UNIVERSITY OF CYPRUS
ECONOMICS RESEARCH CENTRE

Economic Analysis Papers

Forecasting economic activity in sectors of the Cypriot economy using VAR models

Nicoletta Pashourtidou
Economics Research Centre,
University of Cyprus

Christos Papamichael
Economics Research Centre,
University of Cyprus

01-08

April 2018

Publication Editor: Christos Koutsampelas

ERC Sponsors (in alphabetical order)

Central Bank of Cyprus

Economics Department, University of Cyprus

Ministry of Finance

University of Cyprus

Forecasting economic activity in sectors of the Cypriot economy using VAR models*

Nicoletta Pashourtidou and Christos Papamichael**

ABSTRACT

This paper explores the use of Vector Autoregressions (VARs) for the construction of short-term growth forecasts for the production-side components of GDP as well as for GDP growth via aggregation of its component forecasts. The production-side components of GDP are given by the gross value added in six sectors of economic activity (i.e. agriculture, industry, construction, trade and transport, financial and business activities, and other services), and import duties plus value added tax. Apart from simple VARs, we also consider VAR models augmented with exogenous variables represented by common factors. The common factors are computed from separate blocks of series in the dataset so that the resulting factors represent different aspects of the economy.

The results show that, in general, component forecasts obtained from VARs with exogenous factors outperform simple VAR forecasts. Forecast gains, however, are mainly attained for short horizons. The exceptions are the construction sector in which some gains are found for longer horizons, and the sector of financial and business activities in which significant gains are attained for both short and long horizons. Component forecasts computed from single equations are at least as accurate as VAR forecasts. However, the contraction in real activity during the period 2012 – 2013 was better predicted by VAR models in the case of large components such as trade, financial and business activities, and other services. Similarly, bottom-up GDP growth forecasts from single equation models for components are at least as precise as indirect forecasts using VARs to compute the component forecasts. Direct GDP growth forecasts are the best performers over the period examined.

Keywords: VAR, forecasting, combination forecasts, GDP, gross value added, bottom-up forecasts.

* The Economics Research Centre gratefully acknowledges funding from the Central Bank of Cyprus and the Ministry of Finance. The views expressed in this paper are those of the authors.

** Corresponding author: n. pashourtidou@ucy.ac.cy

CONTENTS

Συστήματα εξισώσεων για πρόβλεψη του ρυθμού ανάπτυξης τομέων της κυπριακής οικονομίας

Νικολέττα Πασιουρτίδου και Χρίστος Παπαμιχαήλ

ΠΕΡΙΛΗΨΗ

Το δοκίμιο εξετάζει την εφαρμογή μοντέλων τύπου Vector Autoregressions (VAR) για βραχυπρόθεσμη πρόβλεψη του ρυθμού ανάπτυξης των συνιστωσών του ΑΕΠ, από την πλευρά της προσφοράς, καθώς και του ρυθμού μεταβολής του ΑΕΠ αθροίζοντας τις προβλέψεις των συνιστωσών του. Οι συνιστώσες του ΑΕΠ αποτελούνται από την ακαθάριστη προστιθέμενη αξία έξι τομέων οικονομικής δραστηριότητας (γεωργία, βιομηχανία, κατασκευές, εμπόριο και μεταφορές, χρηματοπιστωτικές και επιχειρηματικές δραστηριότητες και άλλες υπηρεσίες) και από τους εισαγωγικούς δασμούς συν φόρο προστιθέμενης αξίας. Επίσης, εκτιμώνται μοντέλα VAR που περιέχουν ως εξωγενείς μεταβλητές κοινούς παράγοντες. Οι κοινοί παράγοντες αντιπροσωπεύουν διάφορες πτυχές της οικονομίας όπως (α) πραγματική οικονομική δραστηριότητα και αγορά εργασίας, (β) εγχώριες χρηματοπιστωτικές και δημοσιονομικές μεταβλητές, τιμές και δείκτες οικονομικού κλίματος και (γ) ξένες/διεθνείς χρηματοοικονομικές μεταβλητές, ευρωπαϊκούς δείκτες τιμών και δείκτες εμπιστοσύνης.

Τα αποτελέσματα δείχνουν ότι, γενικά, οι προβλέψεις συνιστωσών του ΑΕΠ από μοντέλα VAR με εξωγενείς παράγοντες είναι ακριβέστερες από αυτές που υπολογίζονται από απλά μοντέλα VAR. Ωστόσο, η αυξημένη ακρίβεια των προβλέψεων επιτυγχάνεται κυρίως για πολύ κοντινό ορίζοντα. Εξαίρεση αποτελεί ο τομέας των κατασκευών, στον οποίο η βελτίωση στην ακρίβεια των προβλέψεων καταγράφεται για μακρινό ορίζοντα, και ο τομέας των χρηματοπιστωτικών και επιχειρηματικών δραστηριοτήτων, στον οποίο το σφάλμα της πρόβλεψης μειώνεται σημαντικά τόσο για κοντινό όσο και για πιο μακρινό ορίζοντα. Οι προβλέψεις συνιστωσών που υπολογίζονται από μοντέλα μιας εξίσωσης (αντί συστήματα) είναι τουλάχιστον εξίσου ακριβείς με τις προβλέψεις από VAR μοντέλα. Παρά ταύτα, η συρρίκνωση της πραγματικής οικονομικής δραστηριότητας κατά την περίοδο 2012 - 2013 είχε προβλεφθεί με μεγαλύτερη ακρίβεια από τα μοντέλα VAR, στην περίπτωση σημαντικών τομέων όπως το εμπόριο, οι χρηματοοικονομικές και επιχειρηματικές δραστηριότητες και άλλες υπηρεσίες. Οι προβλέψεις του ρυθμού ανάπτυξης του ΑΕΠ αθροίζοντας προβλέψεις για τις συνιστώσες του, οι οποίες υπολογίζονται από μοντέλα μιας εξίσωσης, είναι τουλάχιστον το ίδιο ακριβείς όσο και οι έμμεσες προβλέψεις για τον ρυθμό μεταβολής του ΑΕΠ που βασίζονται σε προβλέψεις των συνιστωσών του από μοντέλα VAR. Η άμεση πρόβλεψη του ρυθμού μεταβολής του ΑΕΠ μέσω μοντέλων μιας εξίσωσης αποτελεί τη μέθοδο με την καλύτερη επίδοση για την υπό εξέταση περίοδο.

Η αξιολόγηση της ακρίβειας των προβλέψεων που υπολογίζονται από τις διάφορες μεθόδους, θα πρέπει να διεξάγεται σε συστηματική βάση καθώς το μέγεθος των διαθέσιμων χρονοσειρών επηρεάζει την αξιοπιστία των αποτελεσμάτων, ειδικότερα σε ορίζοντα πέραν των τεσσάρων τριμήνων. Ωστόσο, η έμμεση πρόβλεψη του ΑΕΠ μέσω συστημάτων τύπου VAR αποτελεί ένα ολοκληρωμένο πλαίσιο για την ανάλυση του ρόλου των τομέων στους οικονομικούς κύκλους και την εκτίμηση των επιπτώσεων εξωγενών παραγόντων σε κλάδους και στο σύνολο της οικονομίας.

1. Introduction

The dynamic interrelations among different sectors of the economy have been extensively used in the literature to evaluate the impact of various shocks and their role in generating business cycles (e.g. Atalay 2017; Foerster et al. 2011; Forni and Reichlin 1998; Lee et al. 1992; Long and Plosser 1983, 1987; Pesaran et al. 1993). For practitioners, the systematic analysis of developments and outlook for sectors of economic activity can provide timely insights into the drivers of output fluctuations.

Econometric techniques, such as common factors and forecast combinations, exploit the richness of information in large databases of economic and financial indicators for forecasting macroeconomic variables, such as growth and inflation (e.g. Artis et al. 2005 for the UK; Giannone et al. 2008, and Stock and Watson 2002a, 2008 for the US; Stock and Watson 2004 for OECD countries). In the context of forecasting macroeconomic aggregates, the literature is also concerned with the level of disaggregation at which the forecasts are computed and typically distinguishes between direct and bottom-up forecasting approaches. Lütkepohl (2010) provides theoretical results on the relative efficiencies of aggregate and disaggregate forecasts under some assumptions and offers some guidelines for applied work. Theoretically, forecasting the disaggregate components using a multivariate model is at least as efficient in terms of mean squared error as directly forecasting the aggregate. However, in practice, issues such as specification and estimation uncertainty, non-linear transformations of the variables of interest and time-varying aggregation weights lead to departures from the theoretical assumptions, and empirical findings could deviate from theoretical results. For example, computing bottom-up forecasts for the aggregate by modelling the disaggregates using a high dimensional multivariate model or a large number of disaggregate single equation models may result in higher estimation uncertainty than directly modelling and forecasting the aggregate variable.

The aim of this paper is to explore the use of multivariate models, and more specifically Vector Autoregressions (VARs), for the construction of short-term forecasts for sectors of economic activity in Cyprus. The growth forecasts fully cover the production-side components (in constant prices) of GDP, namely the Gross Value Added (GVA) of sectors, and import duties plus Value Added Tax (VAT). The growth rate of GVA in the following six sectors is forecasted: (i) agriculture, (ii) industry, (iii) construction, (iv) trade and transport, (v) financial and business activities, and (vi) other services. Apart from simple VARs, we also consider VAR models augmented with exogenous variables represented by common factors. The common factors are computed from separate blocks of series in the dataset so that the resulting factors represent different aspects of the economy. The sectoral forecasts obtained from different VAR models are combined using forecast combinations. Subsequently, the sectoral forecasts are used to compute bottom-up or, equivalently, indirect, GDP growth forecasts; the aggregate forecasts are compared with those obtained from alternative methods of forecasting GDP growth.

Empirical evidence on the forecasting performance of direct aggregate forecasts and bottom-up forecasts computed by aggregating predictions for the disaggregate components is mixed. Based on a literature search, Bermingham and D'Agostino (2014) conclude that papers which do not find evidence in favour of bottom-up forecasts conduct their analysis using (a) short time series and therefore higher estimation error, and (b) a rather small number of disaggregates.

Drechsel and Scheufele (2013) forecast GDP growth in Germany using the direct approach as well as the bottom-up approach, via both supply and demand sides. They employ mixed-data sampling regressions together with model averaging techniques. Their findings suggest that aggregating sector-specific forecasts results in limited forecasting gains compared to forecasting aggregate GDP growth directly, whereas both approaches outperform forecasts produced from the demand side. Dias et al. (2017) use factor-augmented models to compute short-term forecasts for the growth rate of GDP and its demand components in the case of Portugal. They use a targeted diffusion index which incorporates information from all the factors in the database, the variable forecasted and the forecast horizon. They find that models with the targeted index generate more accurate GDP growth forecasts through the bottom-up approach than via forecasting GDP growth directly. The empirical findings of the comparison of direct and bottom-up forecasts for the euro area GDP growth in Foroni and Marcellino (2014) are less strong, but the authors conclude that there is scope for forecasting the components to gain better understanding of the aggregate. They employ a large dataset of monthly indicators and different modelling approaches (e.g. bridge equations, MIDAS and mixed frequency VAR models) for nowcasting both expenditure and production GDP components. They also find that the use of forecast combinations and factor models reduces forecast errors.

Other studies explore the accuracy of bottom-up GDP growth forecasts without providing comparisons with direct forecasts. Barhoumi et al. (2012) forecast GDP growth in France from both the supply and demand sides, using component-specific bridge models. Their results suggest that forecasting GDP growth from the supply side is superior to following a demand-side, bottom-up approach; their finding is likely driven by the availability of more relevant indicators for sectors of economic activity than for expenditure components. Hahn and Skudelny (2008) compute euro area GDP growth forecasts using the bottom-up approach from the production side together with sector-specific bridge equations that vary across the forecast cycle. Their results suggest that the importance of individual predictors varies substantially over the forecast cycle, with survey data being more valuable at earlier stages of the forecast cycle and hard data being more useful at later stages of the cycle.

In the case of inflation, the role of the level of disaggregation in forecast precision has been more thoroughly explored due to the availability of a larger number of disaggregates compared to GDP. Hendry and Hubrich (2011) find that direct forecasts of aggregate US inflation computed using information from selected disaggregate variables or factors representing disaggregate information, are superior to bottom-up forecasts and direct forecasts of the aggregate using only past aggregate information. Hubrich (2005) finds that for horizons relevant for monetary policy, direct euro area inflation forecasts are more accurate than indirect forecasts obtained by aggregating forecasts for inflation sub-indices. Bermingham and D'Agostino (2014) investigate the case of US and euro area inflation and find that the accuracy of the aggregate forecast is enhanced when a large number of disaggregates are used. Duarte and Rua (2007) compare direct inflation forecasts for the Portuguese CPI with bottom-up forecasts computed from intermediate (five-component) and higher (59-component) disaggregation levels; their results indicate an inverse relationship between the forecast horizon and the level of disaggregation. Bruneau et al. (2007) and Moser et al. (2007) use large datasets of economic indicators to forecast HICP inflation in France and Austria, respectively; both papers find that bottom-up forecasts through the construction of forecasts for HICP components are superior to direct inflation forecasts.

Previous studies in the case of Cyprus which used large datasets of predictors, factor-augmented single equation models and combination forecasts found evidence in favour of direct GDP growth forecasts (Papamichael and Pashourtidou 2014; Pashourtidou et al. 2017). The direct forecasts were compared to bottom-up forecasts obtained from the expenditure components or the production side of GDP. This paper uses the production-side components in a less detailed way than in previous work, in order to model and forecast them jointly using a single model, which can add to the mutual consistency of the resulting sectoral forecasts.

The structure of the paper is as follows. Section 2 describes the forecast methods. Section 3 presents the data. Section 4 examines the forecasting performance of growth forecasts for supply-side components of GDP as well as for GDP. Section 5 concludes.

2. Forecasting methods

We use simple VAR models for the growth rates of the production-side components of GDP as well as VAR models that also include exogenous variables in the form of common factors. The inclusion of factors in the VARs allow us to assess the forecasting gains resulting from utilising information from different macroeconomic and financial predictors beyond that contained in history of the component growth rates.

Factor models summarise the information from a large number of economic/financial time series by a small number of estimated indices known as common factors. Thus, the dynamics of a dataset of many economic time series can be driven by a small number of common shocks and a set of idiosyncratic components, i.e. one series-specific shock for each variable in the dataset. The factors are estimated using the principal components method. [1]

2.1 Models

The variables to be forecasted are expressed as annualised percentage changes and as a function of the forecasting horizon. Let $y_{t+h}^{h,s} = (400/h)(\ln Z_{t+h}^{s} - \ln Z_{t}^{s})$ denote the annualised growth rate of component s, $s = 1, \ldots, 7$, over the next h quarters, for $h = 1, \ldots, 8$. Then, collecting all the components in a vector, $y_{t+h}^{h} = [y_{t+h}^{h,1} \; y_{t+h}^{h,2} \; \cdots \; y_{t+h}^{h,7}]'$, the h-step ahead VAR model used for computing the forecasts for $h = 1, \ldots, 8$ is given by

$$y_{t+h}^{h} = A + \sum_{i=0}^{q} B_i y_{t-i} + e_{t+h}^{h}\ , \qquad t = 1, \ldots, T \tag{1}$$

where $y_{t-i} = [y_{t-i}^{1} \; y_{t-i}^{2} \; \cdots \; y_{t-i}^{7}]'$ is the vector of past growth rates of components, A is a 7×1 vector of constants, B_i is a 7×7 matrix of coefficients of lagged growth rates and e_{t+h}^{h} is the vector error term.

Equation (1) gives the VAR model of order q. For $B_i = 0$, equation (1) reduces to the random walk model for the log-level of component s, which is a constant growth model for component s.

The simple VAR model can be extended by including estimated quarterly factors which are treated as exogenous variables, i.e.,

$$y_{t+h}^{h} = A + \sum_{i=0}^{q} B_i y_{t-i} + \sum_{i=0}^{l} \Phi_i \hat{f}_{t-i} + \varepsilon_{t+h}^{h} \tag{2}$$

where $\hat{f}_t$ is one of the estimated factors summarising one of three different blocks of series in the dataset, representing aspects of the domestic economy and international economic environment; Φ_i is a 7×1 matrix of factor coefficients. Due to the large number of parameters in the VAR model, the factors are introduced one at a time and their joint information content is exploited through the computation of forecast combinations. More details about the factors and forecast combinations are given in section 3 and section 2.2, respectively.

The estimation of the parameters and the lag length selection in models (1) and (2) are carried out in a pseudo out-of-sample setup using recursive OLS and recursive determination of lag order. We use three alternative criteria for lag determination, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the Hannan–Quinn Information Criterion (HIC). For comparison purposes we also estimate the random walk model for each (log) component. The forecast constructed in period t, for period $t + h$, uses

[1] For details about the factor model and estimation see Stock and Watson (2002a, 2002b).

data up to t, thus no additional projections for predictors are required. This type of forecast is known as 'direct', unlike the case of iterated forecasts (see e.g. Stock and Watson 2004, 2008). The choice of the number of lags for the endogenous variables in vector y_t and factor $\hat{f}_t$ is between one and four.

2.2 Forecast combinations

The h-quarter ahead forecast for component s, $y_{t+h}^{h,s}$, computed in period t is given by $\hat{y}_{t+h}^{h,s}$. Then the Mean Squared Forecast Error (MSFE) used for evaluating the forecasting performance of each component equation in the VAR, for horizon h is given by

$$MSFE_s = \frac{1}{(T-h)-(T_0+h)+1} \sum_{t=T_0}^{T-2h} (y_{t+h}^{h,s} - \hat{y}_{t+h}^{h,s})^2 \quad (3)$$

where T_0 is the number of observations used for estimation only (i.e. the first T_0 observations in the sample) and T denotes the number of available observations in the sample.

The inclusion of one factor at a time in the model, allow us to estimate a number of different VARs and to obtain many alternative forecasts for the components of interest. The forecasts for each component can be further exploited by constructing combinations of forecasts. Forecast combinations can provide more accurate forecasts by taking into account evidence from all the models considered rather than relying on a specific model (e.g. Stock and Watson 2004, 2008). Forecast combinations reduce the uncertainty resulting from the specification of individual models due to different set of predictors, lag structures and modelling approaches.

There are different methods to construct forecast combinations depending on how the forecast weights are formed. Given M models and associated forecasts, a combination forecast for component s denoted by $\hat{F}_{t+h}^{h,s}$, is the weighted average of individual forecasts, with fixed or time-varying weights,

$$\hat{F}_{t+h}^{h,s} = \sum_{i=1}^{M} w_{i,t}^{s} \, \hat{y}_{i,t+h}^{h,s} \quad (5)$$

where $\hat{y}_{i,t+h}^{h,s}$ is the $h-$step ahead forecast from model i computed in period t and $w_{i,t}^{s}$ is the weight assigned to that forecast. In general, the weights ($w_{i,t}^{s}$) depend on the historical forecasting performance of model i, however weights can be fixed, leading to simple forecast combinations such as the mean ($w_{i,t}^{s} = 1/M$) or the median. In cases in which the weights depend on models' past forecasting performance, the resulting combination forecasts are known as discounted MSFE forecasts (Stock and Watson 2004, 2008). The weights can be inversely proportional to the discounted MSFE (or the square of the discounted MSFE) of the individual models, i.e.

$$w_{i,t}^{s} = \frac{\epsilon_{i,t}^{s}}{\sum_{j=1}^{M} \epsilon_{j,t}^{s}} \text{, or,} \quad (6)$$

$$w_{i,t}^{s} = \frac{(\epsilon_{i,t}^{s})^2}{\sum_{j=1}^{M}(\epsilon_{j,t}^{s})^2} \quad (7)$$

where $\epsilon_{i,t}^{s} = \left[\sum_{k=T_0}^{t-h} \delta^{t-h-k}(y_{k+h}^{h,s} - \hat{y}_{i,k+h}^{h,s})^2\right]^{-1}$;

δ is the discount factor so that forecast errors made in the distant past are of smaller importance. Larger weights are assigned to forecasts from models with lower MSFE (i.e. better historical forecasting performance).

3. Data

The data on the variables of interest are obtained from the quarterly national accounts and cover the period 1995Q1 – 2017Q2. Forecasts are constructed for the growth rate of seven production-side components of GDP, namely GVA in constant prices in the following six sectors:

1. agriculture, forestry and fishing (NACE code A),
2. industry (NACE codes B, C, D, E),
3. construction (NACE code F),
4. trade, transport, accommodation and food services (NACE codes G, H, I),
5. financial services and business activities (NACE codes J, K, L, M, N),
6. other services (NACE codes O, P, Q, R, S, T),

and

7. import duties plus VAT.

The seven components add up to GDP and thus forecasts for GDP growth are also computed via the projected growth rates of these production-side components.

Some statistics on the variables of interest are shown in Table 1. The broader services sectors comprise about 70% of GDP. The value added of financial services and business activities accounts for 25% of GDP, while trade, transport, accommodation and food services is the second largest sector with a contribution of 23%. Over the last 20 years, the sector of financial and business activities has registered the fastest growth rate, while the primary and secondary sectors have contracted. The sectors of construction and agriculture exhibit the highest volatility in activity growth.

TABLE 1

GDP and production-side components

		GDP	GVA: Agriculture, forestry, fishing	GVA: Industry	GVA: Construction	GVA: Trade, transport, accommodation and food services	GVA: Financial and business activities	GVA: Other services	Import duties plus VAT
Component share to GDP	Mean		2.6	8.4	6.7	22.6	24.9	22.2	12.6
	St. Dev.		0.8	1.7	1.8	0.6	3.3	0.7	0.1
Year-on-year changes (%)	Mean	2.1	-1.8	-0.5	-0.6	2.0	4.1	2.2	2.1
	St. Dev.	3.1	10.9	5.2	12.9	4.3	3.1	2.6	3.1
Quarter-on-quarter changes (%)	Mean	0.5	-0.5	-0.2	-0.1	0.5	1.0	0.5	0.6
	St. Dev.	1.0	8.6	1.7	5.5	1.5	0.9	0.8	1.6

Table 2 presents the results of Granger causality tests which are used for checking whether the lagged growth rates of a given component contain information for the current growth rate of another component, beyond that conveyed by its own past values.

TABLE 2

Granger causality tests, p-values of Chi-square tests

Equation [3]	Lagged terms excluded [1]: Agriculture, forestry, fishing	Industry	Construction	Trade, transport, accommodation and food services	Financial and business activities	Other services	Import duties plus VAT	All lagged terms excluded (except own lags) [2]
Agriculture	—	0.0000	0.0900	0.6480	0.0000	0.1070	0.1500	0.0000
Industry	0.0270	—	0.0530	0.0040	0.0480	0.3710	0.2640	0.0060
Construction	0.4710	0.0950	—	0.0000	0.0480	0.3870	0.3780	0.0030
Trade, transport, accommodation and food services	0.2270	0.0030	0.0150	—	0.1120	0.3330	0.9990	0.0070
Financial and business activities	0.3290	0.0660	0.0260	0.4850	—	0.2690	0.0990	0.0230
Other services	0.9390	0.0330	0.0190	0.4120	0.0000	—	0.9650	0.0000
Import duties plus VAT	0.0010	0.0000	0.0000	0.0000	0.0000	0.0420	—	0.0000

Notes: [1] The null hypothesis is that the coefficients on all the lags of the endogenous variable j are jointly equal to zero in the equation for component i.
[2] The null hypothesis is that the coefficients on all the lags of endogenous variables in the equation for component i, except the coefficients on the lags of variable i, are jointly equal to zero.
[3] The tests are carried out in a VAR model of order four estimated using 85 observations. The p-value for the Lagrange Multiplier test for the absence of first order autocorrelation in the VAR residuals is 0.05. The p-value for the Jarque-Bera test for normality in the VAR disturbances is 0.78.

When the endogenous variables in each VAR equation are tested separately, the hypothesis that the j-th endogenous variable does not Granger-cause the i-th dependent variable (i.e. in the i-th equation) cannot be rejected at 10% significance level for 24 out of the 42 cases. For example, the hypothesis that past changes in trade GVA are not informative for the current growth rate in agriculture cannot be rejected. However, when the joint significance of all the lagged terms, except own lags, is tested, the hypothesis of no Granger causality is rejected in all equations. Therefore, there is evidence that all components are jointly useful for predicting each production-side component of GDP.

The dataset contains additional quarterly variables (about 290) over the period 1995Q1 – 2017Q2 that represent many aspects of the domestic economy and the external economic environment. Domestic data include variables such as economic activity indicators (e.g. volume indices of retail trade and manufacturing, building permits, tourist arrivals, etc.), labour market series (employment, unemployment, vacancies), fiscal data and public debt, banking sector data (loans, deposits, interest rates), price indices, Cyprus Stock Exchange indices and data on business and consumer confidence. Foreign/international data are comprised of euro exchange rates to different currencies (US dollar, British pound, Russian rouble, etc.), foreign activity and labour market indicators, European interest rates and spreads, foreign price indices and international commodity prices (e.g. oil, gold, wheat, etc.), stock market indicators and European economic confidence/sentiment indicators.[2]

The dataset is split in three blocks: (a) domestic real economic activity and labour market indicators, (b) domestic financial and fiscal variables, prices and economic confidence indicators, and (c) foreign/international financial variables, prices and European confidence indicators. The dataset is summarised by extracting a small number of common factors from each block, representing three distinct aspects of the economy. The factors are estimated via the application of principal component analysis. The factors are used as predictors in the VAR models and therefore are estimated recursively over the pseudo-out-of-sample period spanning from 2008Q3 to 2017Q1.

Table 3 presents the number of factors chosen (out of a total of five factors per block) by three alternative information criteria —ICP1, ICP2, ICP3— (Bai and Ng 2002). These criteria suggest at most three factors for each block of data. The first three factors computed recursively over the pseudo-out-of-sample period account for about a third of the variation in the block of domestic real activity and labour market variables, and about two fifths of the total variance in the block of other domestic indicators. The first three factors capture about half of the cross-section variation in the block of foreign/international economic and financial indicators.

[2] Data are obtained from the following sources: Statistical Service of Cyprus, Central Bank of Cyprus, Eurostat, European Central Bank, European Commission, Datastream, Global Financial Data and other local sources (e.g. Cyprus Stock Exchange, Department of Land and Surveys, Department of Registrar of Companies).

TABLE 3

Estimation of factors

	Blocks of variables		
	Domestic real economic activity and labour market variables	Other domestic variables	Foreign/international variables
Number of series	124	66	101
Pseudo-out-of-sample period			
Number of factors estimated by each information criterion (ICP)			
ICP1	1 – 2	3	3
ICP2	1 – 2	2 – 3	3
ICP3	3	3	3
Percentage of block variance explained by the first three factors	30% – 36%	41% – 45%	47% – 54%
Full sample			
Number of factors estimated by each information criterion (ICP)			
ICP1	2	3	3
ICP2	2	3	3
ICP3	3	3	3
Percentage of block variance explained by the first three factors	34%	42%	47%

4. Forecasting performance

The VAR models described in Section 2 are employed for forecasting the growth rates of the production-side components of GDP i.e. GVA (constant prices) in six sectors of economic activity, and import duties plus VAT. We also consider forecast combinations of the component forecasts obtained from VAR models with factors as the forecasting performance of models based on individual predictors could be unstable over time (e.g. Stock and Watson 2004). A recursive pseudo-out-of-sample forecasting exercise is carried out using data over the period 1995Q1 – 2017Q2; the first estimation period consists of 50 observations and the pseudo-out-of-sample forecasts are constructed for horizons of one to eight quarters ahead.

4.1 Component forecasts

Table 4 presents the results of the forecasting exercise for the growth rate of the seven production-side GDP components. The forecasts are computed using the following methods: (a) simple VAR models for the seven components, and (b) combinations of forecasts obtained for each component from VAR models which include exogenous factors. The table presents the square root of MSFE (RMSFE) of the different methods relative to that of the random walk benchmark for each component. The entries in bold indicate that the forecasts from a model/method are more accurate than the naïve benchmark forecasts, and the difference in the performance is statistically significant.

Simple VAR models yield some forecasting gains, especially for short horizons. However, the difference in the precision between VAR and benchmark forecasts is not always statistically significant. The largest gains in predictive accuracy resulting from VAR forecasts are achieved for the sectors of financial and business activities, and other services as well as for the component of import duties and VAT. The forecasting performance of VAR models is further improved with the inclusion of estimated factors. The factors summarise a large number of domestic and foreign/international macroeconomic and financial indicators. The combinations of forecasts computed from VARs with exogenous factors outperform naïve and simple VAR forecasts in most cases. For industry, trade and other services as well as for duties, the increase in predictive accuracy from the inclusion of factors is significant vis-à-vis the benchmark for very short horizons only. For the construction sector forecast gains are larger for long horizons, while for the financial sector significant improvements over the benchmark are achieved for both short and long horizons.

We also compare the performance of sectoral forecasts obtained through VAR models to those computed using the single equation methods employed in previous work. As the number of unknown parameters in single equation models is smaller compared to that in VARs, the former are used to compute forecasts for more disaggregates than the VARs. The single equation methods explored in previous work provide forecasts for all of the 11 components that appear on the production side of the national accounts.[3] Five out of 11 components used in single equation models, namely agriculture–forestry–fishing, industry, construction, trade–transport–accommodation-food, and duties–VAT, are identical to those in VARs. The remaining six production-side components, which are forecasted individually using single equations, are aggregated into two broader services sectors, i.e. financial and business activities, and other services, when they are forecasted via VAR models.

[3] The single equation dynamic models for the production-side components consist of autoregressive distributed lag (ADL) models and factor augmented ADL (FADL) models. The dataset used in the estimation of the single equation models is the same as that used in VARs. The factors in the FADL models summarise a large set of domestic and foreign indicators on real economic activity as well as the domestic and European labour market. The predictors (other than own lags) in the ADL models include a large number of different macroeconomic and financial indicators, other than those used for extracting the factors. The predictors are included in the ADL models one at a time, together with their leads, if available. The FADL models are an extension of the ADL models by including factors (one at a time) as additional predictors, together with leads of variables included in the factors, if available. For details of sectoral forecasts based on single equation models, see Pashourtidou et al. (2017).

TABLE 4

Relative RMSFE vis-à-vis the random walk, production-side components

Forecast horizon (quarters)	1	2	3	4	5	6	7	8
1. Agriculture, forestry, fishing								
Random walk (log-levels) benchmark, RMSFE	9.06	9.33	9.60	11.23	11.75	12.19	12.69	13.91
VAR models								
VAR(1)	0.90	1.14	1.06	1.04	1.10	1.15	1.03	1.03
VAR(4)	1.05	1.14	1.06	1.04	1.10	1.12	1.03	1.03
VAR(AIC)	0.90	0.95	1.06	1.04	1.10	1.15	1.03	1.03
VAR(BIC)	1.17	1.53	1.84	1.76	1.51	1.37	1.48	1.69
VAR(HIC)	1.17	1.46	1.84	1.76	1.51	1.37	1.48	1.69
VAR models with exogenous factors								
Forecast combinations								
Median	0.91	0.96	1.06	1.05	1.10	1.15	1.01	1.02
Mean	0.92	0.96	1.08	1.06	1.09	1.15	1.01	1.02
Discounted MSFE (0.90)	0.94	0.97	1.08	1.07	1.09	1.16	1.01	1.00
Squared discounted MSFE (0.90)	0.98	0.97	1.09	1.09	1.09	1.16	1.01	0.99
2. Industry								
Random walk (log-levels) benchmark, RMSFE	2.57	4.47	6.68	8.94	7.42	5.94	4.92	5.09
VAR models								
VAR(1)	0.89	0.86	0.98	1.10	1.25	1.65	2.23	2.60
VAR(4)	0.92	0.86	0.98	1.10	1.25	2.10	2.23	2.60
VAR(AIC)	0.89	0.89	0.98	1.10	1.25	1.65	2.23	2.60
VAR(BIC)	1.03	1.10	1.34	1.57	2.13	3.15	4.37	4.96
VAR(HIC)	1.03	1.07	1.34	1.57	2.13	3.15	4.37	4.96
VAR models with exogenous factors								
Forecast combinations								
Median	0.88	0.87	0.98	1.09	1.24	1.64	2.22	2.59
Mean	**0.87**	0.86	0.96	1.07	1.21	1.62	2.19	2.57
Discounted MSFE (0.90)	**0.87**	0.86	0.96	1.07	1.21	1.62	2.18	2.56
Squared discounted MSFE (0.90)	**0.86**	0.87	0.96	1.07	1.21	1.61	2.18	2.56
3. Construction								
Random walk (log-levels) benchmark, RMSFE	6.60	10.23	14.23	19.07	19.90	19.46	15.67	14.79
VAR models								
VAR(1)	1.09	0.98	0.97	0.93	1.01	0.95	0.86	0.95
VAR(4)	1.17	0.98	0.97	0.93	1.01	1.06	0.86	0.95
VAR(AIC)	1.09	0.96	0.97	0.93	1.01	0.95	0.86	0.95
VAR(BIC)	1.28	1.31	1.27	1.34	1.46	1.63	2.02	2.96
VAR(HIC)	1.28	1.25	1.27	1.34	1.46	1.63	2.02	2.96
VAR models with exogenous factors								
Forecast combinations								
Median	1.09	0.97	0.97	0.92	0.99	**0.93**	0.85	0.94
Mean	1.08	0.97	0.97	0.92	0.98	**0.92**	0.82	0.88
Discounted MSFE (0.90)	1.08	0.96	0.97	0.92	0.98	**0.92**	0.82	0.87
Squared discounted MSFE (0.90)	1.08	0.96	0.97	0.93	0.98	**0.92**	0.83	0.86
4. Trade, transport, accommodation and food services								
Random walk (log-levels) benchmark, RMSFE	1.50	2.58	3.92	5.48	5.89	5.99	4.73	3.52
VAR models								
VAR(1)	0.79	1.00	0.81	0.89	0.96	1.02	0.97	0.95
VAR(4)	1.14	1.00	0.81	0.89	0.96	1.19	0.97	0.95
VAR(AIC)	**0.79**	0.82	0.81	0.89	0.96	1.02	0.97	0.95
VAR(BIC)	1.08	1.19	1.19	1.27	1.35	1.30	1.15	1.91
VAR(HIC)	1.08	1.07	1.19	1.27	1.35	1.30	1.15	1.91
VAR models with exogenous factors								
Forecast combinations								
Median	**0.76**	0.79	0.79	0.88	0.94	1.00	0.98	0.93
Mean	**0.72**	0.78	0.78	0.88	0.94	0.99	0.95	0.89
Discounted MSFE (0.90)	**0.75**	0.79	0.78	0.90	0.93	0.98	1.03	0.88
Squared discounted MSFE (0.90)	**0.77**	0.78	0.78	0.91	0.93	0.97	1.07	0.88

TABLE 4 (continued)

Forecast horizon (quarters)	1	2	3	4	5	6	7	8
5. Financial and business activities								
Random walk (log-levels) benchmark, RMSFE	1.14	2.17	3.29	4.56	4.91	5.06	4.33	3.27
VAR models								
VAR(1)	**0.64**	**0.66**	0.73	0.80	0.91	0.98	**0.89**	0.80
VAR(4)	**0.73**	**0.66**	0.73	0.80	0.91	1.11	**0.89**	0.80
VAR(AIC)	**0.64**	**0.63**	0.73	0.80	0.91	0.98	**0.89**	0.80
VAR(BIC)	**0.76**	0.74	0.92	1.00	1.21	1.32	1.38	1.86
VAR(HIC)	**0.76**	0.71	0.91	1.00	1.21	1.32	1.38	1.86
VAR models with exogenous factors								
Forecast combinations								
Median	**0.63**	**0.63**	0.72	0.80	0.91	0.97	**0.89**	0.80
Mean	**0.62**	**0.62**	0.71	0.79	0.90	0.97	**0.89**	**0.81**
Discounted MSFE (0.90)	**0.63**	**0.62**	0.72	0.79	0.89	0.96	**0.89**	0.79
Squared discounted MSFE (0.90)	**0.63**	**0.63**	0.72	0.79	0.89	0.96	**0.90**	0.78
6. Other services								
Random walk (log-levels) benchmark, RMSFE	1.04	2.08	3.19	4.26	4.46	4.71	4.02	2.05
VAR models								
VAR(1)	**0.68**	0.64	0.68	0.65	0.66	0.69	0.77	0.73
VAR(4)	**0.72**	0.64	0.68	0.65	0.66	0.61	0.77	0.73
VAR(AIC)	0.64	0.63	0.73	0.80	0.91	0.98	0.89	0.80
VAR(BIC)	0.76	0.74	0.92	1.00	1.21	1.32	1.38	1.86
VAR(HIC)	**0.78**	0.93	0.93	0.86	0.71	0.64	0.69	1.30
VAR models with exogenous factors								
Forecast combinations								
Median	**0.67**	**0.62**	0.68	0.65	0.66	0.69	0.76	0.73
Mean	**0.67**	**0.61**	0.67	0.65	0.65	0.68	0.75	0.73
Discounted MSFE (0.90)	**0.68**	**0.61**	0.67	0.64	0.65	0.68	0.73	0.71
Squared discounted MSFE (0.90)	**0.69**	**0.61**	0.66	0.64	0.65	0.68	0.68	0.73
7. Import duties plus VAT								
Random walk (log-levels) benchmark, RMSFE	1.79	2.53	3.81	5.02	5.29	5.27	4.07	2.33
VAR models								
VAR(1)	**0.78**	**0.62**	0.69	0.80	0.89	0.95	0.97	1.08
VAR(4)	**0.70**	**0.62**	0.69	0.80	0.89	1.03	0.97	1.08
VAR(AIC)	**0.68**	**0.62**	0.68	0.65	0.66	0.69	0.77	0.73
VAR(BIC)	**0.78**	0.94	0.94	0.86	0.71	0.64	0.69	1.30
VAR(HIC)	**0.76**	0.79	0.87	0.99	1.19	1.32	1.76	3.91
VAR models with exogenous factors								
Forecast combinations								
Median	**0.77**	**0.78**	0.68	0.79	0.87	0.93	0.96	1.06
Mean	**0.77**	**0.77**	0.67	0.78	0.85	0.92	0.93	0.98
Discounted MSFE (0.90)	**0.76**	**0.76**	0.68	0.79	0.86	0.92	0.94	0.96
Squared discounted MSFE (0.90)	**0.76**	**0.76**	0.68	0.79	0.86	0.92	0.95	0.94

Notes: Entries in bold denote statistical significance at 10% level of the modified Diebold-Mariano test of equal forecast accuracy (Diebold and Mariano 1995; Harvey et al. 1997). The test compares the benchmark model (random walk) forecasts to VAR and combination forecasts.

VAR(AIC), VAR(BIC) and VAR(HIC) denote the vector autoregressive models with lag lengths selected using the Akaike, Bayesian and Hannan – Quinn information criteria, respectively; VAR(1) and VAR(4) are the vector autoregressive models of order one and four, respectively.

For the discounted and squared discounted MSFE forecast combination method the discount factor is given in parentheses.

The results of the forecasting performance of single equation methods are given in Table 1A (Appendix). We focus on comparing the accuracy of forecast combinations only, as they typically outperform autoregressive forecasts in both VAR and single equation settings. For sectors in which there is a larger number of relevant predictors, such as industry, construction and trade, combination forecasts from single equations are associated with higher precision than those from VAR models in which exogenous factors summarise the information in the predictors. The growth forecasts for financial and business activities, and other services are computed as bottom-up forecasts, using combinations of single equation model forecasts for four and two sub-components, respectively. For these two broad services sectors the performance of single equation and VAR forecasts is similar.

Overall, the differences in the precision of component forecasts from single equations and VAR models are subtle as can be gathered from the evolution of forecasted values in Figure 1A. However, for large components such as trade, financial and business activities, other services, and to a lesser extent duties, the contraction in activity during the period 2012 – 2013 was better predicted by VAR models.

4.2 GDP growth forecasts

We compute GDP growth forecasts by aggregating the component forecasts whose performance was discussed in section 4.1. The resulting GDP growth forecasts are known as bottom-up or indirect forecasts, i.e. they are constructed by adding up the forecasts for all the production-side components of GDP. We compute a bottom-up GDP growth forecast by aggregating the component forecasts obtained via simple VAR models or combinations of forecasts from VARs with exogenous factors, i.e.

$$\hat{Z}_{t+h}^{(GDP,k)} = \sum_{s=1}^{S} \hat{Z}_{t+h}^{(s,k)}$$

where $\hat{Z}_{t+h}^{(s,k)}$ is the forecasted level of component s implied by the corresponding growth forecast for period $t+h$ constructed with data up to period t. The superscript k denotes the VAR model or forecast combination that gave the component forecast. $\hat{Z}_{t+h}^{(GDP,k)}$ is the resulting forecast for the level of GDP, which is transformed into growth rate prior to evaluating the forecasting performance. The component forecasts obtained from v distinct VARs with exogenous factors can be aggregated into v different forecasts for GDP, i.e. $\hat{Z}_{t+h}^{(GDP,l)}$ for $l = 1,2,\ldots,v$; subsequently, the resulting v bottom-up growth forecasts can be combined into a single forecast by applying combination methods.

Table 5 compares the performance of bottom-up GDP growth forecasts obtained from VARs with the following types of GDP growth forecasts:

(a) bottom-up GDP growth forecasts based on combinations of component growth forecasts computed from single equation models for each one of the 11 production-side components;

(b) direct GDP growth forecasts based on combinations of forecasts computed from single equation models for GDP growth.[4]

The benchmark for comparisons is the random walk model for GDP (log-level). The entries in bold indicate a superior performance vis-à-vis the naïve benchmark, with the difference in forecast accuracy being statistically significant.

[4] The single equation dynamic models for the growth rate of GDP consist of autoregressive distributed lag (ADL) models and factor augmented ADL (FADL) models as described in footnote 3.

TABLE 5

Relative RMSFE vis-à-vis the random walk, GDP growth

Forecast horizon (quarters)	1	2	3	4	5	6	7	8
I. DIRECT FORECASTS								
Single equation models for GDP growth								
I.1. Random walk benchmark, RMSFE	1.29	2.37	3.66	5.08	5.33	5.28	3.94	2.28
I.2. Combinations of GDP growth forecasts								
Median	**0.61**	**0.56**	**0.60**	**0.67**	0.74	0.79	0.73	0.59
Mean	**0.59**	**0.55**	**0.59**	**0.65**	0.71	0.76	0.70	0.59
Discounted MSFE (0.90)	**0.59**	**0.55**	**0.59**	**0.64**	0.70	0.74	0.69	0.60
Squared discounted MSFE (0.90)	**0.58**	**0.55**	**0.59**	**0.63**	0.67	0.71	0.66	0.63
II. BOTTOM-UP FORECASTS								
II.1. VAR models for seven components								
Random walk	1.06	1.03	1.02	1.02	1.03	1.03	1.04	1.08
VAR(1)	**0.79**	**0.68**	0.76	0.81	0.89	0.94	0.92	1.05
VAR(4)	**0.80**	**0.68**	0.76	0.81	0.89	1.03	0.92	1.05
VAR(AIC)	**0.79**	**0.70**	0.76	0.81	0.89	0.94	0.92	1.05
VAR(BIC)	**0.85**	0.83	0.92	0.97	1.11	1.21	1.51	3.28
VAR(HIC)	**0.85**	0.80	0.92	0.97	1.11	1.21	1.51	3.28
II.2. VAR models with exogenous factors								
II.2.1. Combinations of component growth forecasts								
Median	**0.78**	**0.70**	0.75	0.80	0.87	0.93	0.91	1.03
Mean	**0.75**	**0.69**	0.74	0.79	0.86	0.91	0.89	0.97
Discounted MSFE (0.90)	**0.77**	**0.69**	0.74	0.80	0.85	0.91	0.90	0.93
Squared discounted MSFE (0.90)	**0.78**	**0.69**	0.74	0.80	0.85	0.90	0.90	0.90
II.2.2. Combinations of bottom-up GDP growth forecasts								
Median	**0.79**	**0.70**	0.75	0.80	0.88	0.92	0.91	1.03
Mean	**0.75**	**0.69**	0.74	0.79	0.86	0.91	0.89	0.97
Discounted MSFE (0.90)	**0.71**	**0.68**	0.74	0.79	0.86	0.91	0.89	0.95
Squared discounted MSFE (0.90)	**0.69**	**0.66**	0.74	0.80	0.86	0.92	0.90	0.94
II.3. Single equation models for 11 components								
Combinations of component growth forecasts								
Median	**0.70**	**0.66**	**0.71**	**0.79**	0.85	0.88	0.86	0.74
Mean	**0.69**	**0.65**	**0.70**	**0.77**	0.83	0.86	0.84	0.72
Discounted MSFE (0.90)	**0.68**	**0.64**	**0.69**	**0.76**	0.81	0.85	0.82	0.71
Squared discounted MSFE (0.90)	**0.68**	**0.63**	**0.68**	**0.74**	0.80	0.83	0.81	0.71

Notes: Entries in bold denote statistical significance at 10% level of the modified Diebold-Mariano test of equal forecast accuracy (Diebold and Mariano 1995; Harvey et al. 1997). The test compares the benchmark model (random walk) forecasts to the different bottom-up and direct GDP growth forecasts shown in the table.

VAR(AIC), VAR(BIC) and VAR(HIC) denote the vector autoregressive models with lag lengths selected using the Akaike, Bayesian and Hannan – Quinn information criteria, respectively; VAR(1) and VAR(4) are the vector autoregressive models of order one and four, respectively.

For the discounted and squared discounted MSFE forecast combination method the discount factor is given in parentheses.

A number of different forecasts for the aggregate can also be computed using forecasts for the eleven components obtained from different single equation models; subsequently the bottom-up GDP growth forecasts are combined into a single forecast using forecast combinations. The results are almost indistinguishable from those obtained using the forecast combinations of the components (shown in section II.3 of the table) and are therefore omitted.

Bottom-up GDP growth forecasts computed from VAR forecasts for the seven production-side components lead to significant gains over the benchmark only for very short horizons. The improvements over the benchmark are larger, albeit not always significant, when information from exogenous factors is included in the VARs. Nevertheless, single equation methods for forecasting GDP growth either directly or indirectly via its 11 production-side disaggregates outperform the bottom-up VAR-based forecasts throughout. Moreover, combinations of direct GDP growth forecasts exhibit the highest accuracy with gains vis-a-vis the benchmark ranging between 21% and 45% over one to eight quarters. Improvements associated with bottom-up forecasts from single equations and VARs vary between 17% and 37%, and between 6% and 34%, respectively, over the forecast horizon, when the best-performing squared discounted combination method is applied.[5]

Figure 1 plots the cumulative RMSFE of direct and bottom-up GDP growth forecasts for the squared discounted MSFE method which is associated with superior performance. In the case of VARs, bottom-up GDP growth forecasts are computed using the component forecasts from the different VARs with exogenous factors; subsequently the bottom-up GDP growth forecasts are combined into a single forecast using the abovementioned combination method. Throughout the forecasting period, direct GDP growth forecasts are more accurate than bottom-up forecasts. In the pre-crisis period the performance of single equation bottom-up forecasts was superior to that of VAR-based predictions, but during the 2012 – 2014 recession the error of VAR forecasts declined below that of single equation predictions for one-quarter ahead horizon. Subsequently, the errors of one-quarter ahead bottom-up forecasts from the two methods become indistinguishable. For four-quarter ahead forecasts, VAR bottom-up predictions are associated with the highest error over the entire period, despite the fast reduction in their error during and after the crisis.

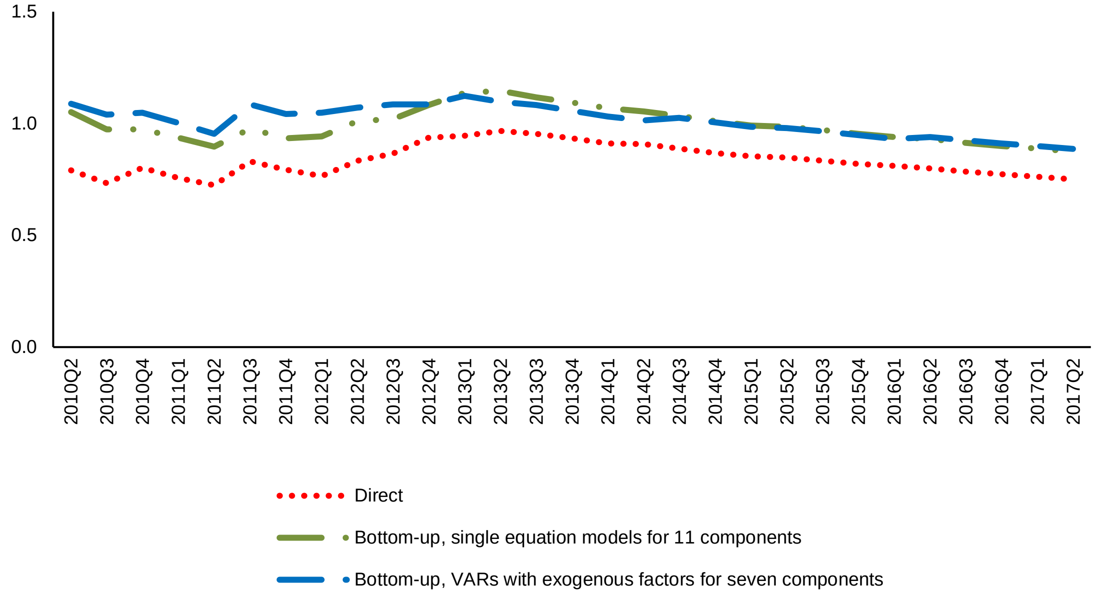
Figure 2 shows the evolution of one- and four-quarter ahead GDP growth forecasts obtained from the different methods with superior performance presented in Figure 1. VAR-based forecasts are the most volatile due to higher estimation uncertainty, but it seems that they captured the depth of the recession in 2013 and the moderation of the recession in 2014 more accurately than the predictions produced by the other two methods.

[5] The predictive accuracy of direct GDP growth forecasts obtained through the squared discounted MSFE (0.90) combination method is statistically superior to the bottom-up forecasts computed using single equation or VAR component forecasts for horizons of up to four quarters ahead.

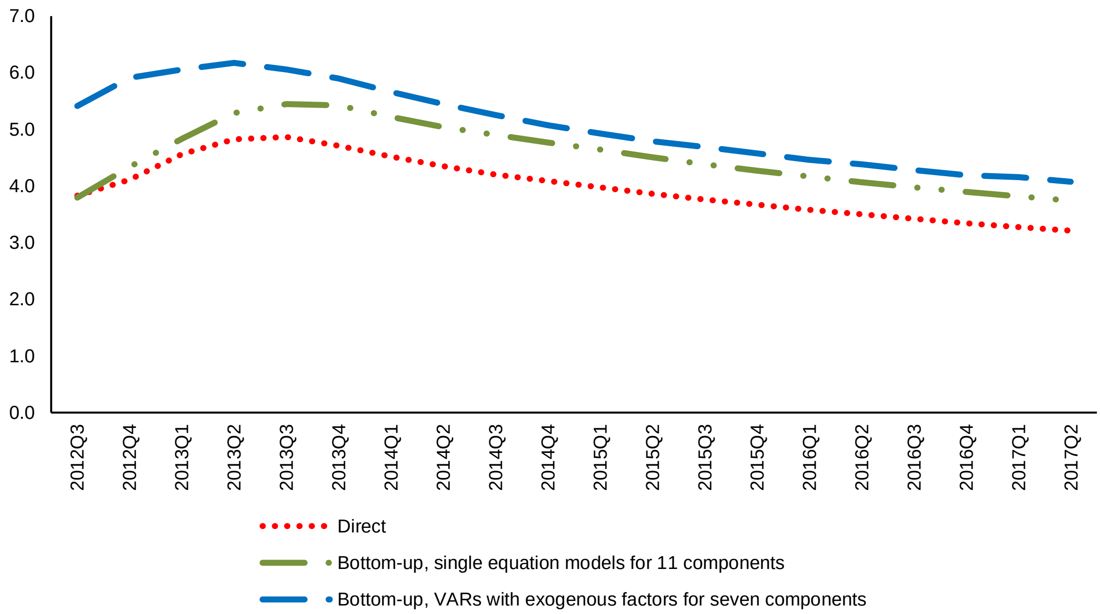
FIGURE 1

Cumulative RMSFE, GDP growth forecasts

A. One-quarter ahead forecasts



B. Four-quarter ahead forecasts



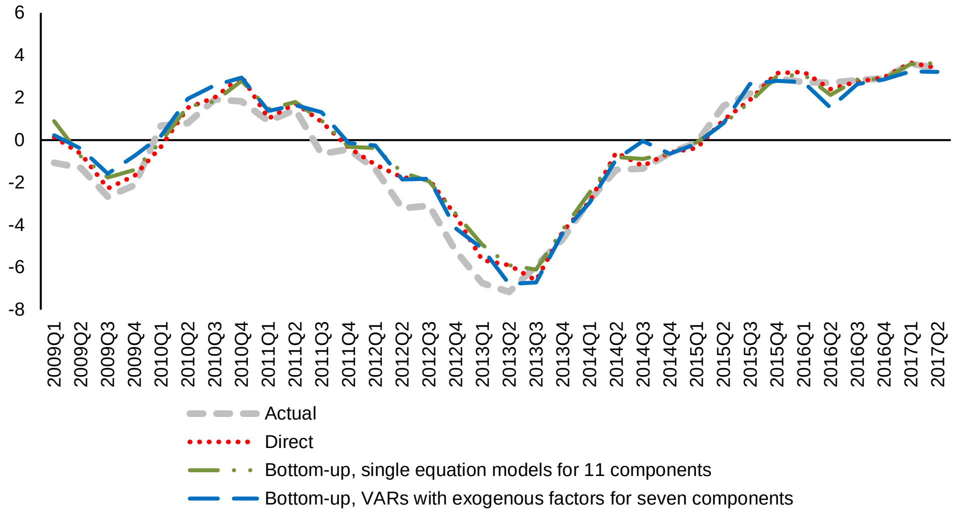
Notes: In the case of single equations, the bottom-up GDP growth forecast is computed using the squared discounted MSFE combination (with discount factor equal to 0.9) of forecasts for the growth rate of each one of the 11 components obtained from single equation models for each component (section II.3 of Table 5). In the case of VARs, bottom-up GDP growth forecasts are computed using the component forecasts from the different VARs; subsequently the bottom-up GDP growth forecasts are combined into a single forecast using the squared discounted MSFE combination with discount factor equal to 0.9 (section II.2.2 of Table 5).

The dates on the horizontal axis are adjusted so that they correspond to the reference dates of the actual GDP growth rates.

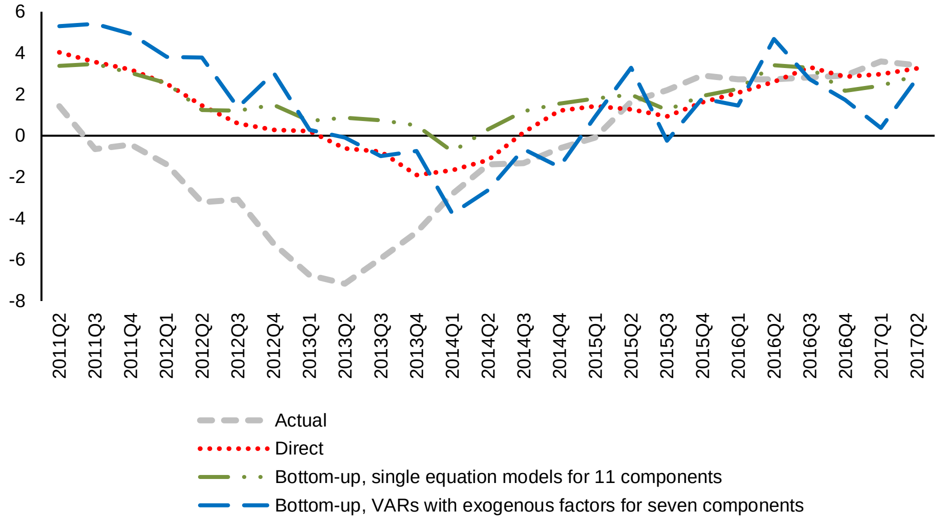
FIGURE 2

GDP growth forecasts (y-o-y percentage change)

A. One-quarter ahead forecasts



B. Four-quarter ahead forecasts



Note: The dates on the horizontal axis are adjusted so that they correspond to the reference dates of the actual GDP growth rates.

5. Summary and conclusions

This paper explores the use of VAR models for the construction of growth forecasts in the case of production-side GDP components and GDP via aggregation of its component forecasts. This paper extends previous work on methods for estimating sectoral forecasts by employing single equation dynamic models and common factors.

The results show that, in general, component forecasts obtained from VARs with exogenous factors outperform simple VAR forecasts. Forecast gains, however, are mainly attained for short horizons. The exceptions are the sector of construction in which there are minimal gains for longer horizons, and the sector of financial and business activities in which significant gains are attained for both short and long horizons. Component forecasts computed from single equations are at least as accurate as VAR forecasts. However, the contraction in activity during the period 2012 – 2013 was better predicted by VAR models in the case of large components such as trade, financial and business activities, and other services. Similarly, bottom-up GDP growth forecasts from single equation models for components are at least as precise as indirect forecasts using VARs to compute the component forecasts. Nevertheless, direct GDP growth forecasts are the best performers over the period considered.

The findings of the paper resonate with the considerations in Lutkepohl (2010), namely that in practice, estimation uncertainty in a multivariate model could be more dominant than in a single equation model for the aggregate, therefore impacting negatively on the forecast precision of the bottom-up aggregate forecast vis-à-vis its direct counterpart, despite the theoretical optimality of the former. In the empirical analysis in this paper, estimation uncertainty could play a key role in the superiority of direct forecasts as the available time series are not particularly long. Also, the use of factors as exogenous variables in the VARs appears to be a useful extension of simple VARs as it incorporates valuable information in the models without increasing the parameter space considerably. Bayesian VARs constitute an alternative parameter space reduction technique which is under investigation for sectoral and GDP growth forecasting.

The evaluation of the forecasting performance of the different methods considered in this paper should be repeated on a systematic basis as the size of the available time series influences the reliability of the results, especially for longer forecast horizons. Nonetheless, forecasting GDP growth indirectly through VAR-type systems provides an integrated framework for analysing the role of sectors in economic cycles and assessing the impact of shocks on sectors and the economy as a whole.

Appendix

TABLE A1

Relative RMSFE vis-à-vis the random walk, single equation forecasts for production-side components

Forecast horizon (quarters)	1	2	3	4	5	6	7	8
1. Agriculture, forestry, fishing								
Random walk (log-levels) benchmark, RMSFE	9.06	9.33	9.60	11.23	11.75	12.19	12.69	13.91
Single equation models for the growth rate of component								
Forecast combinations								
Median	1.02	0.99	1.00	1.00	1.00	1.00	0.99	1.00
Mean	1.02	0.99	1.01	1.02	1.01	1.01	1.00	1.00
Discounted MSFE (0.90)	1.02	0.99	1.01	1.02	1.01	1.01	1.00	1.00
Squared discounted MSFE (0.90)	1.02	0.99	1.01	1.02	1.01	1.01	1.00	1.01
2. Industry								
Random walk (log-levels) benchmark, RMSFE	2.57	4.47	6.68	8.94	7.42	5.94	4.92	5.09
Single equation models for the growth rate of component								
Forecast combinations								
Median	**0.83**	**0.82**	**0.86**	**0.90**	**0.86**	**0.86**	0.81	0.85
Mean	**0.82**	**0.81**	**0.84**	**0.88**	**0.82**	**0.80**	0.76	0.81
Discounted MSFE (0.90)	**0.82**	**0.80**	**0.83**	**0.87**	**0.82**	**0.80**	0.75	0.81
Squared discounted MSFE (0.90)	**0.82**	**0.79**	**0.82**	**0.85**	**0.81**	**0.79**	**0.75**	0.81
3. Construction								
Random walk (log-levels) benchmark, RMSFE	6.60	10.23	14.23	19.07	19.90	19.46	15.67	14.79
Single equation models for the growth rate of component								
Forecast combinations								
Median	**0.88**	**0.79**	0.80	0.79	0.77	0.78	0.72	0.90
Mean	**0.88**	**0.80**	0.78	0.77	**0.75**	0.76	0.72	0.90
Discounted MSFE (0.90)	**0.88**	**0.80**	0.79	0.77	**0.75**	0.77	0.73	0.91
Squared discounted MSFE (0.90)	**0.88**	**0.80**	0.79	0.78	**0.75**	0.77	0.75	0.94
4. Trade, transport, accommodation and food services								
Random walk (log-levels) benchmark, RMSFE	1.50	2.58	3.92	5.48	5.89	5.99	4.73	3.52
Single equation models for the growth rate of component								
Forecast combinations								
Median	**0.69**	**0.75**	**0.81**	**0.89**	0.95	0.95	1.00	1.03
Mean	**0.66**	**0.73**	**0.79**	**0.87**	0.92	0.93	0.99	1.03
Discounted MSFE (0.90)	**0.66**	**0.73**	**0.79**	**0.86**	0.91	0.92	0.98	1.02
Squared discounted MSFE (0.90)	**0.66**	**0.73**	**0.77**	**0.84**	0.91	0.92	0.97	1.02
5. Financial and business activities								
Random walk (log-levels) benchmark, RMSFE	1.14	2.17	3.29	4.56	4.91	5.06	4.33	3.27
Single equation models for the growth rate of component								
Forecast combinations								
Median	**0.60**	**0.67**	**0.72**	**0.76**	0.83	0.90	0.91	0.80
Mean	**0.58**	**0.62**	**0.72**	**0.75**	0.82	0.89	0.89	0.80
Discounted MSFE (0.90)	**0.57**	**0.62**	**0.71**	**0.75**	0.81	0.88	0.88	0.79
Squared discounted MSFE (0.90)	**0.56**	**0.61**	**0.71**	**0.74**	0.80	0.86	0.87	0.77
6. Other services								
Random walk (log-levels) benchmark, RMSFE	1.04	2.08	3.19	4.26	4.46	4.71	4.02	2.05
Single equation models for the growth rate of component								
Forecast combinations								
Median	**0.66**	**0.67**	**0.75**	0.85	0.92	0.96	0.94	0.84
Mean	**0.65**	**0.66**	**0.74**	0.84	0.90	0.94	0.91	0.78
Discounted MSFE (0.90)	**0.65**	**0.65**	**0.73**	0.81	0.88	0.92	0.89	0.75
Squared discounted MSFE (0.90)	**0.64**	**0.65**	**0.72**	0.78	0.85	0.90	0.87	0.70
7. Import duties plus VAT								
Random walk (log-levels) benchmark, RMSFE	1.79	2.53	3.81	5.02	5.29	5.27	4.07	2.33
Single equation models for the growth rate of component								
Forecast combinations								
Median	**0.74**	**0.70**	**0.68**	**0.73**	0.83	0.87	0.94	0.89
Mean	**0.72**	**0.68**	**0.67**	**0.71**	0.81	0.84	0.91	0.88
Discounted MSFE (0.90)	**0.71**	**0.67**	**0.66**	**0.69**	0.80	0.81	0.89	0.89
Squared discounted MSFE (0.90)	**0.71**	**0.65**	**0.66**	**0.67**	0.78	0.76	0.83	0.88

Notes: Entries in bold denote statistical significance at 10% level of the modified Diebold-Mariano test of equal forecast accuracy (Diebold and Mariano 1995; Harvey et al. 1997). The test compares the benchmark model (random walk) forecasts to combination forecasts.

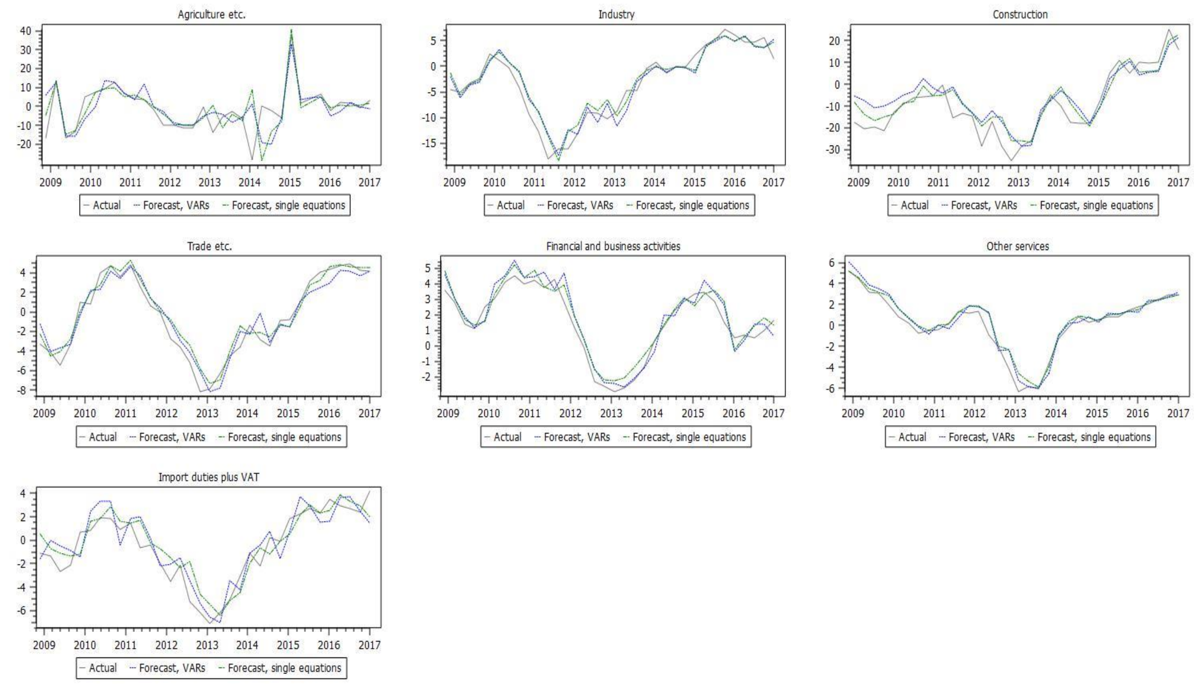
For the discounted and squared discounted MSFE forecast combination method the discount factor is given in parentheses.

Forecasts for the sectors of financial and business activities, and other services are computed as bottom-up forecasts using combinations of single equation model forecasts for four and two sub-components, respectively.

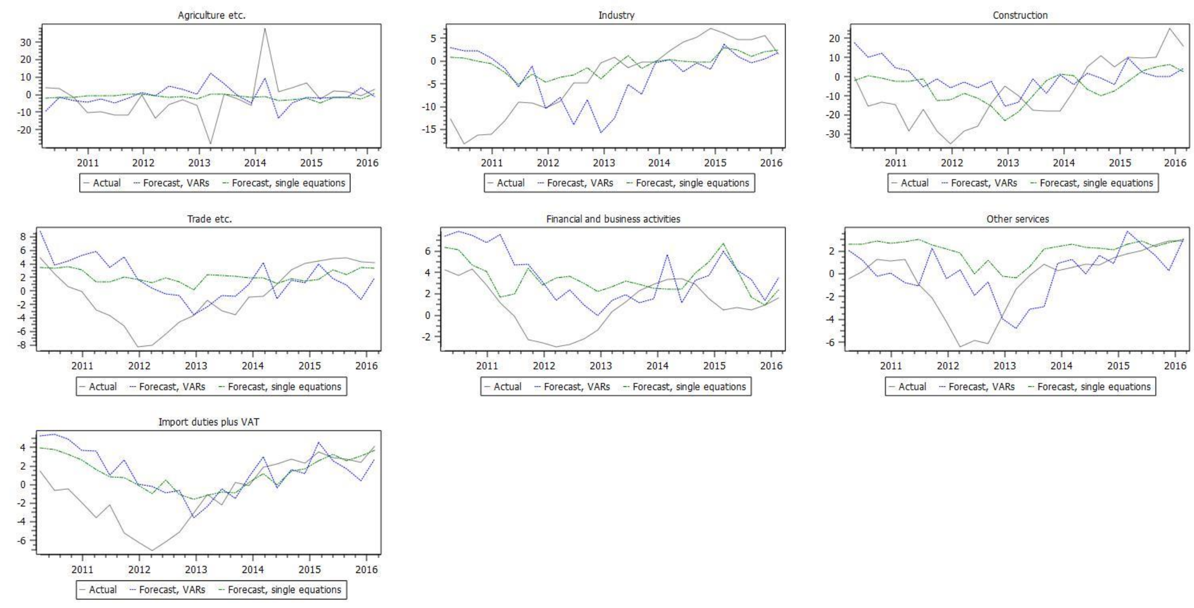
FIGURE A1

Forecasts for the growth rate of production-side components (y-o-y percentage change)

A. One-quarter ahead forecasts



B. Four-quarter ahead forecasts



Note: The dates on the horizontal axis are adjusted so that they correspond to the reference dates of the actual GDP growth rates.

References

Artis, M. J., Banerjee A. and Marcellino M. (2005), 'Factor forecasts for the UK', Journal of Forecasting 24, 279-298.

Atalay, E. (2017), 'How important are sectoral shocks?', American Economic Journal: Macroeconomics 9, 254–280.

Bai, J., Ng S. (2002), 'Determining the number of factors in the approximate factor models', Econometrica 70, 191-221.

Barhoumi, K., Darne O., Ferrara L., and Pluyaud B. (2012), 'Monthly GDP forecasting using bridge models: Application for the French economy', Bulletin of Economic Research 64:S1, 3307-3378.

Bermingham, C. and D'Agostino A. (2014), 'Understanding and forecasting aggregate and disaggregate price dynamics', Empirical Economics 46, 765–788.

Bruneau, C., De Bandt O., Flageollet A. and Michaux E. (2007), 'Forecasting inflation using economic indicators: The case of France', Journal of Forecasting 26, 1-22.

Dias, F., Pinheiro M. and Rua A. (2017), 'A bottom-up approach for forecasting GDP in a data-rich environment', Applied Economics Letters, DOI: 10.1080/13504851.2017.1361000.

Diebold, F. X. and Mariano R. S. (1995), 'Comparing predictive accuracy', Journal of Business and Economic Statistics 13, 134-144.

Duarte, C. and Rua A. (2014), 'Forecasting inflation through a bottom-up approach: How bottom is bottom?', Economic Modelling 24, 941–953.

Drechsel, K. and Scheufele R. (2013), 'Bottom-up or direct? Forecasting German GDP in a data-rich environment', IWH-Diskussionspapiere, No. 2013, 7.

Foroni, C. and Marcellino M. (2014), 'A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates', International Journal of Forecasting 30, 554–568.

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

Foerster, A. T., Sarte P. G. and Watson M. W. (2011), 'Sectoral versus aggregate shocks: A structural factor analysis of industrial production', Journal of Political Economy 119, 1-38.

Forni, M. and Reichlin L. (1998), 'Let's get real: A dynamic factor analytical approach to disaggregated business cycle', Review of Economic Studies 65, 453–474.

Hahn, E. and Skudelny F. (2008), 'Early estimates of euro area real GDP growth: A bottom up approach from the production side', Working Paper Series No. 975, European Central Bank.

Harvey, D., Leybourne S. and Newbold P. (1997), 'Testing the equality of prediction mean squared errors', International Journal of Forecasting 13, 281-291.

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

Hubrich, K. (2005), 'Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?', International Journal of Forecasting 21, 119-136.

Moser, G., Rumler, F. and Scharler J. (2007), 'Forecasting Austrian inflation', Economic Modelling 24, 470–480.

Lee, K. C., Pesaran H. M. and Pierse R. G. (1992), 'Persistence of shocks and their sources in a multisectoral model of UK output growth', Economic Journal 102, 342-356.

Long, J. B. and Plosser C. I. (1983), 'Real business cycles', Journal of Political Economy 91, 39-69.

Long, J. B. and Plosser C. I. (1987), 'Sectoral vs. aggregate shocks in the business cycle', American Economic Review 77, 333-336.

Lütkepohl, H. (2010), 'Forecasting aggregated time series variables: A survey', OECD Journal of Business Cycle Measurement and Analysis 2, 37-62.

Papamichael, C. and Pashourtidou N. (2014), 'Forecasting Cyprus GDP and its demand components: Applications of dynamic factor models', Economic Analysis Paper No. 16-14, Economics Research Centre, University of Cyprus.

Pashourtidou, N., Papamichael C. and Karagiannakis C. (2017), 'Forecasting economic activity in sectors of the Cypriot economy', Economic Analysis Paper No. 17-03, Economics Research Centre, University of Cyprus.

Pesaran, M. H., Pierse R. G and Lee K. C. (1993), 'Persistence, cointegration, and aggregation: A disaggregated analysis of output fluctuations in the U. S. economy' Journal of Econometrics 56, 57-88.

Stock, J., and Watson M. (2002a), 'Macroeconomic forecasting using diffusion indexes', Journal of Business and Economic Statistics 20, 147-162.

Stock, J., and Watson M. (2002b), 'Forecasting using principal components from a large number of predictors', Journal of the American Statistical Association 97, 1167-1179.

Stock, J. and Watson M. (2004), 'Combination forecasts of output growth and in a seven-country dataset', Journal of Forecasting 23, 405-430.

Stock, J. H., and Watson M. W., (2008), 'Phillips curve inflation forecasts', NBER Working Paper 14322. Available at: http://www.nber.org/papers/w14322.

RECENT ECONOMIC POLICY/ANALYSIS PAPERS

10-17 Giannakis, E. and Th. P. Mamuneas, “Sectoral linkages and economic crisis: An input-output analysis of the Cypriot economy”, December 2017.

09-17 Koutsampelas, C., “Non-standard employment in Cyprus: Trends and policy responses”, December 2017.

08-17 Andreou, E. and A. Kourtellos, “Scoring rules for simple forecasting models: The case of Cyprus GDP and its sectors”, December 2017.

07-17 Eliophotou, M., S. N. Andreou and E. Markadjis, “The expectations of prospective students regarding the economic returns to higher education: evidence from Cyprus”, October 2017.

06-17 Zachariadis, T., A. Michopoulos, Y. Vougiouklakis, B. Struss, K. Piripitsi and C. Ellinopoulos, “In-depth Assessment of the Energy Efficient Potential in Cyprus”, October 2017.

05-17 Koutsampelas, C. and P. Pashardes, “Social protection in Cyprus: Overview and challenges”, September 2017.

04-17 Mamuneas, T. and E. Ketteni, “Rates of return to public infrastructure”, August 2017.

03-17 Pashourtidou, N., C. Papamichael and C. Karagiannakis, “Forecasting economic activity in sectors of the Cypriot economy”, May 2017.

02-17 Ketteni, E., T. Mamuneas and P. Pashardes, “Factors affecting the productivity of European economies”, May 2017.

01-17 Savva, S. C. and N. A. Michail, “Modelling House price volatility states in Cyprus with switching ARCH models”, January 2017.