect such measurement errors: full sample tests for autocorrelation will have low power, and often congruency is anyway imposed in the selection procedure to ensure no full-sample autocorrelation. Hence, residual analysis at the end of sample is all that is available to decipher whether an 'extreme' value for the last observation is a measurement error or a break. Section 4.1 outlines the characteristics of the last observation, $y_{i,T}$, when there is measurement error or a break, including the use of IIS. Section 4.2 discusses the inconsistency between nowcasting methods designed to mitigate measurement errors and those designed to be adaptable to structural breaks, and section 4.3 considers whether pooling can mitigate forecast failure when there is a break.

4.1   Measurement errors and breaks

Let $\delta < 1$ denote a high-frequency sub-period, and $\widetilde{y}_{i,T|T-\delta}$ a forecast of $y_{i,T}$ based on information available at time $T - \delta$. There are N disaggregate series, of which J are known at T and $N - J$ are unknown at T. Then if $\widetilde{y}_{i,T|T-\delta}$ differs significantly from $y_{i,T}$ for $i = 1, \ldots, J$, there are three plausible interpretations; either the discrepancy is due to a structural break or a serious measurement error, or a combination of both at T.

First, consider the case of measurement error. Following Clements and Hendry (1998, ch.8), assume the estimation period consists of $t = 1, \ldots, T^e$, during which data are measured without error, and $s = T^e, \ldots, T$, when data are measured with error. This captures the idea that near the forecast origin, data are measured in real-time, whereas data further back in the sample have been revised as more information is accumulated. Assume a simple autoregressive DGP for each disaggregate:

$$y_{i,t} = \psi_i y_{i,t-1} + v_{i,t} \quad \text{for } t = 1, \ldots, T; \; i = 1, \ldots, J \tag{1}$$

where $v_{i,t} \sim \mathsf{IN}\left[0, \sigma_{v_i}^2\right]$ and $|\psi_i| < 1$. For the period $1, \ldots, T^e$, we assume a correctly specified model with known parameters, so the model coincides with (1), and for $s = T^e, \ldots, T$ the 'true' data, $y_{i,s}^*$, are measured with error, so we observe:

$$y_{i,s} = y_{i,s}^* + \eta_{i,s} \tag{2}$$

For convenience, assume $\eta_{i,s} \sim \mathsf{IN}\left[0, \sigma_{\eta_i}^2\right]$ and $\mathsf{E}\left[y_{i,s}^* \eta_{i,t}\right] = 0$ for all s, t. In practice, these assumptions may not hold; revisions often have a systematic and predictable component, which implies that the measurement error is also systematic. For example, a possible mistake is that $\mathsf{E}\left[\eta_{i,s}\right] = \mu_{\eta_i} \neq 0$, which could be calculated from historical revisions.

For the period $T^e, \ldots, T$, the model in (1) becomes:

$$y_{i,s} = \psi_i y_{i,s-1} + \left(v_{i,s} + \eta_{i,s} - \psi_i \eta_{i,s-1}\right) = \psi_i y_{i,s-1} + e_{i,s}$$

where $e_{i,s}$ is the orthogonal component of the residual relative to $y_{i,s-1}$. When $\psi_i = 1$, so the models for the disaggregates are expressed for simplicity as:

$$\Delta y_{i,s} = v_{i,s} + \eta_{i,s} - \psi_i \eta_{i,s-1} = e_{i,s}$$

the residuals should exhibit the following properties over the period s (usually just the last few observations depending on data frequency):

$$
\begin{aligned}
\mathsf{E}\left[e_{i,s}\right] &\simeq 0 \\
\mathsf{V}\left[e_{i,s}\right] &\simeq \sigma_{v_i}^2 + \left(1 + \psi_i^2\right)\sigma_{\eta_i}^2 \\
\mathsf{E}\left[e_{i,s}e_{i,s-1}\right] &= \rho_i \simeq \frac{-\psi_i \sigma_{\eta_i}^2}{\sigma_{v_i}^2 + \left(1 + \psi_i^2\right)\sigma_{\eta_i}^2}
\end{aligned}
$$

where ρ_i denotes late-onset error autocorrelation, assuming homoskedasticity such that $\sigma_{\eta_i}^2$ is constant as s approaches T, again, a stringent assumption. Hence, if the last few residuals exhibit an increase in variance and autocorrelation, more weight can be placed on the hypothesis that the observation at T can be partly attributed to measurement error.

The alternative hypothesis is that a location shift has occurred. In some cases, future changes in economic policy may be anticipated and so the error at T could be justified *ex ante*. Examples include changes in VAT rates that are announced in advance, or disaggregate specific policies such as prescription charges. Such breaks can be accounted for in the nowcast using an intercept correction. Other location shifts at T may be unpredictable, but IIS enables such shifts to be detected after they occur. By including an impulse indicator at T in the general model and selecting a specific model, the indicator will be retained if the final observation exhibits a location shift relative to the full sample mean. It is essential that impulse indicators are included for all observations to account for shifts that occur in sample, otherwise late onset shifts will be harder to detect. The final observation is evaluated against the in-sample mean,

7

so if the mean for $t = 1, \ldots, T - 1$ is incorrectly estimated due to unmodelled breaks, selection will have incorrect retention properties, and $\widehat{e}_{i,T}$ will have the incorrect mean and variance.

If an impulse is detected at T in one or more of the disaggregate models, the $J + 1, \ldots, N$ unknown disaggregates that are nowcast may need to be adjusted, depending on the correlation across disaggregate series and the correlation of measurement errors, both of which are almost certainly non-constant, particularly when a break occurs. Unfortunately, the two explanations for an 'extreme observation' at T require quite different forecasting models, leading to a tension between alternative solutions. Exponentially weighted moving average (EWMA) schemes work well when there are 'regular' noise-like measurement errors. They can be shown to be optimal for a random-walk process subject to random measurement errors, in contrast to the intercept correction and differencing approaches discussed above which offset location shifts, but exacerbate the impact of measurement errors.

4.2 EWMA and intercept correction

The EWMA forecasting formula for a scalar time series $\{y_{i,t}\}$ when $h > 0$ is:

$$\widehat{y}_{i,T+h|T} = (1 - \lambda_i) \sum_{j=0}^{T} \lambda_i^j y_{i,T-j}, \tag{3}$$

where $\lambda_i \in (0, 1)$, so with start-up value $\widehat{y}_{i,1} = y_{i,1}$ and $y_{i,0} = 0$:

$$\widehat{y}_{i,T+1|T} = (1 - \lambda_i) y_{i,T} + \lambda_i \widehat{y}_{i,T|T-1} = y_{i,T} - \lambda_i \left(y_{i,T} - \widehat{y}_{i,T|T-1} \right), \tag{4}$$

This is a 'random-walk' forecast modified by the proportion λ_i of the previous forecast error $(y_{i,T} - \widehat{y}_{i,T|T-1})$. EWMA approximates an ARIMA(0,1,1):

$$\Delta y_{i,t} = \epsilon_{i,t} - \lambda_i \epsilon_{i,t-1} \tag{5}$$

which arises from a random walk in the true variable $y_{i,t}^*$ that is measured with an error $u_{i,t}$ so $y_{i,t}$ is observed:

$$
\begin{aligned}
y_{i,t}^* &= y_{i,t-1}^* + e_{i,t} \ \text{ where } \ \mathsf{E}\left[e_{i,t}^2\right] = \sigma_{e_i}^2 \\
y_{i,t} &= y_{i,t}^* + u_{i,t} \ \text{ with } \ \mathsf{E}\left[u_{i,t}^2\right] = \sigma_{u_i}^2
\end{aligned} \tag{6}
$$

and hence is a special case of a 'structural time series' model (see Harvey, 1993, and Koopman, Harvey, Doornik and Shephard, 1999), where λ_i is determined by the relative magnitudes of the innovation error variance, $\sigma_{e_i}^2$, and the measurement error variance, $\sigma_{u_i}^2$.

A central problem of nowcasting is balancing measurement errors and location shifts, and these 'conflict' in EWMA: (4) dampens recent forecast errors, whereas rapid adjustment is essential for location shifts, so the closer λ_i is to zero the more rapidly a location shift is assimilated. Adding an impulse indicator to (6) at the forecast origin:

$$
\begin{aligned}
y_{i,t}^* &= \delta_i 1_{\{t=T\}} + y_{i,t-1}^* + e_{i,t} \\
y_{i,t} &= y_{i,t}^* + u_{i,t}
\end{aligned}
$$

leads to:

$$
\begin{aligned}
y_{i,T} &= y_{i,T-1} + \delta_i + \epsilon_{i,T} - \lambda_i \epsilon_{i,T-1} \\
y_{i,T+1} &= y_{i,T} + \epsilon_{i,T+1} - \lambda_i \epsilon_{i,T}
\end{aligned}
$$

8

then for simplicity letting $\left(y_{i,T} - \widehat{y}_{i,T|T-1}\right) = \epsilon_{i,T-1}$ with known λ_i, the forecast error at T is $y_{i,T} - \widehat{y}_{i,T|T-1} = \delta_i + \epsilon_{i,T}$. Hence, the next forecast becomes:

$$\widehat{y}_{i,T+1|T} = y_{i,T} - \lambda_i \left(\delta_i + \epsilon_{i,T}\right) \tag{7}$$

with error:

$$y_{i,T+1} - \widehat{y}_{i,T+1|T} = \lambda_i \delta_i + \epsilon_{i,T+1}$$

so large λ_i will induce a sequence of systematic errors and an MSFE of:

$$\mathsf{E}\left[\left(y_{i,T+1} - \widehat{y}_{i,T+1|T}\right)^2\right] = \lambda_i^2 \delta_i^2 + \sigma_{\varepsilon_i}^2$$

An alternative predictor can be obtained by adding an impulse indicator to the last observation to estimate δ_i as $\delta_i + \epsilon_{i,T}$, but then not extrapolating the indicator so that:

$$\widetilde{y}_{i,T|T-1} = y_{i,T-1} - \lambda_i \left(y_{i,T-1} - \widehat{y}_{i,T-1|T}\right) + \left(\delta_i + \epsilon_{i,T}\right)$$

and hence $\widetilde{y}_{i,T+1|T} = y_{i,T}$ with a forecast error $(\epsilon_{i,T+1} - \lambda \epsilon_{i,T})$. The final-period impulse indicator essentially takes the forecast back to a random walk, with MSFE:

$$\mathsf{E}\left[\left(y_{i,T+1} - \widetilde{y}_{i,T|T-1}\right)^2\right] = \left(1 + \lambda_i^2\right) \sigma_{\varepsilon_i}^2$$

so will yield a lower MSFE when $\delta_i^2 > \sigma_{\varepsilon_i}^2$, which is a standardized location shift greater than one error standard deviation, independently of λ_i. Such an impulse indicator is akin to an intercept correction (IC), which adds back a function of recent errors to offset a location shift, and can be efficacious in situations in which the form, magnitude, timing and sign of any shift are unknown (see e.g., Clements and Hendry, 1998). The key difference here is that the IC is not extrapolated into the future since the process is integrated, so is close to the approach discussed in Hendry and Santos (2005).

4.3 Pooling of forecasts

The importance of pooling lies both in revealing divergences between methods and averaging performance across methods. Since some methods are less affected by structural changes than others, alternative devices can deliver rather different forecasts in such a setting. In particular, large differences from robust methods should prompt action. Since the best forecasting strategy is inevitably unknown, pooling across forecasts from a number of models, methods, transformations and sub-sample periods can reduce serious inaccuracy: see, e.g., Newbold and Harvey (2002), Fildes and Ord (2002), and Stock and Watson (1999). Nevertheless, care is required in selecting the set of models to pool as shown by Hendry and Reade (2008b): an analogy would be the inadvisability of mixing a glass of poison with glasses of pure water, then drinking the combination.

Our proposed strategy uses an element of pooling. We suggest averaging within blocks across the impulse indicators retained at T for the known disaggregates to obtain an estimate of the break direction and magnitude, which is incorporated into the nowcast for the unknown disaggregates. Even if there is only weak evidence of a break, averaging across the intercept-corrected model, the differenced model, the EWMA and the unadjusted model will provide an 'insurance policy' forecast as in Hendry (2004).

5 Breaks and measurement errors in nowcasting

Successfully resolving the conflict between the opposite responses entailed to location shifts and measurement errors is central to accurate nowcasting. A location shift necessitates a positively-signed intercept correction (IC) for the nowcast period (see e.g., Clements and Hendry, 1996). Conversely, a measurement error entails a negative correction to offset that error, which should be ignored thereafter. But a sudden unexpected shift at a point in time could correspond to either happening. Given the inconsistency between methods to address measurement errors and those that are robust to breaks, we now consider whether additional information can help distinguish between these two hypotheses.[4]

From a policy perspective, the issue is also serious: a location shift (in inflation, say) usually necessitates a policy reaction, whereas a measurement error will not. Thus, accurate discrimination between these would be invaluable. Any analysis needs to allow for location shifts and measurement errors before and after nowcasting in real time. Several features suggest it may be possible to distinguish a location shift from a measurement error shortly after the nowcast is made, specifically within two periods. First, a measurement error at time T does not 'carry forward' into the next period, although its effects will do so in a dynamic process. Secondly, that data point will usually be revised and should then be more accurate: thus, revisions to nowcast errors at $T + 1$ should be informative about the source of the error from T to $T + 1$. Thirdly, the next nowcast error from $T + 1$ to $T + 2$ will again be large if the source is a location shift but not for a one-off measurement error: thus, repeated mis-forecasting is more indicative of a location shift. Fourth, extraneous contemporaneous data may help in discriminating within a couple of periods by whether a discrepancy from a pre-existing model persists (likely to signal a break) or rapidly disappears (more likely to be a measurement error) Finally, the variance of measurement errors usually changes as the forecast origin approaches, and such an effect should be capable of being modelled into the nowcasting procedure.

The literature on measurement errors in nowcasting is not large, unlike that on the impacts of measurement errors on estimation (summarized in e.g., Hendry, 1995). Some related approaches for forecast scenarios are considered in Clements and Hendry (1998, 1999) and Hendry and Clements (2003). Harrison, Kapetanios and Yates (2005) consider the induced negative moving average in dynamic models as well as the error variance effects caused by measurement errors, so argue for down-weighting parameters estimated from recent data when forecasting.

We discuss the simplest model formulated in section 5.1, then consider the impact on it of measurement errors only in subsection 5.2. Subsection 5.3 discusses location shifts, and contrasts their effects them with those of measurement errors. Subsection 5.4 briefly records the impacts of each incorrectly using the other's solution, and subsection 5.5 notes the way ahead.

5.1 The model

Denote the aggregate $y_t = \sum_{i=1}^{N} w_i y_{i,t}$, where $y_{i,t}$ are the disaggregates and w_i are the weights, which could be changing over time. The simplest setup that can adequately characterize the problem is:

$$y_t = y^* + \epsilon_t \tag{8}$$

where ϵ_t is an IID measurement error which we take to be:

$$\epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{9}$$

---

[4]We are indebted to Rob Engle for help in formulating this section.

independently of y^*. Intermittently, and unpredictably, y^* alters, which represents the location shift. We first consider the setting where y^* is constant; then where $\epsilon_t = 0 \; \forall t$ but y^* alters; and finally, combine these. The nowcast is for time T following an estimation sample from $1, \ldots, T - \delta$, for small δ which may be zero for some variables, but is unity for y, and we imagine a sequence of nowcasts as if in real time, at $T + 1$, $T + 2$ etc. Abstracting from dynamics and from stochastic shocks greatly simplifies the problem, but does not seem to sacrifice the essential features.

5.2 Measurement errors only

When y^* is constant in-sample, the nowcasting problem can be resolved by using the conditional expectation of y_{T+1} given information on it dated T or earlier:

$$\mathsf{E}_{T+1}\left[y_{T+1} \mid \mathcal{I}_{T+1}\right] = \mathsf{E}_{T+1}\left[(y^* + \epsilon_{T+1})\right] = y^* \tag{10}$$

Here $\mathcal{I}_{T+1}$ only includes $(y_1 \ldots y_T)$, but more generally may also include ancillary data series, some of which may be contemporaneous at $T + 1$.

Even though ϵ_t is a measurement error, regression of y_t on a constant delivers the 1-step minimum MSE forecasting device:

$$\widehat{y}_{T+1|T} = \overline{y}_{(T)} = \frac{1}{T} \sum_{t=1}^{T} y_t$$

where unconditionally from (8) and (9):

$$\mathsf{E}_{T+1}\left[\overline{y}_{(T)}\right] = y^* \text{ and } \mathsf{V}_{T+1}\left[\overline{y}_{(T)}\right] = \mathsf{E}_{T+1}\left[\left(\overline{y}_{(T)} - y^*\right)^2\right] = \mathsf{E}_{T+1}\left[\left(\frac{1}{T} \sum_{t=1}^{T} \epsilon_t\right)^2\right] = \frac{\sigma_\epsilon^2}{T} \tag{11}$$

Thus:

$$\mathsf{E}_{T+1}\left[\widehat{y}_{T+1|T} \mid \mathcal{I}_{T+1}\right] = y^*$$

and:

$$\mathsf{V}_{T+1}\left[\widehat{y}_{T+1|T}\right] = \mathsf{E}_{T+1}\left[(y_{T+1} - \overline{y})^2\right] = \mathsf{E}_{T+1}\left[\epsilon_{T+1}^2 + \left(y^* - \overline{y}_{(T)}\right)^2\right] = \sigma_\epsilon^2 \left(1 + \frac{1}{T}\right). \tag{12}$$

so the full-sample estimate $\overline{y}_{(T)}$ of y^* is an unbiased predictor, with the smallest variance of any sample size choice, which is just $\frac{1}{T}$ larger than the nowcast from the known location y^*.

5.3 Location shifts only

When $\epsilon_t = 0 \; \forall t$ but y^* is not constant, then the best predictor of y_{T+1} is $\widetilde{y}_{T+1|T} = y_T$. When no location shifts occur, then $\widetilde{y}_{T+1|T} = y_T = y^*$ is exact; and when one does occur at T, but not at either $T - 1$ or $T + 1$, denoted $\nabla_{(T)}y^*$, a one-period mistake is made:

$$y_{T+1} - \widetilde{y}_{T+1|T} = y_{T+1} - y_T = \nabla_{(T)}y^*$$

whereas the next nowcast at $T + 2$ is unbiased if the process now remains constant:

$$y_{T+2} - \widetilde{y}_{T+2|T+1} = y_{T+2} - y_{T+1} = 0.$$

Thus, the best 1-step predictor here is simply the lagged value, namely a random walk, which uses the smallest possible sample size of unity, ignoring all earlier observations.

The contrast between the outcomes in sections 5.2 and 5.3 is stark: in the former setting of just measurement error, the *full sample estimates are best*; and in the latter of just a location shift, *only the final observation* is used. It would be unsurprising to find that when both problems occur, a compromise is required, and we first illustrate the conflict by considering using the 'best solution' to each problem when it is applied to the other.

5.4 'Cross' solutions

First, consider $\widetilde{y}_{T+1|T}$ when there are no location shifts, and only measurement errors. Then:

$$\mathsf{E}_{T+1}\left[\widetilde{y}_{T+1|T}\right] = \mathsf{E}_{T+1}\left[y_T\right] = y^*$$

and:

$$\mathsf{V}_{T+1}\left[\widetilde{y}_{T+1|T}\right] = \mathsf{E}_{T+1}\left[(y_{T+1}-y_T)^2\right] = \mathsf{E}_{T+1}\left[((y_{T+1}-y^*)+(y^*-y_T))^2\right] = \mathsf{E}_{T+1}\left[\epsilon_{T+1}^2 + \epsilon_T^2\right] = 2\sigma_\epsilon^2$$
(13)

Thus, the predictor remains unbiased, but the variance increases by $(T-1)/T$, essentially doubling.

Next, consider $\widehat{y}_{T+1|T}$ when there are no measurement errors, but location shifts occur. Now we need to specify all in-sample location shifts, and let these be $\nabla_{(\tau)}y^*$ for $\tau \in \mathcal{T}$, so that after a shift:

$$y_t = y^* + \nabla_{(\tau)}y^*,$$

and hence:

$$\overline{y}_{(T)} = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}\sum_{\tau \in \mathcal{T}}\left(y^* + \nabla_{(\tau)}y^*\right) = y^* + \frac{1}{T}\sum_{\tau \in \mathcal{T}}\nabla_{(\tau)}y^*.$$

The simplest case is that of one in-sample location shift at τ^* (say), so that:

$$\overline{y}_{(T)} = y^* + \frac{(T-\tau^*)\nabla_{(\tau^*)}y^*}{T}.$$

When no further breaks occur, $\overline{y}_{(T)}$ is a biased predictor of $y_{T+1} = y^* + \nabla_{(\tau^*)}y^*$:

$$\mathsf{E}_{T+1}\left[y_{T+1}-\widehat{y}_{T+1|T}\right] = \mathsf{E}_{T+1}\left[y^* + \nabla_{(\tau^*)}y^* - \overline{y}_{(T)}\right] = \frac{\tau^*\nabla_{(\tau^*)}y^*}{T},$$

with a squared prediction error of:

$$\mathsf{E}_{T+1}\left[(y_{T+1}-\widehat{y}_{T+1|T})^2\right] = \left(\frac{\tau^*\nabla_{(\tau^*)}y^*}{T}\right)^2.$$

The impact of the in-sample break is under-estimated when not modelled, because the sample spans periods before and after the break. Either a congruent model is needed that models the break, or a shorter sample period that post-dates the most recent break is required, as would occur with a rolling-sample estimator. Assuming one of these routes is followed, we need only consider the case where no previous break contaminates the model in use, and focus on the measurement error and location shift at T.

12

5.5 Measurement errors and location shifts

A number of interacting features suggest it may be possible to distinguish a break from a measurement error after about two periods. Of course, a systematic measurement error will also act like a break, and a one-off shift like a measurement error, which will 'confuse' the analysis, as would a measurement error followed by a break of a similar magnitude and sign. Nevertheless, there are four sources for distinguishing breaks from measurement errors:

a] the measurement error does not 'carry forward' into the next period;

b] the previous data point will usually be revised, and should then be more accurate;

c] the second nowcast error using $\overline{y}_{(T+1)}$ at $T+2$ will again be large for a break, but not for a one-off measurement error;

d] the variance of measurement errors changes as the nowcast origin recedes.

Thus, we assume (8) and (9) operate till time T, then letting $\breve{y}$ etc. denote a preliminary measurement:

$$\breve{y}_{T+1} = y_a^* + \breve{\epsilon}_{T+1} \quad \text{and} \quad \breve{y}_{T+2} = y_a^* + \breve{\epsilon}_{T+2}$$

where there are no unmodelled in-sample shifts, but:

$$y_a^* = y^* + \nabla_{(T+1)} y^*$$

and after data revision at $T+2$:

$$y_{T+1} = y_a^* + \epsilon_{T+1} \tag{14}$$

Then, from above, $\overline{y}_{(T)}$ provides the minimum MSE estimator of y^*, with variance $\sigma_\epsilon^2 T^{-1}$ on average; $y_{T+1} - \breve{y}_{T+1}$ estimates the revision error, independently of y^* being constant or shifting; $\breve{y}_{T+1} - y_T$ and later $y_{T+1} - y_T$ both estimate $\nabla_{(T+1)} y^*$, albeit imprecisely; and $\mathsf{V}_{T+1}[\breve{\epsilon}_{T+1}]$ and $\mathsf{V}_{T+2}[\breve{\epsilon}_{T+2}]$ are likely to be larger than σ_ϵ^2. Since $(y_{T+1} + \breve{y}_{T+2})/2$ estimates y_a^* with a variance of:

$$\mathsf{V}_{T+2}\left[\frac{1}{2}(y_{T+1} + \breve{y}_{T+2})\right] = \frac{1}{4}\left(\sigma_\epsilon^2 + \mathsf{V}_{T+2}[\breve{\epsilon}_{T+2}]\right),$$

combining this with $\overline{y}$ estimated up to T, and the estimate of $\nabla_{(T+1)} y^*$ gained from $y_{T+1} - y_T$ yields

$$\widetilde{y}_a^* = \frac{1}{2}\left(\overline{y}_{(T)} + (y_{T+1} - y_T) + \frac{1}{2}(y_{T+1} + \breve{y}_{T+2})\right) \tag{15}$$

which also estimates y_a^*, with a variance of:

$$\mathsf{V}_{T+2}[\widetilde{y}_a^*] = \frac{1}{4}\left(\sigma_\epsilon^2\left(2 + \frac{1}{T}\right) + \frac{1}{4}\left(\sigma_\epsilon^2 + \mathsf{V}_{T+2}[\breve{\epsilon}_{T+2}]\right)\right)$$

The weighted estimator (15) has a lower variance than $(y_{T+1} + \breve{y}_{T+2})/2$ if $\mathsf{V}_{T+2}[\breve{\epsilon}_{T+2}] > 5\sigma_\epsilon^2/3$, which could be satisfied, given point b] above. Signal extraction suggests using both the latest y_{T+1} and y_T, where the weights vary with the measurement error variance; Fawcett (2008) finds that this is the case for dynamic models, in a similar setting.

6 Nowcasting Euro-area GDP growth

In this section, we undertake a nowcasting exercise using a variant of the nowcasting strategy in Castle and Hendry (2009). That paper proposed a framework to handle a large set of disaggregate variables

with non-synchronous release dates. Here, we compute nowcasts of Euro-area GDP growth using a small subset of data that is used to nowcast GDP growth directly. The data have varying release dates, and we exploit the timely release of survey data to test for location shifts at the nowcast origin. Ferrara *et al.* (2008) and Angelini, Camba-Méndez, Giannone, Rünstler and Reichlin (2008) propose the use of bridge equations to exploit monthly data in forecasting quarterly time series. In these models, each bridge equation is a forecasting model for GDP growth conditioning on different information sets. We utilize the bridge equations in a different way, instead forming blocks of variables to forecast the disaggregate monthly series directly. The blocks group together variables with close linkages such that a break in one series is likely to spread through to the other series, enabling tests of breaks within each block. The disaggregate monthly models are selected using *Autometrics* and the forecasts from these models are used to compute nowcasts of GDP growth. Section 6.1 reviews the data, §6.2 outlines the methodology, §6.3 discusses the model specifications, §6.4 reports the nowcast results, and §6.5 considers the use of judgement to distinguish between location shifts and measurement errors.

6.1 Data

The appendix lists the data and sources, along with the transformations undertaken to remove stochastic trends, and the lag in releasing the data after the reference month. We compute nowcasts for Euro-area quarterly GDP growth, (Δy_t), using the real-time data provided by EABCN.[5] Quarterly observations are indexed by t, and τ denotes the monthly index, with $t = 3\tau$. There are three releases of quarterly GDP: the first flash estimate of Euro-area GDP growth is published approximately 43 days after the end of the reference quarter, with two further revised releases in the following consecutive months. Hence, quarterly GDP estimates for y_t are released at $\tau + 2$, $\tau + 3$ and $\tau + 4$. We denote these 3 vintages of data as y_t^{v1} for the flash estimate, y_t^{v2} as the release in the month following the flash estimate and y_t^{v3} as the third release for the reference quarter.[6] Survey data are more timely, published at the end of, or just a few days after, the reference month. We assume all survey data are available at the end of the reference month, so the publication lag is 0 for these indicators. Other data such as industrial production, financial variables, unemployment, etc. are released with an intermediate lag of between 1 and 3 months.

We calculate the nowcasts over the period 2003Q2–2008Q1, giving a total of 20 nowcast observations. For each quarter, we compute 3 nowcasts of quarterly GDP growth, denoted horizon 1 (H_1) for the month at the end of the reference quarter, horizon 2 (H_2) for the month after the end of the reference quarter, and horizon 3 (H_3) for the second month after the end of the reference quarter. At the end of the reference quarter, H_1, survey data for the reference quarter will be available, but additional conditioning information may not be available (although Google Trends suggests a potential source as discussed earlier). By H_3, the flash estimate of GDP will be available. Table 1 summarizes the nowcast horizons and the timing of GDP releases.

6.2 Nowcasting methodology

The nowcasts of GDP growth are computed in 2 steps. First, forecasts of the monthly variables that are released with a lag are obtained using a block selection approach based on bridge equations. The standard procedure in the literature is to use univariate time-series models, but we use the automatic general-to-specific selection procedure with IIS to allow for covariates as well as breaks that have occurred during

---

[6]We abstract from intra-period data in this application, but there is scope for using such data to update. For example, the flash estimate of GDP is typically available 6 weeks after the end of the reference quarter so weekly updates would incorporate this information faster than monthly data.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
Q1			H_1	H_2	H_3 $y_t^{v_1}$	$y_t^{v_2}$	$y_t^{v_3}$					
Q2						H_1	H_2	H_3 $y_t^{v_1}$	$y_t^{v_2}$	$y_t^{v_3}$		
Q3	$y_t^{v_3}$								H_1	H_2	H_3 $y_t^{v_1}$	$y_t^{v_2}$
Q4	H_2	H_3 $y_t^{v_1}$	$y_t^{v_2}$	$y_t^{v_3}$								H_1

Table 1: Timing of nowcasts and GDP data releases

the in-sample period. The monthly variables are available with varying publication lags so we use direct forecasts to compute the required step-ahead forecasts, resulting in a set of monthly forecasts for the nowcast horizon. The forecast models are selected and estimated recursively using a 1% significance level, so result in 60 monthly forecasts for the period 2003M6–2008M5.

Impulse-indicator saturation is undertaken jointly with selection of the forecasting models. If an impulse indicator is found to be significant for the final observation in-sample, it is recorded. Although the data are not as timely as survey data, significant impulses could be an indication of either a location shift or measurement error. Furthermore, for each of the surveys that are available without a lag, we use the actual realization, but we also test for impulse indicators at the final in-sample observation, again to detect possible structural breaks. Adjustments can be made to the nowcasts of GDP growth if there is strong evidence of a break due to significant final-observation indicators. Adjustments will depend on judgement as to whether the break detected in one or more disaggregate series is likely to be correlated with a break in GDP growth: see section 6.5.

The second step is to compute nowcasts of GDP growth. Three quarterly time series are created for the monthly variables. Denoting $\mathbf{x}_t = (x_t, \ldots, x_0)'$ as the past history of series x_t, the three quarterly series are given by:

$$
\begin{aligned}
\mathbf{x}_t^{r_1} &= x_{\tau-2}, x_{\tau-5}, x_{\tau-8}, \ldots \\
\mathbf{x}_t^{r_2} &= x_{\tau-1}, x_{\tau-4}, x_{\tau-7}, \ldots \\
\mathbf{x}_t^{r_3} &= x_{\tau}, x_{\tau-3}, x_{\tau-6}, \ldots
\end{aligned}
$$

such that $\mathbf{x}_t^{r_1}$ corresponds to data in the first month of the quarter, $\mathbf{x}_t^{r_2}$ corresponds to data in the second month of the quarter and $\mathbf{x}_t^{r_3}$ for the final month of the quarter. All three quarterly series are included in the general unrestricted model (GUM) for GDP growth, thus avoiding the aggregation issue for monthly data, and allowing all monthly data to be included in the quarterly model if all three series are retained in the selection process. The nowcast models include all in-sample data for the conditioning variables and the latest available vintage of GDP growth. The models are selected recursively using *Autometrics* with IIS for each nowcast origin and each nowcast horizon, resulting in 60 model specifications. The nowcasts are then computed by plugging in the forecasts for those variables that are selected but still unknown at the nowcast horizon.

The nowcasts are compared to the actuals which are given by the final available vintage of data, released in December 2008. Clements and Galvão (2008) discuss which vintage of data to use for the actuals depending on whether the forecaster's aim is to forecast an early release, or later revised releases

of the data, also see Corradi, Fernandez and Swanson (2008). For the purpose of nowcasting which aims to estimate the current position of the economy, evaluating against the final available vintage seems appropriate. However, revisions to the actuals will be ongoing and this may impact on nowcast accuracy more towards the end of the nowcast sample.

6.3 Empirical Specification

We consider three nowcast model specifications including two quarterly benchmark models; [A] a univariate model; [B] a model with covariates; and [C] the disaggregate approach outlined above. As all models are selected recursively using *Autometrics*, the model specifications and parameter estimates will change with each nowcast origin. Hence, we report the GUM specifications for each model.

[A] As a benchmark, we compute the forecasts from univariate models using just the quarterly GDP growth data. The GUM for each horizon is given by:

$$
\begin{aligned}
H_1 : \Delta \widehat{y}_t &= f\left(\Delta y_{t-1}^{v_1}, \dots, \Delta y_{t-4}^{v_1}; \Delta y_{t-1}^{v_2}, \dots, \Delta y_{t-4}^{v_2}; \Delta y_{t-2}^{v_3}, \dots, \Delta y_{t-4}^{v_3}\right) \\
H_2 : \Delta \widehat{y}_t &= f\left(\Delta y_{t-1}^{v_1}, \dots, \Delta y_{t-4}^{v_1}; \Delta y_{t-1}^{v_2}, \dots, \Delta y_{t-4}^{v_2}; \Delta y_{t-1}^{v_3}, \dots, \Delta y_{t-4}^{v_3}\right) \\
H_3 : \Delta \widehat{y}_t &= f\left(\Delta y_{t}^{v_1}, \dots, \Delta y_{t-4}^{v_1}; \Delta y_{t-1}^{v_2}, \dots, \Delta y_{t-4}^{v_2}; \Delta y_{t-1}^{v_3}, \dots, \Delta y_{t-4}^{v_3}\right)
\end{aligned} \tag{16}
$$

Lags of vintages 1 and 2 are included in the GUM despite vintage 3 data being available, as this will pick up any benefits to forecasting the revision process. The vintages of data will be highly correlated, but Doornik (2009a) demonstrates the viable properties of *Autometrics* for correlated data. Note that only the GUM for H_3 includes contemporaneous data as the flash estimate is available by the third horizon. The nowcasts are computed as 1-step ahead forecasts from the recursively selected models for each horizon.

[B] As a second benchmark, we augment the information sets in (16) with conditioning variables. The set of conditioning variables includes $\mathbf{x}_t = (\Delta ip_t, \Delta ipc_t, \Delta SCI_t, \Delta rs_t, \Delta CARS_t, \Delta MCI_t,$ $\Delta ESI_t, \Delta CCI_t, \Delta RCI_t, \Delta eer_t, \Delta eurox_t, \Delta oecdli_t, \Delta EUROCOIN_t, \Delta ur_t, \Delta m1_t,$ and $SPREAD_t)$, and we use the quarterly time series of the monthly data $(x_t^{r1}, x_t^{r2}, x_t^{r3})$ for each variable.[7] As all variables in $\mathbf{x}_t$ are released with a 3-month lag or less, we include all variables at $t-1, \dots, t-4$. The nowcasts are computed as 1-step ahead forecasts from the recursively-selected models for each horizon.

[C] The disaggregate nowcasts are computed as outlined in section 6.2. The first step is to compute forecasts for the monthly variables that are available with a lag. Table 2 records the GUM specification, in-sample equation standard error, and root mean square error (RMSE) for each monthly variable. The conditioning variables are based on the bridge specifications in Ferrara *et al.* (2008).

The second stage is to select the nowcast models for quarterly GDP growth. The GUM specifications for the three-horizon nowcasts of GDP growth include $SCI_{t-i}^{r_j}, MCI_{t-i}^{r_j}, ESI_{t-i}^{r_j}, CCI_{t-i}^{r_j},$ $RCI_{t-i}^{r_j}, SPREAD_{t-i}^{r_j}, \Delta ip_{t-i}^{r_j}, \Delta ipc_{t-i}^{r_j}, \Delta rs_{t-i}^{r_j}, \Delta CARS_{t-i}^{r_j}, \Delta eer_{t-i}^{r_j}, \Delta eurox_{t-i}^{r_j}, \Delta oecdli_{t-i}^{r_j},$ $\Delta ur_{t-i}^{r_j}, \Delta EUROCOIN_{t-i}^{r_j}, \Delta m1_{t-i}^{r_j},$ for $i = 0, \dots, 2$ and $j = 1, 2, 3,$ an intercept and impulse indicator dummies. This GUM is augmented with the lagged dependent variables given in (17) for

---

[7]Third differences of the monthly covariates ($\Delta_3 x_\tau = x_\tau - x_{\tau-3}$) correspond to first differences of the quarterly series ($\Delta x_t^{r_j} = x_t^{r_j} - x_{t-1}^{r_j}, \forall j = 1, 2, 3$).

Forecast	GUM Variables	ESE	RMSE
$\widehat{\Delta ip}_\tau$	$\Delta ip_{\tau-3,\ldots,\tau-12}; \Delta ipc_{\tau-3,\ldots,\tau-12}; SCI_{\tau,\ldots,\tau-12};$C;S;I	1.02%	1.32%
$\widehat{\Delta ipc}_\tau$	$\Delta ip_{\tau-3,\ldots,\tau-12}; \Delta ipc_{\tau-3,\ldots,\tau-12}; SCI_{\tau,\ldots,\tau-12};$C;S;I	2.04%	3.83%
$\widehat{\Delta rs}_\tau$	$\Delta rs_{\tau-2,\ldots,\tau-12}; \Delta CARS_{\tau-2,\ldots,\tau-12}; \Delta ip_{\tau-3,\ldots,\tau-12};$ $\Delta ipc_{\tau-3,\ldots,\tau-12};$C;I	0.59%	0.63%
$\widehat{\Delta CARS}_\tau$	$\Delta rs_{\tau-2,\ldots,\tau-12}; \Delta CARS_{\tau-2,\ldots,\tau-12}; \Delta ip_{\tau-3,\ldots,\tau-12};$ $\Delta ipc_{\tau-3,\ldots,\tau-12};$C;I	2.81%	3.41%
$\widehat{\Delta eer}_\tau$	$\Delta eer_{\tau-1,\ldots,\tau-12}; \Delta eurox_{\tau-1,\ldots,\tau-12}; SPREAD_{\tau-1,\ldots,\tau-12};$C;I	1.30%	1.08%
$\widehat{\Delta eurox}_\tau$	$\Delta eer_{\tau-1,\ldots,\tau-12}; \Delta eurox_{\tau-1,\ldots,\tau-12}; SPREAD_{\tau-1,\ldots,\tau-12};$C;I	2.22%	3.33%
$\widehat{SPREAD}_\tau$	$\Delta eer_{\tau-1,\ldots,\tau-12}; \Delta eurox_{\tau-1,\ldots,\tau-12}; SPREAD_{\tau-1,\ldots,\tau-12};$C;I	0.14%	0.16%
$\widehat{\Delta ur}_\tau$	$\Delta ur_{\tau-2,\ldots,\tau-12}; \Delta rs_{\tau-2,\ldots,\tau-12}; \Delta ip_{\tau-3,\ldots,\tau-12};$ $\Delta ipc_{\tau-3,\ldots,\tau-12}; ESI_{\tau,\ldots,\tau-12};$ C;I	0.35%	0.56%
$\widehat{\Delta m1}_\tau$	$\Delta m1_{\tau-2,\ldots,\tau-12}; SPREAD_{\tau-1,\ldots,\tau-12}; \Delta ip_{\tau-3,\ldots,\tau-12};$ $\Delta ipc_{\tau-3,\ldots,\tau-12}; \Delta rs_{\tau-2,\ldots,\tau-12};$ C;I	0.47%	0.89%

Table 2: GUM specification for monthly variables. C=intercept, S=monthly seasonal dummies, I=impulse indicator, ESE=in-sample equation standard error of selected model, RMSE=root mean square error. Sample is 1996M5–2008M5, with 60 recursive forecasts.

each horizon, resulting in 150, 151 and 152 variables in the GUM for H_1, H_2 and H_3 respectively, plus an additional T indicator dummies.

$$
\begin{aligned}
H_1 &: \quad \Delta y_{t-1}^{v_1}, \Delta y_{t-2}^{v_1}; \Delta y_{t-1}^{v_2}, \Delta y_{t-2}^{v_2}; \Delta y_{t-2}^{v_3} \\
H_2 &: \quad \Delta y_{t-1}^{v_1}, \Delta y_{t-2}^{v_1}; \Delta y_{t-1}^{v_2}, \Delta y_{t-2}^{v_2}; \Delta y_{t-1}^{v_3}, \Delta y_{t-2}^{v_3} \\
H_3 &: \quad \Delta y_{t}^{v_1}, \Delta y_{t-1}^{v_1}, \Delta y_{t-2}^{v_1}; \Delta y_{t-1}^{v_2}, \Delta y_{t-2}^{v_2}; \Delta y_{t-1}^{v_3}, \Delta y_{t-2}^{v_3}
\end{aligned}
\tag{17}
$$

The models are selected recursively at the 1% significance level and the nowcasts are computed using:

$$
\Delta \widehat{y}_t^{H_k} = \widehat{\boldsymbol{\beta}}' \overline{\mathbf{x}}_t + \widehat{\boldsymbol{\gamma}}' \widetilde{\mathbf{x}}_t + \widehat{\boldsymbol{\delta}}' \boldsymbol{\iota}_t
\tag{18}
$$

where $\overline{\mathbf{x}}_t$ denotes the retained variables that are known at t for horizon k, $\widetilde{\mathbf{x}}_t$ denotes the forecasts for the retained variables that are unknown at t for horizon k and $\boldsymbol{\iota}_t$ denotes the retained impulse indicators. For example, if the nowcast model for H_2 retained Δip_t this would be replaced with the forecast $\widehat{\Delta ip}_t$ as industrial production figures are released with a 3 month lag and horizon 2 is computed for the month after the nowcast horizon.

We compare 3 different disaggregate forecast methods for $\widetilde{\mathbf{x}}_t$. First, we compute *ex post* nowcasts taking the disaggregates as known, denoted [C1]. Hence, all retained variables will be in the set $\overline{\mathbf{x}}_t$. This provides a benchmark against which to assess the disaggregate nowcasts, but will be infeasible in a real-time setting. Second, we use the forecasts from the forecasting models reported in table 2, [C2]; and finally, we use a robust forecasting device to compute the disaggregate forecasts given by the double differenced device (DDD), [C3]. The DDD is given by:

$$
\widetilde{x}_\tau = x_{\tau-l}
\tag{19}
$$

for data released with lag l. The forecast device in (19) is robust to location shifts once l periods have elapsed after a break, but differencing will usually increase the error variance of the forecast.

17

6.4 Real-time nowcast results

Table 3 reports the ratio of RMSE for methods [B] and [C] relative to the univariate benchmark model [A] (actual RMSE for [A] reported in italics).

	A	B	C1	C2	C3
H_1	*0.2951*	0.9823	0.4340	0.7994	0.7531
H_2	*0.3056*	1.2245	0.4808	0.6372	0.6222
H_3	*0.1486*	1.0372	0.8320	0.8320	0.8320

Table 3: Actual RMSE for [A] and ratio of RMSE to the univariate benchmark model [A] for [B] and [C].

There are substantial benefits to using the disaggregated approach at all horizons. If the monthly indicators were known at the nowcast origin, the RMSE more than halves for the horizons before the flash estimate is released. However, as the monthly indicators are unknown they must be forecast. Despite the additional uncertainty in forecasting the indicators, the RMSE is reduced by more than 20% at H_1 and 35% at H_2 when more monthly indicators are known. By H_3 all monthly indicators are known (other than industrial production which was not retained in the nowcast models) so all disaggregate methods are equal. However, augmenting the flash estimate with monthly indicators still yields a 17% improvement in RMSE. The differenced device for forecasting the monthly indicators yields a small improvement over the forecasting models for [C2], despite the forecasting models including IIS and recursive selection, both of which should account for location shifts. There is no benefit to augmenting the univariate model with additional covariates at the quarterly frequency, and at H_2 it is costly, suggesting that the main benefit of the disaggregate approach comes from incorporating higher-frequency data that is available with more timely release dates.

Figure 2 reports the mean absolute nowcast error (MANE) and RMSE for each method at all three horizons. The MANE for the univariate models is large for horizons before the flash estimate is released, but conditioning on covariates, even at the quarterly frequency, yields a substantial improvement. The MANE is slightly larger for the disaggregate nowcasts by H_3, but the smaller variance yields a smaller RMSE. The RMSE of the disaggregate model with known covariates [C1] at H_1 is as small as any nowcast at H_3 which suggests that the flash estimate does not yield additional information over the monthly indicators; indeed, the flash estimate is likely to condition on the same information set.

Figure 3 records the nowcast errors over the 5-year nowcast period for each horizon (all panels have the same scale). There is no systematic nowcast error for any of the models. Models [A] and [B] have a higher variance, with model [A] producing downward biased nowcasts on average, but there is no predictable component to the nowcast errors. By H_3, all nowcast errors follow a similar pattern.

To summarize, we find that using higher-frequency monthly data to produce more timely nowcasts of Euro-area GDP growth yields substantial benefits in terms of RMSE, particularly over short horizons. The main benefits come from rapid incorporation of information available at the monthly frequency, accounting for outliers in the GDP growth models using impulse-indicator saturation, and the use of recursive selection and estimation to update the models rapidly. We also find that there is little benefit to using well-specified econometric models to forecast the monthly indicators as opposed to using a robust forecasting device such as (19).
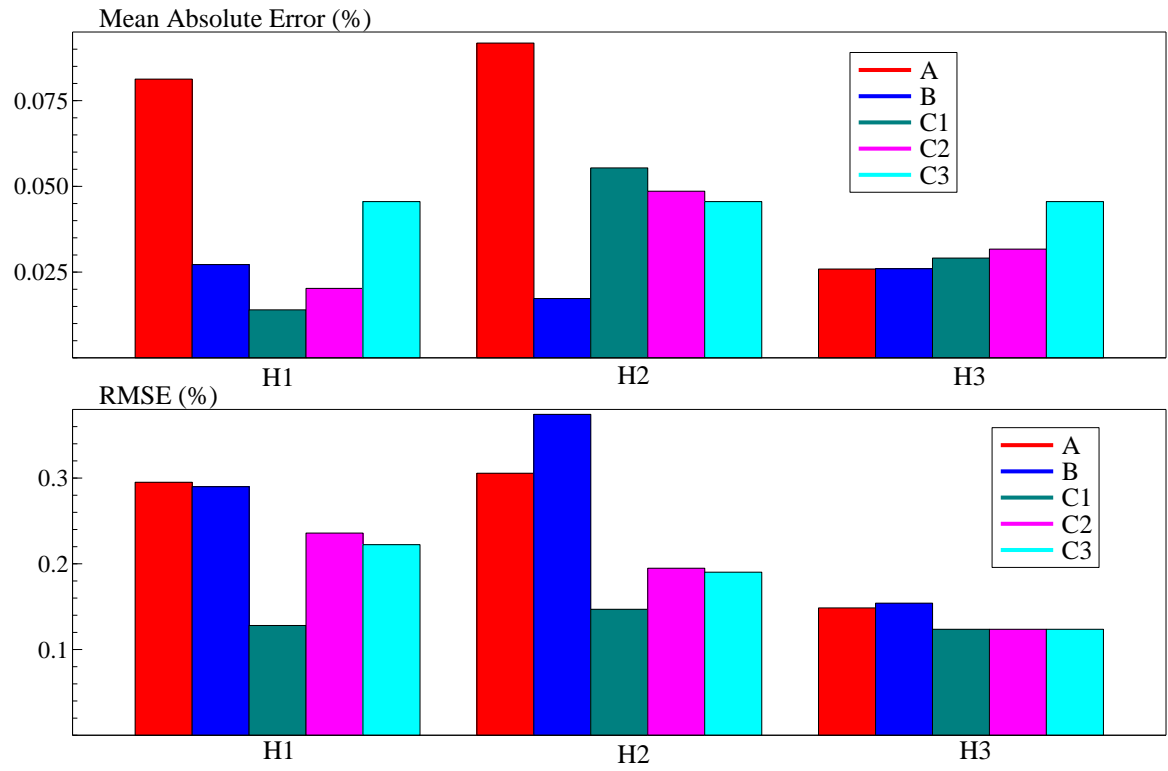
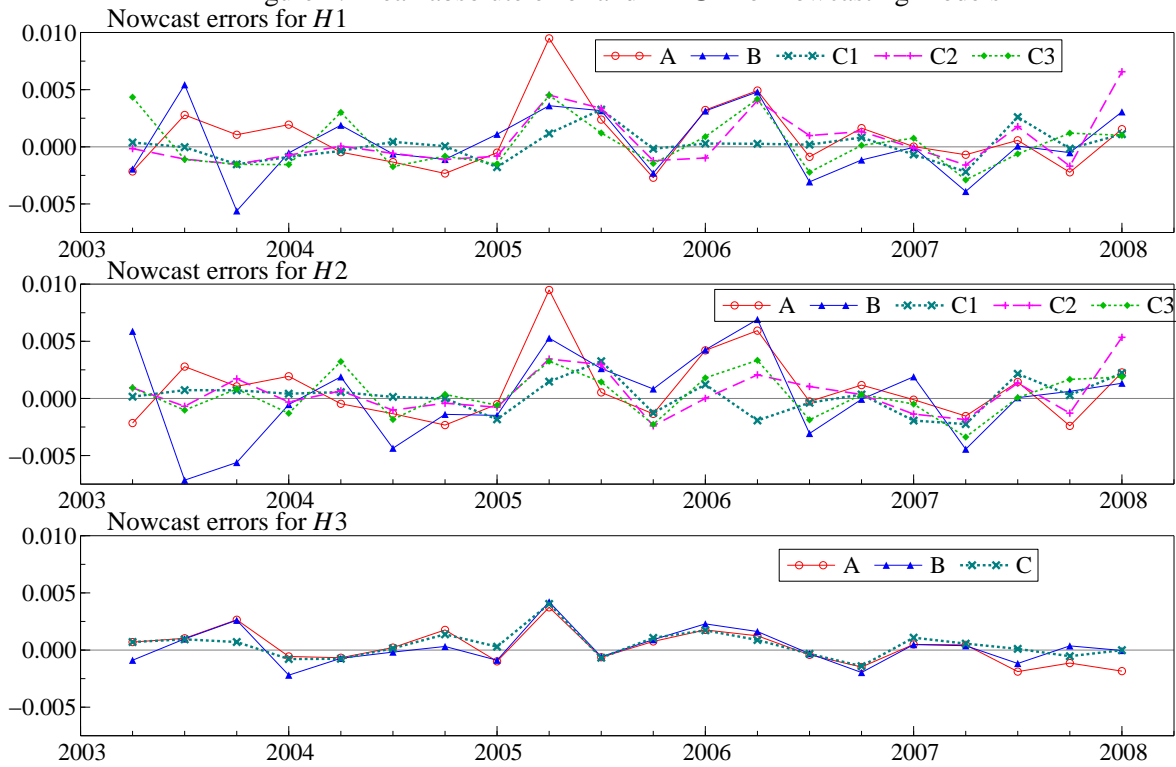Figure 2: Mean absolute error and RMSE for nowcasting models

Figure 3: Nowcast errors

19

6.5 Measurement errors and structural breaks: the role of judgement

We now consider whether judgement can improve the nowcast performance of the disaggregate model. Two possible explanations for poor nowcast performance could be measurement errors or location shifts: section 5 discussed under what conditions it is possible to distinguish between these two possible explanations. In our nowcast exercise, the flash estimate of GDP is only available in H_3, so detecting a measurement error *ex ante* is very difficult as there is no revision process upon which to base the decision. However, when it comes to forecasting GDP growth, both the new flash estimate and the past revision will help to determine whether a shift is due to a temporary measurement error or a more permanent structural break. *Ex post*, one can look at the revisions process to see how informative revisions are, to more rapidly detect measurement errors versus breaks. Mitchell (2009) demonstrates that the weights placed on 'hard' and 'soft' indicators shift in times of structural change, so detection of breaks is essential for the nowcasting model specifications.

Table 4 reports the magnitudes of revisions over the nowcast horizon, where v_f is the final available vintage (December 2008). Revisions between the flash estimate and vintages 2 and 3 are small. However, revisions between the flash estimate and the final vintage are more than 10 times larger in absolute value (in keeping with figure 1), suggesting that there are substantial measurement errors in the first few vintages of data, but there is little information content in the revisions in the first few months. This implies that even in a nowcasting context, detecting measurement errors close to the nowcast origin is very difficult, and there is little additional information in the second and third vintage of GDP data. Figure 4 provides more detail to the table, recording GDP growth over the nowcast period (given by the final vintage) along with three revisions to the flash estimate; the second vintage, third vintage and final vintage. There is no systematic component to the revisions process over this period, so again the information content to detect measurement errors using the revisions process is limited.

	$\Delta y_t^{v_1} - \Delta y_t^{v_2}$	$\Delta y_t^{v_1} - \Delta y_t^{v_3}$	$\Delta y_t^{v_1} - \Delta y_t^{v_f}$
ME (%)	0.0045	-0.0059	-0.0674
SD (%)	0.0214	0.0391	0.1583

Table 4: Mean error (ME) and standard deviation (SD) of the revisions to GDP growth. Results for 2003Q2–2008Q1, reported as percentages.

We next assess whether we can use breaks in the monthly indicators to detect structural breaks in GDP growth. Impulse-indicator saturation at 1% is applied recursively to models for all the monthly variables, including the surveys known at the forecast horizon, and any significant final-observation indicators are recorded. No significant impulse indicators were found over the nowcast horizon for Δip_τ, SCI_τ, Δrs_τ, CCI_τ, RCI_τ, Δeer_τ, $\Delta eurox_\tau$, $\Delta oecdli_\tau$ and $\Delta EUROCOIN_\tau$, but significant impulses were found for:

$$
\begin{aligned}
\Delta ipc_\tau &= 2005(12); 2006(1) \\
\Delta CARS_\tau &= 2005(5); 2005(6) \\
MCI_\tau &= 2003(12) \\
ESI_\tau &= 2003(12) \\
\Delta ur_\tau &= 2006(4); 2007(7); 2007(9); 2008(1) \\
\Delta m1_\tau &= 2005(1); 2005(6); 2006(12) \\
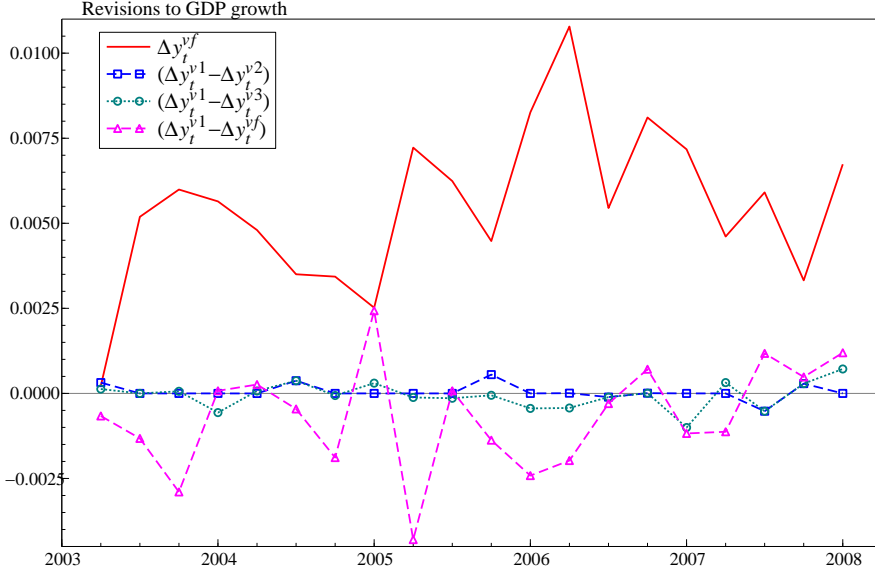SPREAD_\tau &= 2007(8)
\end{aligned}
$$

Figure 4: Revisions to GDP growth

Castle and Hendry (2009) propose a tighter significance level ($0.5\% \leq \alpha \leq 0.1\%$) for selecting impulse indicators. Using a 0.1% significance level we find significant impulse indicators for:

$$\Delta ur_\tau = 2006(4); 2007(7); 2008(1)$$
$$\Delta m1_\tau = 2005(6)$$

Figure 5 plots GDP growth against the retained impulse indicators at $\alpha = 1\%$ for all covariates aggregated to the quarterly scale. Visual inspection suggests that there are no significant location shifts in GDP growth over the nowcast period, so we cannot anticipate that judgemental adjustments will improve the nowcast performance in this example. Further, the retained dummies are not clustered together; if they were this would be stronger evidence for a location shift in GDP growth. We check whether adjustments yield an improvement in the nowcast performance using the rule:

$$\Delta \widehat{y}_t = (1 - \mathsf{I}_k) \Delta \widetilde{y}_t + \mathsf{I}_k \Delta \widetilde{y}_t^* \tag{20}$$

I_k is an indicator function taking the value 1 when an indicator dummy is retained for the covariates, $\Delta \widetilde{y}_t$ are the nowcasts obtained from [C2], and $\Delta \widetilde{y}_t^*$ is the DDD given by:

$$\Delta \widetilde{y}_t^* = \begin{cases} \Delta y_{t-1}^{v_2} & \text{for } H_1 \\ \Delta y_{t-1}^{v_3} & \text{for } H_2 \\ \Delta y_t^{v_1} & \text{for } H_3 \end{cases}$$

Results are reported in table 5. The adjusted model has worse nowcast performance at longer horizons. This suggests that a tighter criterion is needed for the adjustment rule. Castle and Hendry (2009) propose the rule that I_k takes the value 1 when $k\alpha J \geq p$ for a small integer k where $p = \sum_{i=1}^{J} 1_T$ is the number of impulse indicators retained at the nowcast origin for the J known monthly covariates. In our example this would be the surveys and as no surveys retain dummies at $\alpha = 0.1\%$, no adjustment should be made.

In summary, distinguishing between measurement errors and location shifts in a nowcasting context is extremely difficult given the limited information set available. In the empirical example we consider,
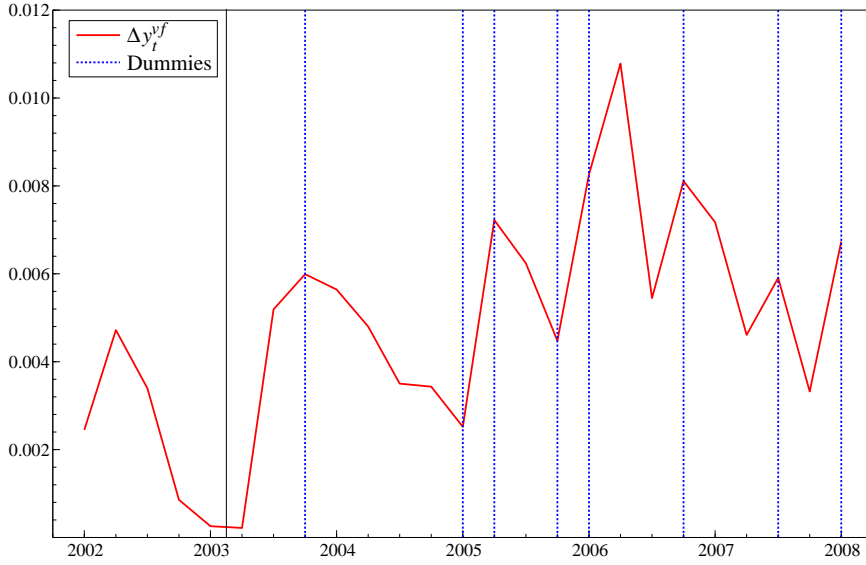
21

Figure 5: GDP growth and retained dummies for monthly indicators at $\alpha = 1\%$

	C2	C2-adjust
H_1	0.2359	0.2327
H_2	0.1948	0.2146
H_3	0.1236	0.1603

Table 5: RMSEs for disaggregate nowcast model C2 and adjusted model, C2-adjust.

the nowcast period does not span a period with large measurement errors or structural breaks. Extending the analysis to incorporate 2008/2009 data may well yield more interesting results given the increased uncertainty and location shifts evident in the most recent data.

7 Conclusion

Nowcasting is not just contemporaneous forecasting since the use and timing of contemporaneous data introduces importantly different aspects compared to *ex ante* forecasting. We first consider the four main reasons for nowcasting, and their associated problems of missing data, measurement errors, changing database, and breaks. Location shifts can induce nowcast failure, but interact with measurement errors to make discrimination difficult in the available time horizon.

We develop the nowcasting strategy proposed in Castle and Hendry (2009) of building models by automatic model selection methods (here *Autometrics*), using all available measures (possibly including disaggregates, surveys, and on-line data), testing for shifts using impulse-indicator saturation. *Autometrics* can allow for location shifts both in-sample and at the nowcast origin while still including all available covariate information by handling more variables than observations. Although data are released at varying times, this can be accommodated in our approach, so the method ensures that the largest available information set is used at each nowcast origin.

Next forecasts are made of all variables that are released with a publication lag each period, switching to robust forecasts of missing series when breaks are detected, but using a conservative strategy that requires strong evidence of breaks. The available information and forecasts then combine to produce

nowcasts of the desired variables. We apply a variant of this nowcasting strategy to nowcast Euro-area GDP growth. The application is illustrative of potential gains to reducing RMSEs of flash estimates, and we anticipate further gains by using more extensive information sets as discussed in the paper, especially surveys and Google Trends data.

Initial estimates of economic data can have a substantive impact on expectations, plans, future activity and economic policy, driving the need for accurate nowcasts. The apparent 'break down' of the ONS's current models, reported by the *Financial Times* on 25 July 2009, shows the difficulty of nowcasting in times of economic uncertainty and structural change. A methodology that is robust to such breaks is needed in the current climate, and this paper demonstrates that automatic model selection with higher-frequency covariates, break detection, and robust forecasts can yield gains in nowcast accuracy.

References

Angelini, E., Camba-Méndez, G., Giannone, D., Rünstler, G., and Reichlin, L. (2008). Short-term forecasts of Euro-area GDP growth. Working paper no. 949, European Central Bank, Frankfurt.

Ashley, J., Driver, R., Hayes, S., , and Jeffery, C. (2005). Dealing with data uncertainty. *Bank of England Quarterly Bulletin*, **Spring**, 23–29.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2009a). Model selection when there are multiple breaks. Working paper, Economics Department, University of Oxford.

Castle, J. L., Fawcett, N. W. P., and Hendry, D. F. (2009b). Forecasting with equilibrium-correction models during structural breaks. *Journal of Econometrics*, **forthcoming**, –.

Castle, J. L., and Hendry, D. F. (2008). Forecasting UK inflation: Empirical evidence on robust forecasting devices. In Rapach, D. E., and Wohar, M. E. (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty: Frontiers of Economics and Globalization Volume 3*, pp. 41–92. Bingley, UK: Emerald Group.

Castle, J. L., and Hendry, D. F. (2009). Nowcasting from disaggregates in the face of location shifts. *Journal of Forecasting*, **forthcoming**, –.

Castle, J. L., and Shephard, N. (eds.)(2009). *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*. Oxford: Oxford University Press.

Choi, H., and Varian, H. (2009). Predicting the present with google trends. Unpublished paper, Economics Research Group, Google.

Chow, G., and Lin, A. L. (1971). Best linear unbiased interpolation, distribution and extrapolation of time series by related series. *The Review of Economics and Statistics*, **53**, 372–5.

Clements, M. P., and Galvão, A. B. (2008). Macroeconomic forecasting with mixed frequency data: Forecasting US output growth. *Journal of Business and Economic Statistics*, **26**, 546–554.

Clements, M. P., and Hendry, D. F. (1996). Intercept corrections and structural change. *Journal of Applied Econometrics*, **11**, 475–494.

Clements, M. P., and Hendry, D. F. (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

Clements, M. P., and Hendry, D. F. (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, Mass.: MIT Press.

Clements, M. P., and Hendry, D. F. (2003). Forecasting in the National Accounts at the Office for National Statistics. Report no 12, Statistics Commission.

Corradi, V., Fernandez, A., and Swanson, N. R. (2008). Information in the revision process of real-time data. *Journal of Business and Economic Statistics*, forthcoming.

Croxson, K., and Reade, J. J. (2008). Information and efficiency: Goal arrival in soccer betting. Mimeo, Economics Department, Oxford University.

Diebold, F. X., and Rudebusch, G. D. (1991). Turning point prediction with the composite leading index: An ex ante analysis. In Lahiri, K., and Moore, G. H. (eds.), *Leading economic indicators: New approaches and forecasting records*, pp. 231–256. Cambridge: Cambridge University Press.

Diron, M. (2006). Short-term forecasts of Euro-area real GDP growth: an assessment of real-time performance based on vintage data. Working paper No. 622, European Central Bank.

Doornik, J. A. (2009a). Autometrics. in Castle, and Shephard (2009), pp. 88–121.

Doornik, J. A. (2009b). Improving the timeliness of data on influenza-like illnesses using Google Trends. Typescript, Department of Economics, University of Oxford.

Emerson, R. A., and Hendry, D. F. (1996). An evaluation of forecasting using leading indicators. *Journal of Forecasting*, **15**, 271–91.

Faust, J., Rogers, J. H., and Wright, J. H. (2007). News and noise in G-7 GDP announcements. *Journal of Money, Credit and Banking*, **37**, 403–419.

Fawcett, N. W. P. (2008). Essays on econometrics and forecasting. D.Phil Thesis, Economics Department, University of Oxford.

Ferrara, L., Guegan, D., and Rakotomarolahy, P. (2008). GDP nowcasting with ragged-edge data: A semi-parametric modelling. CES working paper, 82, Paris School of Economics.

Fildes, R., and Ord, K. (2002). Forecasting competitions – their role in improving forecasting practice and research. In Clements, M. P., and Hendry, D. F. (eds.), *A Companion to Economic Forecasting*, pp. 322–353: Oxford: Blackwells.

Garratt, A., and Vahey, S. P. (2006). UK real-time macro data characteristics. *Economic Journal*, **116**, F119–135.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The MIDAS touch: Mixed Data Sampling regression. Mimeo, Chapel Hill, N.C.

Ghysels, E., Sinko, A., and Valkanov, R. (2006). MIDAS regressions: Further results and new directions. *Econometric Reviews*, **26**, 53–90.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, **55**, 665–676.

Gil, R., and Levitt, S. D. (2006). Testing the efficiency of markets in the 2002 world cup. Mimeo, Economics Department, University of California at Santa Cruz.

Harrison, R., Kapetanios, G., and Yates, T. (2005). Forecasting with measurement errors in dynamic models. *International Journal of Forecasting*, **21**, 595–607.

Harvey, A. C. (1993). *Time Series Models* 2nd (first edition 1981) edn. Hemel Hempstead: Harvester Wheatsheaf.

Harvey, A. C., and Pierse, R. G. (1984). Estimating missing observations in economic time series. *Journal of the American Statistical Association*, **79**, 125–31.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F. (2004). Forecasting long-run TV advertising expenditure in the UK. Commissioned report, Ofcom, London. http://www.ofcom.org.uk/research/tv/reports/tvadvmarket.pdf.

Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics*, **135**, 399–426.

Hendry, D. F., and Clements, M. P. (2003). Economic forecasting: Some lessons from recent research. *Economic Modelling*, **20**, 301–329. European Central Bank, Working Paper 82.

Hendry, D. F., Johansen, S., and Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, **33**, 317–335. Erratum, 337–339.

Hendry, D. F., and Reade, J. J. (2008a). Economic forecasting and prediction markets. Working paper, Economics Department, Oxford University.

Hendry, D. F., and Reade, J. J. (2008b). Modelling and forecasting using model averaging. Working paper, Economics Department, Oxford University.

Hendry, D. F., and Santos, C. (2005). Regression models with data-based indicator variables. *Oxford Bulletin of Economics and Statistics*, **67**, 571–595.

Johansen, S., and Nielsen, B. (2009). An analysis of the indicator saturation estimator as a robust regression estimator. in Castle, and Shephard (2009), pp. 1–36.

Kitchen, J., and Monaco, R. M. (2003). Real-time forecasting in practice: the U.S. treasury staff's real-time GDP forecast system. *Business Economics*, 10–19.

Koopman, S. J., Harvey, A. C., Doornik, J. A., and Shephard, N. (1999). *Structural Time Series Analysis, Modelling, and Prediction using STAMP* 2nd edn. London: Timberlake Consultants Press.

Mitchell, J. (2004). Revisions to economic statistics. Commission report no. 17, volume 2, april, Review by National Institute of Economic and Social Research. http://www.statscom.org.uk/uploads/files/reports/Revisions-vol-2.pdf.

Mitchell, J. (2009). Where are we now? The UK recession and nowcasting GDP growth using statistical models. *National Institute Economic Review*, **209**, 60–69.

Mitchell, J., Smith, R. J., Weale, M. R., Wright, S. H., and Salazar, E. L. (2005). An indicator of monthly GDP and an early estimate of quarterly GDP growth. *Economic Journal*, **115(501)**, F108–29.

Newbold, P., and Harvey, D. I. (2002). Forecasting combination and encompassing. In Clements, M. P., and Hendry, D. F. (eds.), *A Companion to Economic Forecasting*, pp. 268–283: Oxford: Blackwells.

Reed, G. (2000). How the preliminary estimate of GDP is produced. *Economic Trends*, **556**, 53–61.

Reed, G. (2002). How much information is in the UK preliminary estimate of GDP?. *Economic Trends*, **585**, 1–8.

Salazar, E., and Weale, M. (1999). Monthly data and short-term forecasting: an assessment of monthly data in a VAR model. *Journal of Forecasting*, **18(7)**, 447–62.

Stock, J. H., and Watson, M. W. (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In Engle, R. F., and White, H. (eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive Granger*, pp. 1–44. Oxford: Oxford University Press.

Walton, D. R. (2006). Dealing with data uncertainty. MFE lecture slides, Department of Economics, Oxford University.

Wolfers, J., and Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, **18**, 107–126.

Appendix

Label	Description	Source	Period	Transformation	Lag release
Y_t	real GDP, million, constant prices	[1] (yer.xls)	1995Q1-2008Q4	2	2,3,4
IP_t	industrial production index (NSA)	[1] (MI-B)	1990M1-2008M9	2	3
IPC_t	industrial production index for construction (NSA)	[1] (MI-L)	1990M1-2008M9	2	3
SCI_t	service sector confidence indicator (survey)	[1] (MI-GE)	1995M4-2008M11	3	0
RS_t	retail trade index, except motor vehicles, constant prices	[1] (MI-P)	1995M1-2008M10	2	2
$CARS_t$	new passenger car registrations, total	[1] (MI-V)	1990M1-2008M10	1*	2
MCI_t	industry confidence indicator (survey)	[1] (MI-FO)	1990M1-2008M11	3	0
ESI_t	economic sentiment indicator (base=100)	[1] (MI-FN)	1990M1-2008M11	3*	0
CCI_t	consumer confidence indicator (survey)	[1] (MI-FS)	1990M1-2008M11	3	0
RCI_t	retail trade confidence indicator (survey)	[1] (MI-GA)	1990M1-2008M11	3	0
EER_t	real effective exchange rate, CPI deflated, core group	[1] (MI-BK)	1993M1-2008M11	2	1
$EUROX_t$	Dow Jones Euro Stoxx Broad stock exchange index	[1] (MI-DP)	1994M1-2008M11	2	1
$OECDLI_t$	OECD composite leading indicator, Euro area, trend restored	[2] (CLI-AF)	1990M1-2008M12	2	0
$EUROCOIN_t$	Eurocoin indicator	[3]	1990M1-2008M12	1	0
UR_t	unemployment rate, total	[1] (MI-X)	1995M1-2008M10	2	2
$M1_t$	Index of notional stock, money M1	[1] (MI-FA)	1990M1-2008M10	2	2
$3MTB_t$	3-month interest rate, Euribor	[1] (MI-DD)	1994M1-2008M11	1	1
$10YB_t$	10-year Euro area bond yield	[1] (gbr10y.xls)	1994M1-2008M8	1	1
$SPREAD_t$	=(3MTB-10YB)/100	-	1994M1-2008M8	-	1

Table 6: Sources: [1] EACBN real-time database (www.eacbn.org), [2] OECD monthly indicators database (http://stats.oecd.org/mei), [3] CEPR database (from Bank of Italy) (http://eurocoin.cepr.org/). Notes: MI-? refers to the monthly indicators spreadsheet, CLI-? refers to the composite leading indicators spreadsheet and ? denotes the column reference. NSA - not seasonally adjusted. Transformations: $(1) = \Delta X_t = (X_t - X_{t-1})$; $(1^*) = \Delta (X_t/1000000)$; $(2) = \Delta x_t = (ln(X_t) - ln(X_{t-1}))$; $(3) = X_t/10$; $(3^*) = X_t/100$. Upper case denotes levels and lower case denotes logs.