Globalisation and density forecasts of Euro-area inflation from Phillips Curve models

GIAN LUIGI MAZZI AND JAMES MITCHELL

2016 edition
Globalisation and density forecasts of Euro-area inflation from Phillips Curve models
GIAN LUIGI MAZZI AND JAMES MITCHELL
2016 edition
Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*):

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).


ISSN: 2315-0807
doi: 10.2785/395084
Cat. No: KS-TC-16-015-EN-N

© European Union, 2016

Reproduction is authorised provided the source is acknowledged.
For more information, please consult: http://ec.europa.eu/eurostat/about/our-partners/copyright
Copyright for the photograph of the cover: ©Shutterstock.
For reproduction or use of this photo, permission must be sought directly from the copyright holder.

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.
# Table of contents

Abstract ............................................................................................................................................. 4  
1 Introduction........................................................................................................................................ 5  
2 Globalisation, output gaps and inflation density forecasting.......................................................... 6  
  2.1 Introduction ...................................................................................................................................... 6  
  2.2 In search of an open-economy Phillips curve .................................................................................... 7  
  2.3 Indicators of globalisation .................................................................................................................. 8  
    2.3.1 KOF indicator .................................................................................................................................. 8  
    2.3.2 Foreign output: and the global output gap ....................................................................................... 10  
    2.3.3 Import and oil prices ..................................................................................................................... 11  
  2.4 The VAR model: common trends in the Euro area ......................................................................... 11  
    2.4.1 Cointegration .................................................................................................................................. 14  
    2.4.2 The multivariate BN decomposition ............................................................................................... 14  
    2.4.3 Bivariate Phillips Curve models ................................................................................................... 16  
    2.4.4 Combination forecasts .................................................................................................................. 17  
    2.4.5 Occam’s Window: excluding bad models and grand ensembles ............................................... 19  
  2.5 Density forecast evaluation .............................................................................................................. 20  
    2.5.1 The continuous ranked probability score (CRPS) ........................................................................... 21  
    2.5.2 Tail scores ..................................................................................................................................... 21  
  2.6 Empirical Results .............................................................................................................................. 22  
    2.6.1 Estimating the Euro area cycle ...................................................................................................... 22  
    2.6.2 Forecasting inflation in a globalised world ..................................................................................... 26  
  2.7 AR and Output Gap Ensembles ....................................................................................................... 28  
    2.7.1 Forecasting the GDP deflator ........................................................................................................ 28  
    2.7.2 Controlling for changes in import prices and oil prices ............................................................... 35  
    2.7.3 Grand ensemble ......................................................................................................................... 39  
    2.7.4 Forecasting the import price deflator ........................................................................................... 41  
  2.8 Conclusion ....................................................................................................................................... 45  

References .......................................................................................................................................... 46  

3 Appendix: Real-time data vintages .................................................................................................... 50  
  3.1 Real-time GDP data ......................................................................................................................... 50  
    3.1.1 Creation of real-time GDP data triangles ..................................................................................... 50
Abstract

In this paper we test whether globalisation has affected the relationship between the output gap and inflation in the euro area. Then we investigate how the traditional Phillips Curve relating the inflation and the domestic output gap can be enhanced incorporating globalisation effects. We analyse the effect of different specification of the Phillips Curve in an out-of-sample real-time simulation exercise, focusing on the inflation density forecast.

Keywords: globalisation, input output gap, business cycles

Acknowledgements:

Paper first presented at the 31st CIRET conference in Vienna (September 2012). The paper has benefitted from the outcomes of the multi-annual PEEIs project financed by Eurostat.

Authors:

Gian Luigi Mazzi (1) and James Mitchell (2)

(1) Gian Luigi Mazzi, Eurostat, Luxembourg; email: gianluigi.mazzi@ec.europa.eu.
(2) James Mitchell University of Warwick; email: James.Mitchell@wbs.ac.uk.
1 Introduction

We first consider how inference about the euro area business cycle, specifically the output gap, is affected by globalisation. Examination of traditional indicators of globalisation, such as the KOF indicator which is based on the share of trade in GDP, indicates that globalisation has increased dramatically in the last three decades. We consider contrasting means of accommodating any globalisation effects - namely of accommodating increased cross-sectional dependence across countries - when estimating and using business cycles. In particular, we focus on the alleged effect globalisation is said to have on the dynamics of Euro area inflation and its relationship with traditional (domestic) measures of the output gap.

Traditionally, inflation is modelled by macroeconomists with respect to the (domestic) output gap, often estimated using a univariate filter. But this relationship is well-known to be unstable and furthermore many believe it has weakened over time. Globalisation is often cited as one explanation for this. If a multivariate filter is estimated, it tends to be one which in addition to domestic GDP data considers other domestic variables, such as inflation and say unemployment. But increasingly, globalisation has led to interest in examining whether domestic inflation also has foreign determinants.

Accordingly in this paper we test whether globalisation has affected the relationship between the output gap and inflation in the Euro area. Has it led to the breakdown of the traditional Phillips curve relationship and the emergence of an open-economy (global) Phillips Curve where global output gaps, or other indicators of globalisation, affect domestic inflation? As well as constructing a global output gap estimate, and considering the informational content of US output gaps when forecasting Euro area inflation, we also consider whether import and oil prices have had an effect. Importantly we test for globalisation effects not just by looking at the central tendency of inflation, but its whole density.

As well as testing for globalisation effects by augmenting the traditional Phillips Curve with these globalisation indicators, and testing their significance in out-of-sample simulations, we also consider re-estimating the Euro area cycle itself but allowing for cross-country dependence. This is achieved by de-trending output using a multivariate estimator which allows for cross-country linkages. Specifically, we use a possibly co-integrating VAR model, in the GDP of each Euro area country. We also allow oil prices, US GDP or global output to enter the VAR as exogenous I(1) regressors. We look at the output gap in the European VAR as a function of the rank of the VAR - i.e. as a function of the number of exactly identified co-integrating vectors, the number of common trends.

We then propose a density forecast combination approach to forecast Euro area inflation using (a) Euro area output gaps; (b) Global output gaps and (c) import/oil prices. Thereby we integrate out uncertainty about both the preferred regressor when forecasting inflation and the form of its relationship with inflation.

Our analysis shows that domestic inflation over the last decade has clearly been influenced by foreign economic conditions, working principally through import and oil prices rather than global output gaps. But the domestic output gap has played a role since late 2006. Consideration of import and oil prices, in conjunction with domestic and global output gap data, does help deliver better fitting density forecasts of inflation than autoregressive models. Therefore it is important to allow for the process determining inflation to vary over time, as documented in other studies; but there is not clear evidence that it is important to consider global output gaps when forecasting Euro area inflation. This contrasts the widely cited findings of Borio & Filardo (2007).
2 Globalisation, output gaps and inflation density forecasting

2.1 Introduction

Traditionally, inflation is modelled by macroeconomists with respect to the (domestic) output gap, often estimated using a univariate filter. But this relationship is well-known to be unstable and furthermore many believe it has weakened over time. Globalisation is often cited as one explanation for this. If a multivariate filter is estimated, it tends to be one which in addition to domestic GDP data considers other domestic variables, such as inflation and say unemployment. But increasingly, globalisation has led to interest in examining whether domestic inflation also has foreign determinants. This has led Borio & Filardo (2007), Calza (2008) and others to add global output gaps to the inflation equation in an attempt to find a more stable relationship between the output gap and inflation. However, focus has remained on testing for a relationship under quadratic loss, usually based on tests of in-sample predictability based on OLS estimation. Studies do not seem to test for out-of-sample predictability by looking at RMSE ratios. Moreover, no attempt has been made to model the domestic output gap explicitly accommodating the cross-country dependencies implied by globalisation. In this paper we therefore look at density forecasts, and also in addition to forecasting inflation using domestic and global output gaps, based on considering domestic and foreign GDP data separately, consider estimating the domestic output gap using a multivariate model. This model allows domestic GDP to depend on foreign GDP data, and indeed on additional explanatory variables. We concentrate on the use of the Beveridge-Nelson decomposition as a means of identifying and estimating the output gap. The BN offers a unified means of extracting the cycle in univariate and multivariate models, and involves conditioning on different information sets when taking the long-run forecast of output which is used to define the trend component to GDP.

The motivation for considering this multivariate estimator of the output gap is based on the belief that these foreign variables may help explain domestic GDP growth. If they do, and output growth, to some extent, is forecastable, then there is a temporary component, i.e. there is an “output gap”, which in due course reverts to its unconditional mean (of zero). The output gap is larger the greater the degree of long-run forecastability: only when output growth is unforecastable does output follow a random walk and there is no output gap. When output growth, to some extent, is forecastable there is a temporary component, i.e. there is an output gap, which in due course reverts to its mean. When the output gap is positive (output is above trend) we are requiring future growth at a below trend rate. Analogously, if the output gap is negative recovery requires future growth at an above trend rate. As Nelson (2008) says: “predictability is the essence of transitory variation”. So the finding, in some studies, that global output gaps do matter may simply be a result of not estimating the domestic output gap allowing for the fact that domestic output depends on foreign output. That is, globalisation may mean that global output gaps affect domestic inflation; but it might mean that the domestic output gap does influence domestic inflation when one allows for linkages between domestic and foreign output. By considering both global and multivariate output gap estimates we attempt to distinguish these two effects.

Empirically we focus on testing whether Euro area inflation, as an aggregate, has been affected by the globalisation process. Previously Calza (2008) studied the Euro area, as did Borio & Filardo (2007) in their panel study. Our paper tests out-of-sample predictability using real-time data. In contrast the extant literature has considered in-sample tests for predictability using final-vintage data, despite the fact that we know the output gap in particular is subject to considerable measurement error in real-time due to both data revisions and the one-sided nature of the detrending filter when applied at the end of the sample. We also extend previous work by looking not just at the central forecast for inflation – but its uncertainty. This is important because it may well be that the output gap, whether measured domestically or globally, matters more for inflation uncertainty than the central estimate of inflation. Indeed globalisation is also often claimed to have reduced macroeconomic uncertainty. That is, globalisation may matter more for second moment rather than first moment inflationary dynamics. Importantly our methodology lets us test whether foreign
Globalisation, output gaps and inflation density forecasting

Factors are affecting inflation to an increasing extent over time. We ascertain whether the dynamics of inflation have changed over time.

We distinguish between forecasts of the GDP deflator and of the import price deflator. Globalisation might be thought to mean that the prices of domestically produced and imported goods are determined differently, with the traditional Phillips curve model only useful in explaining the evolution of the prices of domestically produced goods. The prices of imported goods are instead determined on world markets.

Our econometric methodology explicitly accommodates the fact that there is uncertainty about the determinants of domestic inflation. We adopt a density forecast combination approach which involves combining inflation density forecasts which exploit domestic, multivariate and global output gaps. While we do not seek to identify structurally the channels through which global factors may affect domestic inflation, from a practical perspective we seek to identify whether consideration of global variables helps deliver improved density forecasts in an out-of-sample context.

We find that global output gaps play little, if any, role in the determination of Euro area inflation, certainly independent of any role it might play in the determination of the Euro area output gap. But import and oil price do matter, though to varying degrees over time. Importantly, we find that consideration of Euro area gaps does help deliver improved density forecasts for inflation relative to benchmark AR models.

2.2 In search of an open-economy Phillips curve

Romer (1993) documents that there is a negative correlation between openness and inflation; he argues that as the degree of openness of an economy increases, the cost of inflationary surprises rises, thereby reducing the incentive to inflate. Borio & Filardo (2007) suggest that in an integrated global economy where many goods can be produced and consumed anywhere in the world, it is the global output gap, defined as the deviation of world demand from world potential supply, and measured by aggregating output gaps across different economies, that should be considered when assessing global influences on domestic inflation. Similarly Lucas Papademos, Vice President of the European Central Bank has said: “One conclusion that has some empirical support is that domestic inflation is no longer determined predominantly by domestic demand and supply constraints, but seems to depend more on the degree of global economic slack” (2006, page 6). In contrast, Ball (2006) writes “there is little reason to think that globalization has changed the structure of the Phillips curve or the long-run level of inflation” (Page 15). Ihrig et al. (2010) also find little evidence that globalization has increased the role of international factors and decreased the role of domestic factors in the inflation process in industrial economies.

Measures of the domestic output gap should already contain some information about global influences on domestic costs and prices in the sense that net exports are part of GDP. So strong global demand for goods/services produced in the domestic country should affect the domestic output gap. But a global output gap measure might contain additional information in two ways. First, a positive global output gap could contain information about future import price inflation; e.g. a positive global output gap could signal rising cost pressure in these economies, which could be transmitted to the EA via higher import prices. Secondly, the global output gap may affect the labour market, and potential supply in the domestic economy. For instance, spare capacity overseas could imply a low cost of replacing domestic workers with foreign workers and therefore be associated with weak bargaining power of domestic workers. These effects would make domestic labour supply and wages respond to the cyclical position of the world economy. They also imply that, with migrant labour flowing into and out of the domestic labour force, domestic potential output and the domestic output gap are less well defined.

As the introduction notes, there have been a number of attempts to isolate globalization’s effects on U.S. inflation. An early study was Tootell (1998), who asks whether globalisation could account for the missing inflation of the late 1990s in the US. Using a standard Phillips curve approach, he finds little evidence that globalisation—specifically, measures of foreign slack—helps determine U.S. inflation. Previously Calza (2008) and Borio & Filardo (2007) have studied the Euro area, with Calza (2008) again finding little evidence that global gaps matter.
2.3 Indicators of globalisation

However measured it is clear that globalisation has increased in recent decades. The production of many goods and services has become increasingly internationalised and the level of trade in goods and services has increased. The most commonly used measure of openness, which we might equate with globalisation, is the sum of imports and exports divided by gross domestic product; this is picked up by the KOF indicator, discussed further below.

Another aspect of globalisation is the growth in commodity prices estimated to have resulted from strong output growth in developing economies, and a decline in the average rate of non-commodity import price inflation that is estimated to have resulted from higher levels of trade with developing countries. More generally these price variables pick up supply-side shocks. In the analysis below we accordingly also consider measures of both oil and commodity prices.

2.3.1 KOF INDICATOR

To measure the degree of economic globalisation KOF measure actual flows: trade, foreign direct investment and portfolio investment (all in percent of GDP). Income payments to foreign nationals and capital employed (in percent of GDP) are included to proxy for the extent a country employs foreign people and capital in its production processes.

They also measure restrictions on trade and capital using hidden import barriers, mean tariff rates, taxes on international trade (as a share of current revenue) and an index of capital controls. Given a certain level of trade, a country with higher revenues from trade taxes is less globalised. Openness is measured based on the IMF’s Annual Report on Exchange Arrangements and Exchange Restrictions and includes 13 different types of capital controls. The index is constructed by subtracting the number of restriction from 13 and multiplying the result by 10.

Figure 1 plots the KOF indicators of (economic) globalisation for the nine European countries; these data are interpolated to the quarterly frequency, from the available annual estimates, by fitting a flexible UC model to each series separately. Across the 9 countries we clearly see that globalisation has increased, and appears to exhibit a stochastic trend. Interestingly, since about 2000 globalisation appears to have slowed according to this estimate.
Since this indicator is a compiled one, in our econometric work we seek to pick up the increased evidence for globalisation and international cross-sectional dependence by including foreign output, the so-called global output gap, in our models. Globalisation is believed to have increased cross-sectional dependence; therefore models which allows for this dependence internationally, perhaps by including a global output gap, are one means of implementing globalisation when modelling.
2.3.2 FOREIGN OUTPUT: AND THE GLOBAL OUTPUT GAP

Following Calza (2008) we measure foreign (as far as the Euro area is concerned) output based on the global VAR model of Dees et al. (2007) Foreign output is obtained by aggregating data on 25 foreign countries (Argentina, Australia, Brazil, Canada, Chile, China, India, Indonesia, Japan, Malaysia, Mexico, New Zealand, Norway, Peru, Philippines, Saudi Arabia, Singapore, South Africa, South Korea, Sweden, Switzerland, Thailand, Turkey, UK and US) using trade weights derived from bilateral trade volumes. These weights can be seen as a measure of economic distance.

Figure 2 provides estimates of the global (foreign) output gap. It can be seen that apart from a sharp downturn in 2008, the global output gap has experienced little persistence and small amplitude.

*Figure 2: The global output gap: BN estimates of the output gap for \( y^* \) using AR(8) and ARIMA(1,1) specifications*

It is interesting to compare the global output gap estimates with those for the EA9, computed using a univariate BN decomposition via both the direct and indirect approaches. In both cases estimates using AR(8) and ARMA(1,1) specifications are provided. Comparison of the top and middle panels shows that the European and global cycles appear quite similar, in the sense that both are dominated by the recession of 2007-9. But it is also seen, and this is clearer in the bottom panel of Figure 2 which excludes the recessionary period, that from 1980-2006 the global cycle and European cycles do exhibit differences, especially in the early 1990s and again in the 2000s when the global output gap is more positive than the European one.
2.3.3 IMPORT AND OIL PRICES

Globalisation is commonly claimed to be reflected in asset price inflation, particularly growth in commodity prices, such as oil. Accordingly, another possible means of accommodating globalisation in business cycle analysis is to incorporate import and oil price variables into the model. Below we describe how this can be done.

2.4 The VAR model: common trends in the Euro area

Consider the following VAR model of order \( p \), \([\text{VAR}(p)]\) in the \( m \)-vector process \( \{z_t\}_{t=1}^T \)

\[
\Phi(L)(z_t - \mu - \gamma t) = u_t, \tag{1}
\]

for \( t = 1, 2, ..., T \), where \( \Phi(L) \) is a \((m \times m)\) matrix polynomial such that \( \Phi(L) = I_m - \sum_{j=1}^{p} \Phi_j L^j \), \( L z_t = z_{t+j} \), \( \mu \) and \( \gamma \) are \( m \)-vectors of intercepts and trend coefficients and \( u_t \) is a serially uncorrelated \( m \)-vector of errors with a zero mean and a constant positive definite variance-covariance matrix, \( \Sigma = (\sigma_{ij}) \).
From (1) the more conventional representation of the VAR model is derived

\[ z_t = a_0 + a_1 t + \Phi_1 z_{t-1} + \ldots + \Phi_p z_{t-p} + u_t, \]  

where the intercept and trend coefficients, \( a_0 \) and \( a_1 \), are given by

\[ a_0 = -\Pi \mu + (\Gamma + \Pi) \gamma, \]  
\[ a_1 = -\Pi \gamma, \]  

such that \( \Gamma = -\Pi + \sum_{i=1}^{p} \Phi_i \) and \( \Pi = \Phi(1) \). \( a_0 \) and \( a_1 \) are restricted when \( \Pi \) is rank deficient. This proves important when there is cointegration.
Two alternative but equivalent representations of the VAR model that prove useful are the Vector Error Correction Model [VECM] and the Vector Moving Average [VMA] representation:

1. The VECM representation of \( \{z_t\}_{t=1}^{\infty} \), is obtained by rewriting (2) as

\[
\Delta z_t = a_0 + a_1 t + \Pi z_{t-1} + \sum_{i=1}^{\infty} \Gamma_i \Delta z_{t-i} + u_t, \quad \text{or}
\]

\[
\Delta z_t = a_0 + \Pi^* z_{t-1}^* + \sum_{i=1}^{\infty} \Gamma_i \Delta z_{t-i} + u_t
\]

where \( \Gamma_i = -\sum_{j=i+1}^{p} \Phi_i \) for \( i = 1, \ldots, p-1 \), \( \Pi^* = \Pi(\gamma, \Pi_m) \) and \( z_{t-1}^* = (t, z_{t-1}') \). The following relations define the mapping between the parameters of the VAR model and that of the VECM formulation:

\[
\Phi_1 = I_m - \Pi + \Gamma_1, \quad (7)
\]

\[
\Phi_i = \Gamma_i - \Gamma_{i-1}, \quad i = 2, 3, \ldots, p-1, \quad (8)
\]

\[
\Phi_p = -\Gamma_{p-1}. \quad (9)
\]

2. The infinite VMA representation of \( \{z_t\}_{t=1}^{\infty} \) is defined as:

\[
\Delta z_t = C(L) (a_0 + a_1 t + u_t) = b_0 + b_1 t + C(L) u_t \quad (10)
\]

where

\[
b_0 = C(1) a_0 + C^*(1) a_1; \quad b_1 = C(1) a_1 \quad (11)
\]

The matrix lag polynomial \( C(L) \) is given by

\[
C(L) = I_m + \sum_{i=1}^{\infty} C_i L^i = C(1) + (1 - L) C^*(L), \quad C^*(L) = \sum_{i=0}^{\infty} C_i^* L^i \quad (12)
\]

\[
C(1) = \sum_{i=1}^{\infty} C_i, \quad C^*(1) = \sum_{i=0}^{\infty} C_i^*. \quad (13)
\]

The matrices \( \{C_i\} \) derive from the recursion

\[
C_i = \sum_{j=1}^{P} C_{i-j} \Phi_j, \quad i > 0, \quad C_0 = I_m \quad \text{and} \quad C_1 = 0 \quad \text{for} \quad i < 0 \quad (14)
\]

Similarly for the matrices \( \{C_i^*\} \)

\[
C_i^* = C_{i-1} + C_{i-1}^*, \quad i > 0, \quad C_0^* = C_0 - C(1). \quad (15)
\]

---

3 See, for example, Hamilton (1994) [p.549] for this alternative representation of a VAR(p) model.

4 See Johansen [(1995a), p.36] or Pesaran et al. (1998). Equation (12) is analogous to a multivariate Beveridge and Nelson decomposition of \( z \) into its permanent and transitory components. See Section 2.4.2 for further discussion in the multivariate case.
It follows from (2) and (10) that \( C(L) \Phi(L) = \Phi(L)C(L) = (1 - L)_m \), implying

\[
\Pi C(1) = C(1) \Pi = 0, \tag{16}
\]

\[
C(1) a_1 = 0. \tag{17}
\]

Therefore \( b_1 = 0 \) and (10) is given by

\[
\Delta z_t = b_0 + C(L) u_t, \tag{18}
\]

### 2.4.1 COINTEGRATION

The cointegration rank hypothesis is defined by\(^5\)

\[
H_0 : \text{Rank } [\Pi] = r, \ r < m. \tag{19}
\]

Under \( H_0 \) of (19), \( \Pi \) may be expressed as

\[
\Pi = \alpha \beta', \tag{20}
\]

where \( \alpha \) and \( \beta \) are \((m \times r)\) matrices of full column rank.\(^6\) Let \( \alpha_\perp \) and \( \beta_\perp \) denote \((m \times m - r)\) matrices defined as the bases for the orthogonal complements of \( \alpha \) and \( \beta \), such that \( \alpha' \alpha_\perp = 0 \) and \( \beta' \beta_\perp = 0 \). Furthermore, \( C(1) \) is a singular matrix which can be written as

\[
C(1) = \beta_\perp (\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp, \tag{21}
\]

The cointegration rank hypothesis, (19), can therefore be equivalently expressed as

\[
H_0 : \text{Rank } [C(1)] = m - r, \ 0 < r < m. \tag{22}
\]

### 2.4.2 THE MULTIVARIATE BN DECOMPOSITION

Substituting (12) into the VMA representation, (18), and solving for the level, \( z_t \), yields the common trends representation or the multivariate BN decomposition\(^7\)

\[
z_t = z_0 + b_0 t + C(1) s_t + C(L)(u_t - u_0), \tag{23}
\]

\[
z_t = \mu + y_t + C(1) s_t + C(L) u_t, \tag{24}
\]

---

\(^5\) See Engle & Granger (1987).

\(^6\) Note that \( \text{rank}(\Pi) = \text{rank}(\Pi') \).

\(^7\) See King et al. (1987) and Stock & Watson (1988) for seminal contributions deriving common stochastic trend representations.
where \( \{ z_t \} \) is initialised by \( z_0 = \mu + C(1)u_t \), \( s_t^u \) is the partial sum defined by, \( s_t^u = \sum_{j=1}^{t} u_j, t = 1, 2, \ldots \) and \( y = b_0 \) since \( C(1) \Pi = 0, C(1) \Pi = I_m \) and \( C(1) \Gamma - C(1) \Pi = I_m \).

Substituting (21) in (24) yields

\[
\begin{align*}
\dot{z}_t &= \mu_t + \gamma_t + \psi \alpha' \cdot s_t^u + C(L)u_t,
\end{align*}
\]

where \( \alpha' \cdot s_t^u \) are \((m \times r)\) independent “common stochastic trends” and \( \psi = \beta \cdot (\alpha' \cdot \Gamma \beta')^{-1} \) is a factor loading matrix of order \((m \times m - r)\) measuring the long run response on \( z_t \) of innovations to the common stochastic trends. The \((m-r)\) independent common stochastic trends are supplemented by \( r \) linearly independent deterministic trends, see (2) and (4). Crucially there are fewer common stochastic trends than variables, the difference being the number of co-integrating relations, \( r \).

\( C(1)s_t^u = \psi \alpha' \cdot s_t^u \), represents the permanent or long run component of \( z_t \); the shocks that drive the common stochastic trends are permanent since

\[
\psi \alpha' \cdot s_t^u = \psi \tau_t
\]

where \( \tau_t \) is a \((m-r)\) vector of random walks given by \( \tau_t = \tau_{t-1} + u_\tau \), with \( u_\tau = \alpha' \cdot u_t \) denoting \((m-r)\) permanent shocks. The common stochastic trends are consequently also interpreted as a characterisation of the “long run” in cointegrating VAR models. Note that in (25) the permanent shocks, \( u_\tau \), and the transitory shocks, \( u_t \), are collinear. But (25) is only one possible permanent/transitory decomposition of the shocks; see King et al. (1991), pp. 40-2 for an alternative decomposition based on a Jordan decomposition of \( C(1) \). The lack of uniqueness of any particular decomposition is only resolved by imposing restrictions on \( (\psi, \alpha') \) motivated by a priori theory.

Importantly as shown say by Garratt et al. (2006), BN trends are the limiting trends of any decomposition method.
2.4.3 BIVARIATE PHILLIPS CURVE MODELS

Having extracted the (many) cyclical estimates they are related to inflation using a bivariate VAR model in the cycle and inflation. Following van Norden & Orphanides (2005), we consider linear Phillips curve forecasting models of the form:

\[
\pi_{t+h} = \alpha_1 + \sum_{p=0}^{P} \beta_{1p} \pi_{t-p} + \sum_{p=0}^{P} \gamma_{1p} y_{t-p}^j + \epsilon_{1t+h}
\]

(27)

where inflation is denoted \(\pi_t\) and the various output gap measures are denoted \(y_t^j\), with \(j = 1 \ldots J; P + 1\) denotes the maximum number of lags in inflation and the output gap measures, and \(h\) is the forecast horizon.

We augment this specification with the corresponding output gap equation:

\[
y_{t+h}^j = \alpha_2 + \sum_{p=0}^{P} \beta_{2p} \pi_{t-p} + \sum_{p=0}^{P} \gamma_{2p} y_{t-p}^j + \epsilon_{2t+h}
\]

(28)

to create a bivariate VAR system. For simplicity, we assume that the lag structure is identical in the two equation VAR system.

We emphasise that although each VAR component uses a particular output gap measure, our aim is not to find the “best” single measure of the output gap. Rather, we wish to build ensemble forecast densities for inflation, conditional on a number of candidate measures.

The two-equation recursive structure described by these equations is common to many more detailed models of inflation determination. For example, Rudebusch and Svensson (2002) and Laubach & Williams (2003) start with similar bivariate structures, and add additional explanatory relationships and restrictions; Garratt, Koop, Mise & Vahey (2008) consider bigger VARs to assess whether money causes inflation and output. In principle, additional equations and the imposition of identifying conditions pose no conceptual problems for our methodology, although the computational burden would increase. We prefer to restrict our attention to a two-equation model space which, as Sims (2008) notes, lies at the heart of many explanations of inflation determination.

If it is believed that inflation is determined by a New Keynesian Phillips Curve such that

\[
\pi_t = \beta E(\pi_{t+1} \mid \Omega_t) y_t^j
\]

(29)

then if it is assumed that the output gap follows an AR(1) process

\[
y_t^j = y_{t-1}^j + \epsilon_{2t}
\]

(30)

Then
Globalisation, output gaps and inflation density forecasting

\[ \pi_t = \gamma \sum_{k=0}^{\infty} \beta^k E (Y_{t+k} | \Omega_t) \]  

(31)

\[ \pi_t = \frac{\gamma}{1-\beta \gamma} y_t^f \]  

(32)

implying inflation and the output gap follow identical univariate process, subject to a scalar. So the 1-step ahead forecast of inflation would be \( \frac{\gamma}{1-\beta \gamma} y_{t-1}^f \), showing that the VAR model is perfectly consistent with forward as well as backward looking Phillips Curves. It offers a reduced-form.

Armed with our many VAR specifications, it is straightforward to produce forecast densities for both variables of interest. Given non-informative priors, the predictive densities for both inflation and the output gap, for a given VAR specification, (32), are multivariate Student-t; see Zellner (1971), pp. 233-236 and, for a more recent application, Garratt, Koop, Mise & Vahey (2008). Appendix 2 provides details of how the predictive densities are constructed from our VARs.

These forecasts are then combined as discussed below.

2.4.4 COMBINATION FORECASTS

In statistics, Bayesian Model Averaging (BMA) offers a conceptually elegant means of dealing with ‘model uncertainty’. BMA forecasts condition not on a single ‘best’ model but take a weighted average over a range of candidate models; see Hoeting et al. (1999). This follows from appreciation of the fact that, although one model may be ‘better’ than the others, we may not select it with probability one. We may not be sure that it is the best forecast. Therefore, if we considered this single forecast alone, we would be overstating its precision.

Similarly in practical macroeconomic forecasting exercises, whether within a Bayesian context or not, it is a stylised fact that combination forecasts are hard to beat. The estimated parameters of a single forecasting model are commonly found to exhibit instabilities and these can be difficult to identify in real-time. In the presence of these so-called ‘uncertain instabilities’ it can be helpful to combine the evidence from many models. For example, Clark & McCracken (2009) examine the scope for taking linear combinations of point forecasts in real time, motivated by the desire to circumvent the uncertain instabilities in any particular specification. In a series of influential papers, Stock & Watson (2004) have documented the robust performance of point forecast combinations using various types of models for numerous economic and financial variables. Selecting a single model has little appeal under ‘uncertain instabilities’ when the single best model suffers from instability. This might happen either if the ‘true’ model is not within the model space considered by the modeller, or if the model selection process performs poorly on short macroeconomic samples. We may better approximate the truth, and account for the uncertainty in model selection, by combining forecasts.

While methods for combining point forecasts are well established and much exploited, less direct attention has been given in econometrics to the combination of density forecasts.\(^8\) Although, as Wallis (2005) has noted, density forecasts, in fact, have been combined in the ASA-NBER Survey of Professional Forecasters since 1969. The forecasters’ densities are combined by taking a linear average, a so-called “linear opinion pool”, as in BMA. Mitchell & Hall (2005) and Hall & Mitchell (2007) used this combination rule to combine and then analyse density forecasts from the Bank of England and the National Institute of

---

\(^8\) Outside the econometrics literature, the benefits of producing density forecasts by combining information across different models have been recognised for some time. For reviews see Genest & Zidek (1986) and Clemen & Winkler (1999).
Economic and Social Research. They considered how, in practice, the densities in the combination might be weighted, considering alternatives to equal weights. Jore et al. (2010) examine linear combinations of densities from VAR models, and Bache et al. (2009) take a linear combination of VAR and DSGE densities. Alternatives include consideration of logarithmic pooling rules; see Kascha & Ravazzolo (2010) and Wallis (2010). Using several examples, and theoretical analysis, Geweke & Amisano (2008) demonstrate the scope for pooled forecast densities to produce superior predictions, even if the set of components to be combined excludes the ‘true’ model. Geweke (2010), Chapter 5, stresses the efficacy of density forecast combination in the case where the ‘true’ model is absent from the model space—sometimes referred to as an ‘incomplete model space’.

In this report we follow in the spirit of this forecasting literature and use the linear opinion pool to combine density nowcasts.

We formalise density combination in a way that extends the commonly-adopted convex mix of point forecasts by utilising the linear opinion pool approach; see Timmermann (2006), (p.177), and the references described therein.

Given \( i = 1 \ldots N \) component models, the combination densities for GDP growth are given by the linear opinion pool:

\[
p(\Delta y_t) = \sum_{i=1}^{N_j} \omega_{i,t,j} g(\Delta y_t \mid \Omega_{j}^t), \quad \tau = \tau_0, \ldots, \tau,
\]

where \( N_j \) (j = 1, ..., 5) where \( N_{j+1} > N_j \); \( g(\Delta y_t \mid \Omega_{j}^t) \) are the nowcast forecast densities from component model \( i \), \( i = 1 \ldots, N_j \) of \( \Delta y_t \) conditional on the information set \( \Omega_{j} \). These densities, as we discuss below, are obtained having estimated (33). The non-negative weights, \( \omega_{i,t,j} \), in this finite mixture sum to unity. Furthermore, the weights may change with each recursion in the evaluation period \( \tau = \tau_0, \ldots, \tau \).

The predictive densities for \( \Delta y_t \) (with non-informative priors), \( g(\Delta y_t \mid \Omega_{j}^t) \), allowing for small sample issues, are Student-t; see Zellner (1971), pp. 233-236 and, for a more recent application, Garratt, Koop, Mise & Vahey (2008). Since each component model, (29), considered produces a forecast density that is \( t \), the combined density defined by equation (33) will be a mixture — accommodating skewness and kurtosis. That is, the combination delivers a more flexible distribution than each of the individual densities from which it was derived. As \( N_j \) increases, the combined density becomes more and more flexible, with the potential to approximate non-linear specifications.

We construct the weights \( \omega_{i,t,j} \) in three ways.

First, we consider equal weights (EW). The EW strategy attaches equal (prior) weight to each model with no updating of the weights through the recursive analysis: \( \omega_{i,t,j} = \omega_{i,j} = 1/N_j \). We present results for the EW strategy without (prior) truncation of the set of models to be included. The EW strategy is often recommended when combining point forecasts, although its effectiveness for density forecasts has been questioned (see Jore et al. (2010) and Garratt et al. (2010)).

---

9 The restriction that each weight is positive could be relaxed; for discussion see Genest & Zidek (1986). Note that in (33) the only unknown parameters to be estimated are the \( \omega_{i,t,j} \). The N component densities are taken as given. Somewhat confusingly, in “mixture models” these weights are interpreted on the basis of a latent binary random variable, which is often assumed to have a Markov structure; see Geweke & Amisano (2008) and Mitchell & Wallis (2010). But in these models the parameters of the component models are often estimated simultaneously with \( \omega_{i,t,j} \). In so-called BMA for ensemble forecasting models (see Raftery et al. (1995)), the component densities \( g(.) \) are centered on the point forecasts from the competing component models, but the variance of the component density forecasts is assumed common across \( N \) and estimated simultaneously with \( \omega_{i,t,j} \).
Secondly, we construct the weights $\omega_{t,r,j}$ based on the fit of the individual model forecast densities: the Recursive Weight (RW) strategy. Following Jore et al. (2010) and Garratt et al. (2009), we use the logarithmic score to measure density fit for each model through the evaluation period. The logarithmic scoring rule is intuitively appealing as it gives a high score to a density forecast that assigns a high probability to the realised value.\(^{10}\) Specifically, the recursive weights for the nowcast densities take the form:

$$
\omega_{t,r,j} = \frac{\exp \left[ \sum_{t=1}^{r-8} \ln g(\Delta y_t | \Omega_j) \right]}{\sum_{j=1}^{N} \exp \left[ \sum_{t=1}^{r-8} \ln g(\Delta y_t | \Omega_{j}) \right]}, \quad r = r, ..., R,
$$

(34)

where the $r - 8$ to $r$ comprises the two-year training period, since we employ quarterly data, used to initialise the weights. Computation of these weights is feasible even for large $N$. Given the uncertain instabilities problem, the recursive weights should be expected to vary across $r$.

From a Bayesian perspective, density combination based on recursive logarithmic score weights has many similarities with an approximate predictive likelihood approach (see Raftery & Zheng (2003), and Eklund & Karlsson (2007)). Given our definition of density fit, the model densities are combined using Bayes’ rule with equal (prior) weight on each model—which a Bayesian would term non-informative priors. (Koop (2003) (chapter 11) and Geweke & Whiteman (2006) provide recent general discussions of Bayesian model averaging methods.) Andersson & Karlsson (2007) propose Bayesian predictive likelihood methods for forecast combination with Bayesian VARs but do not consider forecast density evaluation. Hall & Mitchell (2007) and Geweke & Amisano (2008) consider iterative algorithms to select weights that maximize the logarithmic score, suitable for small $N$. We note that instead of looking at fit over the entire density the component models, with a larger out-of-sample window than available in our application, could be scored according to their ability to forecast tail events. Nevertheless, there are important differences with (predictive) BMA as Geweke (2010) explains. For example, since we assume that our component models represent an incomplete model space, the conventional Bayesian interpretation of the weights as reflecting the posterior probabilities of the components is inappropriate. Accordingly, we do not consider model selection using the combination weights; nor do we consider strategies averaging a selection of component models. We are interested in the performance of the density from the linear opinion pool.

Thirdly, as discussed below, we consider combining the models using their continuous ranked probability score (CRPS) - instead of their log score. The CRPS is believed to be a more robust measure of density fit.

### 2.4.5 OCCAM’S WINDOW: EXCLUDING BAD MODELS AND GRAND ENSEMBLES

There is always a question about how one should choose the set of models over which one combines. Our strategy is to select a model space, namely VAR models linking inflation and the output gap that is commonly used. We then consider a range of output gap estimators distinguished by the information set which they use to define the long-run output forecast used to define the BN cycle. We start by employing an uninformative prior on all models and use the data (Bayes rule) to update the weight on each model as evidence accumulates.

There may be empirical benefits, however, to excluding some bad models prior to taking the ensemble. Occam’s window is one option; it is an empirically driven strategy to improve forecast performance by excluding bad models from the set of model. Madigan & Raftery (1994) propose an algorithm which is based on selecting a small set of models by using posterior sampling and treats all the worst fitting models outside the subset as having zero posterior probability. We have not explored this strategy to-date.

---

\(^{10}\) The logarithmic score of the density forecast, $\ln g(\Delta y_t | \Omega_j)$, is the logarithm of the probability density function $g(\cdot | \Omega_j)$, evaluated at the outturn $\Delta y_t$.\[\text{[0x0]}\]
2.5 Density forecast evaluation

In constructing the combined densities using the linear opinion pool, we evaluate the density forecasts using the logarithmic score at each recursion. We emphasise that in deriving the weights based on this measure of density fit, the component models are repeatedly evaluated using real-time data. These weights provide an indication of whether the support for the component models is similar, or not, based on the score of the individual densities. A finding of similar weights across component models would be consistent with the equal-weight strategy.

A common approach to forecast density evaluation provides statistics suitable for tests of (absolute) forecast accuracy, relative to the “true” but unobserved density. A popular method, following Rosenblatt (1952), Dawid (1984) and Diebold et al. (1998), evaluates using the probability integral transforms (pits) of the realisation of the variable with respect to the forecast densities. A density forecast can be considered optimal (regardless of the user’s loss function) if the model for the density is correctly conditionally calibrated. We gauge calibration by examining whether the pits \(z_t\), where \(z_t = \int_{-\infty}^{\tau_t} p(u)du\), are uniform and independently and identically distributed (see Diebold et al. (1998)). In practice, therefore, density evaluation with the pits requires application of tests for goodness-of-fit and independence at the end of the evaluation period. Mitchell & Wallis (2010) refer to this two component condition as “complete calibration”.

The goodness-of-fit tests employed include the Likelihood Ratio (LR) test proposed by Berkowitz (2001). We use a three degrees-of-freedom variant with a test for independence, where under the alternative \(z_t\) follows an AR(1) process. We also follow Berkowitz (2001) and report a censored LR test which focuses on the 10% top and bottom tails. This is designed to detect forecast failure in the tails of the forecast density. We also consider the Anderson-Darling (AD) test for uniformity, a modification of the Kolmogorov-Smirnov test, intended to give more weight to the tails (and advocated by Noceti et al. (2003)). We also follow Wallis (2003) and employ a Pearson chi-squared test which divides the range of the \(z_t\) into eight equiprobable classes and tests whether the resulting histogram is uniform.

Turning to the test for independence of the pits, we use a Ljung-Box (LB) test, based on autocorrelation coefficients up to four. To investigate possible higher order dependence we also undertook tests in the second and third powers of the pits in the histogram. Turning to the test for independence of the pits, we use a Ljung-Box (LB) test, based on (up to) fourth-order autocorrelation.

We also investigate relative predictive accuracy by considering a Kullback-Leibler information criterion (KLIC)-based test, utilizing the expected difference in the log scores of candidate densities; see Bao et al. (2007), Mitchell & Hall (2005) and Amisano & Giacomini (2007). Suppose there are two forecast densities, \(p(\tau_t \mid I_{1,t})\) and \(p(\tau_t \mid I_{2,t})\), so that the KLIC differential between them is the expected difference in their log scores: \(d_t = \ln p(\tau_t^1 \mid I_{1,t}) - \ln p(\tau_t^1 \mid I_{2,t})\). The null hypothesis of equal forecast performance is \(H_0 = E(d_t) = 0\). A test can then be constructed since the mean of \(d_t\) over the evaluation period, \(\bar{d}_t\), under appropriate assumptions, has the limiting distribution: \(\sqrt{T}\bar{d}_t \rightarrow N(0,\Omega)\), where \(\Omega\) is a consistent estimator of the asymptotic variance of \(d_t\).

Mitchell & Wallis (2010) discuss the value of information-based methods for evaluating forecast densities that look well-calibrated from the perspective of the pits.

---

11 Given the large number of component densities under consideration, we do not allow for estimation (parameter) uncertainty when evaluating the pits. Corradi & Swanson (2006) review pits tests computationally feasible for small \(N\).

12 When evaluating the forecast densities we abstract from the method used to produce them. Amisano and Giacomini (2007) and Giacomini and White (2006) discuss more generally the limiting distribution of related test statistics.
2.5.1 THE CONTINUOUS RANKED PROBABILITY SCORE (CRPS)

Scoring rules can be used to assess the quality of weather forecasts by assigning a numerical score based on the forecast and the value or event that materializes. A proper scoring rule maximizes the expected reward (or minimizes the expected penalty) for forecasting one’s true beliefs, thereby discouraging hedging or cheating. For more details see Gneiting & Raftery (2007).

Although the logarithmic score is trivial to calculate and is also proper, it allocates large penalties to low probability events, making it more sensitive to outliers than CRPS. If \( F \) is the cumulative distribution function of the forecast distribution and \( x \) verifies, the CRPS is defined as

\[
CRPS (F, x) = \int_{-\infty}^{\infty} [F(y) - 1\{y \geq x\}]^2 dy
\]

where \( 1\{y \geq x\} \) denotes a step function along the real line that attains the value 1 if \( y \geq x \) and the value 0 otherwise. The continuous ranked probability score can be written equivalently as

\[
CRPS (F, x) = E\{|X - x|\} - 0.5E\{|X - X^*|\}
\]

where \( X \) and \( X^* \) are independent copies of a linear random variable with distribution function \( F \), and \( E\{\cdot\} \) denotes the expectation operator. The CRPS is proper and is expressed in the same unit as the observed variable. The CRPS generalizes the absolute error, and therefore provides a direct way of comparing various deterministic and probabilistic forecasts using a single metric.

The linear CRPS is defined in terms of the forecast cumulative distribution function \( F \) and the verifying observation \( x \) of a linear, real-valued variable as in (35). Traditionally, this score is interpreted as the integral of the Brier score for binary probability forecasts at all real-valued thresholds \( y \). The alternative, yet equivalent, representation given by (36) makes the interpretation of CRPS lucid. The first term on the right-hand side of (36) is the expected value of the absolute error, and the second term is a correction factor that measures the sharpness of the probabilistic forecast \( F \) and renders the score proper. The linear CRPS generalizes the absolute error, to which it reduces if \( F \) is a deterministic forecast.

An analytic expression for the integral in (35) when the forecast distribution function \( F \) is Gaussian with mean \( \mu \) and variance \( \sigma^2 \), using (36), is given as:

\[
CRPS (N[\mu, \sigma^2], x) = \sigma \left( \frac{x-\mu}{\sigma} \right) \left[ 2 \Phi \left( \frac{x-\mu}{\sigma} \right) - 1 \right] + 2 \phi \left( \frac{x-\mu}{\sigma} - \frac{1}{\sqrt{\pi}} \right)
\]

where \( \phi \) and \( \Phi \) represent the standard Gaussian probability density and cumulative distribution functions, respectively.

2.5.2 TAIL SCORES

We use a new scoring rule proposed by Diks et al. (2011) based on conditional and censored likelihood for assessing the predictive accuracy of competing density forecasts over a specific region of interest, specifically the left tail (deflationary events).
2.6 Empirical Results

Before turning to forecasting inflation we consider estimation of the cycle - the output gap.

2.6.1 ESTIMATING THE EURO AREA CYCLE

Reflecting the fact that Euro area GDP involves the aggregation of GDP data for the member countries, we consider estimating the Euro area cycle using three different approaches:

1. The “direct” approach consists of decomposing the raw data of the aggregate itself.

2. The “indirect” approach consists of decomposing the raw data corresponding to the sub-components (national GDP) and then aggregating the county-series using their current quarter shares of aggregate GDP as weights.

3. The “multivariate” or simultaneous approach involves decomposing the disaggregates simultaneously.

In general, these approaches will deliver different aggregate cycles. Only for so-called uniform decomposition filters does the order of decomposition and aggregation not matter.

But when the filters differ, as they will when an optimal signal extraction method is used like an unobserved components model, the order is crucial. For further discussion see Ghysels (1997) and Planas & Campolongo (2001). The multivariate approach has certain optimality properties, although since it is computationally demanding and requires hard choices to be made about the appropriate information set this approach is rarely considered; see Geweke (1978). We however do consider it using the VAR model mentioned above.

Since we wish to use real-time GDP data, availability at the national level limits us to the nine EA countries (Austria, Belgium, Germany, Spain, Finland, France, Greece, Italy and The Netherlands), for which real-time GDP data are available on a consistent basis. The data are available in Euros; therefore the Euro area aggregate (in levels) can be obtained through simple aggregation of the nine countries data. See the Appendix for details on the real-time data from EuroIND.

The data are then backcast, when necessary, to 1980q1.\(^\text{13}\) In total we have 40 out-of-sample data vintages containing successive first releases of GDP, starting in May 2001 and ending in November 2010, over which to evaluate the quality of our estimates. The first vintage, contains quarterly GDP data from 1980q1 until 2000q4. The last vintage contains GDP data from 1980q1 to 2010q3.

---

\(^{13}\) German data backdated with data available from the Bundesbank (http://www.bundesbank.de/statistik/statistik_zeitreihen.en.php?lang=en&open=&func=list&tr=www s311 lr bip).
Direct, indirect and multivariate cyclical estimates are all produced via the Beveridge-Nelson decomposition. For the direct and indirect approaches this involves recursively estimating univariate models in GDP growth. Given the lag order of the underlying ARMA model in GDP growth is known to affect inference we consider two variants both of which have been used in the literature. First we consider AR(8) models as in Garratt, Lee, Mise & Shields (2008). Secondly, we consider ARMA(1,1) models as suggested by Robertson & Wright (2009). The multivariate estimates are the multivariate BN cycles from a 9-variable VAR (of lag order 2) in the 9 European countries’ GDP with as an exogenous I(1) regressor either: (i) global output, as measured by Dees et al. (2007)\textsuperscript{14}; (ii) US output as measured in real-time by the Federal Reserve Bank of Philadelphia; (iii) oil prices.

If this additional series has significant explanatory power for European GDP, then the Euro area output gap should be larger than otherwise expected.

\textsuperscript{14} These data are not available in real-time; and we therefore consider only the most recent vintage of data.
To provide an indication of what the cyclical estimates for Euro area GDP look like from these three approaches, Figure plots direct, indirect and multivariate estimates using the final vintage data. The multivariate estimates are those from the VAR with global GDP as an exogenous I(1) regressor.

Figure 4: Direct, Indirect and Multivariate based estimates of the Euro area (9 countries) Beveridge-Nelson cycle using final vintage data

Figure 4 shows that while the direct and indirect approaches give a similar picture of the European cycle, with correlation coefficients at 0.83 or higher, the multivariate approach can paint quite a different picture. But this does depend on the number of cointegrating vectors imposed in the VAR. Given inspection of information criteria clearly indicates uncertainty about the number of cointegrating vectors, and since this choice does affect the look of the Euro area cycle, in the analysis below we accommodate uncertainty about the cointegrating rank.
Globalisation, output gaps and inflation density forecasting

Table 1 shows quite clearly that the correlation between the VAR-based cycle and the direct and indirect estimates declines as the number of cointegrating vectors rises above 3 or 4. That is, imposing more and more commonality on the VAR reduces the similarity with the direct and indirect cycles. But there is evidence that imposing a degree of commonality does increase correlation against the direct and indirect estimates, since a cointegrating rank of about 3 trends to deliver increased correlation relative to the case of assuming less or no cointegration.

<table>
<thead>
<tr>
<th>r=0</th>
<th>r=1</th>
<th>r=2</th>
<th>r=3</th>
<th>r=4</th>
<th>r=5</th>
<th>r=6</th>
<th>r=7</th>
<th>r=8</th>
<th>r=9</th>
<th>Direct:</th>
<th>Direct:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ARMA(1,1)</td>
<td>AR(8)</td>
</tr>
<tr>
<td>r=0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=1</td>
<td>0.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=2</td>
<td>0.9</td>
<td>0.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=3</td>
<td>0.8</td>
<td>0.8</td>
<td>0.9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.52</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=5</td>
<td>0.7</td>
<td>0.7</td>
<td>0.6</td>
<td>0.51</td>
<td>0.34</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
<td>0.34</td>
<td>0.45</td>
<td>0.8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=7</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.23</td>
<td>0.4</td>
<td>0.8</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r=8</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
<td>-0</td>
<td>-0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>r=9</td>
<td>0.3</td>
<td>0.3</td>
<td>0.1</td>
<td>0.13</td>
<td>-0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Direct:</td>
<td>0.7</td>
<td>0.6</td>
<td>0.7</td>
<td>0.7</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>-0</td>
<td>1</td>
</tr>
<tr>
<td>Direct:</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.58</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.1</td>
<td>0.05</td>
<td>0.94</td>
</tr>
<tr>
<td>Indirect:</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
<td>0.61</td>
<td>0.39</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.2</td>
<td>0.19</td>
<td>0.83</td>
</tr>
<tr>
<td>Indirect:</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>0.36</td>
<td>0.6</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.05</td>
<td>0.95</td>
<td>0.87</td>
</tr>
</tbody>
</table>

But it is important to apply the decomposition methods in real-time. We find that there is a general tendency for both the direct and indirect methods to overstate the size of the (negative) output gap in the recession. The direct and indirect approaches, in real-time, gave very contrasting impressions of the depth of the negative output gap during the recession. The indirect approach suggested an output gap of around 4%; but the direct approach indicated about 10%. Typically they indicated that output was 5% or more below trend in real-time; when the final vintage estimates suggest a more moderate deterioration of around 2%. This is because with the advantage of hindsight much of the dramatic fall in European GDP is attributed to a trend movement, rather than a cyclical movement. It takes some time after the event for the trend estimates to catch up with the actual movements in GDP. However, the multivariate approach does not suffer from this problem, although revisions are in fact often larger as there are more persistent differences between the real-time and final estimates than for the direct and indirect approaches.
2.6.2 FORECASTING INFLATION IN A GLOBALISED WORLD

We now consider the ability of these output gaps to forecast Euro area inflation. An extension will consider their ability to forecast national inflation series. We consider various measures of inflation. First we look at HICP and import price inflation, and the data refer to the EA16. These are final vintage data; although revisions to HICP inflation are minimal. Revisions to the HICP are concentrated in January 2003, when Germany, the Netherlands and Portugal moved their weight reference period from 1995 to 2000. The HICP are published in seasonally unadjusted form. In contrast the GDP deflator are published in seasonally adjusted form in real-time. Following O’Reilly & Whelan (2005) in their analysis of Euro area inflation we therefore focus on the GDP deflator. However, importantly extending O’Reilly & Whelan (2005) who use the ECB’s AWM data our data are real-time. (We found our results to be robust to consideration of HICP inflation (seasonally adjusted).)

HICP inflation data for the Euro area are published monthly by Eurostat at t+15 days. In recent months they have also released a Flash estimate at the end of the month, so at about t-1 days. The deflator data are published with the national accounts at about t+65 days.

Figure 5 plots quarterly HICP (seasonally unadjusted) and import price inflation; we see that import price inflation is much more volatile. While both inflation series trend downwards from the early 1980s, thereafter HICP inflation is pretty stable at around 2%, while import price inflation exhibits some big swings.

We also consider real-time GDP deflator and import price deflator data (seasonally adjusted). Here the data refer to an evolving concept of the Euro area. Figure 6 plots both the first-release and the final estimate, as of Nov 2010. The figure indicates that the real-time data track the general tendency of the final data, but there can be differences. A notable one is in the recession, when in real-time the GDP deflator suggested there was deflation but this was subsequently revised away.

---

15 The GDP deflator and import price deflator data are in fact taken from the Euro Area Real-Time Database (RTDB) which is an experimental dataset that consists of vintages, or snapshots, of time series of several variables, based on series reported in the ECB’s Monthly Bulletin (MoBu). The underlying data are from Eurostat.

We use 40 out-of-sample data vintages containing successive first releases of GDP, starting in May 2001 and ending in November 2010, over which to evaluate the quality of our estimates. The first vintage, contains quarterly GDP data from 1980q1 until 2000q4. The last vintage contains GDP data from 1980q1 to 2010q3. These GDP data are then related to the real-time deflator data such that the first vintage of GDP and deflator data, containing data up to 2000q4, are used to forecast inflation in 2001q1. Since the GDP and import price deflator data are available at about 65-70 days after the end of the quarter, and the Euro area GDP data are published from t+45 to about t+60 days (for the slower countries), this means we are producing about a one-quarter ahead inflation forecast.

Our objective is to forecast the first release value for the deflator.
2.7 AR and Output Gap Ensembles

2.7.1 FORECASTING THE GDP DEFLATOR

As a benchmark we run an AR ensemble, which involves estimating AR(p) models with \( p = 1, \ldots, 4 \) and we take both EW and RW combinations. Since the weights in the RW ensemble can change over time, this AR ensemble is quite flexible and able to capture changes in the persistence of inflation, for example.

We then consider four output gap ensembles.

1. **OG Euro**: The first considers 4 aggregate estimates of the Euro area cycle only. Specifically it considers 2 indirect estimates of the Euro area output gap based on both ARMA(1,1) and AR(8) models and 2 direct estimates of the Euro area output gap based on both ARMA(1,1) and AR(8) models.

2. **OG Agg**: The second considers 11 aggregate estimates of the Euro area and global cycle only. Specifically it considers 2 indirect estimates of the Euro area output gap based on both ARMA(1,1) and AR(8) models; 2 direct estimates of the Euro area output gap based on both ARMA(1,1) and AR(8) models; direct estimates of the global output gap based on both ARMA(1,1) and AR(8) models; estimates of the US output gap based on both ARMA(1,1) and AR(8) models; indirect estimates of the Euro area output gap based on estimation of a VAR(2) model with either global
Globalisation, output gaps and inflation density forecasting

output, US output or the oil price as an exogenous I(1) regressor and the cointegrating rank imposed at r=0, as suggested by the BIC.

3. Disag: The third augments this set of output gap estimates with disaggregate estimates based on consideration of each of the 9 countries individually. Again both ARMA(1,1) and AR(8) models are considered. In total this makes 29 different output gap estimates are considered, so given the lag uncertainty (p=1,...,4), 116 component models are considered. (We did also try a variant where we dropped the 4 Euro area cycles so that Disag contained only global cycles or multivariate-based Euro cycles; this made little difference; e.g. the log score increased from -1.257 to -1.262 indicating that the 4 Euro area cycles in Disag are not affecting performance in any material way).

4. DisagM: The fourth augments Disag with consideration of the VAR based multi-variate estimates of the Euro area cycle for r=1,...,9. Uncertainty about the number of common trends in the Euro area is thereby accommodated.

For each ensemble we consider weights that are equal, based on the CRPS and on the log score.

Before turning to evaluation of the predictive densities from these four ensembles we look in Table 2 at their RMSE. Inspection of the RMSE results indicates that there is not much to choose between the different ensembles. There appear to be modest gains to consideration of output gap information when forecasting inflation relative to the AR benchmark.

Table 2: RMSE

<table>
<thead>
<tr>
<th>RMSE</th>
<th>OG Euro</th>
<th>OG Agg</th>
<th>OG Disag</th>
<th>OG DisagM</th>
<th>AR</th>
<th>Import/ oil price</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td>GDP defl</td>
<td>0.915</td>
<td>0.95</td>
<td>0.938</td>
<td>0.936</td>
<td>0.959</td>
</tr>
<tr>
<td>CRPS</td>
<td></td>
<td>0.902</td>
<td>0.953</td>
<td>0.929</td>
<td>0.932</td>
<td>0.965</td>
</tr>
<tr>
<td>RW</td>
<td></td>
<td>0.872</td>
<td>0.999</td>
<td>0.887</td>
<td>0.9</td>
<td>0.962</td>
</tr>
<tr>
<td>EW</td>
<td>Import price defl</td>
<td>5.072</td>
<td>4.935</td>
<td>4.856</td>
<td>4.833</td>
<td>4.767</td>
</tr>
<tr>
<td>CRPS</td>
<td></td>
<td>5.113</td>
<td>5.64</td>
<td>5.79</td>
<td>5.932</td>
<td>4.847</td>
</tr>
<tr>
<td>RW</td>
<td></td>
<td>5.168</td>
<td>5.953</td>
<td>6.078</td>
<td>6.184</td>
<td>4.798</td>
</tr>
<tr>
<td>HICP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.48</td>
</tr>
<tr>
<td>CRPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.468</td>
</tr>
<tr>
<td>RW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.464</td>
</tr>
</tbody>
</table>

Table 3 then evaluates the densities by reporting their average logarithmic scores over the evaluation period. It shows that the Disag ensemble, with RW, is preferred. Consideration of the national data appears to help, since Disag produces a higher log score than the Agg ensemble which considers Euro area data only. But accommodating uncertainty about the cointegrating rank in the VAR model does not deliver gains. It can pay, consistent with Occam’s Razor, to restrict attention to a subset of the models.
Table 3: Log Score statistics: GDP deflator

<table>
<thead>
<tr>
<th></th>
<th>Prop</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR EW</td>
<td>-1.378</td>
</tr>
<tr>
<td>AR CRPS</td>
<td>-1.385</td>
</tr>
<tr>
<td>AR RW</td>
<td>-1.381</td>
</tr>
<tr>
<td>AR1</td>
<td>-1.433</td>
</tr>
<tr>
<td>AR2</td>
<td>-1.487</td>
</tr>
<tr>
<td>AR3</td>
<td>-1.372</td>
</tr>
<tr>
<td>AR4</td>
<td>-1.355</td>
</tr>
<tr>
<td>OG Euro EW</td>
<td>-1.298</td>
</tr>
<tr>
<td>OG Euro CRPS</td>
<td>-1.287</td>
</tr>
<tr>
<td>OG Euro RW</td>
<td>-1.262</td>
</tr>
<tr>
<td>OG Agg EW</td>
<td>-1.304</td>
</tr>
<tr>
<td>OG Agg CRPS</td>
<td>-1.295</td>
</tr>
<tr>
<td>OG Agg RW</td>
<td>-1.29</td>
</tr>
<tr>
<td>OG Disag EW</td>
<td>-1.312</td>
</tr>
<tr>
<td>OG Disag CRPS</td>
<td>-1.299</td>
</tr>
<tr>
<td>OG Disag RW</td>
<td>-1.257</td>
</tr>
<tr>
<td>OG DisagM EW</td>
<td>-1.309</td>
</tr>
<tr>
<td>OG DisagM CRPS</td>
<td>-1.298</td>
</tr>
<tr>
<td>OG DisagM RW</td>
<td>-1.275</td>
</tr>
</tbody>
</table>
Globalisation, output gaps and inflation density forecasting

Table 4 shows that the densities from the ensembles which consider the output gap in at least one form appear well-calibrated according to the pits tests. It is only the AR models and their ensemble which fail at least one of the pits tests at a 95% significance level; this appears to be in the tail.

Table 4: PIT statistics: GDP deflator

<table>
<thead>
<tr>
<th>Method</th>
<th>KLIC</th>
<th>LR2</th>
<th>tail score</th>
<th>AD</th>
<th>walls</th>
<th>LB1</th>
<th>LB2</th>
<th>LB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR EW</td>
<td>0.045</td>
<td>0.179</td>
<td>0.061</td>
<td>0.47</td>
<td>0.31</td>
<td>1.208</td>
<td>0.933</td>
<td>0.229</td>
</tr>
<tr>
<td>AR CRPS</td>
<td>0.05</td>
<td>0.15</td>
<td>0.025</td>
<td>0.365</td>
<td>0.283</td>
<td>1.05</td>
<td>0.899</td>
<td>0.416</td>
</tr>
<tr>
<td>AR RW</td>
<td>0.053</td>
<td>0.131</td>
<td>0.008</td>
<td>0.329</td>
<td>0.248</td>
<td>1.04</td>
<td>0.899</td>
<td>0.496</td>
</tr>
<tr>
<td>AR1</td>
<td>0.068</td>
<td>0.076</td>
<td>0.119</td>
<td>0.418</td>
<td>0.064</td>
<td>1.863</td>
<td>0.768</td>
<td>0.041</td>
</tr>
<tr>
<td>AR2</td>
<td>0.079</td>
<td>0.05</td>
<td>0.017</td>
<td>0.895</td>
<td>0.082</td>
<td>1.509</td>
<td>0.565</td>
<td>0.091</td>
</tr>
<tr>
<td>AR3</td>
<td>0.048</td>
<td>0.164</td>
<td>0.009</td>
<td>0.372</td>
<td>0.3</td>
<td>1.086</td>
<td>0.86</td>
<td>0.385</td>
</tr>
<tr>
<td>AR4</td>
<td>0.05</td>
<td>0.15</td>
<td>0.017</td>
<td>0.315</td>
<td>0.261</td>
<td>1.112</td>
<td>0.86</td>
<td>0.559</td>
</tr>
<tr>
<td>OG Euro EW</td>
<td>0.009</td>
<td>0.723</td>
<td>0.416</td>
<td>0.523</td>
<td>0.405</td>
<td>0.634</td>
<td>0.515</td>
<td>0.157</td>
</tr>
<tr>
<td>OG Euro CRPS</td>
<td>0.003</td>
<td>0.885</td>
<td>0.484</td>
<td>0.507</td>
<td>0.591</td>
<td>0.498</td>
<td>0.342</td>
<td>0.203</td>
</tr>
<tr>
<td>OG Euro RW</td>
<td>0.002</td>
<td>0.917</td>
<td>0.555</td>
<td>0.906</td>
<td>0.668</td>
<td>0.364</td>
<td>0.718</td>
<td>0.225</td>
</tr>
<tr>
<td>OG Agg EW</td>
<td>0.011</td>
<td>0.652</td>
<td>0.355</td>
<td>0.528</td>
<td>0.435</td>
<td>0.754</td>
<td>0.666</td>
<td>0.122</td>
</tr>
<tr>
<td>OG Agg CRPS</td>
<td>0.005</td>
<td>0.827</td>
<td>0.388</td>
<td>0.508</td>
<td>0.85</td>
<td>0.608</td>
<td>0.214</td>
<td>0.165</td>
</tr>
<tr>
<td>OG Agg RW</td>
<td>0.004</td>
<td>0.857</td>
<td>0.381</td>
<td>0.508</td>
<td>0.885</td>
<td>0.581</td>
<td>0.515</td>
<td>0.136</td>
</tr>
<tr>
<td>OG Disag EW</td>
<td>0.014</td>
<td>0.586</td>
<td>0.292</td>
<td>0.514</td>
<td>0.491</td>
<td>0.832</td>
<td>0.899</td>
<td>0.131</td>
</tr>
<tr>
<td>OG Disag CRPS</td>
<td>0.008</td>
<td>0.749</td>
<td>0.31</td>
<td>0.488</td>
<td>0.742</td>
<td>0.709</td>
<td>0.565</td>
<td>0.159</td>
</tr>
<tr>
<td>OG Disag RW</td>
<td>0.007</td>
<td>0.773</td>
<td>0.342</td>
<td>0.531</td>
<td>0.87</td>
<td>0.589</td>
<td>0.815</td>
<td>0.12</td>
</tr>
<tr>
<td>OG DisagM EW</td>
<td>0.015</td>
<td>0.558</td>
<td>0.476</td>
<td>0.489</td>
<td>0.497</td>
<td>0.873</td>
<td>0.815</td>
<td>0.152</td>
</tr>
<tr>
<td>OG DisagM CRPS</td>
<td>0.006</td>
<td>0.798</td>
<td>0.239</td>
<td>0.444</td>
<td>0.875</td>
<td>0.782</td>
<td>0.242</td>
<td>0.187</td>
</tr>
<tr>
<td>OG DisagM RW</td>
<td>0.008</td>
<td>0.733</td>
<td>0.272</td>
<td>0.439</td>
<td>0.722</td>
<td>0.778</td>
<td>0.306</td>
<td>0.238</td>
</tr>
</tbody>
</table>

Table 5 shows that results are robust to focusing attention on the tail of the forecast density for inflation.

Table 5: PIT statistics: GDP deflator

<table>
<thead>
<tr>
<th>GDP Deflator</th>
<th>Tail score</th>
<th>OG Euro</th>
<th>OG Disag</th>
<th>OG agg</th>
<th>OG DisagM</th>
<th>Import/ oil price</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td>-1.33</td>
<td>-1.417</td>
<td>-1.476</td>
<td>-1.449</td>
<td>-1.404</td>
<td>-1.331</td>
</tr>
<tr>
<td>RW</td>
<td>-1.242</td>
<td>-1.244</td>
<td>-1.179</td>
<td>-1.236</td>
<td>-1.198</td>
<td>-1.182</td>
</tr>
<tr>
<td>CRPS</td>
<td>-1.259</td>
<td>-1.343</td>
<td>-1.296</td>
<td>-1.313</td>
<td>-1.211</td>
<td>-1.214</td>
</tr>
<tr>
<td>AR1</td>
<td>-1.603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR2</td>
<td>-1.591</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR3</td>
<td>-1.228</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR4</td>
<td>-1.197</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7 shows that in the Disag ensemble the German output gap has most explanatory power for Euro area inflation. But if we focus on the aggregate ensemble, which excludes the country-level output gap estimates, in Figure 8 we see that prior to the recession the US output gap was the most informative. Over the recession itself its informational content increased further, but this abruptly ends in 2009 when both of the indirect Euro area output gaps become important. Figure 9 shows that the different VARs do receive a weight in the DisagM ensemble of around 0.4, although this does vary over time.

*Figure 7: Weights on the 29 different output gap estimates in the disaggregate Output Gap ensemble*
Figure 8: Weights on the 11 different output gap estimates in the aggregate Output Gap ensemble
Figure 9: Weights on the multivariate BN cycle in the DisaggM ensemble; and weight on import prices and oil prices when forecasting the GDP deflator
To give a clearer idea of how consideration of the output gap estimates deliver improved inflation forecasts in Figure 10 we plot the probability of deflation from the recursively weighted AR ensemble and both the aggregate and disaggregate output gap ensembles. This shows clearly that the two output gap ensembles pick up the deflationary tendencies ahead of the AR ensemble which unsurprisingly can pick up the deflation only at a lag.

*Figure 10: Probability of deflation from the AR ensemble and the output gap ensembles, alongside the (scaled) outturn for Euro area inflation*

2.7.2 CONTROLLING FOR CHANGES IN IMPORT PRICES AND OIL PRICES

The analysis above establishes that when forecasting Euro area inflation information about the output gap does help. But there is at best only weak evidence that global gaps matter. There are gains to considering disaggregate Euro area gaps, specifically examination of the German output gap, over and above consideration of only Euro area gaps. But foreign gaps are found to matter only when these disaggregate gaps are not considered.

Previous work, such as Pain et al. (2006), has shown that any increases in the sensitivity of inflation to foreign economic conditions might work through import prices. Therefore, we consider an alternative ensemble which seeks to forecast inflation using both lagged import price deflator data and contemporaneous oil price inflation data, since these data unlike the deflator data are published at the end of the quarter concerned.
Table 6 shows that this RW ensemble, which makes no recourse to output gap data, performs almost as well as the Disag ensemble with a log score of -1.270 instead of -1.257. It also passes the pits tests. This shows that consideration of import/oil price data is informative. Inspection of Figure 9 shows that the RW weight on import prices and oil prices fluctuates over time but is always non-zero for both series.

**Table 6: Log Score statistics: GDP deflator using import price and oil price ensemble**

<table>
<thead>
<tr>
<th></th>
<th>EW</th>
<th>CRPS</th>
<th>RW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>-1.304</td>
<td>-1.281</td>
<td>-1.270</td>
</tr>
<tr>
<td>Score</td>
<td>0.750</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Table 7: PIT statistics: GDP deflator using import price and oil price ensemble**

<table>
<thead>
<tr>
<th></th>
<th>KLIC</th>
<th>LR2</th>
<th>tail</th>
<th>LR3</th>
<th>AD</th>
<th>wallis</th>
<th>LB1</th>
<th>LB2</th>
<th>LB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td>0.017</td>
<td>0.533</td>
<td>0.296</td>
<td>0.474</td>
<td>0.649</td>
<td>0.846</td>
<td>0.515</td>
<td>0.301</td>
<td>0.925</td>
</tr>
<tr>
<td>CRPS</td>
<td>0.01</td>
<td>0.675</td>
<td>0.315</td>
<td>0.415</td>
<td>0.814</td>
<td>0.744</td>
<td>0.899</td>
<td>0.476</td>
<td>0.867</td>
</tr>
<tr>
<td>RW</td>
<td>0.009</td>
<td>0.721</td>
<td>0.305</td>
<td>0.395</td>
<td>0.876</td>
<td>0.69</td>
<td>0.96</td>
<td>0.571</td>
<td>0.771</td>
</tr>
</tbody>
</table>

Figure 11 and 12 then show the probability of deflation from the AR, OG Euro and OG Disag gap based and import/oil price ensembles. It shows that the AR ensemble, as expected, is slow at picking up deflationary pressures. The gap-based forecasts are quickest and give the highest probabilities; but the import/oil price ensemble does perform much better than the AR even if not quite as well as the Disag ensemble. The OG Euro gap estimate picks up the deflationary tendencies earliest.
Figure 11: Probability of deflation from the recursively weighted ensembles
Figure 12: Probability of deflation from the recursively weighted ensembles with the Euro OG added in too.
2.7.3 GRAND ENSEMBLE

To try and weigh up the relative informational content of forming inflation density forecasts using autoregressive, output gap or import/oil price inflation information, we take the ensembles from each of these three systems and combine them at a second step, in what following Garratt et al. (2009) we call a grand ensemble. Specifically, we consider the Disag output gap ensemble, as this was the best performing.

Specifically, we take the ensemble density forecasts produced from expert AR, expert OG (Disag) and expert M and at a second step seek to combine them. This way each density receives an equal prior weight of 1/3.

Figure 13 plots the weights on the three experts. We lose the first observation as an extra training observation. We see considerable variation in the weights, especially for the optimised weights. Ignoring the first couple of years, to allow the sample period over which the average log score is calculated to increase, we see that the AR component receives a minimal weight. Interestingly, during the recession it does again receive a non-zero weight. But Figure 13 shows that information about both the output gap and import/oil price does help deliver improved inflation densities.

*Figure 13: Weights on the AR, Output Gap (OG Disag) and oil/import price (ensemble) experts when forecasting inflation in the grand ensemble*

It is of note that while the weights in Figure 13 detail the optimal weight on each expert, use of these weights in real-time, due to the fact that we have to wait a quarter to evaluate the densities, would not have delivered better density forecasts, as measured by the log score, than the best of the individual experts -
the OG expert. The OG expert has a log score of -1.188 over this period, the AR expert of -1.300, and the oil expert of -1.210; and the RW and optimised combinations have log scores of -1.203 and -1.215 respectively. But the weights in Figure 13 do indicate the relative informational content of the different experts, even if in real-time due to publication lags one could not exploit this to produce better forecasts.

Then to compare and contrast the ability of European and global output gap estimates, we combine the AR expert, the oil/import price expert and the OG Euro expert; see Figure 14.

*Figure 14: Weights on the AR, Euro Output Gap (OG) and oil/import price (ensemble) experts when forecasting inflation in the grand ensemble*

As before we see that the oil/import price receives the lion’s share of the weight for much of the period. But from 2006 we see the weight on the output gap increase. But the weight on the Euro OG does not get as high as that for OG Disag indicating that there is some extra informational content to the foreign information, but this is quite modest. Indeed the log score of the optimised combination is -1.244 and for the RW combination is -1.202, which are not statistically different from those when we consider OG Disag instead of OG Euro.

Therefore, our analysis shows that domestic inflation over the last decade has clearly been influenced by foreign economic conditions, working principally through import and oil prices rather than global output gaps. But the domestic output gap has played a role since late 2006. Consideration of import and oil prices, in conjunction with domestic and global output gap data, does help deliver better fitting density forecasts of inflation than autoregressive models. In contrast when point forecasting, output gaps at least do not so clearly help deliver improved forecasts.
2.7.4 FORECASTING THE IMPORT PRICE DEFLATOR

Rather than seek to forecast the GDP deflator we now repeat the exercise above and forecast import price inflation. Table shows 8 that there is far less between the different ensembles in terms of their log score statistics. But Table 9 shows that the pits are better for the output gap ensembles than the AR ensemble, though there is little between the Disag and DisagM ensembles.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR EW</td>
<td></td>
<td>-3.106</td>
</tr>
<tr>
<td>AR CRPS</td>
<td></td>
<td>-3.112</td>
</tr>
<tr>
<td>AR RW</td>
<td></td>
<td>-3.119</td>
</tr>
<tr>
<td>AR1</td>
<td></td>
<td>-3.071</td>
</tr>
<tr>
<td>AR2</td>
<td></td>
<td>-3.134</td>
</tr>
<tr>
<td>AR3</td>
<td></td>
<td>-3.163</td>
</tr>
<tr>
<td>AR4</td>
<td></td>
<td>-3.262</td>
</tr>
<tr>
<td>OG Euro EW</td>
<td></td>
<td>-3.077</td>
</tr>
<tr>
<td>OG Euro CRPS</td>
<td></td>
<td>-3.109</td>
</tr>
<tr>
<td>OG Euro RW</td>
<td></td>
<td>-3.103</td>
</tr>
<tr>
<td>OG Agg EW</td>
<td></td>
<td>-2.916</td>
</tr>
<tr>
<td>OG Agg CRPS</td>
<td></td>
<td>-2.975</td>
</tr>
<tr>
<td>OG Agg RW</td>
<td></td>
<td>-2.999</td>
</tr>
<tr>
<td>OG Disag EW</td>
<td></td>
<td>-2.934</td>
</tr>
<tr>
<td>OG Disag CRPS</td>
<td></td>
<td>-3.000</td>
</tr>
<tr>
<td>OG Disag RW</td>
<td></td>
<td>-3.019</td>
</tr>
<tr>
<td>OG DisagM EW</td>
<td></td>
<td>-2.945</td>
</tr>
<tr>
<td>OG DisagM CRPS</td>
<td></td>
<td>-3.015</td>
</tr>
<tr>
<td>OG DisagM RW</td>
<td></td>
<td>-3.026</td>
</tr>
</tbody>
</table>
Table 9: PIT statistics: Import price inflation

<table>
<thead>
<tr>
<th></th>
<th>KLIC</th>
<th>LR2</th>
<th>tail</th>
<th>tail</th>
<th>LR3</th>
<th>AD</th>
<th>wallis</th>
<th>LB1</th>
<th>LB2</th>
<th>LB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR EW</td>
<td>0.052</td>
<td>0.136</td>
<td>0.000</td>
<td>0.460</td>
<td>0.258</td>
<td>0.731</td>
<td>0.565</td>
<td>0.649</td>
<td>0.761</td>
<td>0.850</td>
</tr>
<tr>
<td>AR CRPS</td>
<td>0.064</td>
<td>0.087</td>
<td>0.000</td>
<td>0.451</td>
<td>0.158</td>
<td>0.584</td>
<td>0.423</td>
<td>0.625</td>
<td>0.743</td>
<td>0.416</td>
</tr>
<tr>
<td>AR RW</td>
<td>0.064</td>
<td>0.088</td>
<td>0.000</td>
<td>0.537</td>
<td>0.170</td>
<td>0.614</td>
<td>0.381</td>
<td>0.706</td>
<td>0.717</td>
<td>0.591</td>
</tr>
<tr>
<td>AR1</td>
<td>0.055</td>
<td>0.123</td>
<td>0.000</td>
<td>0.429</td>
<td>0.213</td>
<td>0.714</td>
<td>0.768</td>
<td>0.717</td>
<td>0.662</td>
<td>0.513</td>
</tr>
<tr>
<td>AR2</td>
<td>0.074</td>
<td>0.061</td>
<td>0.000</td>
<td>0.203</td>
<td>0.114</td>
<td>0.826</td>
<td>0.565</td>
<td>0.573</td>
<td>0.736</td>
<td>0.478</td>
</tr>
<tr>
<td>AR3</td>
<td>0.077</td>
<td>0.054</td>
<td>0.000</td>
<td>0.285</td>
<td>0.119</td>
<td>0.786</td>
<td>0.815</td>
<td>0.452</td>
<td>0.810</td>
<td>0.598</td>
</tr>
<tr>
<td>AR4</td>
<td>0.156</td>
<td>0.003</td>
<td>0.000</td>
<td>0.497</td>
<td>0.008</td>
<td>0.820</td>
<td>0.515</td>
<td>0.612</td>
<td>0.640</td>
<td>0.981</td>
</tr>
<tr>
<td>OG Euro EW</td>
<td>0.026</td>
<td>0.375</td>
<td>0.000</td>
<td>0.753</td>
<td>0.563</td>
<td>0.550</td>
<td>0.960</td>
<td>0.341</td>
<td>0.847</td>
<td>0.743</td>
</tr>
<tr>
<td>OG Euro CRPS</td>
<td>0.042</td>
<td>0.201</td>
<td>0.001</td>
<td>0.128</td>
<td>0.361</td>
<td>0.608</td>
<td>0.960</td>
<td>0.642</td>
<td>0.929</td>
<td>0.598</td>
</tr>
<tr>
<td>OG Euro RW</td>
<td>0.032</td>
<td>0.296</td>
<td>0.001</td>
<td>0.599</td>
<td>0.474</td>
<td>0.535</td>
<td>0.933</td>
<td>0.391</td>
<td>0.899</td>
<td>0.710</td>
</tr>
<tr>
<td>OG Agg EW</td>
<td>0.000</td>
<td>0.988</td>
<td>0.026</td>
<td>0.355</td>
<td>0.995</td>
<td>0.618</td>
<td>0.899</td>
<td>0.599</td>
<td>0.706</td>
<td>0.891</td>
</tr>
<tr>
<td>OG Agg EW CRPS</td>
<td>0.002</td>
<td>0.934</td>
<td>0.192</td>
<td>0.194</td>
<td>0.985</td>
<td>0.615</td>
<td>0.565</td>
<td>0.579</td>
<td>0.520</td>
<td>0.636</td>
</tr>
<tr>
<td>OG Agg EW RW</td>
<td>0.003</td>
<td>0.901</td>
<td>0.232</td>
<td>0.679</td>
<td>0.976</td>
<td>0.725</td>
<td>0.381</td>
<td>0.796</td>
<td>0.581</td>
<td>0.937</td>
</tr>
<tr>
<td>OG Disag EW</td>
<td>0.000</td>
<td>0.988</td>
<td>0.011</td>
<td>0.285</td>
<td>0.998</td>
<td>0.579</td>
<td>0.718</td>
<td>0.564</td>
<td>0.757</td>
<td>0.880</td>
</tr>
<tr>
<td>OG Disag CRPS</td>
<td>0.001</td>
<td>0.950</td>
<td>0.170</td>
<td>0.430</td>
<td>0.991</td>
<td>0.498</td>
<td>0.768</td>
<td>0.627</td>
<td>0.759</td>
<td>0.730</td>
</tr>
<tr>
<td>OG Disag RW</td>
<td>0.002</td>
<td>0.940</td>
<td>0.160</td>
<td>0.863</td>
<td>0.982</td>
<td>0.660</td>
<td>0.815</td>
<td>0.843</td>
<td>0.522</td>
<td>0.949</td>
</tr>
<tr>
<td>OG DisagM EW</td>
<td>0.015</td>
<td>0.558</td>
<td>0.476</td>
<td>0.489</td>
<td>0.497</td>
<td>0.873</td>
<td>0.815</td>
<td>0.152</td>
<td>0.916</td>
<td>0.690</td>
</tr>
<tr>
<td>OG DisagM CRPS</td>
<td>0.006</td>
<td>0.798</td>
<td>0.239</td>
<td>0.444</td>
<td>0.875</td>
<td>0.782</td>
<td>0.242</td>
<td>0.187</td>
<td>0.730</td>
<td>0.720</td>
</tr>
<tr>
<td>OG DisagM RW</td>
<td>0.008</td>
<td>0.733</td>
<td>0.272</td>
<td>0.439</td>
<td>0.722</td>
<td>0.778</td>
<td>0.306</td>
<td>0.238</td>
<td>0.806</td>
<td>0.696</td>
</tr>
</tbody>
</table>

Table 10: Tail score for the ensembles: Import price Deflator

<table>
<thead>
<tr>
<th>Tail score</th>
<th>AR</th>
<th>OG Euro</th>
<th>OG Agg</th>
<th>OG Disag</th>
<th>OG DisagM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1</td>
<td>-3.588</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AR2</td>
<td>3.766</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AR3</td>
<td>-3.796</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AR4</td>
<td>-3.999</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Figures 15-16 show that the weights on the competing output gap estimators are quite different when forecasting import price inflation; the global and, to a lesser degree, US output gap now play a big role in both ensembles, showing inference is robust. Indeed the global gaps are the most informative output gap estimates. So, as we might expect, foreign (global or US) output gaps do affect import/oil price inflation; but the evidence above suggests that they do not affect domestic inflation when one controls for disaggregate output gap.

**Figure 15: Weights on the 11 different output gap estimates in the aggregate Output Gap ensemble:**

*Import Price Deflator*
Figure 16: Weights on the 29 different output gap estimates in the disaggregate Output Gap ensemble: Import Price Deflator
2.8 Conclusion

Our analysis using real-time data shows that domestic inflation in the Euro area over the last decade has clearly been influenced by foreign economic conditions, working principally through import and oil prices rather than global output gaps. But the domestic output gap has played a role since late 2006. Consideration of import and oil prices, in conjunction with domestic and global output gap data, does help deliver better fitting density forecasts of inflation than autoregressive models. Therefore it is important to allow for the process determining inflation to vary over time, as documented in other studies; but there is not clear evidence that it is important to consider global output gaps when forecasting Euro area inflation. This contrasts the widely cited findings of Borio & Filardo (2007).
References


References


Robertson, D. & Wright, S. (2009), Stambaugh correlations and redundant predictors. University of Cambridge.


Appendix: Real-time data vintages

Our study uses real-time data for GDP for the Euro Area countries (Germany, France, Italy, Netherlands, Spain, Austria, Belgium, Luxembourg, Finland, Greece, Portugal, Ireland, Slovenia, Slovakia, Malta, Cyprus, and Estonia). Real-time data were downloaded from the EuroIND database of Eurostat. This section explains how we have prepared the real-time datasets of GDP for our analysis.

3.1 Real-time GDP data

Real-time vintages of seasonally adjusted quarterly real GDP at market price in million of euros were downloaded from the EuroIND database at Eurostat. Since the 2008Q1 release, seasonally adjusted data were also adjusted by working days. Eurostat discontinued compilation of National Account “at 1995 constant prices” from early 2008, while still offering figures in chain-linked volume with reference year 2000. Therefore, the set of real-time vintages we downloaded from EuroIND database were comprised of (1) vintages compiled “at 1995 constant prices” prior to the release in 2008Q1; (2) vintages compiled as chain-linked volume series with reference year 2000 since the release in 2008Q1.

3.1.1 CREATION OF REAL-TIME GDP DATA TRIANGLES

Real-time GDP data triangles for Euro Area and for the 17 individual countries in the Euro Area were created for our analysis.

We have taken the vintage first released for each quarterly GDP to create our real-time GDP data triangles. However, some data availability issues that affect the dimension of the real-time data triangles of first release arise:

1. Vintages do not start at the same point in time. That is, the beginning part of the vintages are missing. This issue affects the T dimension of the data triangle.

2. The first availability GDP estimate is either 2000Q4 or 2001Q1 for most EA17 countries and also for EA12 aggregate. But for some other smaller countries, their first available GDP estimate was for a quarter in more recent years. It means there are a smaller number of vintages in their data triangles, an issue affecting the N dimension.

The following sections explain how the real-time GDP data triangles are constructed in the face of these two issues.

Quarterly GDP data triangle for EA aggregate

Quarterly GDP data for the EA12 aggregate have vintages starting in 1990Q1 the earliest, and in the 1995Q1 at the latest. The first available vintage is released for the 2000Q4 estimate, while the last vintage available is for 2010Q1 which was published on 4 June 2010. Since then, the EA12 quarterly GDP in the EuroIND database has not been updated. Since we want our analysis to cover the latest period, i.e. to cover the entire year of 2010, vintages for the first release in the other 3 quarters in 2010 are needed. To resolve this problem, we append the vintages for the first release of quarterly GDP for EA16 from 2010Q2 to 2010Q4 to the vintages of the first release of EA12 from 2000Q4 to 2010Q1.

Moreover, the first release vintages in the data triangle have two start dates. Earlier vintages start in 1990Q1 while later vintages started in 1995Q1. We therefore make the assumption that there has been no revision to the estimates between 1990Q1 to 1994Q4 since the release of 1994Q4 first vintage, and filled up those missing part of vintages using the 1990Q1 to 1994Q4 revised estimated from the 1994Q4 release.
As a result of the above two resolutions, our $T \times N$ GDP matrix for EA aggregate thus have 40 first release vintages in total (i.e. $N = 41$), and covers 80 quarters for the period 1990Q1 to 2010Q4 (i.e. $T = 80$, but only the last vintage for 2010Q4 has all 80 observations).

**Quarterly GDP data triangle for individual countries in the Euro Area**

Our aim is to use disaggregate information from the individual countries to nowcast/forecast the aggregate Euro Area GDP; there is therefore a practical need to match the dimension of individual countries’ data triangle with that of the EA aggregate GDP triangle. The data availability issues mentioned above affects the dimensions of these data triangles.

For the issue that affects the $T$ dimension of the data triangles, we treated it the same way as we did for the aggregate EA GDP triangle. That is, to assume no revision in historical estimates and filled up those missing part of the vintages with the previous with available estimates for those missing quarters. As a result, data triangles for most EA12 countries contain vintages all start from 1991Q1, with $T = 80$ (with only the last vintage for 2010Q4 has all 80 observations). As a result, data triangles for most EA12 countries has 41 vintages in total i.e. $N = 41$. Those countries’ the earliest start date of the vintages is later than 1991Q1 (1995Q1 for Portugal, Luxembourg and Cyprus; 1997Q1 for Ireland; and 2001Q1 for Malta) have their real-time GDP matrices with $T < 80$. Luxembourg, Ireland, Malta, Greece, Estonia, Slovenia, Slovakia and Cyprus also have less vintages as their first available vintage was published for the estimate later than 2000Q4, so these matrices have $N < 41$. 
HOW TO OBTAIN EU PUBLICATIONS

Free publications:

• one copy:
  via EU Bookshop (http://bookshop.europa.eu);

• more than one copy or posters/maps:
  from the European Union’s representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

• via EU Bookshop (http://bookshop.europa.eu).
Globalisation and density forecasts of Euro-area inflation from Phillips Curve models

GIAN LUIGI MAZZI AND JAMES MITCHELL

In this paper we test whether globalisation has affected the relationship between the output gap and inflation in the Euro-area. Has it led to the breakdown of the traditional Phillips curve relationship and the emergence of an open-economy (global) Phillips Curve where global output gaps, or other indicators of globalisation, affect domestic inflation.

For more information
http://ec.europa.eu/eurostat/