Nowcasting to Predict Data Revisions

Eddie Casey

Working Paper No.7
October 2018

Irish Fiscal Advisory Council

Nowcasting to Predict Data Revisions

Eddie Casey[1]

Irish Fiscal Advisory Council and University College Dublin

October, 2018

Abstract

National accounts data are often prone to large revisions. Using data for Ireland—where revisions are among the largest in the OECD—we employ a dynamic factor model to nowcast domestic economic activity. We show that nowcasts can offer a timely, less biased substitute for preliminary national accounts data—one which can be used to help predict subsequent revisions to preliminary estimates. Building on our findings, we develop a simple algorithm that can be used to combine the information content of nowcasts with initial estimates of output so as to arrive at "nowcast-augmented initial estimates". We show evidence that, due to a reduction in bias, the augmented estimates can perform better than initial outturns themselves when it comes to predicting the final and revised outturns.

Keywords: dynamic factor model, state space, kalman filter.

[1] The author is a PhD student of University College Dublin and Chief Economist at the Irish Fiscal Advisory Council (IFAC). E-mail: eddie.casey@fiscalcouncil.ie. Views expressed in this paper are the authors' own. We would like to acknowledge the kind assistance from Prof. Karl Whelan (UCD) and from members of the Irish Fiscal Advisory Council, particularly Sebastian Barnes (OECD) and Íde Kearney (De Nederlandsche Bank).

Section 1: Introduction

In the absence of systematic bias and any alternative national accounts measure, initial estimates of economic activity might be argued to represent the best available indicator of current economic activity. However, revisions may continue for years after initial estimates are released as information from annual surveys and administrative sources become available. Moreover, initial estimates may not always be free from bias relative to final outturns.

Having a timely and reliable information set is essential if policymakers are to understand current economic developments and respond to these effectively. What if there were an alternative source of information that could help to predict revisions to initial estimates – one which were based on a sufficiently comprehensive set of available data?

We produce an additional estimate of current economic activity based on publicly available and high-frequency information. This “nowcast” of economic activity can act as an alternative view of current economic activity – one that can supplement the less timely initial national accounts estimates.

Nowcasts are usually defined as estimates of economic activity in the recent past, present or near future, which avail of information from a large panel of high frequency economic indicators. Since official national accounts measures of economic activity are typically published with a long delay and at quarterly frequency, it can be desirable to exploit monthly information to obtain earlier estimates. In cases where initial quarterly national accounts estimates are prone to large revisions—like in Ireland—nowcasts might also act as a means of inferring the direction and even the size of revisions to initial estimates.

This paper makes two major contributions to the literature on nowcasting.

First, we produce nowcasts of sub-components of national accounts data for Ireland that allow us to develop nowcasts of aggregate domestic economic activity. This focus is warranted for Ireland, given that GDP is especially volatile. Distortions to net exports associated with multinational activities often result in a misleading picture of domestic economic developments when looking solely at GDP or similar aggregates. The focus on disaggregated components of domestic activity also allows us to assess the source of nowcast errors more systematically and to discern the drivers of economic activity.

Second, we show how nowcasts can be used to augment preliminary national accounts estimates in order to better predict final national accounts outturns. We construct real-time nowcasts of domestic economic activity for Ireland based on real-time national accounts data and real-time high frequency indicators. We assess our nowcasts of domestic economic activity and the initial estimates produced by the Central Statistics Office (CSO) as predictors of the final revised national accounts outturns. Comparing the performance to that of initial outturns and estimates from a naive benchmark, we find that the nowcasts perform relatively well on a sub-component basis and better for aggregate domestic demand. The better performance at aggregate level reflects the fact that sub-component errors effectively cancel each other out such that – on average – the nowcasts outperform even initial estimates.

Exploring our results further, we combine the information content in our nowcasts with a simple algorithm to arrive at nowcast-augmented initial estimates of economic activity. We compare the performance of (i) our Initial estimates and (ii) our nowcast-augmented initial estimates against final estimates. Our findings suggest that for all components, aside from government consumption, the errors with respect to final outturn estimates are improved by augmenting initial estimates with information from our nowcasts.

Extending the sample period assessed for personal consumption expenditure, we again find that augmenting initial estimates with information from nowcasts can help to substantially reduce bias present in initial estimates. However, we do not find that it reduces the overall size of errors with respect to final outturns. Unlike, Matheson, Mitchell and Silverstone (2009), we find that the additional information value of nowcasts needs to be considered more carefully. Standard forecast combination regressions that include both nowcasts and initial estimates, for example, do not reveal any addidtional information value to nowcasts that is statistically significant. However, the nowcast-augmented approach that we use suggests that nowcasts – in addition to being timely substitutes – can be used to complement preliminary outturns in a way that reduces bias with respect to final estimates.

Further extensions of this work could seek to explore better ways to predict the exact magnitudes of revisions to initial estimates, while drawing on the information availalble in nowcasts.

Section 2: Relevant Literature

There is a large literature on nowcasting. that is, forecasting of the very recent past, the present, or the very near future of indicators for economic activity, such as GDP (Banbura *et al*, 2013; Giannone, Reichlin and Small, 2008; Doz, Giannone, and Reichlin, 2011). The use of more timely estimates followed the success of such approaches for forecasting (Stock and Watson; 2003; Forni, Hallin, Lippi and Reichlin, 2003).

For Ireland, D'Agostino, McQuinn and O' Brien (2012) present nowcast estimates of Irish GDP. Dynamic factor analysis is used to extract a common factor from a panel data set of 41 different variables, with bridging equations used to relate the monthly data to the quarterly aggregate GDP estimates. An out-of-sample forecasting simulation exercise compares the nowcast results with those of a simple benchmark model (a four-quarter moving average of year-on-year growth rates). The results show that errors for the nowcast model are smaller than those for the benchmark model (roughly 5 – 6 percentage points for the nowcast estimates as compared to over 8 percentage points for the benchmark model).

In terms of our dynamic factor model, we use the same technique as availed of in D'Agostino, McQuinn and O' Brien (2012) and Giannone, Reichlin and Small (2008). However, we focus on disaggregated components of domestic Irish activity rather than an aggregate and we extend the application further to assess its usefulness in predicting data revisions.

Earlier literature including Mankiw and Shapiro (1986) suggests that there is limited predictability of early revisions to US GDP growth. Yet later work by Faust, Rogers and Wright (2005) finds a great deal of predictability for several other G7 countries when using standard forecast efficiency tests. They note that revisions to GDP announcements are quite large in all G7 countries with reversion to the mean tendencies apparent in countries such as the UK.

Faust, Rogers and Wright (2005) test for predictability of revisions more formally. They regress revisions on preliminary estimates along with seasonal dummies and five variables known at the time that the preliminary data were released: lagged preliminary data, the growth rate of equity prices, a 3-month interest rate, oil price inflation, and a dummy variable for national elections. For all G-7 countries aside from the US and France (where the sample was too short), they find predictability of revisions. They note that this is mostly due to the predictive power of the preliminary number: extreme values, large or small, in the preliminary growth rate tend to be revised toward the mean.

Clements and Galvão (2013) show that models of multiple data vintages can predict quarterly output and inflation data revisions by exploiting information on past revisions, and in particular, the annual revisions which take place in the third quarter of each year. Further empirical work in this sphere is provided by Clements and Galvão (2017). The authors determine the predictability of early data revisions to US output growth at short horizons (i.e., the second and the third estimates of US GDP at horizons as short as one week). They avail of a suite of AR models, threshold models, regression models using monthly economic indicators and Mixed Data Sampling (MIDAS) regressions with daily financial variables. They also compare their model-based estimates as predictors of later GDP releases with survey forecasts. They find that survey forecasts are far more accurate than model-based forecasts using for the second US GDP release, but that that the survey forecasts of the third US GDP release fail to draw on sources of information which could be tapped.

Casey and Smyth (2016) show that the revisions to real GDP and its components are found to be among the largest in the OECD, with the structure of the traded sector cited as a key source of the revisions. This analysis is echoed in earlier findings (Ruane, 1975; McCarthy, 2004; Bermingham, 2006; and Quill, 2008). Ireland therefore represents a useful testing ground in which to design effective strategies to overcome substantive data revisions.

The second key contribtion we make in this paper is in terms of using nowcasts as a predictor of revisions. In this respect, the literature is sparse. A brief exploration of this research question is offered in Chamberlin (2007), which outlines the use of principal components analysis as an informal check on the statistics produced by the Office for National Statistics in the UK.

A more rigorous approach to predicting data revisions using nowcasts is provided in Matheson, Mitchell and Silverstone (2009). Using panel data on business survey questions, they examine the out-of-sample GDP forecast performance of nowcasts over 52 quarters (25 quarters for manufacturing output). New Zealand’s national accounts data are used. The accuracy of their nowcasts is assessed with respect to initial (preliminary) and final actual outturns. Diebold–Mariano tests confirm no statistical difference (at

99% confidence) on an RMSE basis for their competing nowcasts with respect to either release.

The authors also explore the capacity of the nowcasts to predict revisions. This is pursued, given that they note their nowcasts show a closer relation to final rather than initial estimates of the official data. They proceed by testing the “news” versus “noise” hypothesis through Mincer–Zarnowitz tests along the lines of Faust, Rogers and Wright (2005). In effect, the tests examine whether the weight on a given nowcast should be zero implying that the nowcast offers no additional informational value relative to other estimates available, such as preliminary outturn data.

A key finding in their paper is that nowcasts are statistically helpful in explaining revisions to GDP growth. However, comparing this benefit against that provided by other simpler and more parsimonious methods, they note that the relative gains made by the nowcasts are small relative to initial estimates.

Our contribution is twofold: (1) we use a disaggregated approach to nowcasting output components based on a dynamic factor model; and (2) we show that the resulting estimates can be combined with preliminary official outturn estimates of output to help predict final output estimates. This allows us to exploit information not available in the initial estimates, but which can get us closer to final estimates.

Section 3: Methodology and Data

Our nowcasts are estimated using the popular two-step approach of Doz, Giannone and Reichlin (2011). This means estimating the dynamic factor model whereby we: first produce a balanced panel monthly dataset using the Kalman filter. The principal components method is then used to estimate the common and idiosyncratic factors in the data. Second, these factors are then used as regressors in an associated bridge equation, thus "bridging" the monthly factors to the quarterly national accounts series.

Starting with our unbalanced panel of time series, we wish to identify a common and an idiosyncratic component to the high frequency indicators so that we may use these to derive an estimate of current economic activity. The principal components method is an appealing dimension-reducing technique in this regard and is consistent as the cross-section and time dimension grow large. However, the principal components method requires a balanced dataset wherein the start and end points of the sample must be identical across all time series.

In practice, macroeconomic time series data are often released at different dates. To deal with this so-called "jagged edge" problem, whereby some time series do not have observations for the most recent months, we apply the standard two-step approach in the nowcasting literature as a means of completing the high frequency dataset available to us. The approach allows us to, first, construct a balanced panel of monthly indicators and, second, to get the monthly data to line up with the quarterly national accounts data.

Step 1: We cast our dynamic factor model in a state space representation. The state-space representation contains two equations: (i) a signal equation that links observed variables to latent states; (ii) a state equation that describes how the unobserved latent states evolve over time. The Kalman filter is a state-space model that can be used to provide mean-square optimal projections for both the signal and state variables. This approach allows us to produce smoothed estimates of the missing values of our high frequency variables, thus ensuring a balanced dataset. This is required so that we can use the method of principal components to extract common and idiosyncratic factors from our high frequency variables.

Step 2: we use the common and idiosyncratic factors produced in step 1 as regressors in a bridge equation. This bridges the monthly time series to the quarterly national accounts series. An overview of the two-step approach applied is provided below.

The Dynamic Factor Model

We wish to use high-frequency (monthly) data to estimate (or "nowcast") quarterly economic activity. We start with a $(n \times 1)$ vector (x_t) of N monthly, stationary and standardised time series available over time t that we believe will help us predict this quarterly economic activity series:

$$x_t = \left(x_{1,t}, x_{2t}, \dots, x_{n,t}\right)', \quad i = 1,2, \dots, N; \ t = 1,2, \dots, T$$

where each x_t variable is the standardised transformation of each raw $(x_{i,t}^{u})$ variables (typically the latter are in terms of year-on-year percentage changes to ensure stationarity):

$$x_{i,t} = (x_{i,t}^{u} - \bar{x}_{i,t}^{u}) / \sigma^{x^{u}_{i,t}}$$

The general idea of the dynamic factor model is one where unobserved orthogonal processes are assumed to drive a set of observable variables (Doz, Giannone and Reichlin, 2011). The observable variables are assumed to be made up of: (i) a common component driven by common shocks, which captures the bulk of the covariation between the time series; and (ii) an idiosyncratic component driven by n shocks, which generate series-specific dynamics.

Step 1: We cast our dynamic factor model in a state space representation.

Specifically, we assume that each observable variable $(x_{i,t})$ is the sum of the two independent and unobservable components: the common component c_t, and the idiosyncratic (or time series-specific) component $\xi_{i,t}$. Note that for simplicity we use vector notation and drop the subscript n for x_t, Λ and ξ_t. The model can be expressed as:

$$\begin{aligned} x_t &= \mathrm{c_t} + \xi_t \\ &= \Lambda F_t + \xi_t \end{aligned} \qquad (1)$$

where $x_t = \left(x_{1,t}, \dots, x_{n,t}\right)'$ is our stationary $(n \times 1)$ time series. The common component (c_t) is the product of the the $(n \times r)$ matrix of factor loadings (Λ) on our stationary $(r \times 1)$ vector of latent common factors (F_t):

$$c_t = (c_{1,t}, \dots, c_{N,t})'$$

$$\Lambda = (\lambda_1', \dots, \lambda_N')'$$

$$F_t = \left(f_{1,t}, \dots, f_{r,t}\right)'$$

The idiosyncratic component ξ_t is a stationary $(n \times 1)$ vector:

$$\xi_t = \left(\xi_{1,t}, \dots, \xi_{n,t}\right)'$$

The left-hand side of Equation (1) is observed; the right-hand side is unobserved. The idiosyncratic component is a multivariate white noise with diagonal covariance matrix Σ_ξ. In general, we assume that every $x_{i,t}$ is a weakly stationary process with mean zero.

Dynamics are introduced in the model via our latent common factors. Factor dynamics may be described by a Vector Autoregression with p lags, a "VAR(p)", of the form:

$$F_t = AF_{t-1} + Bu_t \ ; \qquad u_t \sim WN\left(0, I_q\right) \tag{2}$$

where A is a $(r \times r)$ matrix of parameters with all roots of $\det(I_r - A_z)$ outside the unit circle (i.e., a stable, stationary process); u_t is the q-dimensional white noise process of shocks to the common factors; and B is an $(r \times q)$ matrix of full rank q (i.e., each q column is linearly independent).

The idiosyncratic components are specified as cross-sectionally orthogonal white noise processes:

$$E\left(\xi_t {\xi'}_t\right) = \Psi_{t|v_j} = diag(\tilde{\psi}_{1,t}, \dots, \tilde{\psi}_{n,t}) \tag{3}$$

$$E\left(\xi_t {\xi'}_{t-s}\right) = 0, \quad s > 0 \tag{4}$$

It's also assumed that ξ_t is orthogonal to the common shocks u_t:

$$E(\xi_t {u'}_{t-s}) = 0, \quad \text{for all } s$$

To handle missing observations at the end of the sample, we parameterize the variance of the idiosyncratic component as

$$\tilde{\psi}_{i,t} = \begin{cases} \psi_i & \text{if } x_{it} \text{ is available} \\ \infty & \text{if } x_{it} \text{ is not available} \end{cases} \tag{5}$$

Since we assert that the variance of the idiosyncratic part of the time series with missing observations at the end of its sample is infinite at time t, this implies that no weight is put on the missing variable in the computation of the factors at time t. All errors are assumed as normal.

Using the principal components method on the observed variables, we can estimate the unobserved common factors F_t. The principal components are linear combinations of our original x_t series weighted by their contribution to explaining the variance in a particular orthogonal dimension. The objective here is dimension reduction. We wish to arrive at a smaller number of principal components from our larger number of initial variables. These principal components are uncorrelated with each other. The first principal component accounts for as much of the variability in the initial time series as possible, while succeeding principal components account for as much of the remaining variability as possible.

We combine the principal components method with Kalman filtering techniques. The Kalman smoother is used to recursively compute the expected value of the common factors. This framework allows us to estimate the factors even though we have missing values for data not yet released (i.e., an unbalanced panel). Principal components analysis requires a balanced panel. The Kalman filtering technique enables us to overcome this problem, effectively by smoothing through unavailable observations. In Appendix A, we show how conisistent estimates of the parameters of the model are obtained.

With our consistent estimates, we apply the Kalman filter to our state-space representation, replacing the initially estimated parameters:

$$\hat{F} = proj[F_t | x_1, \dots, x_T;\ \widehat{\Lambda}, \widehat{\mathrm{A}}, \widehat{\mathrm{B}}, \hat{\Sigma}_{\xi}] \tag{6}$$

The state-space representation is obtained by replacing estimated parameters in the factor representation:

$$x_t = \widehat{\Lambda} f_t + \xi_t \tag{7}$$

$$f_t = \sum_{j=1}^{p} \hat{A} f_{t-j} + \zeta_t$$

The Kalman filter can be used to evaluate the degree of precision of the factor estimates; to obtain estimates of the signal; and to obtain their degree of precision:

$$V_k = E\left[(F_t - \hat{F}_t)\left(F_t - \hat{F}_t\right) \mid x_1, \dots, x_T;\ \widehat{\Lambda}, \widehat{\mathrm{A}}, \widehat{\mathrm{B}}, \hat{\Sigma}_{\xi}\right], \tag{8}$$

$$X_t = Proj[\chi_t \mid x_1, \dots, x_T;\ \widehat{\Lambda}, \widehat{\mathrm{A}}, \widehat{\mathrm{B}}, \hat{\Sigma}_{\xi}] = \widehat{\Lambda}, \hat{\mathrm{F}}_t, \tag{9}$$

$$E(\chi_t - \hat{\chi}_t)^2 = \widehat{\Lambda}' V_0 \widehat{\Lambda} \tag{10}$$

Step 2: The Bridging Process:

To move from monthly to quarterly frequency, we define the level of economic activity (Y) for a quarter q as the average of the latent monthly observations of economic activity as measured within that quarter:

$$Y_t^q = \frac{1}{3}(Y_t^m + Y_{t-1}^m + Y_{t-2}^m) \tag{11}$$

where Y_t^m denotes the unobservable latent realisation of economic activity at the monthly frequency. Similarly, we average over the year-on-year monthly factors to obtain quarterly factors:

$$f_t^q = \frac{1}{3}(f_t^m + f_{t-1}^m + f_{t-2}^m) \tag{12}$$

The estimates of year-on-year changes in economic activity, on quarterly variables, are then computed with the following bridge equation:

$$\hat{Y}_t^q = \hat{B}'\hat{f}_t^q \tag{13}$$

where $\hat{Y}_t^q$ denotes the yearly estimated growth rate of economic activity and $\hat{B}$ is a vector of estimated parameters; computed for all the available t months in a certain quarter q. Nowcasts of the economic activity series can be computed every month as soon as new information becomes available.

Model Evaluation

To evaluate our model, we consider two comparators.

1. Naive Benchmark Model:

First, we estimate a naive AutoRegressive (AR(2)) model of output in real time. We recursively predict each component of output based on the latest historical data, adding one period at a time and using the relevant vintage of initial estimates for that period. This is used as our first comparator when assessing the real-time nowcasts.

2. Initial Estimates:

Second, we take actual initial (or real-time) outturn data for quarterly estimates of output published by the CSO. This serves as our second comparator for the real-time Nowcasts.

We compare our two benchmarks against our real-time Nowcasts in terms of their performance in predicting final outturns (i.e., estimates as of the latest available vintage of actual outturns for output).[2] We focus on year-on-year percentage changes and relative performances are considered in terms of the Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\widehat{y_t} - y_t)^2} \quad (14)$$

We also decompose the MSE into that part of the error which is attributable to the standard error (SE) of our estimates and that which is attributable to bias:

$$MSE = SE^2 + Bias^2 \quad (15)$$

where

$$SE^2 = \frac{1}{T}\sum_{t=1}^{T}(\widehat{y_t} - y_t - Bias)^2 \quad (16)$$

$$BIAS^2 = \left(\frac{1}{T}\sum_{t=1}^{T}\widehat{y_t} - y_t\right)^2 \quad (17)$$

Contributions to the RMSE in terms of bias and standard error are then given, respectively, by:

$$RMSE\ (BIAS) = \frac{Bias^2}{MSE}(RMSE)$$

$$RMSE\ (SE) = \frac{SE^2}{MSE}(RMSE)$$

Note that all of the tests above are applied on a real-time basis. For example, at each stage, we produce a real-time nowcast of output for a given quarter based on all of the high frequency data available up to and including the months within that quarter. This information is combined with the quarterly national accounts data published up to the point of the preceding quarter to arrive at our real-time Nowcast. The real-time Nowcast errors are then taken as the difference between each of these real-time Nowcasts and the final outturns.

Data

Given distortions to typical economic activity aggregates, we focus on a disaggregated measure of domestic economic activity. This focus is warranted for Ireland, given that standard measures of

---

[2] Note that we constrain the sample to include only final estimates that are published at least 8 quarters later than the initial release. Note also that we examine year-on-year changes as in D'Agostino, McQuinn and O' Brien (2012) owing to
the particularly volatile nature of Irish quarter-on-quarter GDP changes.

output such as GDP are especially volatile due to various distortions. Distortions to net exports related to multinational activities can often result in a misleading picture of domestic economic developments. The focus on disaggregated components of domestic activity allows us to look through these distortions as well as enabling us to assess the source of nowcast errors more systematically and to discern the drivers of economic activity.

We split out components of the volume (in € millions) of our quarterly underlying domestic demand measure (y_t^q) into the following:

$$y_t^q = c_t^q + ui_t^q + g_t^q \tag{18}$$

where c_t^q is personal consumption; ui_t^q is underlying investment and g_t^q is government consumption. The underlying investment we use is computed as gross fixed capital formation less investment in aircraft and less intangible assets. This adjusted investment measure is particularly important as a means of removing distortions associated with the activities of foreign-owned multinational enterprises.[3]

In estimating the latent factors, we focus on each of these components of domestic economic activity separately. Table 1 lists the variables used in the factor analysis for personal consumption; Table 2 for government consumption; and Table 3 for underlying investment.

[3] In particular, these adjustments account for the high degree of investment in almost wholly imported aircraft and intangible assets (e.g., royalties and licenses linked to intellectual property usage), which are largely neutral from an economic perspective (i.e., the imports laregely offset the associated investment activity).

Table 1: Variables in Factor Analysis for Personal Consumption

Sample: June 2000–Dec 2017 (211 monthly obs; spanning 70 quarters)

Variable	Freq.	Transformation	Source
Consumer Sentiment Index	M	% change y/y	ESRI
Consumer Sentiment (Conditions)	M	% change y/y	ESRI
Consumer Sentiment (Expectations)	M	% change y/y	ESRI
Retail Sales (Bars)	M	% change y/y	CSO
Retail Sales (Books)	M	% change y/y	CSO
Retail Sales (Clothes)	M	% change y/y	CSO
Retail Sales (Department Stores)	M	% change y/y	CSO
Retail Sales (Electricals)	M	% change y/y	CSO
Retail Sales (Food and Beverages)	M	% change y/y	CSO
Retail Sales (Furniture)	M	% change y/y	CSO
Retail Sales (Hardware)	M	% change y/y	CSO
Retail Sales (Motors)	M	% change y/y	CSO
Retail Sales (Non-Specialised Stores)	M	% change y/y	CSO
Retail Sales (Other)	M	% change y/y	CSO
Retail Sales (Pharmaceuticals)	M	% change y/y	CSO
Unemployment Rate (15-24)	M	% labour force 15-24	CSO
Unemployment Rate (25-74)	M	% labour force 25-74	CSO
Vehicles Licensed (2nd Hand)	M	% change y/y	CSO
Vehicles Licensed (New)	M	% change y/y	CSO
PMI Services	M	PMI – 50	Markit

Note: M = Monthly.

Table 2: Variables in Factor Analysis for Government Consumption

Sample: Jan 2004–Dec 2017 (168 monthly obs; spanning 56 quarters)

Variable	Freq.	Transformation	Source
Total Voted Current Expenditure	M	% change y/y	Department of Finance
Voted Current Education Expenditure	M	% change y/y	Department of Finance
Voted Current Health Expenditure	M	% change y/y	Department of Finance
Voted Current Justice Expenditure	M	% change y/y	Department of Finance
Voted Current Social Expenditure	M	% change y/y	Department of Finance

Note: M = Monthly.

Table 3: Variables in Factor Analysis for Underlying Investment

Sample: Jan 2005–Dec 2017 (156 monthly obs; spanning 52 quarters)

Variable	Freq.	Transformation	Source
Ireland Housing Completions	M	% change y/y	Dept. Environment
Ireland New House Guarantee Registrations	M	% change y/y	Dept. Environment
Net Imports of Road Vehicles	M	% change y/y	CSO
Net Imports of Machinery and Transport Equipment	M	% change y/y	CSO
PMI Construction	M	PMI – 50	Markit
Euro Area Industry Survey, Export Order Books	M	-	DG ECFIN

Note: M = Monthly.

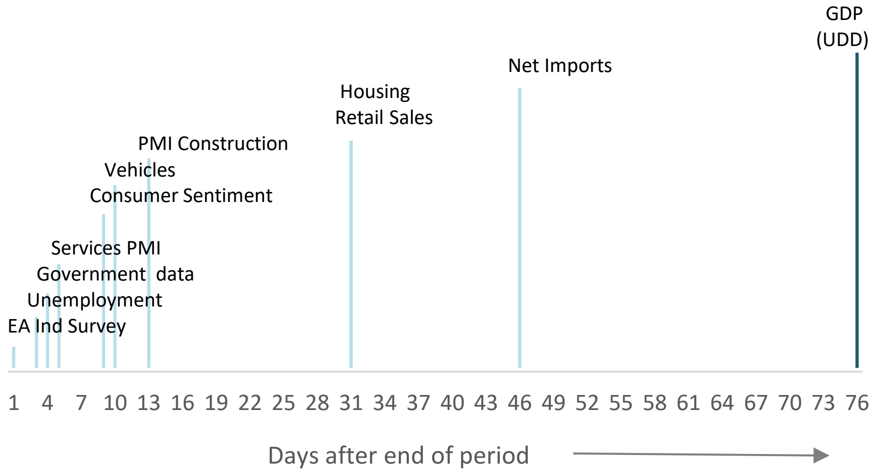
Note that the high-frequency variables that we avail of are not comprehensive. We avail of monthly indicators that bear a strong relationship to the national accounts subcomponents of interest and that are available over a reasonably long time horizon. We could extend the panel to include financial variables as in Hindrayanto, Koopman and De Winter (2016), for example.

In addition to the high frequency (monthly) data, we use quarterly national accounts estimates as published by the CSO. We focus on real Personal Consumption Expenditure (PCE), Government Consumption (Gov), and Underlying Investment (UI). The aggregate of these is our Underlying Domestic Demand measure (UDD). We construct a real-time dataset of the national accounts measures. This gives us both a series of “initial” and “final” estimates for each variable (where final are taken as the last observed series available at the time of writing: Q4 2017).[4]

[4] For UI, we are constrained by not having access to real-time data for investment in planes and intangible assets, such that real-time estimates of UI used prior to 2014 are the same as real-time estimate of overall investment (i.e., without these adjustments). This constraint is mitigated by the fact that divergences between real-time growth rates in the underlying and total invsestment series become most pronounced after 2013.

Figure 1: Timing of Releases Compared to Output

Days after end of period that the data release refers to



In terms of timing of releases, the monthly data that we use for our nowcast estimates are typically all available 46 days after the quarter ends. This compares to a lag of 76 days for preliminary national accounts estimates such as GDP or UDD (Figure 1).

Section 4: Results

In this section, we compare our nowcast estimates against the Benchmark (AR(2)) model estimates, and the initial national accounts estimates. Performance is judged on the basis of how accurate each of these estimates is in terms of predicting the final national accounts estimates.

Our real-time nowcasts are shown in Figure 2 (A–D) alongside each of the other estimates, including the final national accounts outturn estimates. Across the panels, we show results for each sub-component and for aggregate UDD. The aggregate is taken as the sum of each sub-component.[5]

Comparing the performance to that of initial outturns and estimates from a naive benchmark, we find that the nowcasts perform relatively well on a sub-component basis and better in the case of aggregate domestic demand.

For personal consumption expenditure, we find that the nowcasts show smaller errors with respect to final outturns than naive model estimates and even initial outturns themselves. The latter outperformance is explained by smaller bias in the nowcast estimates relative to initial outturns (Figure 4). As shown in Figure 3, initial outturns, by contrast, tend to overstate personal consumption expenditure over the period assessed.

For underlying investment, the nowcasts outperform naive model estimates but not the initial estimates. The RMSE for the nowcasts is 12.2 percentage points as compared to 9.9 percentage points for initial estimates and 14.3 percentage points for the naive AR model. Large bias is apparent in the nowcasts compared to other estimates.

For government consumption, we find that the nowcasts fare more poorly than both naive estimates and initial outturns, yet they display smaller bias than comparators. The weaker overall performance reflects the larger width of standard errors, albeit that the overall errors are not particularly biased in any given direction.

In terms of the aggregate UDD measure, the nowcasts are seen to actually outperform both the naive model estimates and the initial outturns themselves. This finding would suggest that the errors made for sub-components tend to cancel each other out such that –

[5] This is computed in terms of levels for the relevant vintage of data and then converted back to percentage year-on-year changes.

on average – the nowcasts outperform even the initial estimates. Figure 5 highlights the role played by bias correction.

Figure 2: Nowcasts, Benchmark Estimates, and Outturns

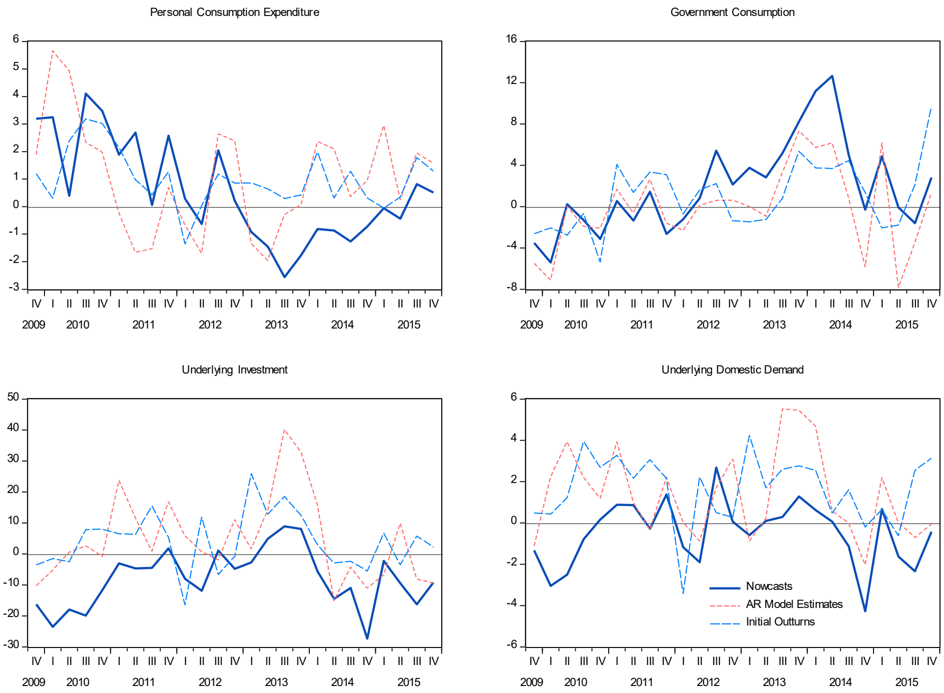
% change year-on-year

Source: Own workings.

Notes: "Final Outturns" are outturns as of Q4 2015 so that these incorporate any revisions over at least 8 quarters after initial estimates become available.

Figure 3: Year-on-Year Growth Rate Errors for Nowcasts, Benchmark Estimates and Initial Outturns with Respect to Final Outturns

Percentage points (forecast or initial estimate minus final outturn)

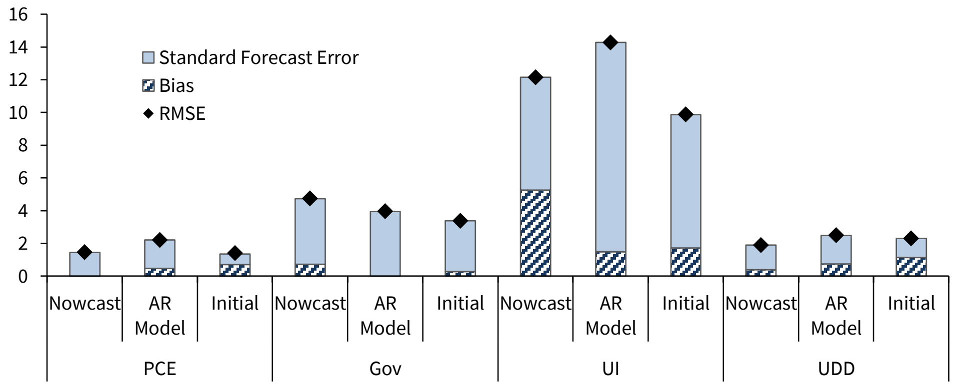


Source: Own workings.

Notes: “Final Outturns” are outturns as of Q4 2015 so that these incorporate any revisions over at least 8 quarters after initial estimates become available.

Figure 4: Decomposition of Errors for Each Component

RMSE (Q4 2009–Q4 2015, spanning 25 quarters)



Source: Own workings.

Figure 5: Distribution of Year-on-Year Growth Rate Errors for Nowcasts, Benchmark Estimates and Initial Outturns with Respect to Final Outturns

Density

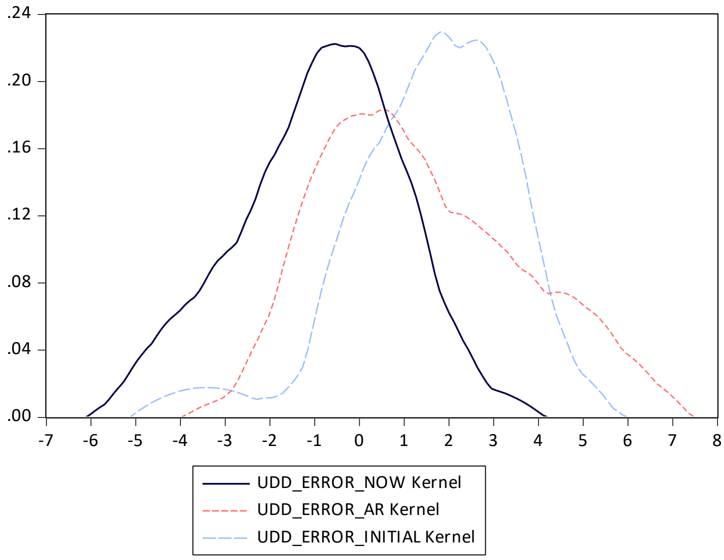


Table 4: Decomposition of Errors

Sample: Q4 2009–Q4 2015 (spanning 25 quarters)

Variable		RMSE	Bias[1]	Standard Error[1]
PCE	Nowcast	1.5	0.0	1.5
	AR Model	2.2	0.5	1.7
	Initial	1.4	0.7	0.7
Gov	Nowcast	4.7	0.7	4.0
	AR Model	4.0	0.0	4.0
	Initial	3.4	0.3	3.1
UI	Nowcast	12.2	5.3	6.9
	AR Model	14.3	1.5	12.8
	Initial	9.9	1.7	8.2
UDD	Nowcast	1.9	0.4	1.5
	AR Model	2.5	0.7	1.7
	Initial	2.3	1.1	1.2

Sources: Own workings.

[1] Contributions to RMSE.

We next explore more formal tests of the informational value of the nowcasts in terms of predicting data revisions. The Mincer–Zarnowitz test is one such test that examines whether there is a predictable or systematic component to revisions that can be exploited. We assert as a null hypothesis that the "news" in the first estimate is an efficient forecast of the final estimate. To test this, we examine, via a Wald or F-test robust to serial correlation and heteroscedasticity, the joint hypothesis that $\alpha = 0$ and $\beta_1 = 0$ in the regression $Rev_t = \alpha + \beta_1 Initial_t + \varepsilon_t$, where Rev_t is the revision relevant for time period t, defined as the difference between the final and initial official outturn. Under the “noise” model, the initial estimate helps predict the subsequent revision, implying a rejection of the null. The “news” model, on the other hand, implies that any extraneous information known at the time the preliminary estimate was formed should be orthogonal to the revision.

As in Matheson, Mitchell and Silverstone (2009), we also examine two additional variants of the Mincer-Zarnowitz test. The second set of regressions simply re-express our first regressions. They take the form of traditional forecast combination regressions, and indicate the optimal weights on the competing nowcasts. This lets us test whether any given nowcast is encompassed by the others. The third regression is also a forecast combination regression. It is relevant for the case when the user is interested in predicting the first estimate rather than its revisions.

This leaves us with the following two Mincer-Zarnowitz test regressions on revisions (MZ1 (a) and (b)):

$$Rev_t = \alpha + \beta_1 Initial_t + \varepsilon_t \qquad \text{MZ1 (a)}$$

$$Rev_t = \alpha + \beta_1 Nowcast_t + \varepsilon_t \qquad \text{MZ1 (b)}$$

along with the extended Mincer-Zarnowitz test regressions on final outturns (MZ2 (a) and (b)):

$$Final_t = \alpha + \beta_1 Initial_t + \beta_2 Nowcast_t + \beta_3 Naiive_t + \varepsilon_t \qquad \text{MZ2 (a)}$$

$$Final_t = \alpha + \beta_2 Nowcast_t + \beta_3 Naiive_t + \varepsilon_t \qquad \text{MZ2 (b)}$$

Table 5 summarises the results. The Mincer-Zarnowitz test regressions (MZ1 a–b) indicate that initial outturn estimates are, on average, efficient. The *F-test* in each case does not reject the null that α and β_1 are jointly equal to zero. Looking at the nowcasts, equation MZ2(b) suggests that these offer additional informational value over and above that provided by the naive model estimates. Furthermore, equation MZ2(a) suggests that when initial outturn estimates are included in the regression on final estimates, the nowcasts are still statistically significant at the 5 per cent level.

Table 5: Predicting Revisions and Final Outturns

MZ1 shows OLS regressions of the revision (between the initial and final outturns), while MZ2 shows OLS regressions of the final outturns.

	MZ1 (Dependent = Rev_t)		MZ2 (Dependent = $Final_t$)	
	(a)	(b)	(a)	(b)
α	1.47 (0.36)***	1.60 (0.39)***	0.63 (0.36)	-0.15 (0.29)
Initial	-0.10 (0.08)		0.46 (0.14)***	
Nowcasts		0.02 (0.09)	0.38 (0.17)**	0.67 (0.22)***
AR Model			0.16 (0.16)	0.32 (0.18)*
R^2	0.03	0.00	0.89	0.85
F	1.72 (0.20)	0.03 (0.86)	72.0 (0.00)	70.6 (0.00)

Sources: Own workings.

Notes: For coefficients, HAC robust standard errors are shown in parentheses. F-tests of the null hypothesis that the coefficients in each regression are zero are also presented, with p-values in parenthesis: in the MZ1 regressions, this F-test is a test of "news". R^2 reports adjusted r-squared statistics. Statistical significance: *** 1 per cent; ** 5 per cent; * 10 per cent.

Our results suggest that nowcasting may prove an effective tool to help cope with substantive data revisions in addition to providing timely estimates of current economic developments. This is particularly true when it comes to correcting for any bias that might exist in initial estimates. Though they are unlikely to ever represent a strong substitute for the initial estimates provided by statistical offices, nowcasts may act as a useful complement. This may be exploited to improve the information set one has when assessing current economic developments.

Nowcast-Augmented Initial Estimates

Exploring our results further, we examine an approach where one uses the information content in nowcasts to predict the sign of revisions to initial estimates.

As an initial investigation, we examine a simple approach. If the initial estimate is lower than the nowcast, we take it that the forecaster would predict an upward revision to the official initial estimate. Correspondingly, if the nowcast is lower, we assert that the forecaster would predict a downward revision. Using this simple heuristic would yield an 80 per cent success rate in terms of predicting the direction of revisions for estimates of personal consumption expenditure growth, 60 per cent for government consumption, and 68 per cent for underlying investment (Q4 2009 – Q4 2015).

Building on this, we propose a simple algorithm through which we may augment initial estimates wih the information contained in our nowcasts. The resulting “Nowcast-Augmented Initial Estimates” ($NAIE_t$) for each period are computed as:

$$NAIE_t = \begin{cases} Initial_t & if\ \ |Nowcast_t - Initial_t\ | < typerror \\ Initial_t + \dfrac{Nowcast_t - Initial_t}{|Nowcast_t - Initial_t\ |} * typerror & if\ \ |Nowcast_t - Initial_t\ | > typerror \end{cases}$$

where $\mathrm{Initial_t}$ are our initial outturns; $\mathrm{Nowcast_t}$ are our nowcasts and $\mathrm{typerror}$ is the typical error (RMSE) between initial and final outturns for a pre-defined historical period. Effectively, this approach means that we increase or decrease the initial estimate by the typical size of revisions whenever we judge that the nowcast has departed from our initial estimates by an amount greater than the typical revision size. Note that this approach only partially accounts for the magnitude of the nowcast (i.e., the magnitude of the nowcast estimate only matters insofar as it either breaches the threshold of a typical error with respect to initial estimates or not). More complex approaches could account more directly for the values of the nowcasts.

We compare the performance of (i) our Initial estimates and (ii) our nowcast-augmented initial estimates against final estimates as before. We find that for all components, aside from government consumption, the RMSE is improved by augmenting initial estimates with information from our nowcasts. In the case of the aggregate UDD estimates, the RMSE is lowered from 2.3 to 1.6 when using this simple algorithm (Table 6). This performance is better than that for the Initial Estimates. This suggests that even a relatively crude approach to determining the magnitudes of adjustments made to initial outturns as applied in our simple algorithm is not an overwhelming problem. With the standard error contribution to the RMSE rising, the reduction in the RMSE is primarily driven by a reduction in the contribution of bias to errors (falling from 1.1p.p. to 0.1p.p. in the case of the UDD aggregate). Looking at the sub-components, we see that this reduction in bias is reflected in personal consumption expenditure and in underlying investment.

Table 6: Performance of Initial Estimates vs Nowcast-Augmented Initial Estimates

Sample: Q4 2009–Q4 2015 (spanning 25 quarters)

Variable	RMSE	Bias[1]	Standard Error[1]
PCE Initial Estimates	1.4	0.7	0.7
PCE Nowcast-Augmented Initial Estimates	1.3	0.1	1.2
GOV Initial Estimates	3.4	0.3	3.1
GOV Nowcast-Augmented Initial Estimates	3.6	0.5	3.1
UI Initial Estimates	9.9	1.7	8.2
UI Nowcast-Augmented Initial Estimates	7.9	0.0	7.9
UDD Initial Estimates	2.3	1.1	1.2
UDD Nowcast-Augmented Initial Estimates	1.6	0.1	1.5

Sources: Own workings.

[1] Contributions to RMSE.

On an absolute error basis and on an RMSE basis, Diebold–Mariano tests confirm that we can reject the hypothesis that the initial outturns have the same predictive power as the nowcast-augmented initial estimates of UDD at the 5 per cent level of significance.[6] This suggests that the superiority of the nowcast-augmented initial estimates may not be simply the product of statistical chance.

Robustness Check: Extending the Sample Period

As a robustness check, we examine an extended sample period for the largest subcomponent: personal consumption expenditure.

We explore two sample extensions:

- Extended Sample 1: an extension that includes all of the same factor variables for personal consumption as shown in Table 1. This allows for a real-time performance assessment from Q1 2005 to Q4 2015 spanning 44 quarters.

- Extended Sample 2: a longer extension, with a more limited set of variables. For the factor analysis, we drop the

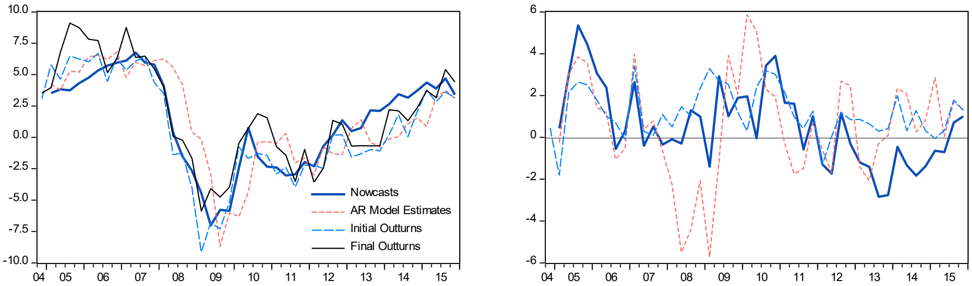
[6] Diebold-Mariano tests examine whether two competing forecasts have equal predictive accuracy. The tests take the null hypothesis of equal expected loss valid under quite general conditions including, for example, wide classes of loss functions and forecast-error serial correlation of unknown form (Diebold, 2013). The test can be summarised as an asymptotic z-test of the hypothesis that the mean of the loss differential is zero. We test the null hypothesis that the nowcast-augmented initial estimates have the same accuracy as the initial outturns, i.e., that the forecast overperformance is not due to statistical chance. The results are also robust to a correction for small-sample bias as in Harvey *et al* (1997).

sentiment indices, the PMI index, the unemployment rates, and the second-hand vehicles statistics. We combine this with longer run quarterly national accounts data.[7] This enables us to extend our data back to the beginning of 1980 and allows for a real-time performance assessment from Q1 1999 to Q4 2015 spanning 68 quarters.

As with our original sample, the nowcasts of personal consumption expenditure have lower errors than the AR model, but slightly larger than that of initial estimates (Table 7). However, the nowcasts can again be seen to substantially reduce bias when compared to the initial estimates. This is true for both of the sample extensions.

Figure 6: Extended Sample 1 Nowcasts and Comparators

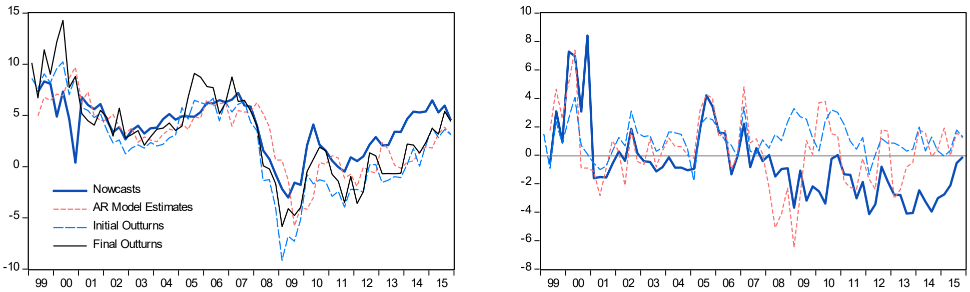
% change year-on-year (left panel); error in percentage points (right panel)



Source: Own workings.
Notes: Training sample: Q2 2000 to Q4 2004; out-of-sample tests with real-time data: Q1 2005 to Q4 2015.

Figure 7: Extended Sample 2 Nowcasts and Comparators

Percentage points (forecast or initial estimate minus final outturn)



Source: Own workings.
Notes: Training sample: Q1 1980 to Q4 1998; out-of-sample tests with real-time data: Q1 1999 to Q4 2015.

[7] National Accounts data sourced from Thomson Reuters Datastream.

Table 7: Decomposition of Errors for Personal Consumption Expenditure with Extended Samples

Variable		RMSE	Bias[1]	Standard Error[1]
Original Sample	Nowcast	1.5	0.0	1.5
	AR Model	2.2	0.5	1.7
	Initial	1.4	0.7	0.7
Extended Sample 1	Nowcast	2.0	0.2	1.8
	AR Model	2.6	0.1	2.5
	Initial	1.6	0.8	0.8
Extended Sample 2	Nowcast	2.6	0.1	2.5
	AR Model	2.5	0.0	2.4
	Initial	1.7	0.8	0.9

Source: Own workings.

Repeating our Mincer Zarnowitz tests as before, the nowcasts are found to not be able to provide significant additional information beyond that contained in the initial estimates or more rudimentary naive models (Table 8). This suggests that a standard forecast combination regression (including initial estimates and nowcasts) might not help us to accurately predict final outturns.

Table 8: Predicting Revisions and Final Outturns

MZ1 shows OLS regressions of the revision (between the initial and final outturns), while MZ2 shows OLS regressions of the final outturns.

	MZ1 (Dependent = Rev_t)		MZ2 (Dependent = $Final_t$)	
	(a)	(b)	(a)	(b)
Extended Sample 1				
α	1.23 (0.21)***	1.31 (0.23)***	1.23 (0.26)***	0.56 (0.43)
Initial	-0.07 (0.06)		0.94 (0.12)***	
Nowcasts		-0.08 (0.05)	-0.10 (0.12)	0.74 (0.13)***
AR Model			0.10 (0.05)***	0.26 (0.16)
R^2	0.07	0.05	0.92	0.79
F	3.18 (0.08)	3.12 (0.08)	161.5 (0.00)	81.6 (0.00)
Extended Sample 2				
α	1.23 (0.20)***	1.44 (0.34)***	1.34 (0.42)***	-1.67 (0.45)***
Initial	-0.04 (0.05)		0.97 (0.11)***	
Nowcasts		-0.08 (0.07)	-0.05 (0.09)	0.79 (0.15)***
AR Model			0.01 (0.09)	0.66 (0.16)***
R^2	0.01	0.02	0.92	0.80
F	1.75 (0.19)	2.16 (0.14)	244.7 (0.00)	129.6 (0.00)

Sources: Own workings.
Notes: For coefficients, HAC robust standard errors are shown in parentheses. F-tests of the null hypothesis that the coefficients in each regression are zero are also presented, with p-values in parenthesis: in the MZ1 regressions, this F-test is a test of "news". R^2 reports adjusted r-squared statistics. Statistical significance: *** 1 per cent; ** 5 per cent; * 10 per cent.

The nowcasts, in and of themselves, might not be sufficient to improve our knowledge of final outturns, given their larger standard errors, but the reduction in bias suggests some useful information may be captured. As before, we explore whether the information contained in the nowcasts might be used to augment the initial estimates.

In this spirit, we can see that the direction of revisions to initial estimates implied by the nowcasts is correct 65.4 per cent of the time for the extended sample 1, and 76 per cent of the time for the extended sample 2 (Table 9). By comparison, the naive AR model is correct 53.8 per cent and 64 per cent of the time, respectively.

Table 9: Revisions to Initial Estimates – % Correct Sign Predictions

Variable		% Correct
PCE (Extended Sample 1)	Nowcast	65.4
	AR Model	53.8
PCE (Extended Sample 2)	Nowcast	76.0
	AR Model	64.0

Source: Own workings.

We apply the same approach as before to augment our initial estimates using our nowcasts (Table 10). The nowcast-augmented initial estimates are shown to reduce bias again over the extended sample periods, though larger standard errors contribute to an overall RMSE that is not improved relative to initial estimates. The results suggest that, while nowcasts can be helpful for inferring the sign of revisions to initial estimates, a means of determining the exact magnitudes of any revisions using nowcasts might need better calibration in future modelling.

Table 10: Performance of Initial Estimates vs Nowcast-Augmented Initial Estimates

Variable	RMSE	Bias[1]	Standard Error[1]
Original Sample			
PCE Initial Estimates	1.4	0.7	0.7
PCE Nowcast-Augmented Initial Estimates	1.3	0.1	1.2
Extended Sample 1			
PCE Initial Estimates	1.6	0.8	0.8
PCE Nowcast-Augmented Initial Estimates	2.1	0.4	1.7
Extended Sample 2			
PCE Initial Estimates	1.7	0.8	0.9
PCE Nowcast-Augmented Initial Estimates	1.5	0.1	1.4

Sources: Own workings.
[1] Contributions to RMSE.

Diebold and Mariano tests indicate that we are able to reject the null hypothesis of the forecasts having identical accuracy for extended sample 2. Further, they indicate that the initial estimates are less accurate than the nowcast-augmented initial estimates at the 5 per cent level of significance. For the shorter sample extension, we are not able to reject the hypothesis that the initial outturns have the same predictive power as the nowcast-augmented initial estimates of personal consumption expenditure (on an absolute error and RMSE basis). Thus, we can reasonably posit that the nowcast-augmented initial estimates are at least as accurate as initial estimates at predicting final outturns, with some evidence pointing to them being more accurate.

Section 4: Conclusions

Initial estimates of economic activity come with fairly large time lags and can be prone to large revisions. Findings by Stark and Croushore (2002) in relation to US forecasts suggest that data revisions may not be just another consideration in forecasting, rather they may be the major source of forecast uncertainty and one which is frequently ignored. Nowcasts present a potentially useful and timely substitute for preliminary national accounts estimates, which we show can be used to help deal with data revisions.

This paper makes two major contributions to the literature on nowcasting:

First, we produce nowcasts of sub-components of national accounts data for Ireland that allow us to develop nowcasts of aggregate domestic economic activity. This focus is warranted for Ireland, given that GDP is especially volatile, with distortions to net exports associated with multinational activities often resulting in a misleading picture of domestic economic developments when looking solely at GDP or similar aggregates. The focus on disaggregated components of domestic activity also allows us to assess the source of nowcast errors more systematically and to discern the drivers of economic activity.

Second, we show how nowcasts can be used to augment preliminary national accounts estimates in order to better predict final national accounts outturns (i.e., the revised "final" estimates that come at least two years later). We construct real-time nowcasts of domestic economic activity for Ireland based on real-time national accounts data and real-time high frequency indicators. We assess our nowcasts of domestic economic activity and the initial estimates produced by the Central Statistics Office (CSO) as predictors of the final revised national accounts outturns. Comparing the performance to that of initial outturns and estimates from a naive benchmark, we find that the nowcasts perform relatively well on a sub-component basis and better for aggregate domestic demand. The better performance at aggregate level reflects the fact that sub-component errors effectively cancel each other out such that – on average – the nowcasts outperform even initial estimates.

Comparing the performance of the nowcasts to that of initial outturns and estimates from a naive benchmark, we find that the nowcasts perform relatively well on a sub-component basis and better when examining aggregate domestic demand. The aggregate result reflects the fact that bias is substantially reduced relative to that present in initial estimates. On average, nowcasts can be seen to outperform even initial estimates.

Exploring our results further, we combine the information content in our nowcasts with a simple algorithm to arrive at nowcast-augmented initial estimates of economic activity. We compare the performance of (i) our Initial estimates and (ii) our nowcast-augmented initial estimates against final estimates. Our findings suggest that for all components, aside from government consumption, the errors with respect to final outturn estimates are improved by augmenting initial estimates with information from our nowcasts.

Extending the sample period assessed for one of the subcomponents studied – personal consumption expenditure – we, again, find that augmenting initial estimates with information from nowcasts can help to substantially reduce bias present in initial estimates. However, we do not find that it reduces the overall size of errors with respect to final outturns.

The results we present suggest that nowcasts can provide an accurate, timely and less biased substitute for preliminary national accounts estimates. Comparing their performance in terms of predicting final revised outturn estimates against the preliminary national accounts estimates, we find some evidence that nowcasts can help to discern a better picture of what final outturns may look like. In particular, nowcasts can help to alleviate any bias that might be contained in preliminary national accounts estimates. In cases where there are large divergences between nowcasts and preliminary estimates, we find that the implied direction of any revision can be helpfully inferred by augmenting the preliminary estimates with the information from our nowcasts.

Further extensions of this work could seek to explore better ways to predict the exact magnitudes of revisions to initial estimates, while drawing on the information available in nowcasts.

References

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-Casting and the Real-Time Data Flow. ECB Working Paper Series, No. 1564, July 2013.

Bermingham, C., (2006). An Examination of Data Revisions in the Quarterly National Accounts. Central Bank of Ireland, Research Technical Paper, 10/RT/06.

Casey, E., and D. Smyth (2016). Revisions to Macroeconomic Data: Ireland and the OECD. The Economic and Social Review, Vol. 47 (Spring), pp. 33–68.

Chamberlin, G. (2007). Forecasting GDP using external data sources. Economic & Labour Market Review, 1, 18-23.

Clements, M. P., and A. B. Galvão (2017). Predicting early data revisions to US GDP and the effects of releases on equity markets. Journal of Business & Economic Statistics, 35 (3), 389-406.

Clements, M. P., and A. B. Galvão (2013). Forecasting with vector autoregressive models of data vintages: US output growth and inflation. International Journal of Forecasting, 29 (4), 698-714.

D'Agostino, A., K. McQuinn, and D. O'Brien (2012). Nowcasting Irish GDP. OECD Journal: Journal of Business Cycle Measurement and Analysis, 2012, Issue 2, 21-31.

Diebold, F.X. (2013). Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests. Journal of Business & Economic Statistics, January 2015, 33:1, 1-1.

Doz, C., D. Giannone, and L. Reichlin (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. Journal of Econometrics, 164(1), 188–205.

Faust, J., Rogers, J. H., and Wright, J. H. (2005). News and noise in G-7 GDP announcements. Journal of Money, Credit and Banking, 37 (3), 403-417.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003). Do financial variables help forecasting inflation and real activity in the euro area?. Journal of Monetary Economics, 50(6), 1243–1255.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics, 55(4), 665–676.

Harvey, D., S. Leybourne, and P. Newbold (1997). Testing the Equality of Prediction Mean Squared Errors. International Journal of Forecasting, 13, pp. 281–91.

Hindrayanto, A. I. W., Koopman, S. J., & de Winter, J. (2016). Forecasting and nowcasting economic growth in the euro area using factor models. International Journal of Forecasting, 32(4), 1284-1305.

Matheson, T.D., J. Mitchell and B. Silverstone (2010). Nowcasting and predicting data revisions using panel survey data. 29 (3), April 2010, 313-330.

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

McCarthy, C. (2004). Volatility in Irish Quarterly Macroeconomic Data. Article in The Quarterly Economic Commentary, Dublin: The Economic and Social Research Institute, Spring.

Quill, P. (2008). An Analysis of Revisions to Growth Rates in the Irish Quarterly National Accounts. Special Article in The Quarterly Economic Commentary, Dublin: The Economic and Social Research Institute, Autumn, 44-59.

Ruane, F. (1975). Statistical Analysis of Revision to Irish National Accounts. The Economic and Social Review, Vol. 6, No. 3.
Stark, T. and D. Croushore, (2002). Forecasting with a Real-Time Data Set for Macroeconomists. Journal of Macroeconomics, Vol. 24 (December), pp. 507−531.

Stock, J. H., and M. W. Watson (2003). Forecasting Output and Inflation: The Role of Asset Prices. Journal of Economic Literature, 41(3), 788–829.

Appendix A

This appendix shows how we obtain consistent estimates of the parameters of the model.

As in D'Agostino, McQuinn and O' Brien (2012), we consider the estimator of common factors as:

$$(\tilde{F}_t, \widehat{\Lambda}) = argmin_{F_t,\Lambda} \sum_{t=1}^{T} \sum_{i=1}^{n} \left(z_{it-\Lambda_i F_t}\right)^2$$

supposing that $z_{it} = y_{it} - \hat{\mu}_i$ and that $x_{it} = 1/\hat{\sigma}_i(y_{it} - \hat{\mu}_i)$, where $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} y_t$ and $\hat{\sigma}_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{\mu}_i)^2}$.

We can define the correlation matrix of observed variables (y_t) as:

$$S = \frac{1}{T}\sum_{t=1}^{T} x_t x'_t$$

If we define D the $r \times r$ diagonal matrix with diagonal elements given by the r largest eigenvalues of S and V the $n \times r$ matrix of the corresponding eigenvectors subject to the normalisation $VV' = I_r$. Factors are estimated as $\tilde{F}_t = \Lambda x_t$ and factor loadings $\widehat{\Lambda}$ are estimated by regressing the variables on the estimated factors:

$$\widehat{\Lambda} = \sum_{t=1}^{T} x_t \tilde{F}_t' (\sum_{t=1}^{T} \tilde{F}_t \tilde{F}_t')^{-1}$$

and the covariance matrix of the idiosyncratic component is estimated as:

$$\hat{\Sigma}_{\xi} = diags(S - VDV)$$

The other parameters $\hat{A}$ and Σ and are estimated by running a VAR on the estimated factors:

$$\hat{A} = \sum_{t=2}^{T} \tilde{F}_t \tilde{F}_{t-1}' \left(\sum_{t=2}^{T} \tilde{F}_{t-1} \tilde{F}_{t-1}'\right)^{-1}$$

$$\hat{\Sigma} = \frac{1}{T-1}\sum_{t=2}^{T} \tilde{F}_t \tilde{F}_t' - \hat{A}(\frac{1}{T-1}\sum_{t=2}^{T} \tilde{F}_{t-1} \tilde{F}_{t-1}')\hat{A}'$$

Finally, P is defined as the $q \times q$ diagonal matrix with entries given by the largest q eigenvalues of $\hat{\Sigma}$ and by M the $r \times q$ matrix of the corresponding eigenvectors, then:

$$\hat{B} = MP^{\frac{1}{2}}$$