2 Explaining nowcast errors

HENRIETTE DRUBA (1),JENNIFER L.CASTLE (2) AND DAVID F. HENDRY (3)

Abstract: Preliminary estimates of macroeconomic aggregates only become available with a time lag and are often unreliable and must be revised, more so during periods of structural change such as the financial crisis and subsequent recession. This article takes the forecast error taxonomy in Hendry and Mizon (2012), and adapts it to nowcasting. The taxonomy provides a framework for thinking about potential problems facing the nowcaster in achieving useful nowcasts. It considers a wide array of sources of nowcast errors, from estimation uncertainty to model mis-specification. Importantly, the taxonomy incorporates unforeseen changes in parameters, and thus allows for a formal analysis into the consequences of structural breaks for nowcasting. Additionally, the taxonomy is applied to evaluate the impact of the different error sources on the expected nowcast error. This evaluation yields seven insights into the sources of nowcast errors, and into sources of nowcast failure. An empirical example nowcasting consumption expenditure illustrates the insights from the taxonomy.

JEL codes: C52, C53, E21

Keywords: Nowcasting, bias correction, model selection, structural breaks, consumption

(1) Department of Economics, University of Oxford.
(2) Magdalen College, and Institute for New Economic Thinking, Oxford Martin School, Un iversity of Oxford.
(3) Department of Economics, and Institute for New Economic Thinking, Oxford Martin School, University of Oxford.

1. Introduction

The current state of an economy is never known with any certainty, as even preliminary estimates of most macroeconomic aggregates only become available with a time lag, and usually are revised later as new information accrues. Periods of rapid or sudden structural change, like the financial crisis and subsequent recession, exacerbate such problems. Unfortunately, incomplete knowledge of the economy's current state can lead to inappropriate economic policies and inaccurate forecasts of its future performance. Lacking accurate measurements, statistical agencies and policy makers face the challenge of 'forecasting' the contemporaneous state of the economy, a process known as nowcasting. The key difference between a nowcast and a forecast is that the former seeks to ascertain what has actually happened, but as yet is not fully observed, whereas the latter concerns what might happen by a future date. This paper outlines an approach to nowcasting that uses statistical forecasting models to exploit high-frequency, real-time information, disaggregated data, and leading indicators published at higher frequencies to provide 'contemporaneous forecasts' of economic activity.

There are two distinct forms of nowcasting; either predicting the macroeconomic aggregate (e.g., GDP) directly using available cognate information, or using the released data on components and supplementing with predictions for the components with no statistical data on the contemporaneous values. For examples of the former see Giannone *et al.* (2008) and Giannone *et al.* (2009). In this paper we propose following the latter approach, with the aim of assisting statistical agencies to construct timely preliminary estimates of the aggregate series. The aggregate series to be nowcast comprises of many disaggregates, some of which are 'known', i.e., the statistical agencies have reported data (which may be revised) on the components contemporaneously, and some are unknown, so data is missing at the current time. The proposed procedure involves a bridge-equation framework, first producing accurate estimates of the disaggregates and then nowcasts of the aggregate are calculated from the disaggregates.

The nowcast problem faces a 'ragged edge' at the nowcast origin, where some disaggregates have statistical releases for the current time, and some do not. The unknown disaggregates are 'forecast', taking account of the data on already reported disaggregates, and other higher-frequency indicators of the state of the economy that are usually available in a timely manner. The higher-frequency data are transformed to remove any unit-root non-stationarity, and to match the frequency of the aggregate. Given the large number of variables involved, automatic model selection offers a viable approach: see Doornik (2009) and Hendry and Doornik (2014). The general approach is flexible to allow for missing data on components to vary over time. Every disaggregate is 'forecast', including those that are already reported, as the contemporaneous forecast errors from the known disaggregates are informative for adjusting the forecasts of the unknown disaggregates. The bridge equation provides the nowcast of the aggregate using these transformed series.

As with forecasting, producing an accurate nowcast is difficult, more so in turbulent times. Despite the key difference noted above, many of the problems that confront forecasting also impinge on nowcasting: see Castle *et al.* (2017). There are a number of taxonomies of the sources of forecast errors: see for example, Clements and Hendry (1998), Clements and Hendry (2006), Hendry and Hubrich (2011) and Hendry and Mizon (2012). This article takes the more general forecast-error taxonomy for open systems of equations in Hendry and Mizon (2012), and adapts it to nowcasting, with extensions that reflect the differences between nowcasts and forecasts, especially contemporaneous information and the ragged edge problem for missing disaggregates. Our taxonomy is designed to provide a framework for thinking about potential problems facing the nowcaster in achieving useful nowcasts. It therefore considers a wide array of sources of nowcast errors, from estimation uncertainty to model mis-specification.

Importantly, the taxonomy incorporates unforeseen changes in parameters, and thus allows for a formal analysis into the consequences for nowcasting of structural breaks. Additionally, the taxonomy is applied to evaluate the impact of the different error sources on the expected nowcast error. This evaluation delivers seven insights into the sources of nowcast errors we discuss below. In particular, the analysis is focused on isolating sources of nowcast errors that cause nowcast failure. Nowcast failure occurs when nowcasts are significantly different from the eventually measured outcome, examples of which are shown in Ericsson (2017). We also record the variance components of the various error sources, but focus on the expected values of the 28 possible errors as these yield seven insights which we believe may be helpful to agencies producing nowcasts.

Section 2 describes the derivation of the nowcast-error taxonomy, and presents the resulting taxonomy table. Section 3 presents seven insights into potential sources of nowcast failure stemming from the taxonomy. Subsequently, Section 4 presents evidence from Monte Carlo simulations for a simplified setting designed to clarify the analysis, illustrated by an empirical example in Section 5. We conclude in Section 6.

2. Nowcast error taxonomy

This section derives the nowcast error taxonomy. In the most general set-up, the variable to be nowcast is a function of disaggregates and other exogenous information that are contemporaneously available, and those that are missing at the nowcast origin. Due to asynchronous release dates of economic data, nowcasting often involves unbalanced panels, here referred to as 'ragged edges'. The missing disaggregates must be forecast in order to avoid ragged edges. The way of handling missing information may vary between different approaches to nowcasting. The derivation of the nowcast error taxonomy is based on a data-generating process (DGP) that is a function of two strongly exogenous vectors, x_t and z_t, which are stationary processes in sample of dimensions $(N_1 \times 1)$ and $(N_2 \times 1)$ respectively ([4]). While x_t is contemporaneously available, z_t includes variables that are missing at the nowcast origin and have to be filled in. Since different variables may be missing at different nowcast origins, the dimensions N_1 and N_2 should be interpreted as time-variant. The DGP thus takes the following form in sample for $t = 1, \ldots, T$, where lags have been omitted for clarity:

$$(1) \quad y_t = \tau + \lambda_1' x_t + \lambda_2' z_t + \epsilon_t = \phi + \lambda_1' (x_t - \rho_1) + \lambda_2' (z_t - \rho_2) + \epsilon_t ,$$

where $\epsilon_t \sim N[0, \sigma_\epsilon^2]$ is the innovation error and $E[\epsilon_t | x_1, \ldots, x_t; z_1, \ldots, z_t] = 0$ is assumed. Further, it holds that $E[y_t] = \phi$, $E[x_t] = \rho_1$ and $E[z_t] = \rho_2$. This gives the relationship

$$(2) \quad \phi = (\tau + \lambda_1' \rho_1 + \lambda_2' \rho_2).$$

Since the DGP is unknown in practice, the researcher may end up falsely including irrelevant variables in the nowcasting model, for example due to their retention in model selection. Thus, suppose equation (1) has been selected in sample over $t = 1, \ldots, T$, starting from a general unrestricted model (GUM) that also includes the vector w_t of irrelevant and strongly exogenous

([4]) The assumption of strong exogeneity is introduced to limit dependencies between mean zero results in the nowcast error taxonomy.

variables, which are assumed to be uncorrelated with ϵ_t. The vector is retained, though its true population parameter, λ_3, equals a vector of zeros. This model is then estimated in-sample, and used to nowcast $T+1$. Since z_{T+1} is missing at the nowcast origin it has to be forecast. The nowcasted value of y_t may be written as

$$(3)\quad \hat{y}_{T+1|T+1} = \hat{\phi} + \hat{\lambda}_1'(x_{T+1} - \hat{\rho}_1) + \hat{\lambda}_2'(\hat{z}_{T+1|T} - \hat{\rho}_2) + \hat{\lambda}_3'(w_{T+1} - \hat{\rho}_3).$$

We consider a situation in which there has been an unanticipated and permanent shift in the DGP of y_t between the nowcast origin, T, and period $T+1$. Allowing for shifts in all terms, the post-shift DGP is:

$$y_{T+h} = \phi^* + \hat{\lambda}_1^{*\prime}(x_{T+h} - \rho_1^*) + \hat{\lambda}_2^{*\prime}(z_{T+h} - \rho_2^*) + \epsilon_{T+h} \quad \text{with } h \geq 1.$$

While mean shifts in the irrelevant variables are incorporated, $\rho_3^* \neq \rho_3$, it is assumed that the vector w_t remains irrelevant following the unanticipated shift, so $\lambda_3^* = \lambda_3 = 0$.

In the nowcast error taxonomy, we also account for estimation uncertainty and for the effect of model selection on expected parameter estimates labelled search bias. Further, we allow for model mis-specification, for example due to omitted variables or in-sample location shifts that were not modelled. The first two factors lead to biased parameter estimates, while the latter two in general entail biased and inconsistent parameter estimates. This is captured by indexing the expected values of the estimators with e, and incorporating that the expected values may not equal the true parameter values provided in the DGP (1), e.g. $\phi \neq \phi^e$.

The expression of the nowcast error can be derived by subtracting (3) from (1), and rearranging terms:

$$\hat{\epsilon}_{T+1|T+1} = y_{T+1} - \hat{y}_{T+1|T+1} = \epsilon_{T+1} + (\phi^* - \hat{\phi}) + \lambda_1^{*\prime}(x_{T+1} - \rho_1^*) - \hat{\lambda}_1'(x_{T+1} - \hat{\rho}_1)$$

$$+ \lambda_2^{*\prime}(z_{T+1} - \rho_2^*) - \hat{\lambda}_2'(\hat{z}_{T+1|T} - \hat{\rho}_2) - \hat{\lambda}_3'(w_{T+1} - \hat{\rho}_3)$$

The key expansion to arrive at the full nowcast error taxonomy is the following:

$$\phi^* - \hat{\phi} = (\phi^* - \phi) + (\phi - \phi^e) + (\phi^e - \hat{\phi}).$$

This expansion is used for all parameters to re-write the nowcast error in terms of separate components for shifts, mis-specification and estimation uncertainty to yield the full nowcast error displayed in Table 1. Because of the large number of terms in Table 1, we will gradually introduce complications starting with the simplest case in equation (4). As noted above, the explanation here focuses on the expected values of the mistakes and shifts, although direct variance terms of individual components are included in the table ([5]).

([5]) For reasons of parsimony we abstract from any covariance terms between the components.

Table 1: Nowcast error taxonomy

	$\hat{\epsilon}_{T+1\mid T+1}=$	Description	Expectation	Variance
(i)	ϵ_{T+1}	Innovation error	0	σ_{ϵ}^{2}
(iia)	$+(\phi^{*}-\phi)$	Equation mean shift	$+(\phi^{*}-\phi)$	0
(iib)	$+(\phi-\phi^{e})$	Equation mis-specification	$+(\phi-\phi^{e})$	0
(iic)	$+(\phi^{e}-\hat{\phi})$	Equation mean mis-estimation	0	$O_p(T^{-1})$
(iiia)	$-\lambda_1'(\rho_1^{*}-\rho_1)$	Mean shift	$-\lambda_1'(\rho_1^{*}-\rho_1)$	0
(iiib)	$+(\lambda_1^{*}-\lambda_1)'(x_{T+1}-\rho_1^{*})$	Slope shift	0	$(\lambda_1^{*}-\lambda_1)'V[x_{T+1}](\lambda_1^{*}-\lambda_1)$
(iiic)	$+\lambda_1^{e\prime}(\rho_1^{e}-\rho_1)$	Mean mis-specification	$\lambda_1^{e\prime}(\rho_1^{e}-\rho_1)$	0
(iiid)	$+(\lambda_1-\lambda_1^{e})'(x_{T+1}-\rho_1^{*})$	Slope mis-specification	0	$(\lambda_1-\lambda_1^{e})'V[x_{T+1}](\lambda_1-\lambda_1^{e})$
(iiie)	$-\lambda_1^{e\prime}(\rho_1^{e}-\hat{\rho}_1)$	Mean mis-estimation	0	$O_p(T^{-1})$
(iiif)	$+(\lambda_1^{e}-\hat{\lambda}_1)'(x_{T+1}-\rho_1^{*})$	Slope mis-estimation	$O_p(T^{-1/2})$	$O_p(T^{-1/2})$
(iiig)	$+(\lambda_1-\hat{\lambda}_1)'(\rho_1^{*}-\rho_1)$	Mean shift covariance	$(\lambda_1-\lambda_1^{e})'(\rho_1^{*}-\rho_1)$	$O_p(T^{-1})$
(iiih)	$+(\hat{\lambda}_1-\lambda_1^{e})'(\hat{\rho}_1-\rho_1)$	Estimation covariance	$O_p(T^{-1})$	$O_p(T^{-1})$
(iva)	$-\lambda_2'(\rho_2^{*}-\rho_2)$	Mean shift	$-\lambda_2'(\rho_2^{*}-\rho_2)$	0
(ivb)	$+(\lambda_2^{*}-\lambda_2)'(z_{T+1}-\rho_2^{*})$	Slope shift	0	$(\lambda_2^{*}-\lambda_2)'V[z_{T+1}](\lambda_2^{*}-\lambda_2)$
(ivc)	$+\lambda_2^{e\prime}(\rho_2^{e}-\rho_2)$	Mean mis-specification	$\lambda_2^{e\prime}(\rho_2^{e}-\rho_2)$	0
(ivd)	$+(\lambda_2-\hat{\lambda}_2)'(z_{T+1}-\rho_2^{*})$	Slope mis-specification	$O_p(T^{-1/2})$	$V[(\lambda_2-\hat{\lambda}_2)'(z_{T+1}-\rho_2^{*})]$
(ive)	$-\lambda_2'(z_{T+1}-E_{T+1}(\hat{z}_{T+1\mid T}))$	Mean mis-forecast	$\lambda_2'(\rho_2^{*}-E_{T+1}[\hat{z}_{T+1\mid T}])$	$(\lambda_2)'V[z_{T+1}](\lambda_2)$
(ivf)	$-\lambda_2^{e\prime}(\rho_2^{e}-\hat{\rho}_2)$	Mean mis-estimation	0	$O_p(T^{-1})$
(ivg)	$+(\lambda_2-\hat{\lambda}_2)'(\rho_2^{*}-\rho_2)$	Mean shift covariance	$(\lambda_2-\lambda_2^{e})'(\rho_2^{*}-\rho_2)$	$O_p(T^{-1})$
(ivh)	$-\lambda_2'(E_{T+1}(\hat{z}_{T+1\mid T})-\hat{z}_{T+1\mid T})$	Mis-forecast	0	$(\lambda_2)'[\hat{z}_{T+1\mid T}](\lambda_2)$
(ivi)	$+(\hat{\lambda}_2-\lambda_2)'(z_{T+1}-\hat{z}_{T+1\mid T})$	Mis-forecast covariance	$E[(\hat{\lambda}_2-\lambda_2)'(z_{T+1}-\hat{z}_{T+1\mid T})]$	$V[(\hat{\lambda}_2-\lambda_2)'(z_{T+1}-\hat{z}_{T+1\mid T})]$
(ivj)	$+(\hat{\lambda}_2-\lambda_2^{e})'(\hat{\rho}_2-\rho_2)$	Estimation covariance	$O_p(T^{-1})$	$O_p(T^{-1})$
(va)	$+\lambda_3^{e\prime}(\rho_3^{e}-\rho_3)$	Mean mis-specification	$\lambda_3^{e\prime}(\rho_3^{e}-\rho_3)$	0
(vb)	$-\lambda_3^{e\prime}(w_{T+1}-\rho_3^{*})$	Slope mis-specification	0	$(\lambda_3^{e})'V[w_{T+1}](\lambda_3^{e})$
(vc)	$-\lambda_3^{e\prime}(\rho_3^{e}-\hat{\rho}_3)$	Mean mis-estimation	0	$O_p(T^{-1})$
(vd)	$-\hat{\lambda}_3'(\rho_3^{*}-\rho_3)$	Mean shift covariance	$-\lambda_3^{e\prime}(\rho_3^{*}-\rho_3)$	$O_p(T^{-1})$
(ve)	$+(\lambda_3^{e}-\hat{\lambda}_3)'(w_{T+1}-\rho_3^{*})$	Slope mis-estimation	$O_p(T^{-1/2})$	$O_p(T^{-1/2})$
(vf)	$+(\hat{\lambda}_3-\lambda_3^{e})'(\hat{\rho}_3-\rho_3)$	Estimation covariance	$O_p(T^{-1})$	$O_p(T^{-1})$

3. Sources of nowcast errors

From this nowcast error taxonomy seven insights into sources for nowcast failure can be derived.

3.1. The exogenous vector: insights 1 and 2

We first focus on the exogenous vector that is contemporaneously available, x_t, and assume that the DGP is a function of this vector only. We therefore effectively set $\lambda_2=\lambda_2^*=0$, and ignore the terms (iva)-(vf) in Table 1. Hence, the nowcast error includes the components (i)-(iiih). If we additionally abstract from search bias and mis-specification so that $\lambda_1^e=\lambda_1$ and $\rho_1^e=\rho_1$, and ignore estimation covariances (6), the terms (i), (iia), (iic), (iiia,b) in Table 1 are the only sources of the nowcast error. Using definition (2), and adding and subtracting the product $\lambda_1'\rho_1^*$, the expected nowcast error thus reduces to:

$$(4)\quad E[\hat{\epsilon}_{T+1|T+1}] \approx (\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*+\lambda_1'(\rho_1^*-\rho_1)-\lambda_1'(\rho_1^*-\rho_1)=(\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*.$$

If the strongly exogenous variable is omitted the nowcast error becomes:

$$\tilde{\epsilon}_{T+1|T+1} \approx (i),(iia),(iic)+\lambda_1^{*\prime}(x_{T+1}-\rho_1^*),$$

with expectation:

$$(5)\quad E[\tilde{\epsilon}_{T+1|T+1}] \approx (\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*+\lambda_1'(\rho_1^*-\rho_1).$$

The *first insight* is that a change in dynamics of the unmodelled exogenous variable, $\lambda_1\neq\lambda_1^*$, alone induces nowcast failure as long as its mean, ρ_1, is different from 0. This nowcast failure is reflected in the expectation of term (iia). Comparing $E[\hat{\epsilon}_{T+1|T+1}]$ with $E[\tilde{\epsilon}_{T+1|T+1}]$ in (4) and (5) gives rise to the *second insight* : incorrectly omitting x_t does not lead to or augment nowcast failure if its mean remains constant, so $\rho_1^*=\rho_1$. Together these first two insights imply that if the mean of the exogenous variable remains constant at 0, the size of the nowcast failure as given by the expectation of the nowcast error is independent of whether or not x_{T+1} is correctly modelled (7). By comparing the above two expectations it can also be seen that if a mean shift $\rho_1^*\neq\rho_1$ takes place, then the expected nowcast error is minimised by correctly including x_{T+1}.

3.2. Filling in ragged edges: insights 3 and 4

With these insights in mind, we include exogenous variables in z_{T+1}, and work with the full DGP (1). Thus, we consider terms (i)-(ivj) in Table 1 as potential sources of the nowcast error. Absent mis-specification and estimation covariances, we can focus attention on terms (i),(iia),(iic),(iiia,b),(iva,b),(ive), and the expectation of the nowcast error equals:

$$(6)\quad E[\hat{\epsilon}_{T+1|T+1}] \approx (\tau^*-\tau)+(\lambda_1^*-\lambda_1)'\rho_1^*+(\lambda_2^*-\lambda_2)'\rho_2^*+\lambda_2'(\rho_2^*-E_{T+1}[\hat{z}_{T+1|T}]).$$

Since we are being agnostic to the method of infilling ragged edges, we allow for the case $\rho_2^*\neq E_{T+1}[\hat{z}_{T+1|T}]$. If z_{T+1} were to be omitted the following nowcast error results:

$$\tilde{\epsilon}_{T+1|T+1} \approx (i),(iia),(iic),(iiia)+\lambda_2^{*\prime}(z_{T+1}-\rho_2^*)$$

(6) The latter can be justified with the argument that in a congruent model, the estimation uncertainty is of order $T^{-1/2}$, and at a sample size of eg. $T = 100$ may be ignored relative to other sources of bias and variance.
(7) The squared nowcast error, however, is increased by the $O_p(1)$ term $(\lambda_1^*)'V[x_{T+1}](\lambda_1^*)$ if x_{T+1} is omitted.

with expectation

$$(7)\ E[\tilde{\epsilon}_{T+1|T+1}] = (\tau^* - \tau) + (\lambda_1^* - \lambda_1)'\rho_1^* + (\lambda_2^* - \lambda_2)'\rho_2^* + \lambda'_2(\rho_2^* - \rho_2).$$

A comparison of the two expectations in (6) and (7) yields the *third insight* : If there are no mean shifts in the exogenous vector z_{T+1}, so $\rho_2^* = \rho_2$, there is no gain in mean nowcast accuracy from accurately forecasting the exogenous variable. Equally, accurately forecasting z_{T+1} does not reduce the nowcast error if the mean remains constant but the slope shifts.

The *fourth insight* results by considering a mean shift in z_{T+1}: If there is a location shift in the missing variable, the forecast of z_{T+1} has to be closer to the new mean, ρ_2^*, than the old mean, ρ_2, in order to reduce the expected nowcast error relative to omitting the missing variable.

$$E[\tilde{\epsilon}_{T+1|T+1}] - E[\hat{\epsilon}_{T+1|T+1}] \approx \lambda'_2(\rho_2^* - \rho_2) - \lambda'_2(\rho_2^* - E_{T+1}[\hat{z}_{T+1|T}]).$$

This highlights how forecast failure at the stage of infilling missing variables at the nowcast origin may affect the final nowcast. This also stresses the importance of updating forecasts of missing disaggregates during times of structural breaks, which may imply an unexpected location shift, and hence systematic forecast failure.

The approximation signs reflect that mis-specification, model selection, and estimation covariances were ignored in the preceding analysis. Mis-specification impacts on the expected nowcast error through (iiic) and (ivc), and can only be avoided when omitting z_{T+1} if x_{T+1} and z_{T+1} are orthogonal. Both mis-specification and estimation covariance are functions of all regressors in the model, and their magnitudes relative to unexpected shifts are difficult to compare analytically. The simulations, however, show that if a congruent nowcasting model capturing in-sample shifts has been selected, mis-specification, and estimation covariance seem to be less detrimental for nowcasting accuracy than unexpected shifts.

3.3. Adding model selection: insights 5, 6, and 7

So far the discussion has ignored model selection. In practice, however, this is an important step, and we therefore consider the full nowcast error taxonomy to evaluate the impact of model selection on nowcasting. As statistical estimation entails non-degenerate null distributions, there is a non-zero probability of retaining irrelevant variables. This is labelled a cost of search.

The *fifth insight* refers to the fact that absent location shifts, falsely retaining w_{T+1} only impacts on the expected nowcast error to the extent that the expected values of the estimated parameters differ from the true values of 0 in (va). Again, the bias introduced by mis-specification is difficult to quantify as it depends on all regressors, and may be negligible. A high false retention rate, however, is not without cost. Component (vb) directly increases the squared nowcast error.

Variables are frequently subject to shifts. This motivates the *sixth insight* that location shifts in any retained irrelevant variables lead to systematic nowcast failure through (vb). Of course, variable selection may result in the omission of relevant variables, too. The impact of omitting x_{T+1} and z_{T+1} has already been presented.

Model selection also affects parameter estimates. All statistics for selecting variables to be kept in the final model have interdependent distributions, which differ under the null and the alternative, and are affected by each modelling decision. Thus, model selection impacts on the expected values of parameter estimates. This effect is termed search bias. Hendry and Krolzig (2005) show that search bias is negligible for highly relevant variables, and positive for the squared magnitude of parameter estimates of irrelevant variables. This provides the *seventh* *insight* that correcting parameter estimates for search bias, for example by using the two-step procedure introduced in Hendry and Krolzig (2005), has the strongest impact on the expected nowcast error through the reduction of term (vd): By eliminating the bias in the parameter estimate of w_{T+1}, correcting for search bias reduces systematic nowcast failure following mean shifts in the falsely retained variable. In addition, it may reduce the expected nowcast error through the terms (iib), (iiic,g), (ivc,g) and (va).

In summary, it may be stated that:

- Absent structural breaks, omitting relevant exogenous vectors, x_{T+1} and z_{T+1}, does not cause nowcast failure in expectation, while retained irrelevant variables in w_{T+1} do so to the extent that there is mean mis-estimation.
- With location shifts, retention of relevant exogenous variables minimises the expected nowcast error, while forecasts of missing disaggregates must reflect any location shifts in real time in order to improve nowcast accuracy. Further, with non-zero means of x_{T+1} and z_{T+1}, slope shifts lead to nowcast failure, which is not attenuated by correctly including the relevant vectors. Equally, retained irrelevant variables may lead to nowcast failure if they shift out of sample.
- Overall, the impact of model selection, which may lead to omission of relevant or retention of irrelevant variables, on the expected nowcast error is most pronounced in the face of structural breaks. As a result, the most important mean accuracy gain due to search bias correction stems from reducing parameter estimates of irrelevant variables, which are likely to be marginally significant, if these variables break out of sample.

4. Simulations

The nowcast error taxonomy highlights the costs associated with various sources of nowcast errors. The simulations complement the taxonomy by consecutively adding ragged edges, model selection, and structural breaks to the simulation design. The simulations abstract from the practical issue of the mismatch between frequencies of the variable to be nowcast and potential explanatory variables. We assume that unit roots in the data have been removed. Across simulations, *Autometrics* (see Doornik (2009) and Hendry and Doornik (2014) for a description of the algorithm) is used for variable selection at the significance level $\alpha = 0.05$, and the intercept term is always retained.

First consider the 'first-best' scenario, in which the researcher has perfect knowledge on the DGPs of the variable to be nowcast, y_t, and all relevant disaggregates are contemporaneously available. Thus, we use the DGP of y_t as the nowcasting model, and exclude ragged edges. Subsequently, we introduce ragged edges, which are filled in using the DGP of missing disaggregates as the forecasting model. Consequently, we are ignoring estimation uncertainty concerning the first stage of the bridge equation framework. This simplication has been introduced since the focus of the taxonomy lies on the second stage of the bridge equation framework.

Building on this, we consider the impact of model selection on nowcast accuracy. First, we introduce model selection starting from the DGP of y_t as the GUM including an intercept term. This isolates omission of relevant variables, which we refer to as costs of inference, from retention of relevant variables. Second, we add irrelevant variables to variable selection, but we force the retention of relevant variables. This means that we select over irrelevant variables only, and consequently consider separately costs of search. Third, we search jointly over all available variables. This replicates that the distinction between relevant and irrelevant variables is not known in practice while it does not allow to differentiate between costs of search and costs of inference.

Last, we add structural breaks in two designs. In the first design all variables are affected by identical breaking dynamics, and in the second, only irrelevant variables break. Table 2 gives an overview of the different simulation designs.

Table 2: Overview of simulation designs

Label	DGP	Ragged edges	GUM 1	GUM 2	GUM 3
Features	**Variable Selection**				
	No	No	Yes	Yes	Yes
	The DGP of y_t (equation 9) is used as the nowcasting model	The DGP of y_t is used as the nowcasting model	Variable selection starts from the DGP of y_t.	Variable selection starts from GUM including $x_1 - x_7$ $x_1 - x_3$ are forced into the final nowcasting model.	Variable selection starts from GUM including $x_1 - x_7$ We select over $x_1 - x_7$
	Missing disaggregates				
	No	Yes			
	All disaggregates are available	Ragged edges in disaggregates according to Table 4. DGP (equation 10) of disaggregates used as forecasting models.			
	Structural breaks				
	a) No location shifts				
	b) $x_1 - x_7$ shift simultaneously and identically				
	c) Irrelevant variables $x_4 - x_7$ shift simultaneously and identically.				

To evaluate the predictive performance of the nowcasts, we consider two statistics: the mean forecast error (MFE), and the root mean squared forecast error (RMSFE) based on the nowcast error in the nowcasting equation of y_t. The MFE is defined as:

$$\frac{1}{M}\sum_{i=1}^{M}\frac{1}{H}\sum_{h=1}^{H}(y_{T+h}-\hat{y}_{T+h|T+h}).$$

The MFE averages the nowcast errors over the nowcast periods, $h = 1, \ldots, H$, and over the M replications performed. If nowcasts are not systematically biased, the MFE should not be statistically significantly different from 0. The RMSFE is defined as:

$$\text{(8)}\quad \frac{1}{M}\sum_{i=1}^{M}\sqrt{\left[\frac{1}{H}\sum_{h=1}^{H}(y_{T+h}-\hat{y}_{T+h|T+h})^2\right]}.$$

It measures the variation of the nowcasted values around the true values averaged over the M replications performed, so that smaller RMSFEs are preferred.

4.1. Formulating the data-generating processes

Here, we introduce the simple DGP for y_t, which may represent an economic aggregate such as GDP. We suppose there is data on seven time series, $x_{i,t}$, $i = 1, \ldots, 7$ available. Out of the monthly time series, the first three are relevant for y_t, and enter the DGP of the aggregate with equal weights. Variables $x_{4,t} - x_{7,t}$ are noise, which are introduced to make variable selection relevant. While the number of irrelevant variables remains small, the design implies that there are more irrelevant than relevant variables. Given the detailed data available to statistical offices this is deemed a realistic set-up. The time series may be interpreted as monthly disaggregated data or leading indicators for y_t. The DGP of y_t takes the following form:

(9) $y_t = 0.5x_{1,t} + 0.5x_{2,t} + 0.5x_{3,t} + \epsilon_t$ with $\epsilon_t \sim IN[0, 1]$ for $t = 0, \ldots, T + h$.

T refers to the number of in-sample periods, which are used to fit the model to the data. h specifies the number of nowcasting periods.

The DGP for the disaggregated time series is specified as a VAR

(10) $x_t = \pi_0 + \pi_1 x_{t-1} + \delta 1_{7,t>T} + v_t$ with $v_t \sim IN_7[0, \Omega_v]$ for $t = 0, \ldots, T + h$

where $x_t = [x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}, x_{5,t}, x_{6,t}, x_{7,t}]'$. π_1 is a (7×7) matrix, and set to equal a diagonal matrix $\pi_1 = \pi I_7$. We assume that $|\pi| < 1$. The intercept π_0 and error term v_t are (7×1) vectors. $1_{7,t>T}$ denotes an indicator variable that enters the DGP for the periods $t > T$ and hence represents an identical location shift across the 7 disaggregates for the out-of-sample periods, $h = 1, 2, 3$. The (7×1) vector δ specifies which disaggregates are affected by the location shift. In the simulations, the initial values for the disaggregates are set equal to zero, $x_0 = 0$. To reduce the dependence on the initial values of the simulated data, we discard the first 20 simulated observations.

Table 3: Parameter values

Number of simulations M	10 000
In-sample period T	75
Nowcasting horizon h	3
Non-centralities of $x_1 - x_3$ in DGP of y_t	3.8
Slope parameter π	0.6
Intercept π_0	0.1
Correlation ρ	0.6
Break coefficient: All δ	(1, 1, 1, 1, 1, 1, 1)
Break coefficient: Irrelevant δ	(0, 0, 0, 1, 1, 1, 1)
Magnitude of break	$4\sigma_x = \frac{4}{\sqrt{(1-0.6^2)}} = 5$

The non-centralities of x_1 - x_3 were computed by averaging their t-statistics computed in the DGP of y_t over all replications.
Source: Authors' calculations.

The correlation structure between the disaggregated time series $x_{i,t}$, $i = 1, \ldots, 7$ must be specified. The 7×7 variance-covariance matrix between the regressors, Σ_x, is induced through the variance-covariance matrix of the error term, Ω_v. The covariance between the error terms entering the DGP of all regressors, ρ, is symmetric, and non-zero. We can derive the correlation structure between regressors based on Ω_v. Since π_1 is a diagonal and symmetric matrix, Σ_X, simplifies to [8]:

$$\Sigma_X = \frac{1}{1-\pi^2}\Omega_v .$$

From the unit variance of the error terms it follows that contemporaneous correlations equal the covariances in this example. Table 3 summarizes the numerical parameter values in the simulations.

4.2. Formulating the general unrestricted models

We consider an in-sample period of length $T = 75$, which is equivalent to over six years of monthly data, and focus on three nowcasting horizons $h = 3$. If ragged edges are incorporated, data on one of the disaggregates is treated as missing in each nowcasting horizon, introducing a ragged-edge dataset as presented in Table 4.

Table 4: Ragged-edge structure

Nowcasting period	x1	x2	x3
$h = 1$	X	✓	✓
$h = 2$	✓	X	✓
$h = 3$	✓	✓	X

Note: ✓ indicates contemporaneous availability. X stands for missing values.

The forecasting models are given by the DGP of the disaggregates (equation 10). We consider three GUMs for model selection. The GUM refers to the most general specification that is the starting point of selection of the final nowcasting model, and should summarize the actual DGP of the variable to be modelled in the space of variables under consideration. The GUMs are chosen to illustrate the different costs associated with model selection on nowcast accuracy. First, we start model selection from the DGP of y_t, equation (9), including an intercept term. We refer to this set-up as 'GUM 1'. Subsequently, we add irrelevant variables. In the set-ups 'GUM 2' and 'GUM 3', the starting point for model selections is:

$$y_t = \beta_0 + \sum_{i=1}^{7} \beta_i x_{i,t} + U_t \qquad \text{for } t = 1, \ldots, T.$$

In scenario GUM 2, the nowcasting model is selected using automatic model selection over disaggregates $x_4 - x_7$ for the in-sample period, $t = 1, \ldots, T$. The nowcast of y_t is computed from the selected model, where $\hat{\beta}_i$ denotes the parameter estimates obtained in-sample, and k_1 is the number of retained variables after variable selection.

$$\hat{y}_{T+h|T+h} = \hat{\beta}_0 + \hat{\beta}_1 x_{1,T+h} + \hat{\beta}_2 x_{2,T+h} + \hat{\beta}_3 x_{3,T+h} + \sum_{k_1} \hat{\beta}_i x_{i,T+h} \qquad \text{for } h = 1, 2, 3 \text{ and } i = 4, \ldots, 7$$

In GUM 3, automatic model selection considers all disaggregates $x_1 - x_7$ over the in-sample period, $t = 1, \ldots, T$. The nowcast of y_t is computed from the selected model with k_2 denoting the number of retained variables after considering all variables for selection:

[8] This assumes that the disaggregates are *I(0)* or that Π_1 has all its eigenvalues within the unit circle.

$$\hat{y}_{T+h|T+h} = \hat{\beta}_0 + \sum_{k_2} \hat{\beta}_i x_{i,T+h} \qquad \text{for } h = 1, 2, 3 \text{ and } i = 1, \ldots, 7$$

4.3. Simulation results

The discussion of the results is structured into four parts, of which the first three refer to the diferent breaking dynamics in Table 2. We begin by evaluating simulations without location shifts. We then discuss simulations with location shifts affecting all disaggregates, and end with an analysis of simulations with location shifts occurring to irrelevant disaggregates $x_{4,t} - x_{7,t}$ only. Additionally, we consider the impact of correcting parameter estimates for search in the last subsection. The numerical simulation results can be found in Table 5, where we distinguish between simulations with or without bias correction. They are illustrated in barplots throughout the analysis.

Table 5: Simulation Results

	DGP	Ragged edges	GUM 1		GUM 2		GUM 3	
Bias correction	**No**	**No**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**
No location shifts								
MFE	0.004	–0.0003	–0.003	0.017	–0.003	0.008	–0.003	0.018
RMSFE	0.922	1.033	1.066	1.091	1.064	1.073	1.077	1.097
Location shifts in $x_1 - x_7$								
MFE	0.004	2.500	2.545	3.085	2.493	2.769	2.510	3.043
RMSFE	0.922	2.667	2.862	3.392	2.786	3.041	2.863	3.371
Location shifts in $x_4 - x_7$								
MFE	0.004	–0.0003	–0.0035	0.017	–0.007	0.006	–0.153	–0.118
RMSFE	0.922	1.033	1.066	1.091	1.365	1.217	1.477	1.330

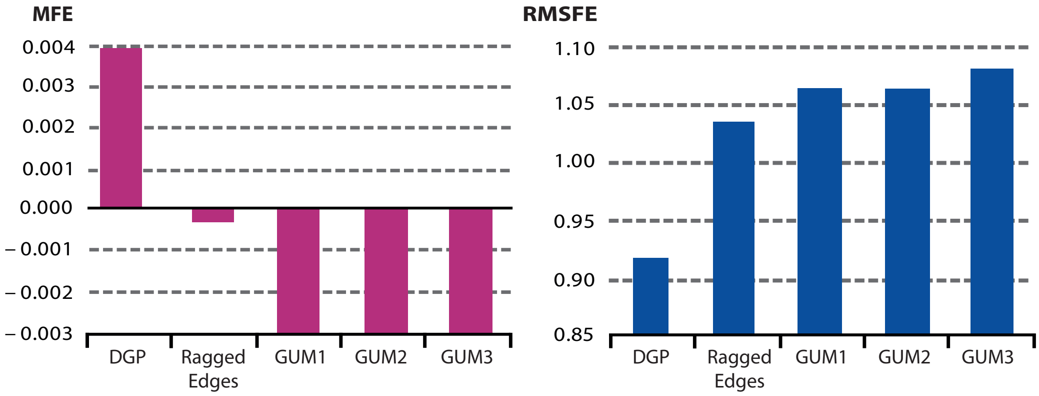
The line 'Bias correction' indicates whether parameters have (Yes) or have not (No) been corrected for search.
Source: Authors' calculations.

4.3.1. WITHOUT STRUCTURAL BREAKS

Across all simulations, the MFE averages to a number close to 0, confirming that there is no systematic nowcast failure as seen in Figure 1. Since their DGP is used to fill in missing disaggregates, and since they do not shift out of sample, their forecasts are correct on average, and the incorporation of ragged edges does not worsen but slightly improves the fit of nowcasts. Equally, starting model selection from GUM 1, and hence introducing costs of inference as well as estimation uncertainty, does not significantly impact on mean nowcast accuracy. This confirms insights 1 and 3: absent location shifts, omitting relevant variables does not cause nowcast failure. Selecting variables in GUM 2 illustrates insight 5, and suggests that costs of search are low if there are no location shifts out of sample. The joint cost of inference, search, and estimation uncertainty as summarised by scenario GUM 3 remains low.

The RMSFEs in Figure 1 show that the magnitude of nowcast errors increases once ragged edges are included, and underline the comparatively small costs related to search and inference. Even if the researcher were able to know the true DGPs for both stages of the bridge equation framework, the fact that there is missing data that have to be forecast increases the RMSFE by around 12 %. Adding costs of inference by introducing model selection starting from GUM 1

Figure 1: MFE and RMSFE: No structural breaks



Source: Authors' calculations.

has a small impact on RMSFEs, which increase by an additional 3 %. Equally, costs of search have negligible effects on the RMSFE, both individually or jointly with costs of inference.

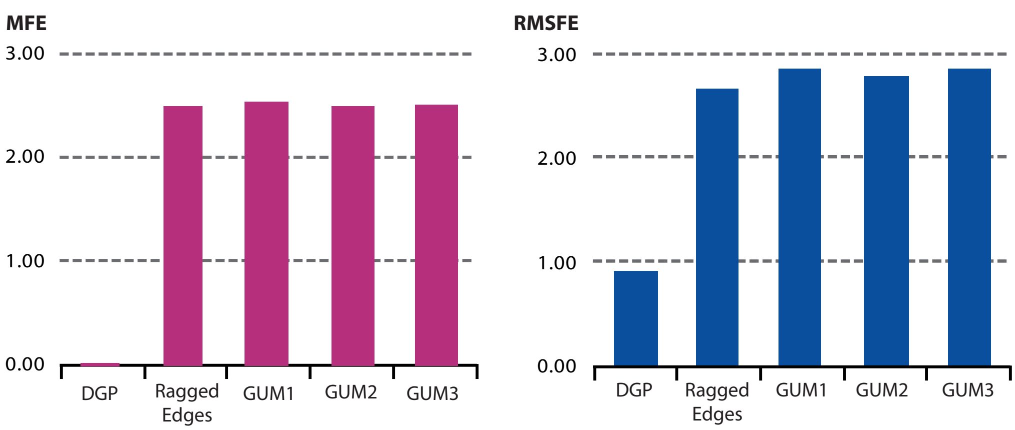
4.3.2. WITH STRUCTURAL BREAKS IN ALL DISAGGREGATES

The initial magnitude of the induced location shift is $\delta_i = 5$ for all disaggregates $i = 1, \ldots, 7$, see Table 3. Since the location shift is permanent, this implies an increase in the long-run mean from $E[x_t] = 0.1/(1-0.6) = 0.25$ for $t = 1, \ldots, 75$ *to* $E[x_t] = 5.1/(1-0.6) = 12.75$ out of sample. Since the in-sample DGP (equation 10) is used as the forecasting model, it is possible to derive that the expected mean forecast error of missing disaggregates equals the size of the location shift:

$$E[\hat{v}_{i,T+h|T+h-1}] = 0.1 + \delta_i + 0.6E[x_{i,T+h-1}] - (0.1 + 0.6E[x_{i,T+h-1}]) = 5. \quad (11)$$

From the DGP of y_t it then follows that the location shift in the aggregate variable amounts to the weighted average of the shifts in the relevant disaggregates, where the weights are

Figure 2: MFE and RMSFE: Location shifts in $x_1 - x_7$



Source: Authors' calculations.

given by the parameters of 0.5 in equation (9). Once ragged edges are included, the expected nowcast error of y_t equals $0.5 \times \delta_i = 2.5$ as a result of systematic failure in forecasting missing disaggregates. This systematic bias introduced by the location shift is indeed visible in the MFE and RMSFE in Figure 2, and is close to the analytical value of 2.5. This underlines insight 4: The forecasts of missing disaggregates has to be close to the post-shift mean to avoid nowcast failure.

As insights 1 and 2 suggest, it is optimal to retain relevant regressors if they shift out of sample. With respect to model selection, this suggests that in the face of structural breaks costs of inference, so the omission of relevant variables, should be particularly high. In this example, all three relevant variables have non-centralities of 3.8 as shown in Table 3. Given this high non-centrality, the retention of the relevant variables at $\alpha = 0.05$ is probable as can be gathered from Table 6, so that costs of inference remain low. Nevertheless, MFE and RMSFE increase in GUM 1 and GUM 3, which involve selection over relevant variables. Both statistics in GUM 2 are close to those in ragged edges in magnitude. As we are forcing all relevant variables, and the irrelevant disaggregates shift by the same amount, costs of search and estimation uncertainty remain of negligible importance for nowcast accuracy.

Overall, simulations with shifts in all disaggregates reveal the detrimental effect of location shifts for nowcast accuracy relative to costs of search, inference or estimation uncertainty.

Table 6: Rejection probabilities as a function of non-centralities, ψ, and selection significance levels, α, for T = 75

ψ	0	2	3	4	5
$\alpha = 0.001$	0.001	0.09	0.35	0.71	0.94
$\alpha = 0.01$	0.01	0.27	0.64	0.91	0.99
$\alpha = 0.05$	0.05	0.51	0.84	0.98	1.00
$\alpha = 0.16$	0.16	0.72	0.94	0.99	1.00

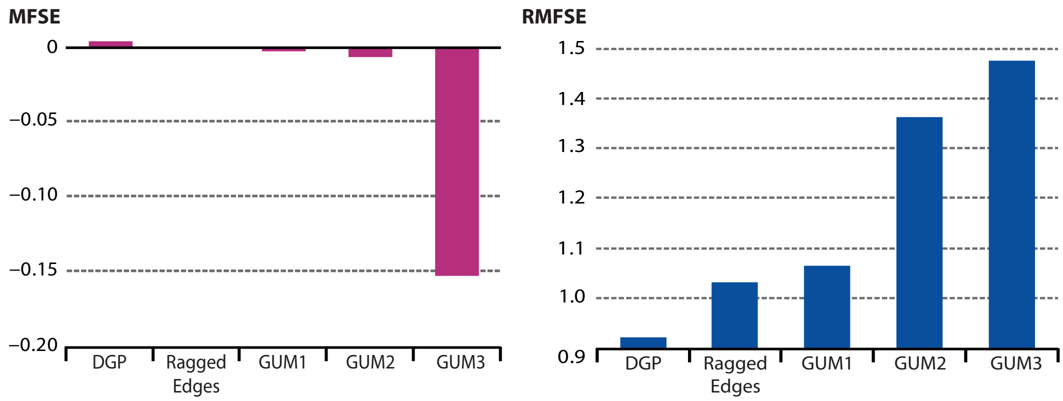
Source: Authors' calculations.

4.3.3. WITH STRUCTURAL BREAKS IN IRRELEVANT VARIABLES

So far, the cost of retaining irrelevant variables has been low. By introducing out-of-sample shifts in irrelevant variables, the distinction between relevant and irrelevant variables in model selection receives more practical importance, and costs of search are increased. Note that the simulated data for y_t, and disaggregates $x_1 - x_3$ are unchanged compared with simulations without location shifts. Consequently, simulation results of DGP, Ragged Edges, and GUM 1, which do not involve irrelevant variables, are identical to those in the first subsection. Figure 3 shows the relevant MFEs and RMSFEs.

Both GUM 2 and GUM 3 involve selection over irrelevant variables. In GUM 2, all relevant variables are forced into the final nowcasting model, and consequently nowcast accuracy does not show a systematic bias. The RMSFE, however, increases substantially, and indicates that magnitudes of nowcast errors have risen. In GUM 3, the example that is closest to model selection in practice, nowcast failure in terms of mean nowcast accuracy becomes apparent. Even in a set-up with a small number of irrelevant variables, so that almost no irrelevant variables will ever be retained in any replication, we can confirm insight 6 that a location shift in irrelevant variables may lead to systematic nowcast failure.

Figure 3: MFE and RMSFE: Location shifts in $x_4 - x_7$



Source: Authors' calculations.

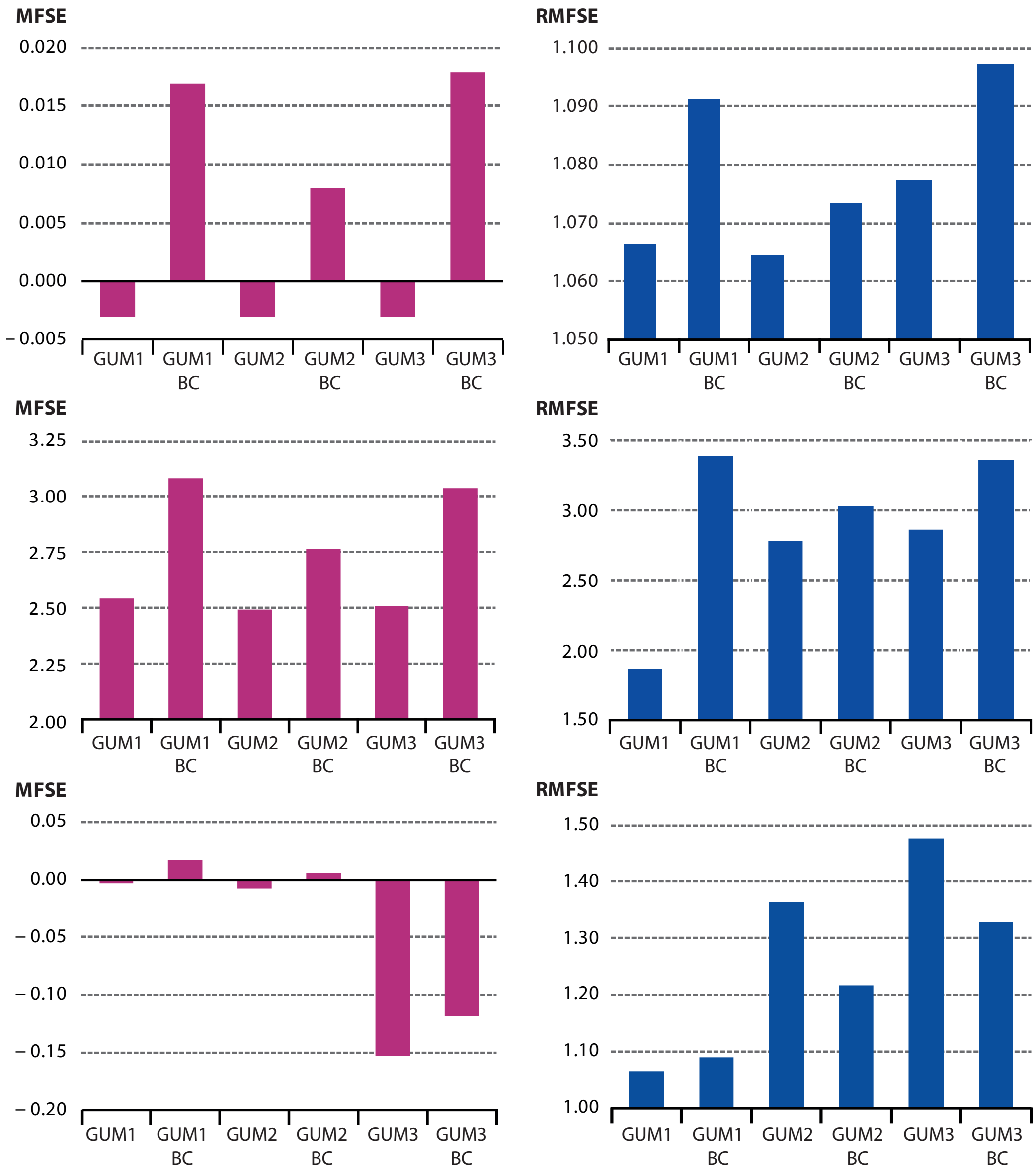
4.3.4. CONSIDERING BIAS CORRECTION

Hendry and Krolzig (2005) present a correction to reduce search bias in parameter estimates. The suggested procedure for bias correction takes into account that conditional on having been selected, the absolute values of parameter estates of irrelevant variables are upward biased. This procedure has been applied to parameter estimates in GUM 1 - GUM 3. Note that bias correction has only been applied to parameter estimates that are subject to search biases so any forced variables are not corrected (i.e., the intercept term and $x_1 - x_3$ in GUM 2).

Insight 7 states that correcting parameters for search is most beneficial if location shifts occur in genuinely irrelevant variables. Across the different breaking dynamics, we observe the potential benefit to be yielded from bias correction. Absent location shifts, correcting parameters for search uniformly worsens nowcast accuracy as displayed in plot (a) and (b) of Figure 4. At $\alpha = 0.05$, and with only 4 irrelevant variables, bias correction mostly acts on the coefficients of the relevant variables. Bias correction of relevant parameters, however, has been found to increase their mean squared errors. Moreover, given the identical DGPs for the disaggregates, an irrelevant variable may in fact function as a close substitute for a relevant one. Together, these arguments can motivate why bias correcting, and hence setting some parameter estimates to 0, is not found to improve nowcast accuracy in this set-up. This also applies to simulations in which relevant and irrelevant disaggregates break by identical amounts, since the latter remain good proxies for relevant disaggregates and search costs are small, as can be seen in plots (c) and (d) of Figure 4. Of course, very high frequency data may reveal that shifts have occurred in some aspects of the economy and allow adjustments thereto, perhaps using a robust device such as intercept correction: see Castle *et al.* (2017).

With different breaking dynamics, and in line with insight 7, bias correction improves nowcast accuracy in terms of MFE and RMSFE as illustrated in plots (e) and (f) of Figure 4. In GUM 2, its impact on the MFE remains marginal, while the RMSFE is lowered by 10 %. With selection over all variables in GUM 3, the impact of bias correction becomes most pronounced. The MFE falls by over 20 % in magnitude, while the RMSFE declines by 10 %. With a higher ratio of irrelevant to relevant variables, and a looser selection criterion the usefulness of bias correction is likely to be higher.

Figure 4: MFE and RMSFE: Bias correction (BC)



Source: Authors' calculations.

5. Nowcasting consumption expenditure

Having demonstrated the relevance of the seven insights from the nowcast taxonomy in a simulation exercise, it is of interest to consider how they are applicable in practice. For this purpose, we present an ex-post nowcast of growth in final consumption expenditure by households and non-profit institutions serving households (NPISH) over two horizons, 2008Q1 -2010Q4 and 2015Q1-2016Q3. For the nowcast of consumption expenditure growth, a small-scale model with two disaggregated series is used. The Office for National Statistics (ONS) uses three approaches to compute GDP: the output, income and expenditure approach. The expenditure approach is obtained from the sum of final consumption expenditure by households, NPISH and government on goods and services, gross capital formation, and net exports of goods and services. This additivity property can be replicated based on current price data on the components of the expenditure approach as published by the ONS. Consequently, the nowcasts of consumption expenditure could be combined with data on the other components of the expenditure approach, or their nowcasts, to arrive at a nowcast of GDP growth. This section commences with a presentation of the data in Section 5.1, and a discussion of seasonals and trends in Section 5.2 to motivate the data transformations used in this empirical example. Subsequently, forecasts of the two disaggregated series are provided in Sections 5.3 and 5.4, and the nowcasts of consumption expenditure growth are presented in Section 5.5. Section 5.6 interprets the empirical findings in light of the nowcast taxonomy.

5.1. The data

For the nowcast of growth in final consumption expenditure, $\Delta c_t = (C_t - C_{t-1})/C_{t-1}$, we consider two disaggregates: the index for retail trade (IR) and the number of newly registered passenger cars (CAR). IR is one of the two primary measures in the computation of final consumption expenditure (9). Approximately 40 % of average weekly household expenditure is on non-durable retail sales, and 4 % on car purchases as computed from ONS (2015), which may proxy for durable consumption expenditure. The index of retail sales and passenger car registrations become available with a lag of one month, and therefore need to be forecast to fill in ragged edges in the dataset. The forecasting models use four survey indicators on consumer confidence (CCI), retail trade confidence (RCI), service sector confidence (SCI) and economic sentiment (ESI). All data used in this empirical application are final estimates, so that we may abstract from measurement errors and revisions in the following analysis. The indicators are released at the end of the month that they refer to, and hence are treated as being available in real-time.

Table 7: Nowcasting horizons

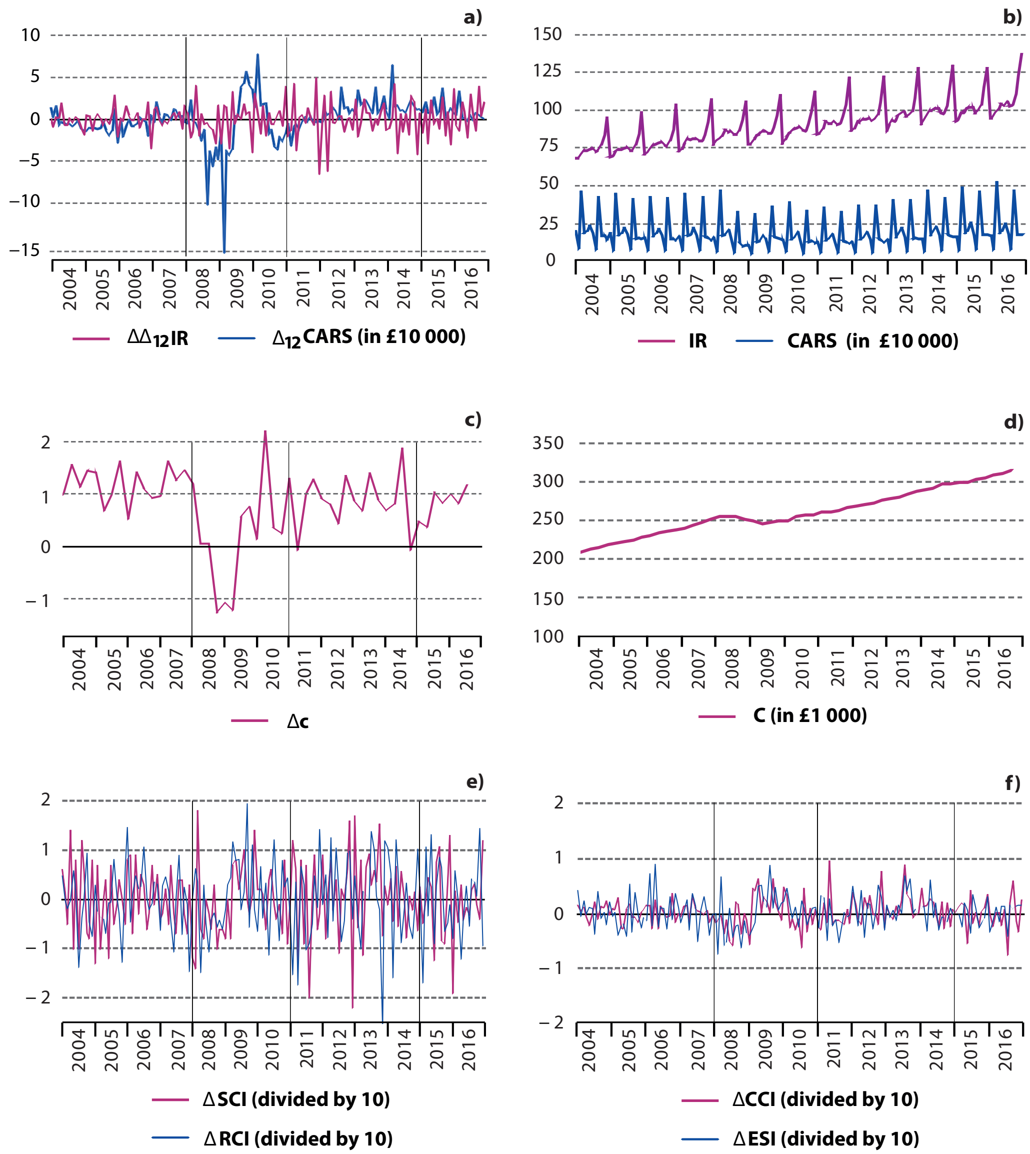
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Q1		$H1_{Q1}$	$H2_{Q1}$	$H3_{Q1}$								
Q2					$H1_{Q2}$	$H2_{Q2}$	$H3_{Q2}$					
Q3								$H1_{Q3}$	$H2_{Q3}$	$H3_{Q3}$		
Q4	$H3_{Q4}$										$H1_{Q4}$	$H2_{Q4}$

Subscripts refer to the respective quarter.

(9) See https://www.ons.gov.uk/economy/nationalaccounts/satelliteaccounts/qmis/consumertrendsqmi.

For the nowcasting period 2008Q1-2010Q4 we consider the in-sample period 2004Q1-2007Q4. In 2008-2009 the UK economy was subject to two large economic shocks, the deterioration in the functioning of financial markets and a fall in international trade, and went into its worst recession since the Great Depression, see Millard (2015). The recession is visible in the plot of consumption growth in panel (c) of Figure 5. The performance of the nowcasting framework during this turbulent period is then compared to nowcasts over the more stable quarters 2015Q1-2016Q3 with in-sample period 2004Q1-2014Q4. We consider three nowcasting origins,

Figure 5: Data series over the period 01/2004–09/2016



Notes: Δc: growth rate of C.
Source: Authors' calculations.

H1-H3, per quarter to be nowcast according to the structure in Table 7. The three nowcast horizons have been chosen to examine the impact of the accumulation of information on retail sales and passenger car registrations throughout the quarter to be nowcast. The first nowcast is estimated at *H1*, the second month of the reference quarter. At this point, data on retail sales and car registrations for the first month of the reference quarter are available, but ragged edges have to be filled in at the nowcast origin *H1*. At *H2*, data on the disaggregates for the first and second month of the reference quarter have been released, and forecast values of retail sales and car registrations for the third month are included. In the month following the end of the reference quarter, at *H3*, data for all three months is contemporaneously available. This coincides with the ONS release date of the initial estimate of GDP growth based on the production approach.

5.2. Trends and seasonals

It is a well established empirical fact that data on consumption is a unit-root process; the paper by Davidson *et al*. (1978) treats the statistical modelling of the aggregate consumption function, and laid the ground for subsequent work on co-integration. As co-integrated data introduces new statistical features in the face of structural breaks such as co-breaking, we choose to model the non-integrated growth rate of final consumption expenditure. For analogous reasons, we decide to model monthly changes of the retail sales index ([10]). Similar to working with first differences, taking growth rates implies that long-run information is dropped from the data. Importantly, structural breaks in the levels are turned into impulses in the first-differenced data. The decision to work with growth rates and first differences therefore ensures that standard asymptotic theory and regression techniques apply, however, it comes at the cost of information loss and makes the detection of structural breaks more challenging.

In addition to unit roots, economic time series exhibit seasonality. Seasonal effects in the data refer to systematic calendar-related fluctuations. Retail sales, for example, rise each year around Christmas. Other examples are effects due to weather, due to administrative measures such as the start of the school year, or variations in the length of months. The data may also include calendar effects which relate to factors that do not occur in the same month/quarter every year such as changing numbers of trading days, or moving holidays, e.g., Easter. Seasonal and calendar effects must be accounted for to make consecutive periods comparable. Seasonality may be modelled explicitly by including seasonal indicators. Alternatively, a simple method for removing seasonality is yearly differencing. Further, statistical offices provide seasonally adjusted data. The ONS uses the software X12-ARIMA to remove seasonality, see ONS (2007). The advantage of X12-ARIMA over seasonal indicators and yearly differencing is that calendar effects are accounted for before removing any systematic seasonality. However, it is also used to adjust for extreme values and outliers, exacerbating the loss of information on dynamics in the data. To preserve the additivity property, we model growth rates of final consumption expenditure by households and NPISHs in current prices using seasonally adjusted data. For retail sales and car registrations we work with non-seasonally adjusted data and take yearly differences. The data appendix summarises the relevant transformations undertaken to remove unit roots and seasonalities from the data series.

([10]) In order to avoid biased results on ADF tests due to the recession during 2008-2010, we perform ADF tests on the small sample 01/2004-12/2007. Given the low power of ADF tests in small samples, we extend the sample to include data up to 1997, however, there is no convincing evidence for the stationarity of the retail sales index.

5.3. Forecasting the retail sales index

In order to fill in missing data in the differenced retail sales index at the nowcast origins *H1* and *H2*, a forecasting model for retail sales has to be specified. *Autometrics* is used for variable selection in-sample based on the following initial GUM:

$$(12)\ \Delta\Delta_{12}\mathrm{IR}_{tm} = \beta_0 + \sum_{i=0}^{12} (\beta_{s,i}\Delta\mathrm{SCI}_{tm-i} + \beta_{R,i}\Delta\mathrm{RCI}_{tm-i} + \beta_{C,i}\Delta\mathrm{CCI}_{tm-i} + \beta_{E,i}\Delta\mathrm{ESI}_{tm-i})$$

$$+ \sum_{i=1}^{12} \beta_{I,i}\Delta\Delta_{12}\mathrm{IR}_{tm-i} + \sum_{k=2}^{Tm} \beta_{1,k} d_k + v_{t_m} \quad \text{for } t_m = 1, ..., T_m.$$

d_k is an indicator taking the value 1 at time $t_m = k$ to implement impulse-indicator saturation (IIS: see Hendry *et al.* (2008) and Johansen and Nielsen (2009)) to automatically detect unknown outliers and structural breaks at any point in the sample. We use the default significance level of 0.01 for diagnostic tests and the intercept is forced to be in the final model specification. Pre-search lag reduction was turned off. In order to account for potential serial correlation, 12 endogenous lags are included in the GUM. Since survey indicators may be leading movements in final consumption expenditure, 12 lags are incorporated into the GUM ([11]).

We allow for bias correction. Since correction for search costs has been shown to reinforce the Hurwicz bias, endogenous lags are excluded from bias correction. Parameter estimates of IIS are unbiased and are therefore not subject to bias correction. IIS and variable selection are applied in several steps. These steps are performed separately for the two in-sample periods, t_m = 01/2004, .., 12/2007 and t_m = 01/2004, .., 12/2014, and the respective nowcast horizons h_m = 01/2008, .., 12/2010 and h_m = 01/2015, .., 09/2016.

(i) Select over indicators at $\alpha = 0.001$, including an intercept.

(ii) Select over variables and any retained indicators from Step (i) at $\alpha = 0.05$.

(iii) For each new observation over the nowcasting horizon, $h_m = 1, .., H_m$, re-estimate the parameters and test for the significance of an indicator for the last observation, $d_{T_m+h_m}$, at $\alpha = 0.05$. If significant, retain the indicator.

The selected model is kept the same across the nowcast horizon to reduce model uncertainty.

The retail sales index was differenced on an annual basis to remove systematic seasonality, with monthly differences of the annual change computed to remove the unit root. The remaining variation captures deviations from the long-run trend in retail sales. These irregular fluctuations may arise due to trend breaks such as the financial crisis. The financial crisis and recession is difficult to spot in Figure 5, panel (b). Indeed, Anagboso and McLaren (2009) confirm that retail sales remained strong over the economic downturn. Consequently, there are erratic fluctuations in the differenced retail sales index over 01/2008–12/2010 in panel (a) of Figure 5, with tentative evidence of location shifts having been turned into impulses over this period.

Equation (13) shows the preferred model specification of retail sales:

$$(13)\ \widehat{\Delta\Delta_{12}\mathrm{IR}}_{tm} = \underset{(0.11)}{0.05} - \underset{(0.10)}{0.25}\Delta\Delta_{12}\mathrm{IR}_{tm-1} - \underset{(0.11)}{0.29}\Delta\Delta_{12}\mathrm{IR}_{tm-5} + \underset{(0.18)}{0.42}\Delta\mathrm{SCI}_{tm-5} - \underset{(0.18)}{0.80}\Delta\mathrm{SCI}_{tm-8}$$

$$+ \underset{(0.17)}{0.71}\Delta\mathrm{SCI}_{tm-10} - \underset{(0.66)}{4.9}\Delta\mathrm{CCI}_{tm-7} - \underset{(0.77)}{3.1}\Delta\mathrm{CCI}_{tm-12}$$

([11]) It was attempted to augment the GUM with lags 1-4 of consumption growth to proxy for real personal disposable income, yet they were not retained and therefore omitted.

$$\hat{\sigma} = 0.63; R^2 = 0.75; \text{Obs.} = 35; \chi^2_N = 0.95; F(7, 27) = 0.00;$$
$$F_{AR}(3, 27) = 0.64; F_{ARCH}(3, 27) = 0.64; F_{Het}(7, 27) = 0.63$$

Numbers in parentheses refer to standard errors, $\hat{\sigma}$ is the standard error of the estimated equation, the p-value of the F-test on the joint significance of the included regressors, denoted F, as well as the R^2. p-values of mis-specification tests are also reported for the F-test of residual autocorrelation, F_{AR}, autoregressive conditional heteroskedasticity, F_{ARCH}, normality, χ^2_N, and heteroskedasticity including squares and cross-products, F_{Het}. Indicators are retained for the period 01/2008-12/2010.

The forecasts and forecast errors are plotted in Figure 6, panels (a) and (b) respectively. The 95 % forecast intervals for conventional forecasts are plotted in panel (a), along with the forecasts resulting from the bias corrected coefficient estimates. Table 8 reports the RMSFEs. As the uncorrected and bias corrected forecasts are very similar, confidence intervals are not reported for the bias corrected forecasts. Bias correction marginally increases the RMSFE despite a large number of retained regressors, and location shifts in retained variables. While bias correction reduces the RMSFE by reducing the bias related to irrelevant variables, its overall impact on the squared nowcast error, or in this context forecast error, may be ambiguous due to its effect on relevant variables. This may serve as a possible explanation for this nevertheless counterintuitive finding. There is evidence of forecast failure over this volatile period.

Table 8: RMSFE in forecasting models

	IR		CAR	
	No BC	BC	No BC	BC
01/2008-12/2010	3.02	3.03	5.47	5.48
01/2015-09/2016	1.57	1.57	1.30	1.19

BC= bias correction

Source: Authors' calculations.

The results can be contrasted with the preferred model selected over 01/2004-12/2014, reported in (14):

$$(14)\quad \widehat{\Delta\Delta_{12}\text{IR}}_{t_m} = \underset{(0.12)}{0.07} - \underset{(0.08)}{0.50}\Delta\Delta_{12}\text{IR}_{t_m-1} - \underset{(0.08)}{0.25}\Delta\Delta_{12}\text{IR}_{t_m-2} + \underset{(0.07)}{0.25}\Delta\text{IR}_{t_m-11}$$
$$+ \underset{(0.17)}{0.51}\Delta\text{SCI}_{t_m-2} + 4 \text{ Impulse Indicators}$$

$$\hat{\sigma} = 1.31; R^2 = 0.57; \text{Obs.} = 119; \chi^2_N = 0.62; F(8, 110) = 0.00;$$
$$F_{AR}(7, 103) = 0.73; F_{ARCH}(7, 105) = 0.81; F_{Het}(8, 106) = 0.90$$

Figure 5 shows that changes in retail sales exhibit more regular movements over 01/2015-09/2016. Only one significant outlier in the period 01/2008-12/2010 is retained at the more conservative significance level of $\alpha = 0.001$, confirming the stability of retail sales over the financial crisis and recession. Significant indicators are found for 12/2011 and 01/2012, in line with more pronounced fluctuations during the end of 2011 and start of 2012. At the time, negative GDP growth indicated that the UK might be heading into a double-dip recession ([12]). In July 2012, the summer olympics started, providing a boost to retail sales. During the nowcast

([12]) See http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171766_263951.pdf.

horizon one positive outlier is retained at 12/2015 and 01/2016, capturing the increase in spending before Christmas.

The final model in Figure 7, panels (a) and (b), does a reasonably good job at predicting the sign of changes in retail sales. In line with the significant outliers during the Christmas period, the fit of forecasts deteriorates during this period though forecast failure can be avoided. Bias correction does not change the RMSFE

5.4. Forecasting passenger car registrations

To forecast yearly differences in passenger car registrations, we use the same steps (i)-(iii) for variable selection and outlier detection as described in the previous section. The initial GUM for CAR includes the following variables:

$$\Delta_{12}\mathrm{CAR}_{t_m} = \gamma_0 + \sum_{i=0}^{12} (\gamma_{s,i}\Delta\mathrm{SCI}_{t_m-i} + \gamma_{R,i}\Delta\mathrm{RCI}_{t_m-i} + \gamma_{C,i}\Delta\mathrm{CCI}_{t_m-i} + \gamma_{E,i}\Delta\mathrm{ESI}_{t_m-i})$$

$$+\sum_{i=1}^{12} \gamma_{CA,i}\Delta_{12}\mathrm{CAR}_{t_m-i} + \sum_{k=2}^{T_m}\gamma_{1,k}d_k + v_{t_m} \qquad \text{for } t_m = 1, \ldots, T_m.$$

In Figure 5, a drop in the twelfth differences of passenger car registrations during 2008-2009 is evident, which is matched by the confidence indicators on the retail trade and service sector. Car registrations also indicates a pronounced outlier in 03/2009. In the second half of 2009, yearly differences in car registrations revert to being positive as a result of the car scrappage scheme implemented by the government. This scheme allowed for the scrappage of 400 000 old vehicles, and provided a £2 000 incentive to buy a new car, see Crossley, Leicester, and Levell (2010). This led to a significant short-term increase in car registrations. The statistical significance of outliers during this period is confirmed in the forecasting model selected at $\alpha = 0.05$. The clustered occurrence of outliers during the second half of 2008 and 2009 as well as during 2010 provides strong evidence for the relevance of structural breaks during this nowcast horizon.

The selected forecasting model is:

$$(15) \quad \widehat{\Delta_{12}\mathrm{CAR}}_{t_m} = \underset{(0.10)}{0.20} + \underset{(0.08)}{0.33}\,\Delta_{12}\mathrm{CAR}_{t_m-3} + \underset{(0.09)}{0.72}\Delta_{12}\mathrm{CAR}_{t_m-6} + \underset{(0.13)}{0.51}\,\Delta\mathrm{SCI}_{t_m-3}$$

$$+ \underset{(0.11)}{0.63}\,\Delta\mathrm{SCI}_{t_m-12} - \underset{(0.46)}{1.5}\,\Delta\mathrm{CCI}_{t_m-2} - \underset{(0.40)}{1.4}\,\Delta\mathrm{CCI}_{t_m-7} + \underset{(0.45)}{1.9}\,\Delta\mathrm{CCI}_{t_m-9}$$

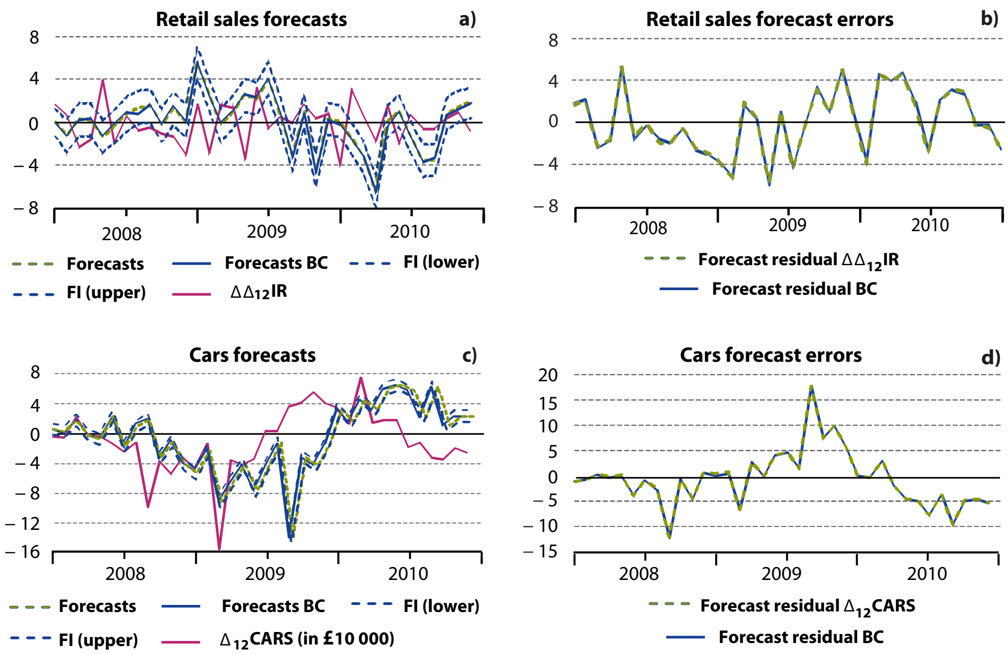
$$- \underset{(0.48)}{1.3}\,\Delta\mathrm{CCI}_{t_m-12} + \underset{(0.22)}{0.76}\,\Delta\mathrm{ESI}_{t_m} + \underset{(0.22)}{1.2}\,\Delta\mathrm{ESI}_{t_m-5} + \underset{(0.23)}{0.94}\,\Delta\mathrm{ESI}_{t_m-11}$$

$$\hat{\sigma} = 0.38;\ R^2 = 0.91;\ \text{Obs.} = 35;\ \chi^2_N = 0.49;\ \mathrm{F}(11, 24) = 0.00;$$

$$\mathrm{F}_{AR}(3, 21) = 0.08;\ \mathrm{F}_{ARCH}(3, 30) = 0.03;\ \mathrm{F}_{Het}(22, 13) = 0.46$$

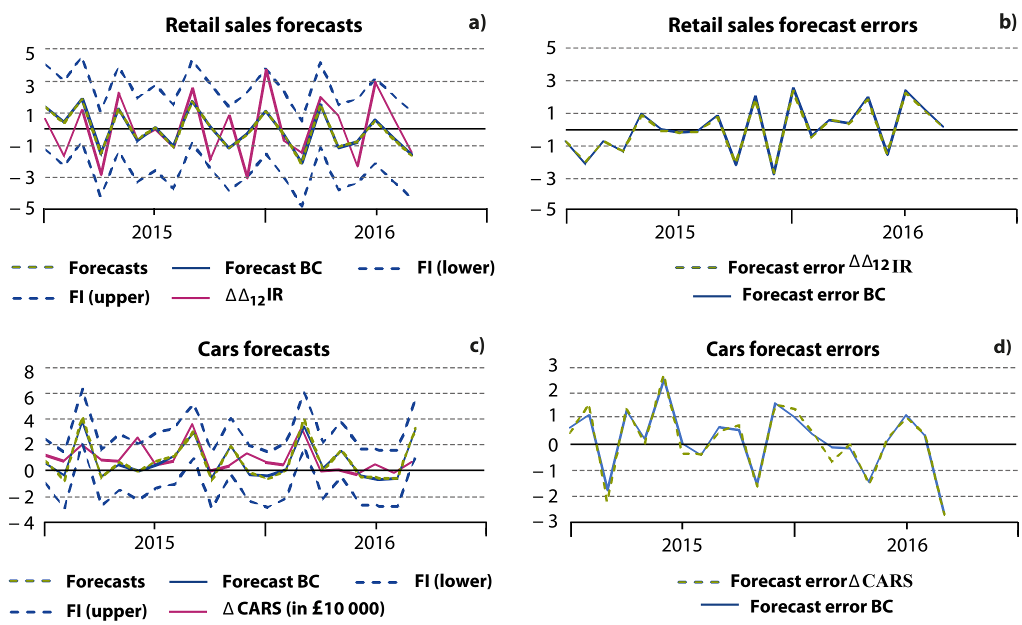
Compared to retail sales, more systematic, rather than purely irregular, fluctuations in the car registrations are retained. The performance of the preferred forecasting model is not enhanced by correcting parameters for search. Figure 6, panels (c) and (d) record uncorrected and bias corrected forecasts, along with the 95% confidence intervals for the forecasts from equation (15). The figure shows evidence of forecast failure over the forecast horizon.

Figure 6: Forecasts and forecast errors for retail sales and cars, 01=2008



Note: IIS is applied recursively as the forecast horizon advances for a xed model formulation, so estimated parameters and error variances do not change over the forecast horizon, hence the small forecast intervals for Cars.
Source: Authors' calculations.

Figure 7: Forecasts and forecast errors for retail sales and cars, 01/2015 - 09/2016



Source: Authors' calcuations.

Looking at the plot of car registrations, it is clear that the period of 01/2015 - 09/2016 is more stable. In order to account for the scrappage scheme, we retain an indicator variable covering the implementation period 05/2009 - 02/2010 in the forecasting model. IIS suggested systematically larger growth rates in car registrations in the months March and September from 2012 onwards. As stated by the ONS, car registrations are usually higher in Q1, Q3 than Q2, Q4, corresponding to the release of new number plates in the months March and September, see Grove (2012). We therefore specify, and retain, the indicator variable *Plate*, which takes the value 1 for March and September in the years 2012 - 2016, in the final forecasting model (16) over the full sample period:

$$(16)\quad \Delta_{12}\widehat{CAR}_{t_m} = \underset{(0.09)}{-0.12} + \underset{(0.04)}{0.30}\,\Delta_{12}CAR_{t_m-1} + \underset{(0.04)}{0.31}\,\Delta_{12}CAR_{t_m-2} + \underset{(0.04)}{0.13}\,\Delta_{12}CAR_{t_m-7}$$

$$+ \underset{(0.11)}{0.27}\,\Delta RCI_{t_m-7} - \underset{(0.12)}{0.21}\,\Delta RCI_{t_m-10} + \underset{(0.31)}{1.65}\,\Delta CCI_{t_m-5} - \underset{(0.35)}{0.92}\,\Delta CCI_{t_m-10}$$

$$+ \underset{(0.35)}{0.68}\,Scrappage_{t_m} + \underset{(0.46)}{3.69}\,Plate_{t_m} + 6\text{ Impulse Indicators}$$

$$\hat{\sigma} = 0.83;\quad R^2 = 0.92;\quad Obs. = 120;\quad \chi^2_N = 0.00;\quad F(17, 102) = 0.00;$$

$$F_{AR}(7, 97) = 0.72;\quad F_{ARCH}(7, 105) = 0.002;\quad F_{Het}(13, 97) = 0.84$$

The large number of retained indicators over 2008 - 2010 underline the importance of structural breaks in yearly car registrations during that period. The bias corrected model provides more accurate forecasts of car registrations. Note the large fall in RMSFEs compared to the horizon 01/2008 - 12/2010 due to the absence of structural breaks with only three forecasts falling marginally outside the 95 % interval in Figure 7.

5.5. Nowcasting final consumption expenditure

To arrive at a nowcast of final consumption expenditure, an equation linking the differenced retail sales and car registrations to consumption expenditure needs to be specified. As described in Section 5.1, the monthly data are separated into three blocks r^1, r^2, r^3 to match the quarterly frequency of final consumption expenditure. These blocks refer to the first, second and third month of the quarter. Due to lack of knowledge of the DGP a more elaborate GUM compared to the Monte Carlo simulations is specified. Beyond the contemporaneous values, the GUM includes two lags of the disaggregates, and four lags of consumption expenditure growth, Δc, to capture potential serial correlation, as well as IIS to help achieve a congruent model specification. The steps (i)-(iii) of variable selection and outlier selection remain unchanged; outlier selection in-sample is performed at $\alpha = 0.001$, retaining an intercept. Variable selection including any retained indicators in-sample is performed at the significance level $\alpha = 0.05$. The same significance level holds for outlier detection over the nowcast horizon. The intercept and all contemporaneous values of the disaggregated components are retained in the final model specification in order to be able to evaluate the performance of bias correction for nowcasting.

At *H1*, the GUM takes the following form, with contemporaneous values of differences in retail sales and in car registrations of blocks r^1, r^2 being forced to be included in the final model. The nowcasting models for the three origins are selected in-sample, so all data may be treated as contemporaneously available.

$$\Delta c^{H1}_{tq} = \pi_0 + \pi_{r^2,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^2}_{tq}\\ \Delta_{12}CAR^{r^2}_{tq}\end{pmatrix} + \pi_{r^1,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^1}_{tq}\\ \Delta_{12}CAR^{r^1}_{tq}\end{pmatrix} + \sum_{k=2}^{Tq}\delta_{0,k}d_k + \sum_{k=1}^{4}\delta_{C,k}\Delta c_{t_q-k}$$

$$+\sum_{j=1}^{2}\left[\pi_{r^1,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^1}_{tq-j}\\ \Delta_{12}CAR^{r^1}_{tq-j}\end{pmatrix} + \pi_{r^2,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^2}_{tq-j}\\ \Delta_{12}CAR^{r^2}_{tq-j}\end{pmatrix} + \pi_{r^3,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^3}_{tq-j}\\ \Delta_{12}CAR^{r^3}_{tq-j}\end{pmatrix}\right] + \epsilon_{tq}$$

At the nowcasting origins there are ragged edges in $\Delta\Delta_{12}IR^{r^2}_{tq}$ and $\Delta_{12}CAR^{r^2}_{tq}$, which have to be forecast. At *H2*, contemporaneous values of retail sales and car registrations in blocks r^1, r^2, r^3, as well as the intercept, are always retained.

(17) $$\Delta c^{H2}_{tq} = \pi_0 + \pi_{r^3,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^3}_{tq}\\ \Delta_{12}CAR^{r^3}_{tq}\end{pmatrix} + \pi_{r^1,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^1}_{tq}\\ \Delta_{12}CAR^{r^1}_{tq}\end{pmatrix} + \pi_{r^2,0}\begin{pmatrix}\Delta\Delta_{12}IR^{r^2}_{tq}\\ \Delta_{12}CAR^{r^2}_{tq}\end{pmatrix}$$

$$+\sum_{k=2}^{Tq}\delta_{0,k}d_k + \sum_{k=1}^{4}\delta_{C,k}\Delta c_{t_q-k} + \sum_{j=1}^{2}\sum_{i=1}^{3}\left[\pi_{r^i,j}\begin{pmatrix}\Delta\Delta_{12}IR^{r^i}_{tq-j}\\ \Delta_{12}CAR^{r^i}_{tq-j}\end{pmatrix}\right] + \epsilon_{tq}$$

At *H2*, it is $\Delta\Delta_{12}IR^{r^3}_{tq}$ and $\Delta CAR^{r^3}_{tq}$ that are missing and have to be filled in at the nowcast origin. The GUM for Δc^{H3}_{tq} is identical to that for Δc^{H2}_{tq} in equation (17). For nowcasts at *H3*, data for all three months of the reference quarter is available. Consequently, the final nowcasting model at *H3* is exclusively based on actual data. Again, the contemporaneous values of retail sales and car registrations of blocks r^1, r^2, r^3 are always retained.

(18) $$\widehat{\Delta c^{H1}_{tq}} = \underset{(0.001)}{0.012} + \underset{(0.001)}{0.0016}\Delta\Delta_{12}IR^{r^1}_{tq} + \underset{(0.002)}{0.0016}\Delta\Delta_{12}IR^{r^2}_{tq} + \underset{(0.001)}{0.0018}\Delta_{12}CAR^{r^1}_{tq} - \underset{(0.002)}{0.004}\Delta_{12}CAR^{r^2}_{tq}$$

$\hat{\sigma} = 0.03$; $R^2 = 0.57$; Obs. = 12; $X^2_N(2) = 0.60$; $F(4, 7) = 0.15$;

$F_{AR}(1, 6) = 0.83$; $F_{ARCH}(1, 10) = 0.17$

(19) $$\widehat{\Delta c^{H2}_{tq}} = \widehat{\Delta c^{H3}_{tq}} = \underset{(0.001)}{0.012} + \underset{(0.001)}{0.0017}\Delta\Delta_{12}IR^{r^1}_{tq} + \underset{(0.002)}{0.0005}\Delta\Delta_{12}IR^{r^2}_{tq} + \underset{(0.001)}{0.0004}\Delta\Delta_{12}IR^{r^3}_{tq}$$

$$+\underset{(0.001)}{0.0007}\Delta_{12}CAR^{r^1}_{tq} - \underset{(0.002)}{0.0033}\Delta_{12}CAR^{r^2}_{tq} + \underset{(0.001)}{0.0015}\Delta_{12}CAR^{r^3}_{tq}$$

$\hat{\sigma} = 0.003$; $R^2 = 0.71$; Obs. = 12; $X^2_N(2) = 0.54$; $F(6, 5) = 0.23$;

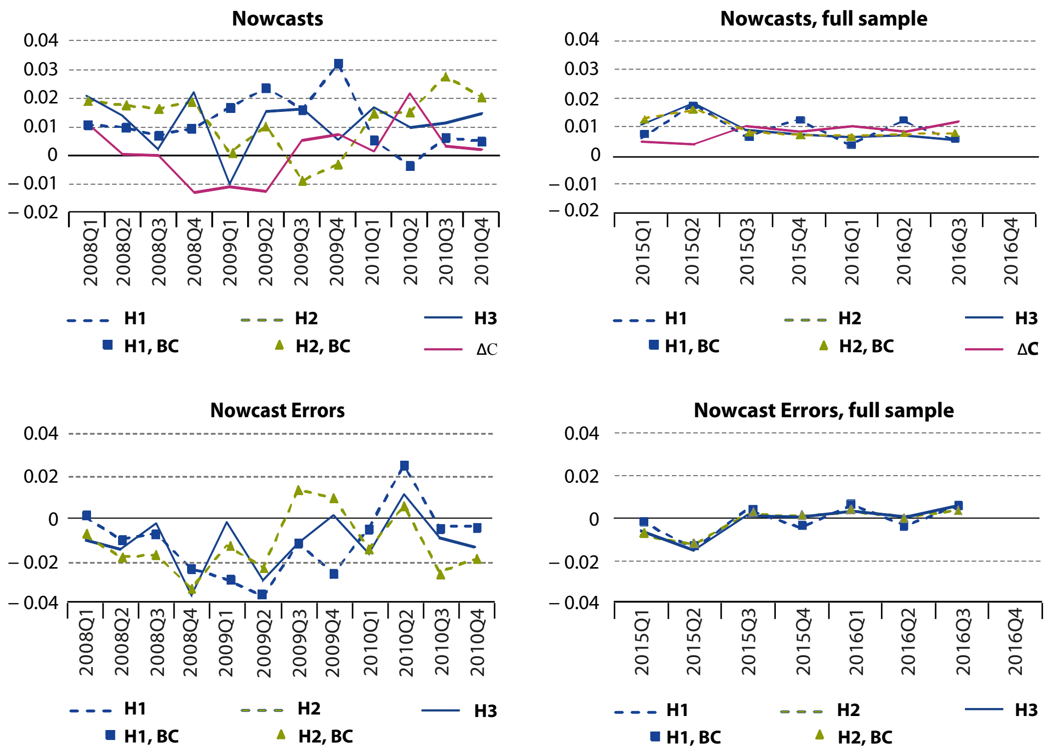
$F_{AR}(1, 4) = 0.55$; $F_{ARCH}(1, 10) = 0.96$

The above nowcasting models were selected over the in-sample period 2004Q1-2007Q4 ([13]). Only forced variables and indicators were retained. Indeed, as can be gathered from the size of parameter estimates relative to standard errors, without forcing the contemporaneous values of car registrations and retail sales into the final model, the preferred specification would have just included the intercept. At *H3*, it was also attempted to include the preliminary estimate of GDP that becomes available at this nowcast origin in the GUM. It was not retained in the final model specification. Recall that at *H3* the nowcasts are based entirely on actual data. Out of the

([13]) There are too few values to compute F_{HET}.

conventional nowcasts, *H3* minimises the RMSFE. Further, nowcasts at *H3* may be taken as a reference point for the explanatory power of the disaggregates used to model growth in final consumption expenditure.

Figure 8: Nowcasts and nowcast errors



Source: Authors' calculations.

Note: BC: with bias correction (excluding endogenous lags); H1, H2, H3 refer to the nowcasting horizon. Since nowcasts at H3 are based on actual data, it is only included once.

Nowcasts in 2015Q1-2016Q3 are more accurate as a result of the smaller fluctuations, and more regular behaviour in consumption expenditure growth, as is evident from Figure 8. Differences in RMSFE become negligible (the figures in Table 9 have been multiplied by 100). Searching over indicators in-sample at $\alpha = 0.001$ in step (i), d_{tq}=*2008Q4*, d_{tq}=*2009Q1*, d_{tq}=*2009Q2* are retained at *H1 – H3*. These are dropped once the disaggregates are included.

$$(20)\quad \widehat{\Delta c}^{H1}_{tq} = \underset{(0.001)}{0.003} + \underset{(0.13)}{0.54}\Delta c_{tq-3} - \underset{(0.001)}{0.001}\Delta\Delta_{12}IR^{r1}_{tq} - \underset{(0.001)}{0.0003}\Delta\Delta_{12}IR^{r2}_{tq} + \underset{(0.001)}{0.0005}\Delta\Delta_{12}CAR^{r1}_{tq} + \underset{(0.001)}{0.002}\Delta_{12}CAR^{r2}_{tq}$$

$$\hat{\sigma} = 0.006;\ R^2 = 0.52;\ Obs. = 40;\ X^2_N(2) = 0.57;\ F(5, 34) = 0.00;$$
$$F_{AR}(3, 31) = 0.13;\ F_{ARCH}(3, 34) = 0.22;\ F_{Het}(10, 29) = 0.99$$

$$(21)\quad \widehat{\Delta c^{H2}_{t_q}} = \widehat{\Delta c^{H3}_{t_q}} = \underset{(0.001)}{0.003} + \underset{(0.13)}{0.58}\Delta c_{t_q-3} - \underset{(0.0006)}{0.0005}\Delta\Delta_{12}IR^{r^1}_{t_q} - \underset{(0.0008)}{0.0001}\Delta\Delta_{12}IR^{r^2}_{t_q} + \underset{(0.0006)}{0.0005}\Delta\Delta_{12}IR^{r^3}_{t_q}$$

$$- \underset{(0.001)}{0.002}\Delta_{12}CAR^{r^1}_{t_q} + \underset{(0.001)}{0.003}\Delta_{12}CAR^{r^2}_{t_q} + \underset{(0.0004)}{0.001}\Delta_{12}CAR^{r^3}_{t_q}$$

$$\hat{\sigma} = 0.005;\ R^2 = 0.58;\ Obs. = 40;\ X^2_N\ (2) = 0.53;\ F(5, 34) = 0.00;$$
$$F_{AR}(3, 29) = 0.24;\ F_{ARCH}\ (3, 34) = 0.28;\ F_{Het}(14, 25) = 0.99$$

In the longer in-sample period, the third lag of consumption growth becomes significant and is retained at all nowcast origins. The size of the parameter estimates on the endogenous lag relative to the disaggregates highlights that this lag dominates the nowcasts. At all three nowcasting origins, the nowcasts miss the lower growth at the start of 2015 in Figure 8. Nowcasts based on bias-corrected forecasts at *H1* marginally improve the fit of the nowcast, while they have adverse effects on nowcasting accuracy at *H2*. Further, the model at *H3* is found to be less accurate in terms of RMSFE than *H2*. In general, findings in the full sample confirm that the disaggregates seem to have little explanatory power for the direction of change in consumption growth.

Table 9: Percentage RMSFE in nowcasting models of final consumption expenditure

	H1		H2		H3
	No BC	BC	No BC	BC	
01/08-12/10	1.845	1.845	1.773	1.773	1.581
01/15-0/16	0.684	0.669	0.589	0.590	0.651

BC= bias correction. Reported figures have been multiplied by 100.

Source: Authors' calculations.

5.6. Interpretation using the 7 insights

This section interprets the empirical example in light of the seven insights from the nowcast error taxonomy, though such an interpretation remains inconclusive without knowledge of the underlying DGP, and hence lack of information on relevant variables, and breaking dynamics. With this caveat in mind, the empirical example, and in particular the comparison of the two nowcasting periods, remain useful in considering how the intuitions of the theoretical insights can be applied in practice.

Over the volatile period 2008Q1-2010Q4, there is nowcast failure. As discussed, retail sales remained strong over the recent recession, while final consumption expenditure was subject to a substantial downward shift. Though taking the first differences reduces the mean of the retail sales index to a small value, the relationship between differenced retail sales and consumption expenditure arguably shifted over the recession. In consideration of this argument, nowcast failure is consistent with the first insight, which states that a shift in dynamics causes nowcasts to deteriorate if the exogenous variable has a non-zero mean.

Taking into account the transformation of the retail sales index, only car registrations include relevant information on the persistent drop in consumption growth rates over 2008-2010. This does not seem to be sufficient to reliably predict consumption expenditure over 2008Q1-2010Q4. However, there is no systematic nowcast failure in the more stable period, 2015Q1-2016Q3, based on the same information set. This comparison is in line with the second insight, which states that incorrect omission of relevant variables matters for the accuracy of nowcasts if the omitted variables are subject to location shifts, while it does not lead to nowcast failure in stable times. The omission of relevant information in this small-scale nowcasting exercise therefore provides an additional explanation for the deterioration of nowcasts over the great recession.

Equally, the discussion on the forecasts of car registrations in the preceding section made clear that there was systematic forecast failure over 01/2008-12/2010, whilst forecasts were more accurate over the later nowcasting period. We know from the third and fourth insight that failing to accurately forecast the post-shift mean may cause systematic nowcast failure, while there is no mean accuracy gain from accurate forecasts absent location shifts in the exogenous variable. Both insights are consistent with the findings in Tables 8 and 9, and forecast failure likely exacerbated the deterioration in nowcasting accuracy over 2008Q1- 2010Q4.

Insights five and six relate to the retention of irrelevant information, and show that absent mean shifts the inclusion of irrelevant variables is not costly, while it is detrimental to the nowcasting performance if there is a shift in irrelevant variables that is not mirrored in the dependent variable. Given the low explanatory power of the included explanatory variables for movements in consumption expenditure growth, it is questionable whether the variables under consideration may be thought of as 'relevant'. If they are deemed to be irrelevant, then the accuracy of nowcasts was reduced not only due to omission of relevant information, but also due to retention of irrelevant variables that were subject to shifts, providing an additional source of nowcast failure over 2008Q1-2010Q4. As alluded to previously, a limitation to this interpretation is that the distinction between relevant versus irrelevant variables becomes unclear when taking into account that in a real-world settings most variables are correlated with each other, so that the retail sales index, and car registrations may proxy for relevant variables.

Insight seven states that bias correction on irrelevant variables is likely to be more beneficial in turbulent periods. In the empirical example, bias correction was applied to the forecasting models, and was found to improve forecasting performance over the stable period, though its overall effect remained negligible. While the confidence indicators in the forecasting models cannot be thought of as causing movements in retail sales or car registrations, and hence are not part of the respective DGPs, insight seven, in line with simulation evidence in Section 4.3.4, would suggest that the included variables in forecasting models are good proxies for relevant information. In particular with respect to car registrations, it seems equally plausible that any gains in forecasting accuracy from bias correcting parameter estimates of irrelevant variables were overshadowed by forecast failure.

While the empirical example offers a rich set-up for applying the nowcast taxonomy, it should be acknowledged that the performance of the nowcasting framework has been unreliable. From insight four, a robust method of forecasting during the great recession would be required. Further, the current model assumes that the shift in consumption growth may be explained by the disaggregates. This assumption is too restrictive, since accounting for shifts at the aggregate level could alleviate the impact of forecast failure from disaggregates, of the impact of dynamic shifts from insight one, or of shifts at the aggregate level that exceed those in disaggregates.

In addition, the empirical exercise underlines the exigencies towards the data that may be used to successfully nowcast with the suggested framework. Taking first differences results in semi-robust data, from which any regular trends have been removed, and location shifts have been turned into impulses. The lack of information contained in semi-robust data increases the difficulty of nowcasting accurately. A natural step is to consider the nowcasting framework in a co-integrated framework to be able to model variables with unit roots. Moreover, the two disaggregates were modelled using survey or 'soft' data only. While surveys are a timely source, the nowcasting literature is undecided about the explanatory content of soft data, see, inter alia, Antolin-Diaz *et al.* (2017) or Mitchell (2009). Based on similar Eurostat surveys to those used here, Mitchell (2009) finds that only in addition to hard data did soft data improve nowcasts at the onset of the recession. Further he found that, due to their projective nature, surveys became less successful further into the quarter to be nowcast. It seems advisable to include hard data beyond endogenous lags, e.g., appropriately proxying for real personal disposable income. In the evaluation of the empirical application, however, it should be taken into account that consumer expenditure is difficult to model empirically. Under certain assumptions on the economic model, an inter-temporal utility maximisation problem with rational expectations predicts that consumption follows a random walk, see Muellbauer (1994).

6. Conclusion

In providing preliminary estimates of economic aggregates, statistical agencies have to select a nowcasting model from a large number of disaggregates, and deal with missing data of mixed frequency. The nowcast error taxonomy presented in this paper offers a framework for thinking about sources of nowcast errors that may be faced by statistical agencies. The nowcast error taxonomy makes clear that the main sources for systematic nowcast failures are unmodelled location shifts, and underlines the importance of accurate infilling of any missing data that is needed to construct the macroeconomic aggregate. Further, the nowcast error taxonomy highlights the costs associated with omitting relevant disaggregates or retaining irrelevant ones, and showed that these costs are most pronounced if the regressors are subject to location shifts. Mis-specification and estimation uncertainty have comparatively small impacts on nowcast accuracy. The theoretical insights have been confirmed by evidence based on a simulation exercise, and have been applied in an empirical nowcasting exercise of final consumption expenditure by households and NPISH using a retail sales index and passenger car registrations to interpret sources of nowcast failure.

Beyond the seven insights, the nowcast error taxonomy helps to interpret trade-offs involved in variable selection. In practice, the distinction between relevant and irrelevant variables is not known. The trade-off in variable selection thus consists of reducing costs of search by specifying a tight significance level for variable selection to reduce the probability of falsely retaining irrelevant variables, whilst simultaneously increasing costs of inference by increasing the likelihood of falsely rejecting relevant variables. The simulations made clear that depending on out-of-sample breaking dynamics, costs of inference or costs of search may dominate. In simulations with location shifts occurring to all relevant disaggregates, and non-zero correlations between variables, costs of inference dominated costs of search with irrelevant variables proxying for relevant ones. Once asymmetric breaking dynamics were specified in the simulations, the adverse impact of costs of search became apparent. The empirical example highlighted additional challenges, such as non-stationarity of real-world data requiring data transformations that lead to information loss, which further complicate the distinction between relevant and irrelevant information in practice. Equally, relevant variables may have low non-centralities in practice due to noise in the data, impacting on the choice of the optimal significance level in variable selection. In line with the difficulty of isolating relevant information in variable selection, results on bias correction after model selection were mixed, both in simulations and in the empirical example. Beyond insights into nowcast failure, the taxonomy may thus inform model selection in nowcasting.

Data Appendix

Table 10: Data Sources

Label	Description	Source	Period	Transfor-mation	Release lag
C_t	Household & NPISH final consumption expenditure in million (CPSA)	[1]	1997Q1-2016Q3	5	4
IR_t	Index for value of retail sales, all retailing incl. automotive fuel	[1]	1997M1-2016M12	1; 2	1
CAR_t	New passenger car registrations UK	[2]	1997M1-2016M12	2; 3	1
SCI_t	Service sector confidence indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0
CCI_t	Consumer confidence indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0
RCI_t	Retail trade confidence indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0
ESI_t	Economic sentiment indicator UK (survey) (SA)	[2]	1997M1-2016M12	4	0

SA - seasonally adjusted, CP - current price.
Transformations: (1) = $\Delta X_t = X_t - X_{t-1}$; (2) = $\Delta_{12} X_t$; (3) = $X_t/10\,000$; (4) = $\Delta X_t/10$; (5) = $\Delta X_t / X_{t-1}$.
The transformed time series were tested for unit roots using Dickey-Fuller tests including available data up to 12/2007. 12 lags were included for the monthly data, and 4 lags for quarterly data..

Sources: [1] ONS (ons.gov.uk), [2] Eurostat database (http://ec.europa.eu/eurostat/data/database).

Acknowledgements

Financial support from the Economic and Social Research Council (award ES/J500112/1), the Robertson Foundation (award 9907422), Institute for New Economic Thinking (grant 20029822) and Statistics Norway through Research Council of Norway Grant 236935 are all gratefully acknowledged.

References

Anagboso, M. and C. McLaren (2009), 'The impact of the recession on retail sales volumes', *Economic and Labour Market Review*, Volume 3. Office for National Statistics.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017), 'Hard vs. Soft Data in Nowcasting GDP: A Comment', https://drive.google.com/file/d/0B7xdKeAL7xDbWUVhMlVJUEtoeG8/view.

Castle, J. L., D. F. Hendry, and O. Kitov (2017), 'Forecasting and nowcasting macroeconomic variables: A methodological overview', *Handbook on Rapid Estimates*, UN and Eurostat.

Castle, J. L. and N. Shephard (Eds.) (2009), *The Methodology and Practice of Econometrics*, Oxford: Oxford University Press.

Clements, M. P. and D. F. Hendry (1998), 'Forecasting economic processes', *International Journal of Forecasting*, 14, pp. 111–131.

Clements, M. P. and D. F. Hendry (2006), 'Forecasting with breaks', in G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Econometrics on Forecasting*, pp. 605– 657, Amsterdam: Elsevier.

Davidson, J. E. H., D. F. Hendry, F. Srba, and S. Yeo (1978), 'Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom', *The Economic Journal*, pp. 661–692.

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

Ericsson, N. R. (2017), 'Economic forecasting in theory and practice: An interview with David F. Hendry', *International Journal of Forecasting,* 33, pp. 523–542.

Crossley, F. T., A. Leicester, and P. Levell (2010), 'Fiscal stimulus and the consumer', in R. Chote, C. Emmerson, and J. Shaw (Eds.), IFS Green Budget. IFS.

Giannone, D., L. Reichlin, and S. Simonelli (2009), 'Nowcasting Euro Area economic activity in Real-Time: The role of confidence indicators', *National Institute Economic Review,* 210, pp. 90–97.

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

Grove, J. (2012), 'Vehicle licensing statistics, Great Britain: Quarter 3 2012', Technical report, Department for Transport.

Hendry, D. F. and J. A. Doornik (2014), *Empirical Model Discovery and Theory Evaluation*, Cambridge, Mass., MIT Press.

Hendry, D. F. and K. Hubrich (2011), 'Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate', *Journal of Business and Economic Statistics*, 29, pp. 216–227.

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

Hendry, D. F. and H.-M. Krolzig (2005), 'The properties of automatic GETS modelling', *The Economic Journal*, 115, pp. C32–C61.

Hendry, D. F. and G. E. Mizon (2012), 'Open-model forecast-error taxonomies', in X. Chen and N. R. Swanson (Eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*, pp. 219–240. New York: Springer.

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

Millard, S. (2015), 'The Great Recession and the UK labour market', *Working Paper,* 566, Bank of England.

Mitchell, J. (2009), 'Where are we now? The UK recession and nowcasting GDP growth using statistical models', *National Institute Economic Review,* 209(1), 60–69.

Muellbauer, J. (1994), 'The assessment: consumer expenditure', *Oxford Review of Economic Policy* 10(2), pp. 1–41.

ONS (2007), 'Guide to seasonal adjustment with X-12-ARIMA', Technical report, Office for National Statistics.

ONS (2015), 'Family spending: 2015', Technical report, Office for National Statistics.