Gabinete de Estratégia e Estudos
Ministério da Economia e da Inovação

GEE Papers
Número 2
Fevereiro de 2007

Nowcasting an Economic Aggregate with Disaggregate Dynamic Factors: An Application to Portuguese GDP

António José Morgado
Luís Catela Nunes
Susana Salvado

Av. da República nº. 79 1050 – 243 Lisboa
Telf: (351) 217998150
Fax: (351) 217998151
Web Site: www.gee.min-economia.pt

ISSN 1647-6212

Nowcasting an Economic Aggregate with Disaggregate Dynamic Factors: An Application to Portuguese GDP†

António José Morgado
Faculdade de Economia, Universidade Nova de Lisboa
Gabinete de Estudos e Estratégia, Ministério da Economia e da Inovação

Luis C. Nunes
Faculdade de Economia, Universidade Nova de Lisboa

Susana Salvado
Faculdade de Economia, Universidade Nova de Lisboa
Gabinete de Estudos e Estratégia, Ministério da Economia e da Inovação

JEL Classification: C53, C82

Keywords:
Forecasting; Dynamic Factor Model; Temporal Disaggregation

Abstract

This paper consists of an empirical study comparing a dynamic factor model approach to estimate the current quarter aggregate GDP with the alternative approach of aggregating the forecasts obtained from specific dynamic factor models for each major expenditure disaggregate. The out-of-sample forecasting performance results suggest that there is no advantage in aggregating the disaggregate forecasts.

.

[0]† Corresponding address: Faculdade de Economia, Universidade Nova de Lisboa, Campus de Campolide, P-1099-032 Lisboa, Portugal. Phone: +351 213801600. Fax: +351 213870933. Email: lcnunes@fe.unl.pt. We are grateful for the support given to this study by Gabinete de Estratégia e Estudos of the Portuguese Ministério da Economia e Inovação. We thank Maximo Camacho for providing us with the Gauss programs used in the paper Camacho and Sancho (2003) on which we based some of our computations. Thanks are also due to Miguel Lebre de Freitas for helpful comments and support. The views expressed in this paper are those of the authors and not necessarily of the institutions with which they are affiliated.

.

Contents

1 Introduction

Reliable estimates of the current state of the economy in the form of real GDP growth rates are essential for policy purposes. However, typically, the first release of quarterly GDP figures are only published several weeks after the end of each quarter. In the EU the delay may range from as few as 3 weeks in the UK to more than 6 weeks in many member states. But since several short-term economic indicators are always available a few days after the end of each quarter and well before the release of the first GDP figures, a preliminary estimate of quarterly GDP growth could be computed from the movements in these indicators. Since any given indicator usually captures only particular aspects of economic activity, it is important to extract the relevant information from a broad set of available indicators. Factor models are a well-known approach to summarize the information coming from a large number of variables into a small number of factors that are linear combinations of all the variables in the dataset. Stock and Watson (2002) propose using a generalized dynamic factor model where the movements in each variable are explained by two sets of mutually orthogonal components, the common factors and the idiosyncratic components. GDP forecasts are then obtained by using a forecasting equation including as regressors the common factors and lagged values of GDP growth.

The theoretical and empirical properties of several variants of the generalized dynamic factor model have been the subject of a number of recent studies. For example, Boivin and Ng (2006) study the importance of the size and composition of the dataset used, while Kapetanios and Marcellino (2003) compare the performance of alternative estimation methods. In this paper we explore how dynamic factor models can be used to forecast GDP by aggregating forecasts of its expenditure disaggregates. There have been a number of recent empirical studies studying the advantages of aggregating forecasts of disaggregates. The majority of these have considered the particular case of inflation forecasting (see Espasa *et al* 2002; Hubrich, 2005; Duarte and Rua, 2005). An application to GDP appears in Fair and Shiller (1990). There are also theoretical arguments in favour of combining the forecasts for each disaggregate over forecasting the aggregate itself when the data generation process is known (see Lutkepohl, 2005). However, when the data generating process is unknown and has to be estimated, such result may no longer hold as it will depend on the true data generating process and the available estimation sample (see Hendry and Hubrich, 2006).

In this paper, we present an empirical study where we compare the out-of-sample forecasting performance of dynamic factor models used to forecast aggregate GDP directly with the alternative approach of aggregating the forecasts obtained from specific dynamic factor models for each major expenditure disaggregate. Our estimated dynamic factor models used to forecast aggregate GDP and each of the expenditure disaggregates are different because when estimating the common factors for each case, we have removed from the dataset those variables that exhibited a small correlation with the variable being forecasted. This choice is in line with the results of recent studies (Schneider and Speitzer, 2004; Boivin and Ng, 2006) suggesting that using smaller datasets may increase forecasting performance.

The rest of the paper is organized as follows. A description of the data is given in Section 2. Section 3 describes the general methodology used to forecast each variable. Estimation results are presented in Section 4 and a comparative analysis of the alternative forecasting methods appears in Section 5. Section

6 concludes.

2 Data Description

The variables of interest in this study are the Portuguese quarterly real GDP and its expenditure disaggregates published by the National Institute of Statistics (INE). These are published 63 days after the end of each quarter. We consider several short-term economic indicators published with a smaller delay that can be used to estimate the variables of interest in advance of their official release. The choice of which indicators to include was dictated by data availability. Moreover, as already mentioned, we further removed indicators exhibiting a small correlation with the variable being forecasted. The final set of variables used in the paper are presented in Tables 1 and 2.

Prior to use, all variables were seasonally adjusted using the TRAMO-SEATS software and year-on-year differences applied where appropriate. For a few monthly series that showed a pronounced seasonal component, we followed Camacho and Sancho (2003) and used trend-cycle, instead of seasonally adjusted transformations because the latter contained a highly erratic component that seriously disturbed the interpolation of the series. Those cases are identified with an asterisk (*) in Tables 1 and 2. The cases identified with a dagger (†) correspond to variables that, by construction, already reflect year-on-year variations and thus were not subject to a difference transformation.

It is important to notice that the variables we want to forecast are quarterly, while the majority of the short-term indicators are monthly. Moreover, as can been seen from Tables 1 and 2, the series are released with different delays and available for different historical periods. This creates two problems we need to solve: (i) the use of an unbalanced monthly dataset when extracting the common factors covering the largest possible time span, and (ii) forecasting the quarterly variables using the information from the monthly common factors. The solutions to these two problems are discussed in the following section.

3 Methodology

3.1 The Dynamic Factor Model

We follow the approximate dynamic factor model described in Stock and Watson (2002). In this model it is assumed that the observed stationary time series variables denoted by the N-dimensional vector x_t can be written as the sum of two mutually orthogonal unobservable components: the r-dimensional vector of common factors f_t and the N-dimensional idiosyncratic component e_t such that

$$x_t = \Lambda f_t + e_t \tag{1}$$

where Λ is the $(N \times r)$ factor loading matrix relating the common factors to the observed series and the N idiosyncratic components in e_t are allowed to be correlated for any given t. The advantage of this model is that the common factors can be estimated using standard principal components analysis. In our application,

Variable description	Frequency	Sample	Release delay	Code
GDP at 2000 prices	Q	1978Q02 - 2006Q02	65 days	Y
Resident Household Consumption Expenditure at 2000 prices	Q	1978Q02 - 2006Q02	65 days	C
Gross Fixed Capital Formation at 2000 prices	Q	1978Q02 - 2006Q02	65 days	I
Cement sales (quantity index)	M	1982M01 - 2006M09	22 days	Y(B), I(B)
Construction and Public Works Survey (activity appraisal) (†)	M	1991M02 - 2006M09	5 days	Y(U), I(U)
Manufacturing Survey (current production) (†)	M	1987M01 - 2006M09	5 days	Y(U)
Retail Trade Survey (sales volume) (†)	M	1989M01 - 2006M09	5 days	Y(U)
Wholesale Trade Survey (sales volume) (†)	M	1989M01 - 2006M09	5 days	Y(U)
Industrial Production Index	M	1968M01 - 2006M09	30 days	Y(B)
Total Vehicles Sales (units)	M	1991M01 - 2006M09	5 days	Y(U)
Foodstuff Retail Trade (index)	M	1991M01 - 2006M09	30 days	Y(U)
Retail Trade of Non-Foodstuff (index)	M	1995M01 - 2006M09	30 days	C(U)
Restaurant and Hotel Services (index)	M	2000M01 - 2006M09	42 days	C(U)
Retail Trade of Furniture, Lighting and other Housing Expenses (index)	M	1991M01 - 2006M09	30 days	C(U)
Imports of Goods and Services	M	1996M01 - 2006M08	50 days	M
Exports of Goods and Services	M	1996M01 - 2006M08	50 days	X
Price deflator for Imports	M	1996M01 - 2006M09	65 days	M
Price deflator for Exports	M	1996M01 - 2006M09	65 days	X
Electricity Consumption	M	1987M01 - 2006M09	5 days	C(B)
Sales Vehicles - Passenger (including 4X4) (units) (*)	M	1989M01 - 2006M09	5 days	C(B)
Sales Vehicles - Commercial under 3.5 ton (excluding 4X4) (units)	M	1989M01 - 2006M09	5 days	I(B)
Sales Vehicles - Commercial above 3.5 ton (units)	M	1989M07 - 2006M09	5 days	I(B)
Credit for Current Consumption (value)	M	1997M12 - 2006M08	40 days	C(U)
Credit for House Investment (value)	M	1997M12 - 2006M08	40 days	C(U)
Bednights (number) (*)	M	1978M01 - 2006M08	45 days	C(U)
Gasoline Consumption	M	1977M01 - 2006M06	82 days	C(U)

Table 1: Variables used (part 1)

Variable Description	frequency	sample	release delay	Code
Industrial Production of Foodstuff	M	1995M01 - 2006M06	30 days	C(U)
Brent	Q	1987Q02 - 2006Q03	0 days	X, M
Euro-Zone Industrial Production Price Index	Q	1995Q01 - 2006Q03	35 days	M
Industrial Production Index of Consumption Goods	M	1990M01 - 2006M06	30 days	C(U)
Industrial Production Index of Durables	M	1995M01 - 2006M06	30 days	C(U)
Industrial Production Index of Machinery and Equipment	M	2000M01 - 2006M06	30 days	I(U)
Industrial Production Index of the Construction Sector (*)	M	2000M01 - 2006M09	30 days	I(U)
Price Index of Industrial Production of Machinery and Equipment (*)	M	1995M01 - 2006M08	30 days	I(U)
Balance of Payments (debit) - Copyright and Patent Royalties	M	1986M01 - 2006M08	50 days	I(U)
Balance of Payments (debit) - Franchising (*)	M	1986M01 - 2006M08	50 days	I(U)
Balance of Payments (debit) - Information Services (*)	M	1986M01 - 2006M08	50 days	I(U)
Balance of Payments (debit) - Computer Services (*)	M	1986M01 - 2006M08	50 days	I(U)
Balance of Payments (debit) - Audiovisual and Related Services(*)	M	1986M01 - 2006M08	50 days	I(U)
Balance of Payments (debit) - Other Personal Cultural and Recreational Services (*)	M	1986M01 - 2006M08	50 days	I(U)
Balance of Payments (debit) - Other Rights of Use (*)	M	1986M01 - 2006M08	50 days	I(U)
Imports of Transport Vehicles	M	2000M01 - 2006M08	70 days	I(U)
Imports of Machinery and Equipment (*)	M	2000M01 - 2006M08	70 days	I(U)
Real Imports of Goods and Services Forecast	M	1996M01 - 2006M08	30 days	Y(U)
Real Exports of Goods and Services Forecast	M	1996M01 - 2006M08	30 days	Y(U)

Table 2: Variables used (part 2)

x_t includes lagged values of the variables used, hence we call f_t dynamic factors as in Stock and Watson (2002). The N variables included in x_t have been seasonally adjusted, stationarized and standardized to have mean zero and variance one. Denoting by X and F the matrices obtained by stacking the x_t and f_t, the estimated factors are given by $\widehat{F} = X\widehat{\Lambda}/N$, where $\widehat{\Lambda}$ is $\sqrt{N}$ times the eigenvectors of $X'X$ corresponding to its r largest eigenvalues.

We also follow Stock and Watson (2002) and use the EM algorithm to deal with the problem that not all the variables in our database are available for the same time span. Following Camacho and Sancho (2003), the EM algorithm can be summarized as follows: (i) an initial estimate of the factors is estimated from the balanced sample (series with no missing data points, which span the whole period of time under analysis); (ii) those initial estimated factors are used to calculate an estimate of the missing observations using equation (1); (iii) new factors are obtained from the whole data set, including the unbalanced series with the missing values replaced by the estimates obtained on stage (ii). Stages (ii) and (iii) are repeated iteratively until convergence is attained, i.e., from one iteration to the next, the estimates do not change significantly.

Finally, as in Stock and Watson (2002), the number of common factors r is chosen according to the forecasting performance of the forecasting equation.

3.2 Forecasting Quarterly Variables Using Monthly Factors

To produce forecasts of the variable of interest we follow Stock and Watson (2002) and use a linear regression of this variable on its own lagged values, a constant and on the estimated common factors. Since we are using monthly estimated factors to forecast a quarterly series, we must consider some method to deal with these different sampling frequencies. A state-space approach as in Nunes (2005) could be used but it would require reformulating the dynamic factor model. Mitchell *et al* (2005) and Mitchell and Weale (2005) propose a specification of a forecasting equation linking the low-frequency to the high-frequency data. We have chosen a simpler version of their method, similar to that used in Koenig *et al* (2000).

In this work we are particularly interested in nowcasting GDP or particular expenditure disaggregates of GDP, that is, infer about the value that a variable took in a given quarter before its official announcement but taking into account all the information contained in other variables whose figures are already available up to that quarter. It follows that in our case the forecasting equation is estimated by regressing the quarterly year-on-year growth rate of the quarterly variable on its own lagged values, a constant, and current and lagged values of the monthly common factors, i.e.:

$$\Delta y_t = \alpha_0 + \sum_{i=1}^{r}\sum_{j=0}^{k}\sum_{m=1}^{3}\beta_{jm}^{i}\hat{f}_{t-j,m}^{i} + \sum_{j=1}^{l}\gamma_j \Delta y_{t-j} + u_t \tag{2}$$

where Δy_t denotes the year-on-year growth rate of the variable in quarter t, $\hat{f}_{t,m}^{i}$ denotes the i-th estimated factor in month m of quarter t. The lag lengths k and l should be chosen large enough to ensure that the error term is white noise and uncorrelated with current and lagged values of the regressors.

3.3 Aggregate and Disaggregate GDP Nowcasting

We compare estimates of current quarterly GDP using two alternative approaches: (i) directly modelling the aggregate variable, and (ii) aggregating forecasts of the expenditure disaggregates.

The first approach of direct estimation starts by using the dynamic factor model to extract the common factors from the set of short-term indicators selected specifically for nowcasting GDP. Afterwards, equation (2) is used to forecast the year-on-year growth rates of quarterly GDP.

The alternative disaggregate approach is performed in two steps. First, forecasts of each expenditure disaggregate are obtained. For private consumption and investment we follow the direct modelling approach. Estimates of the current values of the other expenditure disaggregates are obtained in a different way as explained below. Secondly, the disaggregated forecasts are aggregated to produce a forecast of aggregate GDP.

As already mentioned, the estimated dynamic factors used to forecast GDP and each of the major expenditure aggregates are not the same. This is because the datasets used when estimating the common factors included in the forecasting equation for each variable differ from each other since in each case we exclude those indicators that exhibit a small correlation with the variable being forecasted.

In the disaggregate approach, public consumption is treated as exogenous and set equal to what is stated in the official government budget account. Real exports and imports are obtained by aggregation of the monthly values of real exports and imports. These are calculated, in turn, by deflating the nominal series using an appropriate estimated deflator, as explained below. Given that we have estimated values for the level of quarterly exports, imports and public consumption, we use the estimated year-on-year growth rates of quarterly private consumption and investment to compute the levels of the series, and then obtain the estimate of the aggregate quarterly GDP level. Finally, the year-on-year growth rates of the latter are computed.

Both alternative approaches can then be compared in terms of their out-of-sample performance when forecasting the current year-on-year growth rate of quarterly GDP.

4 Estimation Results

This section presents the estimation results for both the aggregate and disaggregate approaches of nowcasting quarterly GDP.

4.1 Aggregate GDP Nowcasting

The variables used to construct the dynamic factor model for aggregate GDP referred to in Section 3.1 appear in Table 1. As explained in that section, we use the longest series to set a balanced panel (coded Y(B)) with which we estimate the missing points of the unbalanced panel (coded Y(U)). Given that some of the variables used in the panel are expressed in terms of year-on-year monthly growth rates and that we include one lag of each series in the panel to compute the dynamic factors, we end up with estimated factor series ($\hat{f}_t$) that span from 1983M02 to 2006M09. The series of the real imports and exports used in the

Extracted Factors	Explained Variation (%)	Cumulative Explained Variation (%)
1	63.7	63.7
2	7.5	71.2
3	6.3	77.6
4	4.3	81.8
5	3.4	85.3
6	2.6	87.9

Table 3: First six estimated dynamic factors for GDP

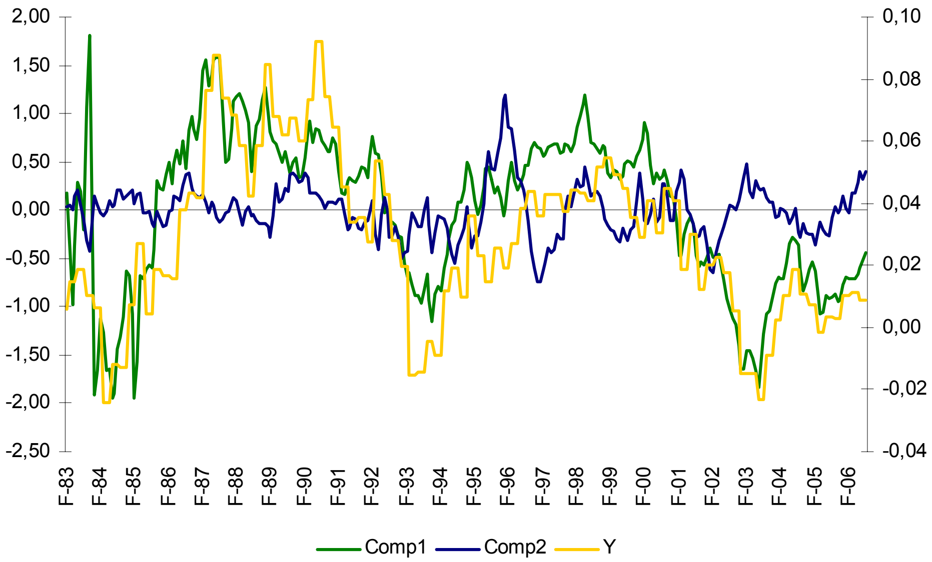

Figure 1: Quarterly real GDP year-on-year growth rate (Y) and corresponding first two monthly factors (Comp1 and Comp2).

unbalanced panel are estimated as described in Section 4.2.2, using variables coded X and M, respectively, appearing in Tables 1 and 2.

As shown in Table 3, the first extracted factor explains 64% of the variation, while the second only accounts for 8%. Figure 1 depicts the two first estimated monthly factors and the quarterly real GDP year-on-year growth rate series where it is evident the relevance of the first factor.

The estimated factors were then used as regressors in estimating the forecasting equation (2) with Δy standing for the year-on-year growth rate of quarterly real GDP. We considered a maximum of two extracted factors ($r = 2$). The results appearing in Table 4 for models including one factor (AGG-1F) and two factors (AGG-2F) suggest that only the first factor should be used. In both models AGG-1F and AGG-2F the lags of the first estimated factor are jointly significant at the 1% level while in model AGG-2F the lags of second factor are not jointly significant (with a p-value equal to 0.12). Table 4 also includes the results of an AR model (denoted as AGG-AR) to be used as a benchmark in the forecasting exercise.

Variable	AGG-1F	AGG-2F	AGG-AR
Constant	0.014222 (0.002344)	0.013486 (0.002245)	0.004125 (0.002060)
Δy_{t-1}	0.587449 (0.088212)	0.626592 (0.086021)	0.878288 (0.104754)
Δy_{t-2}	-0.08447 (0.060785)	-0.096503 (0.056758)	0.015456 (0.138603)
Δy_{t-3}	-	-	0.144349 (0.138460)
Δy_{t-4}	-	-	-0.181859 (0.104537)
$\hat{f}^1_{t,3}$	0.010797 (0.008795)	0.010177 (0.009126)	
$\hat{f}^1_{t,2}$	0.002569 (0.013949)	0.000924 (0.014478)	
$\hat{f}^1_{t,1}$	0.001370 (0.006169)	0.002469 (0.006393)	
$\hat{f}^2_{t,3}$	-	-0.011678 (0.008954)	
$\hat{f}^2_{t,2}$	-	0.020313 (0.011417)	
$\hat{f}^2_{t,1}$	-	-0.000488 (0.005719)	
In-sample RMSE	0.010288	0.009938	0.012339
In-sample MAE	0.007823	0.007665	0.009538
N.Obs.	93	93	93
R^2	0.848364	0.858522	0.781881

Table 4: Estimated forecasting equations for GDP (s.e. in parenthesis)

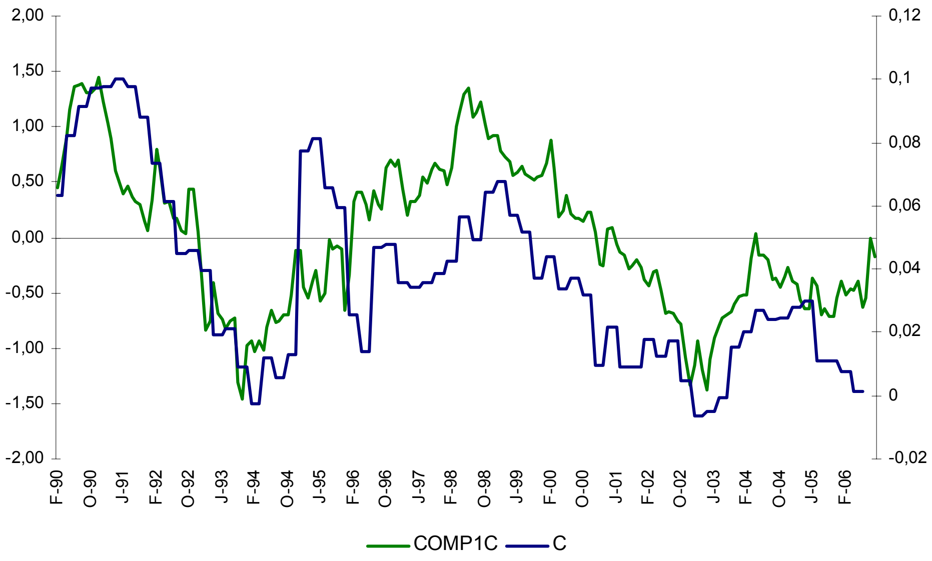

Figure 2: Quarterly real Private Consumption year-on-year growth rate (C) and corresponding first monthly factor (COMP1C).

4.2 GDP Nowcasting by Aggregation of Expenditure Disaggregates

In this subsection, we first present the estimation results for each expenditure disaggregate and at the end the results obtained from aggregating the forecasts of the expenditure disaggregates. As already mentioned, it was not necessary to estimate the current quarter government expenditures.

4.2.1 Private Consumption and Investment Nowcasting

Regarding private consumption, the indicators used to construct the common factors are those indicated in Tables 1 and 2. As explained previously, we use the longest series to set a balanced panel (coded C(B)) with which we estimate the missing data points of the unbalanced panel (coded C(U)). As can be seen in Table 5, the first extracted factor explain 45% of the variance while the second only explains an additional 9%. In our forecasting equation, whose results appear in Table 6, we ended up using only the first factor with its lags being jointly significant at the 1% level. In Figure 2 we compare the observed year-on-year growth rate of quarterly consumption with the first extracted monthly factor, named $COMP1C$.

The indicators used for investment appear in Tables 1 and 2 with codes I(B) and I(U) according to their use in the balanced and unbalanced panel, as before. The first two factors explain 52% of the variance while the third only explains 8%. The best estimated forecasting equation appears in Table 6 and included only the first two factors. The lags of the first estimated factor are jointly significant at the 1% level while the lags of the second estimated factor are jointly significant at the 5% level. In Figure 3 we present the series of the observed year-on-year growth rate of quarterly investment and the first two factors (labeled $COMP1F$ and $COMP2F$).

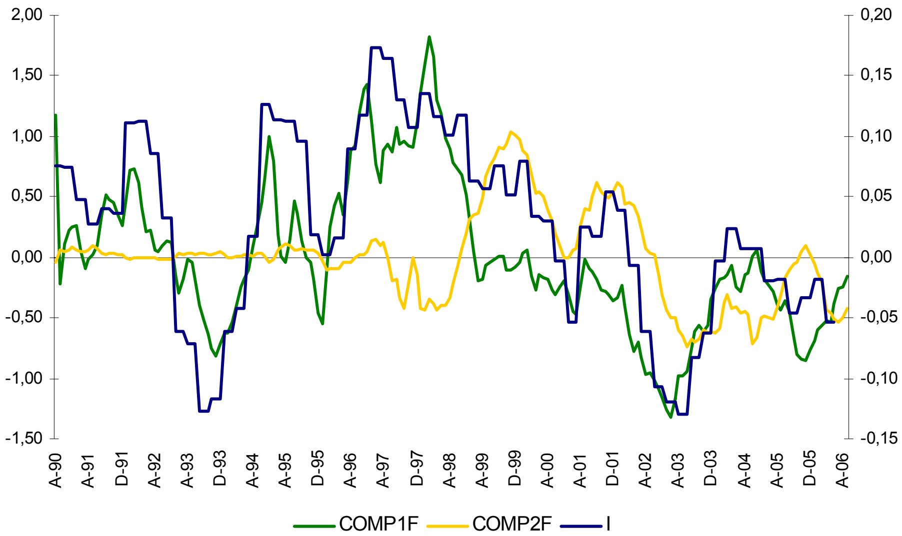

Figure 3: Quarterly real Investment year-on-year growth rate (I) and corresponding first two monthly factors (COMP1F and COMP2F).

Extracted	Private Consumption		Investment	
Factors	Explained Var. (%)	Cumulative (%)	Explained Var. (%)	Cumulative (%)
1	45.4	45.4	38.0	38.0
2	9.2	54.7	13.9	51.9
3	8.5	63.2	8.3	60.3
4	7.0	70.2	6.2	66.5
5	4.8	75.0	5.3	71.7
6	4.2	79.2	3.7	75.4

Table 5: First six estimated dynamic factors for Private Consumption and for Investment

Private Consumption		**Investment**	
Constant	0.008997 (0.003456)	Constant	0.009227 (0.004117)
Δy_{t-1}	0.737841 (0.080782)	Δy_{t-1}	0.584622 (0.090155)
$\hat{f}^1_{t,3}$	0.005874 (0.010406)	$\hat{f}^1_{t,3}$	0.058478 (0.031717)
$\hat{f}^1_{t,2}$	-0.009571 (0.015674)	$\hat{f}^1_{t,2}$	-0.029647 (0.055455)
$\hat{f}^1_{t,1}$	0.018960 (0.015633)	$\hat{f}^1_{t,1}$	0.090632 (0.061497)
$\hat{f}^1_{t-1,3}$	-0.004951 (0.010200)	$\hat{f}^1_{t-1,3}$	-0.070560 (0.035754)
		$\hat{f}^2_{t,3}$	0.122705 (0.062241)
		$\hat{f}^2_{t,2}$	-0.182793 (0.107085)
		$\hat{f}^2_{t,1}$	0.198980 (0.101933)
		$\hat{f}^2_{t-1,3}$	-0.111947 (0.051737)
Out-of-sample RMSE	0.009873	Out-of-sample RMSE	0.026455
In-sample RMSE	0.0122	In-sample RMSE	0.024267
In-sample MAE	0.008587	In-sample MAE	0.018912
N. Obs.	65	N. Obs.	63
R^2	0.819402	R^2	0.897275

Table 6: Estimated forecasting equations for Private Consumption and Investment (s.e. in parenthesis)

Exports Deflator		**Imports Deflator**	
Constant	0.002693 (0.001587)	Constant	-0.00006 (0.001520)
DEXP(-1)	0.246725 (0.107783)	DIMP(-1)	0.324256 (0.095555)
DEXP(-2)	0.280125 (0.109984)	DLBR	0.071424 (0.010561)
DLBR	0.023821 (0.009277)	DLEU	0.006477 (0.002717)
In-sample RMSE	0.010830	In-sample RMSE	0.011539
In-sample MAE	0.008792	In-sample MAE	0.008914
N. Obs.	76	N. Obs.	66
R^2	0.224882	R^2	0.465741

Table 7: Estimated forecasting equations for Exports and Imports deflators (s.e. in parenthesis)

Nominal Exports		**Nominal Imports**	
Constant	0.019600	Constant	0.010321
	(0.008312)		(0.007374)
EXP(-1)	0.263097	IMP(-1)	0.239724
	(0.093392)		(0.086774)
EXP(-2)	0.248747	IMP(-2)	0.182875
	(0.094988)		(0.087912)
EXP(-3)	0.221533	IMP(-3)	0.422615
	(0.094692)		(0.086732)
In-sample RMSE	0.055495	In-sample RMSE	0.054417
In-sample MAE	0.043694	In-sample MAE	0.044980
N. Obs.	114	N. Obs.	114
R^2	0.350949	R^2	0.530005

Table 8: Estimated forecasting equations for nominal Exports and Imports (s.e. in parenthesis)

4.2.2 Exports and Imports

Current quarter real exports and imports were estimated in an alternative way due to data availability. As explained in Section 2, exports and imports price deflators are released with a delay of 65 days and nominal monthly exports and imports with a delay of 50 days. With such a restriction, we decided to follow the following procedure. First, we estimate a forecasting equation for the current year-on-year growth rate of the quarterly price deflator of exports (denoted as DEXP) including as regressors its own lagged values (DEXP(-1) and DEXP(-2)) and the the first difference of the log brent price in euros (denoted as DLBR). For imports, the regressors are one lag of the dependent variable (DIMP(-1)), DLBR and the first difference of the log domestic industrial production price index of the Euro Zone (DLEU). Estimation results appear in Table 7. Forecasts of the year-on-year monthly nominal exports and imports disaggregates are obtained using a simple AR(3) model. Estimation results appear in Table 8. Finally, from the nominal and price deflator forecasts we obtain forecasts for the real series.

4.2.3 Aggregation of the Disaggregate Forecasts

After obtaining estimates of the year-on-year quarterly growth rate of real private consumption and investment as described above, we use them to compute the implicit levels of the real series. Then, we add them to the level of public consumption included in the official government budget accounts to obtain the estimated level of domestic demand. Adding the estimated levels of real exports and imports, we obtain an estimate of the quarterly GDP level. After this step, we compute the year-on-year quarterly growth rates and compare them directly with the actual growth rates to infer about the forecasting performance of the method.

5 Forecast Evaluation

In this section we compare the out-of-sample forecasting performance of the alternative methods. For each quarter in the period 2002Q1-2006Q3, we obtained forecasts using all the information available since the beginning of the available sample up to the first 50 days after the end of that quarter. We compare the results of directly nowcasting aggregate GDP including one or two dynamic factors in the forecasting equation (denoted as AGG-1F and AGG-2F models in Table 4) with the indirect approach of aggregating the forecasts of the expenditure disaggregates (denoted as DIS model). We have also computed forecasts obtained from a simple AR model to serve as benchmark (denoted as AGG-AR).

In Table 9 we present the out-of-sample root mean-squared forecast error and mean absolute forecast error. The aggregation of the disaggregate forecasts approach (DIS) yields the highest values. This suggests that when we aggregate the forecasts of the expenditure disaggregates, the forecast errors do not cancel out, instead, they add-up. In the same table appear the results of pairwise Diebold-Mariano tests comparing the predictive ability of the different models. The loss function used was the absolute value of the forecast error. A cell with a positive value donotes a better performance of the model in column relative to the model in row. These results confirm the superiority of the aggregate approach outperforming the disaggregate approach and the AR model. We also conclude that aggregating the disaggregate forecasts results in significant loss of predictive power.

In Figure 4 are depicted the forecasts from the AGG-1, DIS and AGG-AR models along with the observed values of the GDP year-on-year growth rate series. It is clear that the disaggregate approach performs much worser relative to the other alternatives. To better understand the source of this poor perfomance, Figure 5 graphs the observed values of the expenditure disaggregates against their corresponding forecasts. As can be observed, the major source of forecast error originates in the exports and imports disaggregates.

Diebold-Mariano	AGG-1F	AGG-2F	DIS	AGG-AR
AGG-2	1.52 (0.1461)	-	-	-
DIS	4.29 (0.0005)	4.77 (0.0002)	-	-
AGG-AR	2.49 (0.0233)	1.43 (0.1703)	-3.93 (0.0011)	-
Out-of-sample RMSE	0.0050	0.0062	0.0572	0.0098
Out-of-sample MAE	0.0042	0.0053	0.0433	0.0076

Table 9: Diebold-Mariano tests (p-values in parenthesis)

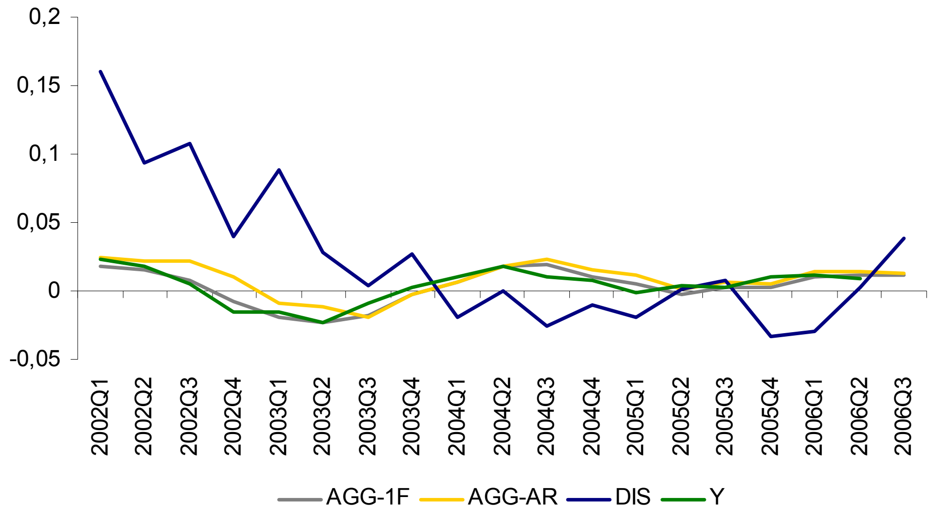

Figure 4: Forecasts and observed values of the quarterly real GDP year-on-year growth rate

Figure 5: Forecasts and observed values of quarterly real Private Consumption, Investment, Exports and Imports year-on-year growth rates.

6 Conclusions

This paper explores how dynamic factor models can be used to forecast aggregate GDP by aggregating forecasts of its expenditure disaggregates in the context of an empirical study. The out-of-sample forecasting performance results obtained for the case of Portuguese GDP suggest that there is no advantage in aggregating the forecasts obtained using specific dynamic factor models for each of the major expenditure disaggregates.

The major source of errors in the disaggregate approach comes from the forecasts of exports and imports. Due to data limitations we decided to follow an alternative approach when forecasting these variables. However, given the results obtained, it might be worthwhile enlarging the dataset and explore forecasting these variables also using a dynamic factor model approach. Nonetheless, the results obtained in a recent work on forecasting exports by Cardoso and Duarte (2006) suggest that even using extra qualitative information still leads to a relatively high RMSE.

There are a number of other possible extensions of this empirical study that could be further explored. It could be interesting to look at the forecasting performance of a model considering quarterly indicators instead of monthly indicators, and also to compare the results using balanced or unbalanced panels. Another important question concerns the choice of method to stationarize the series. We have used year-on-year growth rates of the variables being forecasted as these tend to be highly correlated with the levels of many of the indicators (particularly those coming from qualitative surveys). However it is possible that using simple differences could result in different conclusions.

References

Boivin, J.; Ng, S. (2006) "Are more data always better for factor analysis?", Journal of Econometrics 132, pp. 169-194

Camacho, Maximo; Sancho, Israel (2003) "Spanish Diffusion Indexes", Spanish Economic Review 5 (3) pp. 173-203

Cardoso, Fátima; Duarte, Cláudia (2006) "A utilização de Informação Qualitativa na Previsão das Exportações", Boletim Económico, Banco de Portugal, Inverno 2006, pp. 65-73

Duarte, C.; Rua, A. (2005) "Forecasting inflation through a bottom-up approach: the portuguese case", Bank of Portugal Working Paper 2005-02.

Espasa, A.; Senra, E.; Albacete, R. (2002) "Forecasting inflation in the European Monetary Union: A disaggregated approach by countries and by sectors", European Journal of Finance 8(4), pp. 402-421

Fair, R. C.; Shiller, J. (1990) "Comparing information in forecasts from econometric models", The American Economic Review 80(3), pp. 375-389

Hendry, D.; Hubrich, K. (2006) "Forecasting Economic Aggregates by Disaggregates", European Central Bank Working Paper no. 589

Hubrich, K. (2005) "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?", International Journal of Forecasting 21(1), pp. 119-136

INE (2004) "Contas Nacionais Trimestrais - Documento Metodológico"

Kapetanios, G.; Marcellino, M. (2003) "A comparison of estimation methods for dynamic factor models of large dimensions", Queen Mary University of London, Working Paper No. 489

Koenig, Evan; Dolmas, Sheila; Piger, Jeremy (2000) "The Use and Abuse of Real-Time Data in Economic Forecasting" International Fiance Discussion Papers 684, FED

Lutkepohl, H. (2005) "Forecasting with VARMA processes", in G. Elliott; C. W. J. Granger; A. Timmermann (eds), Handbook of Economic Forecasting, Elsevier

Mitchell, James; Waele, Martin; Smith, Richard; Wright, Stephen; Salazar, Eduardo (2005) "An Indicator of Monthly GDP and an Early Estimate of Quarterly GDP Growth" The Economic Journal, 115 pp. F108-F129

Mitchell, James; Weale, Martin (2005) "A Monthly Indicator of GDP Euro Area: an Application of the Regression Approach to Interpolation" OECD

Nunes, Luis C. (2005) "Nowcasting Quarterly GDP in a Monthly Coincident Indicator Model" Journal of Forecasting 24, pp. 575-592

Schneider M.; Spitzer, M. (2004) "Forecasting Austrian GDP using the generalized dynamic factor model", OeNB Working Paper 89, Vienna

Stock, James; Watson, Mark (2002) "Macroeconomic Forecasting Using Diffusion Indexes" Journal of Business and Economic Statistics, 20(2) pp. 147-162