Banco de Portugal
EUROSISTEMA

Estudos e Documentos de Trabalho

Working Papers

18 | 2008

FORECASTING INVESTMENT: A FISHING CONTEST USING SURVEY DATA

José Ramos Maria
Sara Serra

October 2008

Please address correspondence to
José Ramos Maria
Economics and Research Department
Banco de Portugal, Av. Almirante Reis no. 71, 1150-012 Lisboa, Portugal;
Tel.: 351 21 313 0879, Email: jfrmararia@bportugal.pt

BANCO DE PORTUGAL

Economics and Research Department

Av. Almirante Reis, 71-6th floor

1150-012 Lisboa

www.bportugal.pt

Printed and distributed by

Administrative Services Department

Av. Almirante Reis, 71-2nd floor

1150-012 Lisboa

Number of copies printed

170 issues

Legal Deposit no. 3664/83

ISSN 0870-0117

ISBN 978-989-8061-54-6.

Forecasting investment:

a fishing contest using survey data*

José Ramos Maria

Economic Research Department, Banco de Portugal, 1150-165 Lisboa, Portugal

jfrmaria@bportugal.pt

Sara Serra

Economic Research Department, Banco de Portugal, 1150-165 Lisboa, Portugal

srserra@bportugal.pt

October 24, 2008

DRAFT VERSION

Abstract

This paper assesses the usefulness of business surveys as a source of information for investment developments in Portugal. This will be achieved by what will be named a "fishing contest", where the "participants" are bridge models, models based on principal components (derived from standard and non-standard methods), and models built with the outcome of partial least squares regressions. All models, based on quarterly data, are estimated using a general-to-specific approach and are designed to produce 1 to 4 out-of-sample direct forecasts. The accuracy of these forecasts is then compared with the one of autoregressive processes. The empirical evidence indicates that, in general, there is always a participant in the fishing context that produces a lower out-of-sample Root Mean Squared Error (RMSE) than the one associated with the autoregressive benchmark. In most cases, the combination of autoregressive processes with each participant reduces the RMSE further. A striking outcome is the relative accuracy of bridge models.

Keywords: Investment, business surveys, bridge models, principal components, partial least squares.

JEL Codes: C22, C52, C53, E22

*Without implication, we would like to acknowledge the helpful discussions with Francisco Dias and the comments made by Cláudia Duarte, Rita Duarte, Ricardo Mourinho Félix, Carlos Robalo Marques and Maximiano Pinheiro on earlier versions of this paper. The responsibility for all errors and omissions remains with the authors.

Contents

1 Introduction

Investment is a key variable underlying economic activity developments. However, it is also one the most volatile components of aggregate demand and traditionally difficult to predict. The identification of the structural interactions that explain the investment decisions is always a complex task, in part due to the expectation formation mechanisms of economic agents, which may be time-varying or sector-dependent. Some firms may invest as a reaction to a favourable economic situation, possibly unexpected, while others invest because they expect higher demand over the medium or long run. This may co-exist with firms that do not invest at all simply because they have already achieved their desired capital stock. In addition, the driving forces among different sectors may be rather different, for example between residential and productive investment.

The main objective of this paper is to assess the usefulness of business surveys as a source of information that may be used to capture contemporaneous or leading forces driving investment in Portugal. The goal is therefore not to look for the best functional form, by taking into account standard variables, such as the productivity of capital, adjustment costs or the cost of purchasing new capital, be it in a partial or a general equilibrium framework, or to model the theoretical microfoundations of firms' behaviour, but is instead to use business surveys extensively so as to find valuable empirical comovements between these data and investment over the short run. The use of survey data has several well-known advantages. For instance, besides being in general unrevised, the data is also available in advance of other quantitative indicators, including national accounts data.

The information content of surveys has been widely explored in the literature. The examples include the estimation of bridge-models, usually addressing GDP or private consumption as variables of interest (see Bram and Ludvigson (1997) or Rünstler and Sédillot (2003)). Another branch of the literature uses a wider information environment, combining a larger number of short-term indicators, including surveys, which are used as inputs in static or dynamic factor models (see Stock and Watson (1989),**?** or Hansson, Jansson and Lof (2005)). More recently, Claveria, Pons and Ramos (2007) try to improve forecasts for a relatively large number of macroeconomic variables using the information provided by these surveys. The usefulness of business surveys as an important information source behind investment developments has been the main focus of Larsen (2001) or Barnes and Ellis (2005).

The starting point of the current analysis is to estimate an autoregressive process for the variables of interest. These variables consist of Gross Fixed Capital Formation expenditures (GFCF), and some of its subcomponents. The second step is to estimate a model that only uses specific survey data or some sort of summary indicator of the survey database. The final step is to augment each autoregressive (AR) processes with the information that is solely derived from the survey data. Is the out-of-sample accuracy of the AR model higher then that of the models that only use survey data? Does the augmented AR model outperforms the others? How do the conclusions vary across dependent variables and forecasting horizons? These are the questions that will be addressed below.

The information content of the survey data will be assessed by what will be considered "participants" in a "fishing contest". Previous work regarding the usefulness of surveys focused extensively on bridge models. This will be the natural first participant. The second participant uses the first standard principal components as summary indicators of the database. A recent use of principal components, extracted from survey data, may be found in Claveria et al. (2007). The third participant is also derived from standard principal components but focuses on those components that are more correlated with the variable of interest, which may not necessarily be the first ones, as in the previous case. This is in line with the literature that highlights the importance of having regressors that take into account that the goal is to forecast a specific series and not just the usefulness of summarizing a particular database (see, for instance, Bai and Ng (2007, 2008)). The fourth participant follows the suggestion of Dias, Pinheiro and Rua (2008), who investigated the links between predictors and endogenous variables and suggested the use of a synthetic indicator derived from a particular weighting scheme of all principal components. The fifth alternative is based on a particular weighting scheme of the original data, not the principal components, where more weight is attached to those survey answers that are potentially more important to explain the variable of interest. These weights are defined as the correlation coefficients between each survey data and the variable of interest, where more weight is attached to the variables with higher correlation. Finally, the last participant is based on Partial Least Squares (PLS) regressions, which combines features from principal components and standard OLS regressions. In particular, the components are extracted from the survey data already under the operational restriction that they are also relevant for the variable of interest.

This paper is organized as follows. The next section presents the database. Section 3 introduces the participants of the proposed contest. The empirical evidence is reported in Section 4 and Section 5 concludes.

2 The Database

The survey data used herein is taken from the database of the European Commission (EC) and include information on the following sectors of the Portuguese economy: manufacturing industry (henceforth denominated as the industry survey), construction, retail trade and services.[1]

The majority of the survey responses has a monthly frequency. Besides being in general unrevised and probably less susceptible to sampling and measurement errors, the survey data is also known in advance of national accounts data.[2] The usual survey questions regard recent developments

[1]The data can be retrieved from the Eurostat website http://europa.eu.int/comm/eurostat.

[2]The data is published on the last working day of the month to which it refers. This is about 45 days in advance of the GDP flash estimate and 75 days of the first release of the national accounts, when the investment data is actually disclosed. On the issue of sampling and measurement errors, see Claveria et al. (2007) for references.

in trends in production, order books and stock levels, as well as forward looking questions, regarding production and employment. Answers are usually expressed as being an "unchanged" or "normal" situation, or as a movement "above" or "below" that standard. They are in general of a qualitative nature and are published in the form of seasonally adjusted balances.

By construction, the answer to each question of the surveys usually represents the difference between the percentage of firms which have noted an improvement and those which have reported a deterioration. A typical example is "How do you expect your production to develop for the months ahead?", where the alternatives are "It will... (+) increase; (=) remain unchanged; (-) decrease". Each survey has a composite indicator, calculated as a simple aggregation of a few answers (e.g. "Industry confidence indicator"). Most survey indicators stand between -100 (all firms reported a deterioration) and +100 (all firms reported an improvement). The prospective nature of some questions also predisposes them to be leading indicators of investment, which justifies the investigation of the predictive power of lagged data. Although these features are advantages when compared to other high-frequency information used for short-term forecasting, the subjectivity of these indicators may make them prone to idiosyncratic factors. Given that the usefulness of confidence indicators for business cycle analysis seems to vary from country to country (see Santero and Westerlund (1996)), the generalization of the empirical results may be unwarranted.

The frequency of the data used herein is quarterly, which implies that the series are derived either from monthly averages or directly from quarterly answers that some surveys also report. These quarterly questions include current production capacity, competitive position inside and outside the European Union (EU), expectations regarding export orders, etc. In addition, questions pertaining the factors limiting the production are also included. The answers to these vary from (i) none; (ii) insufficient demand; (iii) shortage of labour force; (iv) shortage of material and/or equipment; (v) financial constraints to (vi) other factors. The current paper will focus exclusively in a balanced database approach, despite the limitations that this imposes on sample size for estimation. In addition, a complete-quarter information context is replicated in the present work, i.e situations where the three months of the quarter are all know.[3]

The surveys database includes several breakdowns. The usefulness of business surveys as a source of information for investment developments concentrates herein in two of them. The first, henceforth denominated as "database of totals", focuses exclusively on the aggregates for whole sectors. It includes, for example, the production trend observed in recent months for the total industry survey, the assessment of order-books levels for the total industry survey, etc. The second, denominated the "database of sectors", breaks down the industry and the construction surveys into several subsectors. It includes, for instance, the production trend observed in recent months for the consumer goods sector, the food and beverages sector, etc.[4] The database

[3]Situations in which the results of the surveys are available for part of the quarter, and the rest has to be forecasted are excluded from the current setup.

[4]The surveys include other breakdowns that will not be explored here, for example by main industrial groups

Table 1: List of survey indicators

Codes		Sectors	Total and subsectors	Frequency	Starts in…
ICI	**Industry**	Industry Confidence Indicator, defined as (I2-I4+ I5)/2	Total Manufacturing	m	Jan 1987
I1		Production trend observed in recent months	Consumer Goods	m	Jan 1987
I2		Assessment of order-book levels	Durable Consumer Goods	m	Jan 1987
I3		Assessment of export order-book levels	Non Durable Consumer Goods	m	Jan 1987
I4		Assessment of stocks of finished products	Food, Beverages	m	Jan 1987
I5		Production expectations for the months ahead	Investment Goods	m	Jan 1987
I6		Employment expectations for the months ahead	Intermediate Goods	m	Jan 1987
Iq1		Assessment of current production capacity		q	Jan 1987
Iq2		Duration of production assured by current order-book levels		q	Jan 1987
Iq3		New orders in recent months		q	Jan 1987
Iq4		Export expectations for the months ahead		q	Jan 1987
Iq5		Current level of capacity utilization		q	Jan 1987
Iq6		Competitive position domestic market		q	Jul 1994
Iq7		Competitive position inside EU		q	Jul 1994
Iq8		Competitive position outside EU		q	Jul 1994
Iq9		Factors limiting the production		q	Jan 1987
Iq9F1		None		q	Jan 1987
Iq9F2		Demand		q	Jan 1987
Iq9F3		Labour		q	Jan 1987
Iq9F4		Equipment		q	Jan 1987
Iq9F5		Other		q	Jan 1987
CCI	**Construction**	Construction Confidence Indicator, defined as (C3+ C4)/2	Construction as a whole	m	Jan 1989
C1		Building activity development over the past 3 months	Building: total	m	Jan 1989
C2		Main factors currently limiting your building activity	Building: residential	m	Jan 1989
C2F1		None	Building: non-residential	m	Jan 1989
C2F2		Insufficient demand	Public works (civil engineering)	m	Jan 1989
C2F3		Weather conditions		m	Jan 1989
C2F4		Shortage of labour force		m	Jan 1989
C2F5		Shortage of material and/or equipment		m	Jan 1989
C2F6		Other factors		m	Jan 1989
C3		Evolution of your current overall order books		m	Jan 1989
C4		Employment expectations over the next 3 months		m	Jan 1989
Cq1		Operating time ensured by current backlog (in months)		q	Jan 1989
RCI	**Retail Trade**	Retail Trade Confidence Indicator, defined as (R1-R2+ R4)/3	Total Retail Trade	m	Jan 1989
R1		Business activity (sales) development over the past 3 months		m	Jan 1989
R2		Volume of stock currently hold		m	Jan 1989
R3		Orders expectations over the next 3 months		m	Jan 1989
R4		Business activity expectations over the next 3 months		m	Jan 1989
R5		Employment expectations over the next 3 months		m	Jan 1989
SCI	**Services**	Services Confidence Indicator, defined as (S1+ S2+ S3)/3	Total Services	m	Jun 1997
S1		Business situation development over the past 3 months		m	Jun 1997
S2		Evolution of the demand over the past 3 months		m	Jun 1997
S3		Expectation of the demand over the next 3 months		m	Jun 1997
S4		Evolution of the employment over the past 3 months		m	Jun 1997
S5		Expectations of the employment over the next 3 months		m	Jun 1997

NOTES: The frequency of the survey releases are indicated by the letter "m" for monthly and "q" for quarterly. The formulæ behind the composite indicators of the different surveys are also indicated, according to the codes presented in the first column.

Table 2: Variables of interest

GFCF Series	Frequency
Overall	Q
Public	Q
Private	Q
Residential	Q
Productive	Q
Construction	Q
Overall excluding construction	Q

of totals has 42 variables and the database of sectors has 185 variables. Information on both databases can be found in Table 1. Due to availability issues, the sample period has 42 observations and ranges from 1997Q3 to 2007Q4. Using these distinct databases allows the evaluation of the potential gains for forecasting purposes of using a richer information environment (in the sense of considering information on the same sectors, but at a more detailed breakdown) given that the information from a given sector seems *a priori* more targeted to forecast a given type of GFCF (as in the case of residential building survey and, eventually, private housing GFCF).

The variables of interest in this work, listed in Table 2, are those of GFCF expenditures and several of its subcomponents, namely Public and Private GFCF, being the latter also disaggregated into residential and productive GFCF. In addition, a disaggregation of GFCF into construction and total excluding construction is also considered.[5] The use of all these variables relies on the possibility that some survey data may contain important interactions that can effectively capture contemporaneous or leading forces over the short run among all agents of the economy. In some situations, as in the case of Public GFCF, although the data depends on Government decisions, one should not neglect the possibility that such decisions may have spill-over effects on the private sector, with an impact on the behavior of some survey data.

All series were tested for the presence of unit roots using standard Augmented Dickey-Fuller (ADF), Philips-Perron (PP) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests, using the longest possible time span of each series. While the survey data is tested without any transformation, the investment series were tested after taking logs.

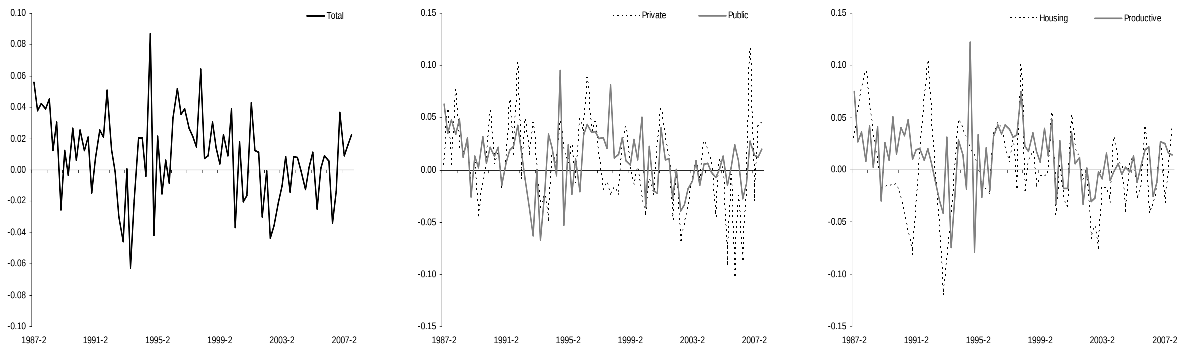
Given the nature of the survey data (whose general movement presumably reflects the different positions of the business cycle, and evolve, in most cases, within a fixed ± 100 interval of possible outcomes), it has been mentioned in the literature that they should be regarded as stationary

or by aggregates collected according to the NACE Rev. 1.1 classification. In addition, some time series were eliminated given that the available time-span was considered too short or are relatively sparse, such as the "Investment survey on the manufacturing sector", which only gathers information on companies' investment plans twice a year. Information on selling prices and the consumer survey were also left out. More information on the survey data can be found in European Commission (2007).

[5]Total GFCF data is taken from the database of Banco de Portugal. See Banco de Portugal (2008) and the website www.bportugal.pt.

Figure 1: Gross fixed capital formation expenditures in Portugal

(log first-quarter differences)



(a) Total (b) Private and public (c) Housing and productive

Figure 2: Gross fixed capital formation expenditures in Portugal

(log four-quarter differences)

(a) Total (b) Private and public (c) Housing and productive

variables (see, for instance, European Commission (2000)). However, the results of the above-mentioned tests do not allow for a clear-cut conclusion in all cases, which is also a result that is not uncommon (see European Commission (2000)) or Artís and Suriñach (2003)). If one assumes that the survey indicator is stationary in the case where at least one of the tests does allow for this interpretation at 10% significance level, only around 20% of the data are found to be non-stationary and should therefore probably be used only after taking first differences. The visual inspection of some of these series, namely against other survey data where the stationary status is less controversial, lead us to conclude that the lack of stationarity may simply be due to the short sample period. Moreover, the use of survey data in levels is quite common in the empirical literature and, to our knowledge, no empirical work has been carried out in first differences of the survey data. If one applies the same criteria for the GFCF series – the variables that we are interested in forecasting –, they can all be considered to be $I(1)$.[6]

The analysis of the quarterly GFCF data will be carried out in quarter-on-quarter and in year-on-

[6] All stationary tests can be made available from the authors.

year changes. On the one hand, the analysis on quarterly rates of change seems more appropriate given that they isolate the innovation of the series in each period, but on the other hand visual inspection suggests that the surveys may be more correlated with the behaviour of investment on year-on-year terms. The two options can be found in the literature.[7] The quarterly evolution of some of the variables of interest is depicted in Figure 1 and the yearly evolution in Figure 2.

The quarter-on-quarter behavior of GFCF does reveal that, on some occasions, the volatility periods are rather striking. For example, total GFCF in 1994Q4 increased almost 10% and in 1995Q1 decreased slightly less than 5%. On the contrary, there are periods, e.g. in 2003-04, where the quarterly rates of total GFCF only oscillate between ±2%. It is also evident that the driving forces among different sectors can also be rather different. For instance, when comparing private and public investment, the volatility of the former during the last part of the sample period is substantially higher than of the later. It is therefore with no surprise that some authors, for instance Barnes and Ellis (2005), classify the investment expenditures as highly volatile and traditionally difficult to predict. The survey data does not depict an evolution with such severe volatility. When the focus is on year-on-year rates, the behavior of GFCF data can still be rather volatile but the degree of volatility is by construction less striking.

3 The methodology

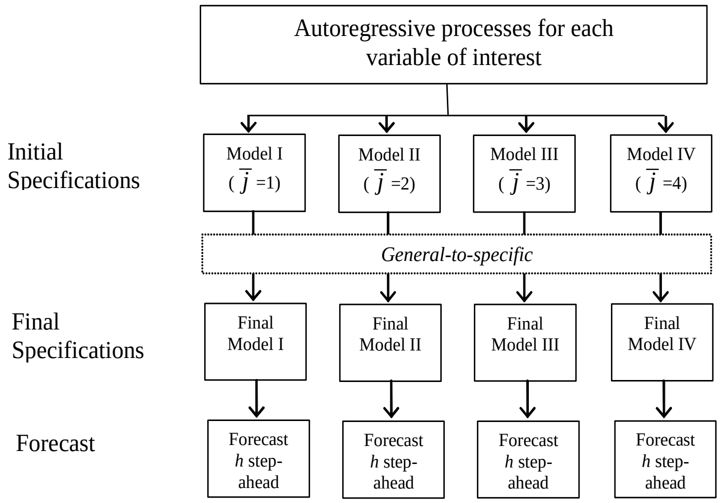
This section clarifies the methodology that will be used to assess the usefulness of business surveys as a source of information that can be used to capture GFCF dynamics in Portugal. The analysis is initiated with the estimation of autoregressive models for each variable of interest. All models, based on quarterly data, are designed to produce 1 to 4 out-of-sample direct forecasts, i.e. the models forecast directly h steps ahead and $h = 1, 2, 3, 4$. When h=1, their "nowcasting" accuracy is being explored, given that the survey data is known in advance of national accounts data.

The forecasting procedure associated with the autoregressive processes is clarified in Figure 3. There are four initial specifications for each variable of interest. Besides the constant, the first specification includes one single regressor ($\overline{j} = 1$), the second includes two regressors ($\overline{j} = 2$), and so on, up to a maximum of 4 regressors. From the initial specifications, a general-to-specific approach is followed. For each model, only one regressor is dropped at a time. This regressor is the one with the lowest level of significance and could be any one of the initial specification (including intermediate lags). After allowing for the sequential exclusion of all regressors that are not significant at 10%, which ensures both in-sample fit and parsimonious features and implies gains in degrees of freedom, there are four final specification that are all used to forecast.

The out-of-sample performance of the autoregressive processes will then be compared with that

[7] Rünstler and Sédillot (2003) use the survey data to forecast quarterly changes of GDP. An analysis based on yearly frequency can be found in Hansson et al. (2005) or Claveria et al. (2007). Artís and Suriñach (2003) and Barnes and Ellis (2005) have analysis in both quarterly and yearly terms.

Figure 3: The forecasting process for autoregressive models



of models that only use survey data, be it with specific time series that are directly available in the database or time series that are previously calculated from the database. The latter are assumed to represent valuable summary indicators of the entire information set (for example, principal components). These models that only use survey data configure what will be considered "participants in a fishing contest", where the objective is to capture the (out-of-sample) GFCF dynamics. The forecasting procedure associated with each participant is equal to the one presented in Figure 3, with the exception that there are 5 initial specifications. More precisely, the initial specifications of those models that only use survey data are constructed will be based on a maximum regressors indexed by $\overline{k}$, where $\overline{k} = 1, 2, 3, 4, 5$. From the initial specifications, the general-to-specific approach presented in Figure 3 remains in place.

As the forward-looking nature of some survey dataseries may imply that they may lead investment to some extent, each one is initially investigated in terms of its correlation against the variables of interest. Therefore, instead of only using contemporaneous data, the correlation of each survey with each variable of interest is computed up to four lags. The original variable is then lagged if the highest correlation is not the contemporaneous one. More precisely, assume that the variable x_t represents a particular time series of the survey database. This variable is then used to create the variable $x^{\star}_{i|t}$, where the subscript $i|t$ highlights the possibility that, conditional on the information available up to t, the time-subscript i can be t or $t-1$,..., up to $t-4$. The variable x_t is then replaced by $x^{\star}_{i|t}$, where i is defined by the highest correlation amongst the contemporaneous and lagged variables. This implies that the original matrix X with survey data is replaced by a matrix X_h, with the $x^{\star}_{i|t}$ variables, which is conditional on the number-of steps h and on the variable of interest. With 4 steps-ahead forecasts and 7 variables of interest, analysed both in quarter-on-quarter and year-on-year rates of change, the total number of matrices X_h is equal to 56. The criteria of using the highest correlation between each survey data and the variable of interest is in line with Barnes and Ellis (2005).

Finally, the participants of the fishing contest are combined with AR models. In this case, $\overline{j}$ is set to 4 in all initial specifications while $\overline{k}$ goes again from 1 up to 5. During the general-to-specific approach, any regressor of the initial specification can be excluded (including therefore intermediate lags of the autoregressive structure or any regressor solely extracted from survey data).

The choice of the maximum number of regressors of all initial specifications takes into account the need to ensure reasonable degrees of freedom. In the case of the models that only use autoregressive terms, the use of quarterly data also contributed to the assumption of 4 as the maximum $\overline{j}$. The options (i) not to choose a particular model with a pre-specified maximum number of regressors (for instance, according to an information criteria) and (ii) to assume all initial specifications as equally important for forecasting purposes is due to the fact that the optimality of an in-sample fit may not be matched by an equivalent performance in terms of out-of-sample accuracy. All final specifications are evaluated by the RMSE of out-of-sample forecasts for the time interval between 2006Q1 and 2007Q4.[8] These last 8 quarters represent around 20% of the entire sample period. The forecasts will be generated using a expanding window where the estimation period starts in 1997Q3.

3.1 The AR model

The AR model (henceforth denominated as "Method 0"), estimated for each variable of interest, provides a naïve benchmark with which all other methods can be compared. This allows to infer if there is any useful information for forecasting purposes contained in the survey data not already inherent to the dependent variable itself.

As already mentioned, the AR process that is used for forecasting purposes is estimated following a general-to-specific approach. The initial specifications have the following form:

$$\begin{aligned} y_{t-1+h} &= \alpha_h + \textstyle\sum_{j=1}^{\overline{j}} \gamma_{jh} y_{t-j} + \epsilon_{h,t-1+h} \\ &\text{where} \quad h = 1 \ldots 4 \text{ and } \overline{j} = 1 \ldots 4 \end{aligned} \tag{1}$$

The variable y_t represents quarter-on-quarter or year-on-year rates of change, depending on the database that is being used, and $\overline{j}$ is the maximum number of lags of each general specification. If no lag is significant during the sequential process embodied in the general-to-specific approach, equation (1) collapses to a constant and this is the model that is used to implement the specific direct forecast. Moreover, if $\overline{j}$ increases but the additional regressor has the lowest level of significance, among all those that are not significant, the model that is actually used to forecast

[8] The superiority of using an out-of-sample in comparison with an in-sample analysis for the purpose of evaluating forecasting methods was investigated, for instance, by Tashman (2000).

remains unchanged against the previous initial specification. Given that the procedure is followed for 4 equations, one for each h step-ahead, note that γ is also indexed by h. To evaluate the "nowcasting" features of the models, it will be considered that if T is the last period of the survey data, then the available sample period of y_t ends at $T-1$, and therefore y_T is not available.

3.2 Method 1: bridge models

The natural first participant of the fishing contest, named "Method 1" henceforth, is based on bridge models, i.e. simple econometric formulations that establish a link, or a bridge, between conjunctural information which is available in advance of other variables and GFCF data, so that the former can be used to forecast the latter. These high-frequency models do not necessarily stem from economic theory, and are thus not behavioural or structural in that sense. Previous work regarding the usefulness of surveys for short-term forecasting has extensively focused on this type of models.

The initial specifications of bridge models without autoregressive terms followed herein are given by:

$$\begin{aligned} y_{t-1+h} &= \beta_h + \textstyle\sum_{k=1}^{\overline{k}} \delta_{kh} z_{k,t} + \eta_{1h,t-1+h} \\ &\text{where} \quad h = 1\ldots4 \ , \ \overline{k} = 1\ldots5 \ \ \mathit{and} \ \ x^\star \equiv x^\star_{i|t} \end{aligned} \tag{2}$$

where $\overline{k}$, which is the maximum number of regressors of each initial specification, and $x^\star_{i|t}$ were already defined in the beginning of this section. $\eta_{1,t}$ is an error term. The actual $x^\star_{i|t}$ variables that are included in equation (2) are the ones with the highest correlation among all $x^\star_{i|t}$ variables.

The initial specifications of bridge models with autoregressive terms are given by:

$$\begin{aligned} y_{t-1+h} &= \mu_h + \textstyle\sum_{j=1}^{4} \theta_{jh} y_{t-j} + \textstyle\sum_{k=1}^{\overline{k}} \psi_{kh} z_{k,t} + \eta_{2h,t-1+h} \\ &\text{where} \quad h = 1\ldots4 \ , \ \overline{k} = 1\ldots5 \ \ \mathit{and} \ \ x^\star \equiv x^\star_{i|t} \end{aligned} \tag{3}$$

where $\eta_{2,t}$ is an error term. The actual $x^\star_{i|t}$ variables that are included in equation (3) are the same as the ones that entered in equation (2). Note that the maximum number of autoregressive terms is now fixed at 4, whereas in equation (1) was allowed to vary between 1 and 4.

The objective of evaluating equations (2) and (3) is to assess whether the combination of autoregressive processes an specific survey data improves the forecast accuracy of the models. As in the forecasting process of autoregressive models depicted in Figure 3, the most parsimonious bridge model is then constructed for equations (2) and (3), following an identical general-to-specific approach. This approach allows more flexibility in the definition of the final specification of each model.[9]

3.3 Method 2: standard PC

Given that the survey database incorporates time series that have a high degree of correlation among them, a standard methodology that explores this feature is the principal components methodology. This participant in the fishing contest, henceforth denominated "Method 2", constructs linear combinations of the time series that can be seen as summary indicators of the entire database. The method allows to take advantage of a large set of indicators without loosing too many degrees of freedom, as would happen in a standard OLS regression framework. The standard approach, used in this section, consists in deriving the components from the correlation matrix of the original variables. In this case, all $x^{\star}_{i|t}$ variables are assumed, using the expression mentioned by Chatfield and Collins (1996), to "arrive on an equal footing". Let Z be a standardized $T \times N$ variables matrix, where T corresponds to the number of observations and N to the number of variables. The components can be obtained from the matrix of second moments of the variables, i.e $(NT)^{-1}Z'Z$[10].

The estimated equations used to forecast have the same form as equations (2) and (3). The only qualitative difference is that the variable $x^{\star}_{i|t}$ does no longer represent specific time series of the survey database, lagged or contemporaneous, and contains, instead, the first principal components. Given that the equation that is used for forecasting purposes is estimated following the same general-to-specific approach, all principal component are excluded if they are not statistically significant at 10%.[11] If the components are available up to T, it is assumed that y_t is only available up to $T-1$.

3.4 Method 3: targeted PC

The possible need of going beyond the standard use of the principal components methodology has received some attention in the empirical literature. Bai and Ng (2007) or Bai and Ng (2008) highlighted the importance of having regressors that take into account that the goal is to forecast a specific series and not necessarily to explain the total variance of a given database. Moreover, there is no reason to think that factors that best explain a particular economic variable are also the same that explain another (completely different) variable. The expression "targeted PC"

[9] Alternative approaches could have been pursued, as for example combining the forecasts from the AR model and each of the fishing contest participants, or setting the the number of autoregressive terms to enter equation (3) to a fixed number.

[10] Alternatively, the principal components can also be extracted from $(NT)^{-1}ZZ'$.The principal component methodology is analysed in detail in Jackson (1991) or Jollife (2002).

[11] The generalization of equations (2) and (3) should allow for lagged principal components. However, the reduction in the degrees of freedom and some preliminary empirical evidence showing that the results did not seem qualitatively superior, lead us to abandon this line of research. Another method that was not pursued is based on a state space approach and on the use of the Kalman filter. An example of this approach may be found in Hansson et al. (2005)).

is taken from the expression "targeted predictors" of Bai and Ng, which highlights exactly this issue of finding adequate components for forecasting purposes.

The new participant of the fishing contest, henceforth denominated "Method 3", starts by computing all principal components, as in the standard approach, but searches for those that are more correlated with the variable of interest, instead of choosing the first ones. Therefore, instead of having the sole objective of explaining the larger percentage of the total variation of the survey data, which is a problem not conditional on the variable of interest, this participant chooses the more correlated components, where the first component is an equal candidate against all others. Moreover, with this approach, different dependent variables and different out-of sample forecasting horizons can bring about different components as regressors, instead of focusing exclusively in the first components.

The estimated equations used to forecast with this participant have the same form as equations (2) and (3). But in this case,$x^{\star}_{i|t}$ contains time series which use the components that are more correlated with the variable of interest for each step ahead. The procedure of finding the most parsimonious model, using the general-to-specific approach, within a expanding window over the last 8 periods remains unchanged. This implies once again that a component may be excluded if it is not statistically significant at 10%.

3.5 Method 4: weighted PC

Dias et al. (2008) suggested that instead of using the standard principal components methodology, the forecasting model could include a synthetic indicator that uses a particular weighting scheme of all components. This will define the next participant in the fishing contest, henceforth denominated "Method 4". In particular, the authors suggested that the n^{th}-principal component (PC_n) should be weighted by ω_n, defined as:

$$\omega_n = \left(\frac{\lambda_n}{\lambda_1}\right) cov(PC_n, y_{t+h}) \tag{4}$$

where λ_n is the n^{th}-eigenvalue associated to the n^{th}-eigenvector with unit length and $cov(PC_n, y_{t+h})$ is the covariance between PC_n and y_{t+h}. The intuition behind ω_n, which is a combination between two forces at work, is rather straightforward. If λ_n is very high, then the associated PC_n is capturing a significant percentage of the total variance present in the survey database and therefore should receive a high weight. However, if the covariance between PC_n and the variable of interest is negligible, then its weight should be low. It's the combination of these forces - alignment with the directions of the common movement of all variables present in the survey database and alignment with the variable of interest - that defines it's effective weight. Note also that to implement direct forecasts, there will exist a different ω_n for each h. However, to simplify the notation, this additional subscript was omitted from (4). Assuming that N principal components exist, the synthetic indicator is defined as:

$$F^{\star} = \sum_{n=1}^{N} \omega_n PC_n \tag{5}$$

3.6 Method 5: correlation-oriented PC

The principal components methodology can be applied on any second-moment matrix of the initial information set. Choosing the correlation matrix, as in Methods 2 to 4, instead of the original variance-covariance matrix involves a definitive, but arbitrary, decision to make all variables "equally important".[12] The next participant of the fishing contest assumes that the survey indicators are not equally important and therefore each survey indicator should have a different weight. Once again, given that the goal is to find adequate regressors, these weights should somehow be linked with the ability of projecting the variable of interest. With this problem, instead of finding summary indicators of the original database, we suggest finding summary indicators of $(Z\Theta)$, which implies that the components are extracted from $(NT)^{-1}(Z\Theta)'(Z\Theta)$. The eigenvectors of this problem are in general not equal to those of the standard approach. The series included in $(Z\Theta)$ will continue to have zero mean, as in the standard approach, but no longer unit variance. The variance is now dependent on Θ. Note that the standard principal components are a special case of this method, given that it can be obtained when $\Theta \equiv I$, the identity matrix.

This participant in the fishing contest opens up a possibility that has an infinite number of alternatives. Any set of weights is potentially usable. One could set some elements of the main diagonal of Θ to zero if some series should be in fact excluded from the computation of the components, or to a very high number in comparison with the other elements of the main diagonal of Θ if their importance should be clearly above the others. The current participant, henceforth denominated "Method 5", explores one single weighting scheme: each element of the main diagonal of Θ is defined as the least square coefficient coming from an univariate regression between each (standardized) $x^{\star}_{i|t}$ variable and the (standardized) variable of interest. Therefore, the original standardized survey variable $x^{\star}_{i|t}$ is replaced by $\beta_{x^{\star}_{i|t}} \times x^{\star}_{i|t}$, which may be seen as their "univariate contribution" to the projection of the variable of interest. $\beta_{x^{\star}_{i|t}}$ is the least square coefficient of the regression of y on $x^{\star}_{i|t}$. The higher this coefficient, the higher the importance of this particular survey data for the computation of the components defined in (**??**). Note that (i) if all $\beta_{x^{\star}_{i|t}}$ were identical, then there would be no qualitative difference against the standard principal components methodology; but (ii) if all $\beta_{x^{\star}_{i|t}}$ are different, which is the current situation, the survey indicators do not arrive anymore on an equal footing and instead receive different weights that are dependent on y_t and on h (the number of steps-ahead forecast). Given that all $x^{\star}_{i|t}$ variables are previously standardized, this procedure is effectively

[12] See Jackson (1991), Chatfield and Collins (1996) and Jollife (2002).

weighting the variables by their correlation with the variables of interest.[13] With this procedure, the different weights are expected to guide the components towards the variable of interest. This feature has been used to name this participant in the fishing contest: “correlation-oriented PC”.

3.7 Method 6: partial least squares

By combining features from principal components and standard OLS regressions, the Partial Least Squares (PLS) regression emerges as an alternative method to compute adequate regressors for forecasting purposes.[14] The variant of PLS used herein is such that the dependent variable is only one (this has been named in the literature as PLS1), which implies that the components of agiven matrix X with exogenous variables will be extracted for each variables of interest and for each step-ahead. More precisely, the goal will be to predict the specific (standardized) dependent variable y from a database of (standardized) $x^{\star}_{i|t}$ variables, while preserving a well-defined structure. This participant in the fishing contest, henceforth denominated “Method 6”, shares with the principal components methodology the well-defined structure that it constructs orthogonal components from the survey database. However, whereas the principal components are defined such that they only explain the variance of the survey data, the components produced by the PLS technique are conditional on the variable of interest. In addition, it should be clarified that when PLS method includes several orthogonal components, the survey data was used to produce a single time series (the $\hat{y}$ produced by the PLS regression).

4 Empirical evidence

This section assesses the out-of sample accuracy of the methodologies introduced in Section 3. This empirical evidence is derived from the database of totals and from the database of sectors, as defined in Section 2. The dependent variables were presented in Table 2 and are analysed both in quarter-on-quarter and in year-on-year terms.

As mentioned in Section 3, all models are estimated using a general-to-specific approach. After neglecting all variables which are not significant at 10%, the autoregressive models and the participants in the fishing contest are used to produce out-of-sample direct forecasts. Appendices B and C are derived from this output but contain only the results for the models with higher forecasting accuracy, defined by the lowest out-of-sample RMSE.[15] The initial rows of each table included in Appendices B and C contain the minimum RMSE that was obtained by using equation (1), i.e models solely based on autoregressive terms. These RMSE are in absolute terms for all steps ahead. The comparison between these figures and the outcome of the various models

[13]Another way to use OLS to produce different Θ is to use a multivariate environment. However, this implies that multivariate regression can be implemented, which is not the case herein, given that the number of variables is higher then the number of observations.

[14]The multivariate PLS methodology is briefly reviewed in appendix A.

[15]All the remaining results are available upon request.

based on equation (2), i.e without autoregressive terms, is reported in Appendix B. The comparison using equation (3), i.e with regressors derived from the combination of the information of the survey database and autoregressive terms, is reported in Appendix C. Both appendices have identical structures. The rows that make the comparison for all methodologies and for $h = 1, 2, 3, 4$ are in relative terms, where a value higher/lower than 1 indicates higher/lower RMSE against the best model solely based on autoregressive terms. If the figure is below 1, then the survey data contains valuable information for forecasting purposes that is not present in the autoregressive process. This situation is highlighted in bold on all the tables and the lowest relative RMSE is highlighted with a shaded area. The average of all minimum RMSE, across $h = 1, 2, 3, 4$, is also reported, and this may be used as an indication regarding the forecast accuracy of each participant for all forecasting horizons. In addition, the situations highlighted with an asterisk ($*$) indicate that, according to the Diebold-Mariano test (see Diebold and Mariano (1995)), the RMSE of that participant is statically different from the one of the benchmark autoregressive process.

The empirical results for the case when the dependent variables are expressed in quarter-on-quarter terms and the models are described by equation (2) are presented in tables 3 and 4 of Appendix B, for the database of totals and of sectors, respectively. For both databases, the best AR model has usually a relative low order, being often composed exclusively of a constant. In the context of models solely based on autoregressive terms, as defined in Section 3, the fact that this simple formulation is found to have the best forecasting performance may be in part related with the specificity of the out-of-sample data, in which periods of positive growth are in most cases followed by periods of negative growth. Among the participants in the fishing contest, bridge models and PLS forecasts are the most accurate against the AR process in both databases, given that they can produce the lowest relative RMSE. The best specification for these models includes in general a relatively large initial number of components, usually between 3 and 5. The remaining methods are many times unable to improve on the naïve AR benchmark, even when an increasing number of regressors is considered in the models' initial specifications. This result, which is based on the number of times that the relative RMSE is higher than 1, is particularly noticeable in the case of the totals database and of the private GFCF and private housing GFCF. Regarding the differences across databases, while PLS seems better for the totals database, in the sense that it is the best model at most horizons for a given dependent variable, bridge models are consistently better for the database of sectors. This database leads also to a relatively broad-based reduction in the minimum RMSE in comparison with the results for each fishing contest participant and forecasting horizon based on the database of totals. Moreover, there are more participants with a lower than 1 relative RMSE. This indicates that using a richer information environment can lead to gains in terms of forecasting performance, and suggests that the information from a given sector may be more targeted to forecast a given type of quarterly GFCF growth.

When the dependent variable is expressed in yearly terms (tables 5 and 6 of appendix B), the performance of all methods is in general also improved by the use of information from the database of sectors, and many participants start to depict lower than 1 relative RMSE. The best-performing relative out-of-sample autoregressive model have in general either one or four lags in their initial specification. In the latter case, this is the maximum number of lags allowed in equation (1). Summing up, in the case of the database of totals, the best methods from the out-of-sample forecasting perspective are methods 1 (bridge model) and 2 (standard PC), and, in the case of construction GFCF, method 3 (targeted PC). In the case of the database of sectors, the best methods are scattered across methods 2 , 3 and 5 (correlation-oriented PC). Therefore, the empirical evidence using the yearly changes shows that the methods that use a summary of the whole database outperform the AR models, which was many times not the case when the dependent variables were in quarterly rates. Given that quarterly rates of change are more volatile (see Section 2), the idiosyncratic information required to forecast them might be more easily provided by specific dataseries. On the other hand, given that the dynamics of the yearly rates of change are smoother, they may be more closely related to an aggregate measure of all surveys. When the participant is a bridge model, although its relative RMSE is not systematically the lowest in the database of sectors, as it was in the database of totals, it continues to depict lower than 1 figures.

The empirical evidence on models with autoregressive terms and information derived from survey data, in line with the initial specification of equation (3), are presented in tables 7 to 10 of Appendix C. The results show that the inclusion of AR terms implies, in general, an improvement in the out-of-sample forecasting accuracy, although this gain is more evident when the model forecasts year-on-year rates of change, possibly because the best specification found for AR models for quarter-on-quarter rates of change was often a constant, as previously mentioned. The reduction in the relative RMSE is more relevant for the database of totals than for the database of sectors. In many cases, particularly for year-on-year forecasts, the equation that led to a reduction in the RMSE includes less survey dataseries, implying that the loss in degrees of freedom is relatively contained. The inclusion of AR terms in the models usually does not change the previous conclusions regarding which are the best performing models for each GFCF component and at each horizon, with the exception of some models that rely on information from the database of sectors for year-on-year forecasts (table 10). In particular, method 3 (targeted PC) shows the lowest RMSEs along with method 2 (standard PC).

Although in many cases the differences between the fishing participants and the corresponding autoregressive benchmark model are found to be statistically significant at a 10% significance level, according to the Diebold-Mariano test, this result is more frequent for the case of quarter-on-quarter forecasts. However, it should be mentioned that these results are conditioned by the small size of the out-of-sample period, which has only 8 observations.

As concerns methods that comprise a summary of the whole survey database, method 2, which aims at explaining the variance of the survey database, is in general outperformed by one of the methods 3 to 6, in the case of quarter-on-quarter forecasts. By taking into account the correlation between the survey data and the dependent variable, these latter models emerge more appropriate for forecasting purposes. On the contrary, when the dependent variables are in year-on-year terms, the standard principal components appear to be more accurate in forecasting, although all methods depict in general lower then 1 relative RMSEs.

Finally, the results across all participants, modeling schemes or forecasting horizons indicate that bridge models are in many cases considered, if not the best method in relative terms, a method that produces a large percentage of lower than 1 relative RMSEs, which is somewhat striking due to their simplicity. This indicates that particular survey dataseries do seem to possess non-negligible leading characteristics, which are isolated in the case of bridge models. In the case of the remaining methods, which try to summarize all or a large part of the information in the database, those leading characteristics may somehow be blurred.

Actual survey dataseries that are included in bridge models with autoregressive terms estimated at each forecasting horizon and for each component are presented in Tables 11 to 14 of Appendix D.[16] The selected survey dataseries are the ones that are used to forecast the last out-of-sample quarter, 2007Q4. In general, questions related to the retail trade survey, and, in the case of the database of sectors, to food and beverages, are frequently included in the specification of the best out-of-sample performing model in the shorter forecasting horizons, that is, for current and one-quarter ahead forecasts. The likely high turnover in these sectors seems to bring along a higher accuracy in the measurement of short-term changes in the economic environment, being more correlated with the volatile changes in GFCF.

For longer forecast horizons, the questions to production and/or export expectations seem to play a relevant role in the equations, either referring to the overall industry, in the case of the database of totals, or to the industries of investment goods and durable consumption goods, in the case of the database of sectors. It is also worth mentioning that the questions relating to factors limiting the production often play a role, particulary those pertaining shortage of equipment and labour. The decision by firms to change their labour force and stock of capital is influenced by either substitution effects (investment may increase the productivity of the capital stock, reducing the need to increase the labour force), or complementarity effects (as new workers may require capital goods to work with). In the present work, the regression coefficients on the questions related to labour as a limiting factor to production are always negative, indicating that the substitution effects seem to be dominant. Finally, questions related to construction do not seem to appear as explanatory variables for the GFCF components as often as it would be expectable, in particular for public, private housing or construction GFCF. However, in the cases

[16]The corresponding results for the models without AR terms, available on request, show no significant changes in the composition of the bridge models, apart from, in general, a reduction in the number of survey dataseries considered in the best final specification. The only exception is in the case of year-on-year forecasts from the database of sectors. In fact, when autoregressive component is added to this model, the best specification of the model in out-of-sample terms includes additional survey dataseries, which in the shorter forecasting horizons relate essencially to the retail and food and beverages surveys.

of tables 11 and 14, they seem to be relevant at some horizons for explaining these variables.

5 Conclusions

The main objective of this paper is to assess the usefulness of business surveys as a source of information for investment developments in Portugal. The analysis is based on two databases of quarterly survey data. The first, denominated "database of totals", focuses exclusively on aggregates for whole sectors. The second, denominated "database of sectors", breaks down the industry and the construction surveys into several subsectors. The investment variables were assessed in quarter-on-quarter and year-on-year terms and refer to Total GFCF, public, private, the later divided further into residential and business, and, finally, construction and Total GFCF excluding construction.

The analysis was implemented on the basis of three approaches. The first approach was to estimate autoregressive processes for all variables of interest, while the second step was to estimate models that only use specific survey data or some sort of summary indicators of the survey database. Finally, each autoregressive processes was augmented with the information that is solely derived from the survey data.

The predictive power of business surveys was implemented by what was named a "fishing contest", where the "participants" are bridge models, models based on principal components (derived from standard and non-standard methods), and models built with the outcome coming from partial least squares regressions. All models were designed to produce 1 to 4 out-of-sample direct forecast and a general-to-specific approach was followed. This criteria ensured both in-sample fit and parsimonious features. Instead of choosing one single model to implement a particular direct forecast, dependent, for instance, on an information criteria, all initial specifications were evaluated assuming that they did not have, *ex ante*, different degrees of importance for forecasting purposes. This option has the likely advantage that the optimality of an in-sample fit may not be matched by an equivalent performance in terms of out-of-sample accuracy. However, it does shift the evaluation of the "participants" towards a high dependency on the out-of-sample period that is used to compute the forecasting errors, which was fixed between 2006Q1 and 2007Q4.

The results of this analysis imply that in general, there is always some participant in the fishing contest that produces a lower out-of-sample RMSE than the one associated with single autoregressive processes. This conclusion is valid for both quarter-on-quarter and year-on-year variables, and as well as for the two databases of surveys. Therefore, there seems to exist useful information for forecasting purposes in the survey data that is not already inherent to the dependent variable itself. In addition, the relative RMSE of the majority of methods show a general tendency to decrease when they are augmented with autoregressive processes, which implies that the surveys, notwithstanding their diversity, are not able to capture all the information contained in the past values of the variables. This is particularly the case when the dependent variables are expressed in year-on-year changes.

For the purpose of forecasting total GFCF and its components, the use of the database of sectors yields in general better results than the database of totals. Although the emprical evidence allows to conclude that no method substantially and systematically outperforms others for each forecasting horizon, or dependent variable, bridge models appear as the best in several cases. In this context, and conditioned on the specific nature of the dependent variable, the information provided by a few particular survey dataseries does seem to possess leading characteristics, which may be blurred in the case of the remaining methods, which try to summarize all or a large part of the information in the database. The composition of bridge models suggests that the retail trade and food and beverage industry surveys are frequently useful for forecasting investment in the short run, while the questions related to production expectations and equipment and labour as a main factor limiting production, particulary in the consumption and investment goods industries play a relevant role in the longer run.

Regarding future lines of research, one may perhaps investigate the predictive power of matrices with lower dimensions. The inclusion of exogenous variables that are not very informative about the dependent variable and the potential reduction in the forecasting ability of methods based on principal components has been mentioned by some authors (see, for instance, Bai and Ng (2007)). Nevertheless, given the relative accuracy of the database of sectors, the selection criteria should perhaps be applied on as many potential regressors as possible. Besides autoregressive terms and time series based on survey data, a natural future line of research could be to investigate whether the out-of-sample RMSE can be further lowered with additional time series. Other methods, such as those based on dynamic principal components (see Kabundi (2004)) or alternative weighting schemes of the survey data, which could include a selection based on out-of-sample criteria or forecast combination of different models, are also areas of possible future research.

References

Abdi, H. (2003), Partial least squares (PLS) regression, *in* Lewis-Beck M., Bryman, A., Futing T., ed., 'The SAGE Encyclopedia of Social Science Research Methods'.

Artís, M. and Suriñach, J. (2003), Forecasting models currently applied to indicators computed on the basis of surveys results, *Final Report, Tender ECFIN/2002/A3-01*, Universitat de Barcelona.

Bai, J. and Ng, S. (2007), 'Forecasting economic time series using targeted predictors', *forthcoming Journal of Econometrics* .

Bai, J. and Ng, S. (2008), 'Boosting difussion indices', *forthcoming* .

Banco de Portugal (2008), Quarterly series for the portuguese economy: 1997-2007, Economic Bulletin - Summer, Banco de Portugal.

Barnes, S. and Ellis, C. (2005), Indicators of short-term movements in business investment, Quarterly Bulletin - Spring, Bank of England.

Bram, J. and Ludvigson, S. (1997), Does consumer confidence forecast household expenditure?: A sentiment index horse race, Research Paper 9708, Federal Reserve Bank of New York.

Chatfield, C. and Collins, A. (1996), *Introduction to Multivariate Analysis*, Chapman and Hall.

Claveria, O., Pons, E. and Ramos, R. (2007), 'Business and consumer expectations and macroeconomic forecasts', *International Journal of Forecasting* **23**(1), 47–69.

Dias, F. C., Pinheiro, M. and Rua, A. (2008), Forecasting using targeted diffusion indexes, Working Paper 7, Banco de Portugal.

Diebold, F. X. and Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **13**(3), 253–63.

European Commission (2000), Business climate indicator for the euro area, Presentation paper, Directorate-General for Economic and Financial Affairs.

European Commission (2007), The joint harmonised eu programme of business and consumer surveys, User guide, Directorate-General for Economic and Financial Affairs.

Geladi, P. and Kowlaski, B. (1986), 'Partial least square regression: A tutorial', *Analytica Chemica Acta* **35**, 1–17.

Hansson, J., Jansson, P. and Lof, M. (2005), 'Business survey data: Do they help in forecasting GDP growth?', *International Journal of Forecasting* **21**(2), 377–389.

Jackson, J. E. (1991), *A User's Guide to Principal Components*, Wiley.

Jollife, I. (2002), *Principal Component Analysis*, Springer.

Kabundi, A. N. (2004), Estimation of economic growth in France using business survey data, IMF Working Papers 04/69, International Monetary Fund.

Larsen, J. (2001), Using surveys of investment intentions, Quarterly Bulletin - Summer, Bank of England.

Rünstler, G. and Sédillot, F. (2003), Short-term estimates of euro area real gdp by means of monthly data, Working Paper Series 276, European Central Bank.

Rosipal, R. and Kramer, N. (2006), Overview and recent advances in partial least squares, *in* Springer Berlin - Heidelberg, ed., 'Subspace, Latent Structure and Feature Selection', pp. 34–51.

Santero, T. and Westerlund, N. (1996), Confidence indicators and their relationship to changes in economic activity, Economics Department Working Papers 170, OECD.

Stock, J. and Watson, M. (1989), 'New indexes of coincident and leading economic indicators', *NBER Macroeconomic Annual* pp. 351–394.

Tashman, L. J. (2000), 'Out-of-sample tests of forecasting accuracy: an analysis and review', *International Journal of Forecasting* **16**(4), 437–450.

Appendices

A Partial least squares

The aim of PLS is to model specific linkages between two sets of observed data by means of unobserved components. It comprises several alternatives that have become popular in various fields, including chemometrics, bioinformatic, food research, medicine, pharmacology, social sciences, etc (see Rosipal and Kramer (2006) for references). PLS is used herein with the sole objective of implementing out-of-sample forecasts. Using a multivariate framework, the PLS regression can be written as:

$$Y_t = X_t \beta_{\text{PLS}} + \xi_t \tag{6}$$

where Y has M endogenous variables (the first set of observed data), X contains N exogenous variables (the second set), β_{PLS} is a $(N \times M)$ matrix with PLS coefficients and ξ_t is an error term. Both information sets have T observations and the PLS estimate of Y is given by $\hat{Y}_t = X_t \widehat{\beta}_{\text{PLS}}$. Following Rosipal and Kramer (2006), this formulation can be derived from the following structural relationships:

$$\underset{(T\times N)}{X} = \underset{(T\times p)}{S} \underset{(p\times N)}{P'} + \underset{(T\times N)}{\epsilon_{\text{X}}} \tag{7}$$

$$\underset{(T\times M)}{Y} = \underset{(T\times p)}{U} \underset{(p\times M)}{Q'} + \underset{(T\times M)}{\epsilon_{\text{Y}}} \tag{8}$$

$$\underset{(T\times p)}{U} = \underset{(T\times p)}{S} \underset{(p\times p)}{B} + \underset{(T\times p)}{\epsilon_{\text{U}}} \tag{9}$$

where S and U are matrices of p components, P and Q are matrices of loadings, B is a diagonal matrix with scalars linking the components and ϵ_i are matrices of residuals, $i=$ X, Y or U. The matrices' dimensions are in brackets. With the exception of Y and X, which are matrices with observed data, all other matrices are unknown.

Taken separately, equation (7) could be seen as a standard representation of X using orthogonal components, which could be principal components, where ϵ_X is an empty matrix if the p vectors perform an exact decomposition of X. Equation (8) has a similar structure and together with (7) build the outer structure of PLS. If equation (9) did not exist, X and Y would have no operational link in the structural relationships. Equation (9) represents the inner structure of PLS. This is nevertheless a partial least squares framework given that instead of regressing Y on X, as in the case of the standard OLS regressions, the linear link is constructed with components. The PLS method consist in directly extracting orthogonal components from X under the operational restriction that they are also relevant to "predict" Y. In particular, let s and u be components belonging to S and U, respectively, and w and c be vectors that weight the two information sets, i.e $s = Xw$ and $u = Yc$. The problem is then to find vectors w and c such that $w'w = c'c = 1$ and $s'u$ is maximal.

To derive an equation that is fully equivalent to an OLS estimation of Y, using S as orthogonal regressors, it is only necessary to note that the expression U of (9) can be replaced in (8).

$$Y = \quad SBQ' + E \tag{10}$$

where $E = (\epsilon_{\mathrm{U}} Q' + \epsilon_{\mathrm{Y}})$ is an error term. To derive an equation that takes the form of (6), equation (7) can be solved for S, for instance, after having post-multiplied it by W, which is a matrix that stacks all vectors w that were previously computed. If one replaces the outcome in (11), this leads to:

$$Y = X[W(P'W)^{-1}BQ'] + \xi \tag{11}$$

where $\xi = E + \epsilon_{\mathrm{X}}(P'W)^{-1}BQ'$ is an error term. Equation (11) defines the PLS regression where $\beta_{\mathrm{PLS}} = W(P'W)^{-1}BQ'$. The expression for β_{PLS} is not the sole possible representation. Rosipal and Kramer (2006) include alternative specifications.

The solution to the estimation of all matrices that are not directly observable can be found by using the outcome of the nonlinear iterative partial least squares algorithm (NIPALS), which is discussed in several papers, including Abdi (2003) or Rosipal and Kramer (2006). An alternative solution can be found by using adequate eigenvalue/eigenvector problems. A classic tutorial on PLS can be found in Geladi and Kowlaski (1986), which also includes several variants of NIPALS, including the one where the outcome reproduces the standard eigenvalue/eigenvector outcome of the principal components methodology. Finally, it will be assumed that $S'S = I$, where I is the identity matrix. Some variants of the PLS technique do not require S to have unit norm.

In the empirical results of this paper, (i) the estimation problem was solved with NIPALS and the program builds on the Matlab code downloaded from the webpage of Hervé Adbi (`www.utdallas.edu/~herve`), which also respects the restriction that $S'S = I$. NIPALS starts with a random initialization of the component $u = u_0$ and proceeds with the computation (up to a negligible numerical error) of all necessary vectors u, s, c and w, which are sequentially stacked in matrices U,S,C and W ; (ii) the decomposition of Y as a matrix of variables was not pursued and M was set to 1. This procedure, which has been named PLS1 in the literature, reduces the complexity of the system and produces a β_{PLS} that is a simple column vector.

B Out-of-sample results without autoregressive terms

Table 3: Out-of-sample RMSE for q-o-q forecasts: Database of totals, AR terms not included in Methods 1 to 6

Out-of-sample RMSE for q-o-q forecasts: Database of Totals, AR terms excluded from models 2 to 6

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.021	1		0.069	1		0.018	1		0.034	2		0.018	1		0.033	1		0.030	1	
	1	0.022	3		0.069	1		0.020	4		0.034	1		0.023	2		0.033	1		0.032	3	
	2	0.021	1		0.061	4		0.019	1		0.034	2		0.019	1		0.033	2		0.029	1	
	3	0.021	4		0.060	4		0.019	4		0.034	1		0.019	1		0.033	1		0.028	1	
	Average	*0.022*	4		*0.070*	1		*0.020*	1		*0.034*	1		*0.020*	1		*0.033*	1		*0.030*	1	
Method 1 -Bridge Model																						
	0	1.02	1	*	**0.71**	3	*	1.05	4	*	**0.97**	5	*	1.19	2		**0.96**	5	*	**0.90**	5	
	1	1.12	3	*	**0.83**	3	*	**0.98**	4	*	1.04	1		**0.87**	5	*	1.03	3	*	**0.81**	2	
	2	**0.88**	3	*	**0.83**	5		**0.87**	3	*	**0.86**	4		1.07	4		1.10	3	*	**0.88**	3	*
	3	1.04	2	*	1.07	4		**0.80**	4	*	**0.85**	4	*	**0.93**	4	*	**0.98**	4	*	**0.89**	2	*
	Average	1.03	3		**0.82**	3		**0.95**	4		**0.95**	4		1.02	4		1.03	5		**0.89**	2	
Method 2 - Standard PC																						
	0	1.00	4	*	**0.87**	5	*	1.08	5		1.05	1	*	1.06	2		**0.92**	4	*	**0.90**	2	
	1	**0.98**	4	*	**0.82**	5		1.01	2		1.09	1	*	**0.98**	1	*	**0.98**	4	*	**0.78**	4	
	2	1.04	3	*	**0.94**	4		**0.99**	2	*	**0.96**	2	*	**0.95**	5	*	1.03	1	*	**0.99**	2	
	3	**0.96**	2		1.05	3		**0.90**	2		**0.89**	2	*	**0.99**	2		1.06	1	*	1.02	5	*
	Average	**1.00**	2		**0.87**	4		**0.96**	2		1.03	2		**1.00**	2		1.02	1		**0.94**	2	
Method 3 - Targeted PC																						
	0	1.08	1	*	**0.75**	5	*	1.13	1		1.04	1	*	1.15	4		**0.98**	5	*	**0.85**	1	
	1	**0.94**	4	*	**0.86**	2		**0.91**	4	*	1.22	1	*	**0.84**	5	*	**0.89**	3	*	**0.92**	3	
	2	1.15	2	*	**0.89**	1		1.01	3	*	1.18	1		**0.68**	3	*	1.03	3	*	1.06	5	*
	3	1.04	3	*	**0.99**	1		**0.84**	1	*	1.02	3		**0.87**	2	*	1.07	1	*	1.09	3	*
	Average	1.07	1		**0.85**	3		1.01	1		1.15	1		**0.94**	2		1.02	1		1.02	1	
Method 4 - Weigthed PC																						
	0	1.07	1	*	**0.96**	1	*	1.15	1		1.05	1	*	1.16	1	*	**0.99**	1	*	**0.96**	1	
	1	1.06	1	*	**0.97**	1	*	1.07	1	*	1.04	1	*	**0.95**	1	*	**0.99**	1	*	1.01	1	
	2	1.09	1	*	1.11	1		1.13	1	*	1.06	1	*	1.17	1		1.03	1	*	**0.99**	1	
	3	**0.99**	1		1.09	1		1.05	1		1.04	1	*	1.04	1	*	1.04	1	*	1.04	1	*
	Average	1.04	1		**0.95**	1		1.06	1		1.05	1		1.07	1		1.01	1		**1.00**	1	
Method 5 - Correlation Oriented PC																						
	0	1.08	1	*	**0.84**	5	*	1.15	1		1.05	1	*	1.09	4		**0.96**	4	*	**0.91**	2	
	1	**0.98**	5	*	**0.89**	3	*	1.08	1	*	1.05	1	*	**0.96**	1	*	**0.95**	4	*	**0.82**	4	
	2	1.08	2	*	**0.87**	5		1.14	2	*	1.09	1		1.12	5		1.02	1	*	1.03	1	
	3	1.08	2	*	1.04	5		1.09	3	*	**0.91**	2	*	1.05	4	*	1.05	1	*	**1.00**	2	*
	Average	1.06	2		**0.84**	5		1.08	3		1.04	2		1.06	4		1.01	1		**0.94**	2	
Method 6 - PLS																						
	0	**0.94**	2	*	1.06	1	*	**0.87**	2		**0.80**	5	*	**0.92**	2		**0.97**	4	*	**0.74**	4	
	1	**0.91**	4	*	1.07	1	*	**0.87**	2	*	**0.81**	5	*	**0.75**	2		**0.87**	4	*	**0.77**	2	
	2	**0.91**	4	*	1.22	1		**0.87**	5	*	**0.85**	5		**0.91**	5	*	**0.98**	3	*	**0.88**	4	
	3	1.04	4		1.26	1		**1.00**	5	*	**0.84**	4	*	**0.98**	5	*	1.04	4	*	**0.94**	1	*
	Average	**0.95**	4		1.06	1		**0.92**	2		**0.82**	5		**0.94**	2		**0.98**	4		**0.88**	4	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

Table 4: Out-of-sample RMSE for q-o-q forecasts: Database of subsectors, AR terms not included in Methods 1 to 6

Out-of-sample RMSE for q-o-q forecasts: Database of Sectors, AR terms excluded from models 2 to 6

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.021	1		0.069	1		0.018	1		0.034	2		0.018	1		0.033	1		0.030	1	
	1	0.022	3		0.069	1		0.020	4		0.034	1		0.023	2		0.033	1		0.032	3	
	2	0.021	1		0.061	4		0.019	1		0.034	2		0.019	1		0.033	2		0.029	1	
	3	0.021	4		0.060	4		0.019	4		0.034	1		0.019	1		0.033	1		0.028	1	
	Average	*0.022*	4		*0.070*	1		*0.020*	1		*0.034*	1		*0.020*	1		*0.033*	1		*0.030*	1	
Method 1 -Bridge Model																						
	0	**0.94**	4		**0.84**	5	*	**0.97**	4		**0.86**	5		1.23	4		**0.74**	4	*	**0.79**	3	
	1	**0.60**	5	*	**0.68**	5	*	**0.82**	5	*	**0.55**	5	*	**0.62**	5	*	**0.79**	5	*	**0.69**	4	
	2	**0.74**	4	*	**0.86**	3		**0.76**	4	*	**0.90**	5		**0.68**	5	*	**0.80**	5	*	**0.90**	3	*
	3	**0.93**	3		**0.99**	3		**0.86**	3	*	**0.96**	3		**0.86**	2	*	**0.66**	5	*	**0.84**	3	*
	Average	**0.84**	5		**0.82**	5		**0.87**	4		**0.83**	5		**0.87**	5		**0.75**	5		**0.83**	3	
Method 2 - Standard PC																						
	0	**0.93**	2		**0.95**	5	*	**0.97**	2		1.03	2		**0.98**	2		**0.98**	1	*	**0.81**	2	
	1	**0.75**	4		**0.90**	4	*	**0.82**	4		1.11	2	*	**0.80**	2		**0.91**	5	*	**0.82**	2	
	2	**0.95**	2		1.06	4		**0.93**	2		**0.92**	3		**0.97**	2	*	**0.91**	4	*	1.00	1	*
	3	**0.88**	2		1.15	4		**0.88**	2		**0.84**	5		**0.89**	2	*	**1.00**	5	*	1.03	2	*
	Average	**0.90**	2		**0.94**	5		**0.89**	2		1.00	3		**0.90**	2		**0.98**	2		**0.91**	2	
Method 3 - Targeted PC																						
	0	1.09	2	*	**0.91**	4	*	1.25	5	*	**0.87**	4		1.03	2	*	**0.98**	1	*	**0.81**	1	
	1	**0.77**	3		**0.88**	2		**0.62**	4	*	1.00	2	*	**0.79**	1	*	**0.86**	4	*	**0.81**	3	
	2	**0.89**	2		1.03	4		**0.93**	1		**1.00**	1		**0.97**	1	*	1.06	1	*	**0.96**	3	
	3	**0.84**	4	*	1.01	4		**0.88**	1		**0.77**	5	*	**0.89**	1	*	1.01	5	*	1.02	2	
	Average	**0.90**	2		**0.90**	4		**0.97**	1		**0.99**	4		**0.93**	1		1.05	1		**0.90**	1	
Method 4 - Weigthed PC																						
	0	1.07	1	*	**0.99**	1	*	1.13	1		1.06	1	*	1.15	1	*	**0.98**	1	*	**0.98**	1	
	1	1.04	1	*	**0.98**	1	*	1.04	1	*	1.05	1	*	**0.92**	1	*	**0.99**	1	*	**0.98**	1	
	2	1.08	1	*	1.12	1		1.16	1	*	1.11	1	*	1.19	1	*	1.02	1	*	**0.94**	1	*
	3	1.12	1	*	1.11	1		1.21	1	*	1.12	1	*	1.22	1	*	1.05	1	*	**0.95**	1	*
	Average	1.06	1		**0.96**	1		1.09	1		1.09	1		1.10	1		1.01	1		**0.96**	1	
Method 5 - Correlation Oriented PC																						
	0	**0.95**	2		**0.99**	1	*	**0.99**	2		1.04	2		1.03	2		**0.98**	4	*	**0.89**	2	
	1	**0.92**	2		**0.97**	5	*	**0.92**	2		1.02	2		**0.84**	2		**0.92**	5	*	**0.85**	2	
	2	**0.99**	2		1.10	2		1.12	2	*	**0.97**	4		1.12	2		**0.98**	4	*	**0.95**	1	*
	3	**0.89**	2		1.14	3		**0.94**	2		**0.86**	4	*	**0.91**	2		**0.90**	2	*	**0.96**	1	*
	Average	**0.93**	2		**0.97**	4		**0.95**	2		**0.98**	4		**0.96**	2		**0.97**	2		**0.93**	2	
Method 6 - PLS																						
	0	**0.88**	2	*	1.09	1	*	**0.86**	2	*	**0.87**	5	*	**0.89**	2		**0.98**	2	*	**0.80**	2	
	1	**0.92**	2	*	1.11	1	*	**0.83**	2	*	**0.90**	4	*	**0.77**	2		1.01	4	*	**0.76**	2	
	2	1.00	2	*	1.28	1		**0.92**	2	*	**0.89**	4	*	**0.94**	2	*	**0.97**	4	*	**0.87**	1	*
	3	**0.96**	4	*	1.40	1	*	**0.94**	2	*	**0.86**	2	*	**0.93**	4	*	**0.95**	4	*	**0.92**	1	*
	Average	**0.98**	2		1.12	1		**0.86**	2		**0.90**	4		**0.90**	2		**0.98**	4		**0.85**	2	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

Table 5: Out-of-sample RMSE for y-o-y forecasts: Database of totals, AR terms not included in Methods 1 to 6

Out-of-sample RMSE for y-o-y forecasts: Database of Totals, AR terms excluded from models 2 to 6

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.026	2		0.094	3		0.022	1		0.030	4		0.022	1		0.033	4		0.029	1	
	1	0.035	4		0.091	3		0.033	4		0.040	2		0.034	4		0.045	3		0.041	1	
	2	0.041	2		0.131	2		0.035	1		0.036	1		0.036	1		0.054	4		0.043	1	
	3	0.040	2		0.143	4		0.032	2		0.034	1		0.034	2		0.055	4		0.044	1	
	Average	*0.036*	3		*0.122*	3		*0.031*	2		*0.036*	1		*0.032*	2		*0.048*	4		*0.039*	1	
Method 1 -Bridge Model																						
	0	1.24	4	*	1.42	3	*	**0.94**	4		**0.87**	3		1.12	1	*	1.20	4		1.13	1	
	1	1.21	2	*	1.33	5		1.15	1		**0.72**	3		1.34	3		**0.78**	3		1.06	5	
	2	**0.98**	2		**0.78**	5		**0.82**	2	*	1.41	4		**0.91**	4	*	**0.72**	3	*	**0.68**	2	*
	3	**0.72**	5	*	**0.68**	3	*	**0.79**	4	*	1.71	5		**0.66**	3	*	1.04	3	*	**0.75**	1	*
	Average	1.05	2		**0.97**	5		1.04	4		1.24	3		1.03	3		**0.90**	3		**0.88**	5	
Method 2 - Standard PC																						
	0	1.16	2		1.30	5	*	1.30	2		1.31	3		1.18	2		1.04	5		1.12	2	
	1	**0.79**	2		1.51	4		**0.75**	2		**1.00**	2		**0.67**	2		**0.86**	1		**0.73**	2	
	2	**0.71**	2		**0.96**	4		**0.79**	2		**0.95**	2		**0.79**	2		**0.75**	5		**0.70**	2	
	3	**0.80**	2		**0.76**	4	*	**0.81**	2	*	**0.97**	2		**0.85**	5	*	**0.79**	4	*	**0.86**	2	*
	Average	**0.83**	2		1.02	4		**0.87**	2		1.04	2		**0.86**	2		**0.87**	5		**0.83**	2	
Method 3 - Targeted PC																						
	0	1.21	5		1.30	5	*	1.43	3	*	1.47	1		1.41	5		1.24	5		1.40	5	*
	1	1.09	1		1.42	3		1.14	1		1.09	2		1.10	1		**0.76**	2		**0.94**	1	*
	2	**0.96**	3	*	**1.00**	4		1.03	3	*	1.07	4		1.15	2		**0.70**	2		**0.85**	2	
	3	**0.87**	3	*	**0.76**	3		**0.81**	2	*	**0.88**	4	*	**0.71**	2	*	**0.65**	2	*	**0.86**	1	*
	Average	1.06	4		1.03	4		1.17	3		1.20	1		1.13	1		**0.81**	2		1.07	1	
Method 4 - Weigthed PC																						
	0	1.39	1		1.54	1	*	1.62	1		1.46	1		1.53	1		1.25	1		1.63	1	*
	1	1.08	1		1.60	1	*	1.12	1		1.11	1		1.09	1		**0.85**	1		1.26	1	
	2	1.02	1	*	1.11	1		1.23	1	*	1.30	1	*	1.22	1		**0.79**	1		1.22	1	
	3	1.11	1	*	1.04	1	*	1.39	1	*	1.33	1	*	1.33	1		**0.81**	1	*	1.10	1	*
	Average	1.11	1		1.19	1		1.30	1		1.26	1		1.26	1		**0.88**	1		1.27	1	
Method 5 - Correlation Oriented PC																						
	0	1.31	5		1.44	4	*	1.36	2		1.39	1		1.18	2		1.27	1		1.15	2	
	1	**0.84**	2		1.36	5		**0.85**	2		**0.99**	2		**0.82**	2		**0.87**	1		**0.84**	2	
	2	**0.81**	2		1.03	4		**0.95**	2		1.07	2		1.00	2		**0.75**	5	*	**0.88**	2	
	3	**0.91**	2		**0.83**	4	*	1.04	2		**0.97**	2		1.19	2		**0.73**	5	*	**0.92**	2	*
	Average	**0.93**	2		1.06	5		1.01	2		1.07	2		1.03	2		**0.88**	1		**0.93**	2	
Method 6 - PLS																						
	0	1.61	4		1.77	1		1.44	2		1.30	2		1.34	4		1.71	1		**0.98**	4	*
	1	1.22	2		1.85	1		**0.82**	2		**0.89**	2		**0.83**	4		1.27	1	*	**0.77**	2	
	2	1.06	4	*	1.30	1		**0.84**	2		1.14	4		**0.76**	2		1.16	1	*	**0.77**	2	
	3	1.17	2		1.21	1	*	1.04	2	*	1.22	4	*	**0.98**	2	*	1.17	1	*	**0.79**	4	*
	Average	1.23	4		1.38	1		**0.99**	2		1.18	4		**0.94**	2		1.27	1		**0.85**	4	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

Table 6: Out-of-sample RMSE for y-o-y forecasts: Database of sector, AR terms not included in Methods 1 to 6

Out-of-sample RMSE for y-o-y forecasts: Database of Sectors, AR terms excluded from models 2 to 6

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.026	2		0.094	3		0.022	1		0.030	4		0.022	1		0.033	4		0.029	1	
	1	0.035	4		0.091	3		0.033	4		0.040	2		0.034	4		0.045	3		0.041	1	
	2	0.041	2		0.131	2		0.035	1		0.036	1		0.036	1		0.054	4		0.043	1	
	3	0.040	2		0.143	4		0.032	2		0.034	1		0.034	2		0.055	4		0.044	1	
	Average	*0.036*	3		*0.122*	3		*0.031*	2		*0.036*	1		*0.032*	2		*0.048*	4		*0.039*	1	
Method 1 -Bridge Model																						
	0	**0.93**	3	*	1.11	2	*	**1.00**	5		**0.87**	4		1.01	4		1.24	2	*	1.22	4	
	1	**0.73**	5	*	**0.85**	1	*	**0.70**	4		1.15	5	*	**0.83**	3		**0.65**	3	*	**0.74**	1	
	2	**0.61**	1	*	**0.70**	2		**0.75**	1	*	1.03	1	*	**0.68**	5	*	1.18	2	*	**0.75**	1	*
	3	**0.58**	1	*	1.03	3	*	**0.88**	5		1.28	2		**0.88**	2	*	1.05	1	*	**0.65**	1	*
	Average	**0.74**	5		**0.91**	1		**0.82**	4		1.21	1		**0.89**	5		1.09	1		**0.81**	1	
Method 2 - Standard PC																						
	0	**0.73**	2		1.04	1	*	**0.71**	5		1.23	2	*	**0.73**	5		1.07	2	*	**0.93**	5	*
	1	**0.45**	4		1.10	1	*	**0.57**	5		**0.86**	2	*	**0.58**	5		**0.68**	4	*	**0.73**	1	*
	2	**0.55**	2		**0.74**	4		**0.62**	4		**1.00**	3	*	**0.59**	4		**0.77**	2		**0.63**	1	*
	3	**0.55**	2		**0.74**	1	*	**0.69**	2	*	1.17	3	*	**0.69**	2	*	**0.79**	2	*	**0.60**	1	*
	Average	**0.56**	2		**0.83**	1		**0.69**	5		1.03	3		**0.64**	5		**0.81**	2		**0.73**	1	
Method 3 - Targeted PC																						
	0	**0.89**	3		1.13	1	*	**0.92**	3		1.59	3		**0.88**	5		1.13	5		**0.97**	4	*
	1	**0.44**	3		1.15	1		**0.57**	4		**0.88**	5		**0.64**	4		1.00	3		**0.75**	5	
	2	**0.56**	4		**0.65**	1		**0.55**	3		1.39	2		**0.63**	4		**0.78**	3		**0.70**	5	
	3	**0.58**	2	*	**0.58**	1		**0.67**	3		1.06	1	*	**0.59**	5		**0.89**	4	*	**0.66**	1	
	Average	**0.63**	3		**0.77**	1		**0.68**	3		1.30	3		**0.69**	5		**0.93**	5		**0.80**	4	
Method 4 - Weigthed PC																						
	0	1.19	1		1.05	1	*	1.26	1		2.11	1		1.06	1		1.46	1		1.05	1	*
	1	**0.92**	1		1.11	1	*	**0.85**	1		1.58	1		**0.71**	1		1.10	1		**0.76**	1	
	2	**0.90**	1	*	**0.78**	1		**0.95**	1	*	1.92	1	*	**0.75**	1		1.05	1	*	**0.86**	1	
	3	**0.98**	1	*	**0.76**	1	*	1.08	1	*	2.09	1	*	**0.85**	1		1.07	1	*	**0.87**	1	*
	Average	**0.97**	1		**0.84**	1		1.00	1		1.87	1		**0.81**	1		1.12	1		**0.87**	1	
Method 5 - Correlation Oriented PC																						
	0	**0.70**	2		1.09	1	*	**0.68**	5		1.17	2	*	**0.95**	5	*	1.11	4	*	**0.93**	5	*
	1	**0.44**	2		1.13	1		**0.54**	5		**0.83**	2	*	**0.58**	4		**0.67**	4	*	**0.73**	4	
	2	**0.60**	2		**0.79**	1		**0.69**	3	*	1.14	3	*	**0.63**	4		**0.75**	4	*	**0.74**	5	
	3	**0.52**	3		**0.76**	1	*	**0.69**	2	*	1.30	2		**0.60**	4		**0.92**	2	*	**0.77**	3	*
	Average	**0.56**	2		**0.85**	1		**0.66**	5		1.08	2		**0.66**	5		**0.84**	4		**0.78**	5	
Method 6 - PLS																						
	0	1.70	2	*	1.83	1		1.35	2		1.46	2	*	1.42	2		1.80	1		1.22	2	
	1	1.25	2	*	1.92	1		**0.85**	2		1.01	4		**0.87**	2		1.31	1	*	**0.79**	2	
	2	1.03	2	*	1.33	1		**0.82**	2	*	1.16	2	*	**0.84**	2		1.12	1	*	**0.77**	2	
	3	1.14	4	*	1.24	1	*	**0.94**	2	*	1.25	2		**0.89**	2	*	1.14	1	*	**0.71**	1	*
	Average	1.23	2		1.43	1		**0.95**	2		1.20	2		**0.95**	2		1.27	1		**0.87**	2	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

C Out-of-sample results with autoregressive terms

Table 7: Out-of-sample RMSE for q-o-q forecasts: Database of totals, AR terms included in all models

Out-of-sample RMSE for q-o-q forecasts: Database of Totals, AR terms included in all models

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.021	1		0.069	1		0.018	1		0.034	2		0.018	1		0.033	1		0.030	1	
	1	0.022	3		0.069	1		0.020	4		0.034	1		0.023	2		0.033	1		0.032	3	
	2	0.021	1		0.061	4		0.019	1		0.034	2		0.019	1		0.033	2		0.029	1	
	3	0.021	4		0.060	4		0.019	4		0.034	1		0.019	1		0.033	1		0.028	1	
	Average	*0.022*	4		*0.070*	1		*0.020*	1		*0.034*	1		*0.020*	1		*0.033*	1		*0.030*	1	
Method 1 -Bridge Model																						
	0	1.06	2	*	**0.89**	1	*	1.14	1		1.00	1	*	1.10	1	*	**0.94**	5	*	**0.93**	4	
	1	**0.96**	1	*	**0.90**	1	*	**0.98**	1	*	**0.88**	1	*	**0.93**	5		1.04	2	*	**0.86**	1	*
	2	**0.73**	2	*	**0.82**	3	*	**0.44**	1	*	**0.88**	1		**0.91**	2	*	1.07	1	*	**0.88**	3	*
	3	**0.59**	1	*	**0.80**	2		**0.38**	1	*	**0.80**	3	*	**0.91**	1		**0.95**	5	*	**0.82**	1	*
	Average	**0.89**	1		**0.83**	1		**0.71**	1		**0.91**	1		1.01	2		1.03	5		**0.90**	1	
Method 2 - Standard PC																						
	0	1.06	2	*	**0.96**	5	*	1.05	2		1.18	1	*	1.17	2		**0.92**	4	*	**0.92**	2	*
	1	**0.97**	4	*	**0.94**	4	*	1.01	2		1.18	1	*	**0.98**	2		**0.96**	4	*	**0.78**	4	
	2	1.05	3	*	**0.66**	4		**0.99**	2	*	1.13	1		**0.97**	4	*	**0.80**	1	*	**0.99**	2	
	3	**0.91**	2		**0.79**	3	*	**0.86**	2		1.00	4		1.03	2		**0.88**	1	*	1.02	5	*
	Average	**0.99**	2		**0.80**	4		**0.94**	2		1.12	1		1.03	2		**0.92**	1		**0.95**	2	
Method 3 - Targeted PC																						
	0	1.08	1	*	**0.71**	5		1.13	1		1.10	1	*	1.26	1	*	**0.99**	1	*	**0.97**	1	
	1	**0.94**	4	*	**0.97**	4		**0.91**	4	*	1.27	1	*	**0.87**	5	*	**0.88**	3	*	**0.93**	3	
	2	1.15	2	*	**0.68**	5		1.01	3	*	1.18	1		**0.66**	3	*	**0.89**	1	*	1.06	5	*
	3	**0.86**	3		**0.78**	1	*	**0.84**	1	*	1.02	3		**0.99**	5	*	1.06	1	*	1.07	3	*
	Average	1.10	3		**0.82**	4		1.03	1		1.18	1		1.02	2		**0.97**	1		1.07	2	
Method 4 - Weigthed PC																						
	0	1.07	1	*	1.08	1	*	1.15	1		1.17	1	*	1.16	1	*	**0.99**	1	*	**0.96**	1	
	1	1.06	1	*	1.08	1	*	1.07	1	*	1.18	1	*	**0.96**	1	*	**0.99**	1	*	1.03	1	*
	2	1.05	1	*	**0.94**	1		1.13	1	*	1.19	1	*	1.17	1		**0.80**	1	*	1.02	1	
	3	**0.96**	1		**0.93**	1	*	**0.99**	1		1.16	1	*	1.10	1		**0.86**	1	*	1.03	1	*
	Average	1.02	1		**0.93**	1		1.04	1		1.17	1		1.08	1		**0.91**	1		1.01	1	
Method 5 - Correlation Oriented PC																						
	0	1.08	1	*	**0.92**	5	*	1.15	1		1.14	4	*	1.24	2		**0.96**	4	*	**0.93**	2	
	1	1.01	5	*	**0.95**	4	*	1.08	1	*	1.19	1	*	**0.99**	2	*	**0.94**	4	*	**0.82**	2	
	2	1.02	2	*	**0.83**	4	*	1.23	1	*	1.26	3		1.12	5		**0.81**	1	*	1.04	2	*
	3	**0.95**	3		**0.76**	3		**0.93**	5		1.12	2		1.04	4	*	**0.83**	1	*	1.02	5	*
	Average	1.01	3		**0.86**	5		1.07	1		1.20	2		1.11	2		**0.91**	1		**0.95**	2	
Method 6 - PLS																						
	0	**0.95**	2	*	1.07	1	*	**0.91**	2	*	**0.85**	5	*	**0.99**	2		**0.95**	4	*	**0.75**	4	
	1	**0.91**	4	*	**0.99**	1	*	**0.87**	2	*	**0.82**	5	*	**0.74**	2		**0.87**	4	*	**0.77**	2	
	2	**0.90**	4	*	1.25	1		**0.87**	5	*	**0.85**	5		**0.90**	5	*	**0.98**	3	*	**0.88**	4	
	3	1.05	4		1.29	1		**1.00**	5	*	**0.87**	4	*	**0.97**	5	*	1.04	4	*	**0.90**	1	*
	Average	**0.95**	4		1.05	1		**0.92**	2		**0.85**	5		**0.94**	2		**0.97**	4		**0.88**	4	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

Table 8: Out-of-sample RMSE for q-o-q forecasts: Database of subsectors, AR terms included in all models

Out-of-sample RMSE for q-o-q forecasts: Database of Sectors, AR terms included in all models

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.021	1		0.069	1		0.018	1		0.034	2		0.018	1		0.033	1		0.030	1	
	1	0.022	3		0.069	1		0.020	4		0.034	1		0.023	2		0.033	1		0.032	3	
	2	0.021	1		0.061	4		0.019	1		0.034	2		0.019	1		0.033	2		0.029	1	
	3	0.021	4		0.060	4		0.019	4		0.034	1		0.019	1		0.033	1		0.028	1	
	Average	*0.022*	4		*0.070*	1		*0.020*	1		*0.034*	1		*0.020*	1		*0.033*	1		*0.030*	1	
Method 1 -Bridge Model																						
	0	1.03	4	*	**0.79**	3	*	1.14	1		**0.91**	1		1.02	1		**0.57**	4	*	**0.80**	2	
	1	**0.57**	3	*	**0.82**	5	*	**0.79**	2	*	**0.61**	5	*	**0.51**	1	*	**0.68**	2	*	**0.72**	4	
	2	**0.73**	5	*	**0.69**	3		**0.76**	4	*	**0.90**	5		**0.67**	5	*	**0.70**	2	*	**0.90**	1	*
	3	**0.81**	1	*	**0.72**	1	*	**0.59**	1	*	**0.80**	1	*	**0.73**	3	*	**0.71**	5	*	**0.78**	1	*
	Average	**0.79**	2		**0.78**	5		**0.84**	2		**0.87**	5		**0.79**	1		**0.73**	4		**0.81**	1	
Method 2 - Standard PC																						
	0	**0.93**	2		1.01	5	*	1.05	2	*	1.15	2	*	1.06	2		**0.98**	4	*	**0.81**	5	
	1	**0.80**	4		**0.90**	4	*	**0.89**	4		1.05	4	*	**0.80**	2		**0.91**	5	*	**0.81**	5	
	2	**0.93**	2	*	**0.95**	1		1.03	1	*	1.01	1	*	1.04	1	*	**0.80**	4	*	1.02	1	*
	3	**0.94**	5		**0.94**	4		**0.88**	2		**0.84**	5		**0.95**	2		**0.81**	2		1.01	1	*
	Average	**0.91**	2		**0.90**	4		**0.96**	2		1.06	5		**0.97**	2		**0.90**	2		**0.94**	2	
Method 3 - Targeted PC																						
	0	1.09	2	*	**0.97**	4	*	1.35	5	*	**0.93**	4		1.06	3	*	**0.98**	1	*	**0.87**	1	*
	1	**0.77**	3		**0.88**	2		**0.66**	4	*	1.00	2	*	**0.81**	2		**0.86**	4	*	**0.82**	1	
	2	**0.89**	2		**0.88**	1		**0.93**	1		1.03	1		**0.97**	1	*	1.05	1	*	**0.94**	3	
	3	**0.85**	4	*	**0.80**	1	*	**0.88**	1		**0.77**	5	*	**0.95**	1		**0.97**	5	*	**0.94**	5	*
	Average	**0.89**	2		**0.86**	1		**0.98**	1		1.01	4		**0.97**	1		1.05	1		**0.95**	1	
Method 4 - Weigthed PC																						
	0	1.07	1	*	1.01	1	*	1.13	1		1.18	1	*	1.15	1	*	**0.98**	1	*	**0.92**	1	
	1	1.04	1	*	1.10	1	*	1.04	1	*	1.18	1	*	**0.92**	1	*	**0.99**	1	*	**0.99**	1	
	2	1.01	1	*	**0.93**	1		1.18	1	*	1.21	1	*	1.21	1		**0.79**	1	*	**0.96**	1	*
	3	1.06	1	*	**0.85**	1		1.27	1	*	1.24	1	*	1.11	1	*	**0.85**	1	*	1.02	1	*
	Average	1.03	1		**0.90**	1		1.11	1		1.20	1		1.08	1		**0.90**	1		**0.97**	1	
Method 5 - Correlation Oriented PC																						
	0	**0.99**	2	*	1.02	1	*	1.06	2	*	1.05	4	*	1.10	5		**0.95**	4	*	**0.92**	2	
	1	1.04	1	*	**0.97**	4	*	1.04	1	*	1.08	4	*	**0.93**	1	*	**0.92**	5	*	**0.86**	2	
	2	1.04	2	*	**0.86**	3		1.19	1	*	1.06	4	*	1.26	1		**0.79**	1	*	**0.98**	1	*
	3	**0.90**	3		**0.94**	3		**0.95**	2		**0.86**	4		**0.92**	2	*	**0.78**	2	*	1.02	4	*
	Average	**0.99**	3		**0.91**	3		1.05	2		1.01	4		1.07	2		**0.90**	1		**0.98**	2	
Method 6 - PLS																						
	0	**0.88**	2	*	1.03	1	*	**0.87**	2	*	**0.84**	2	*	**0.96**	2	*	**0.89**	2	*	**0.85**	2	
	1	**0.94**	2	*	1.03	1	*	**0.83**	2	*	**0.91**	4	*	**0.76**	2		1.02	2	*	**0.76**	2	
	2	1.03	2	*	1.29	1		**0.94**	2	*	**0.90**	4	*	**0.96**	2	*	1.00	4	*	**0.89**	1	*
	3	**0.99**	4	*	1.37	1		**0.95**	2	*	**0.86**	2	*	**0.93**	4	*	**0.97**	4	*	**0.88**	1	*
	Average	**0.99**	2		1.08	1		**0.87**	2		**0.91**	4		**0.91**	2		**1.00**	4		**0.87**	2	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

Table 9: Out-of-sample RMSE for y-o-y forecasts: Database of totals, AR terms included in all models

Out-of-sample RMSE for y-o-y forecasts: Database of Totals, AR terms included in all models

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.026	2		0.094	3		0.022	1		0.030	4		0.022	1		0.033	4		0.029	1	
	1	0.035	4		0.091	3		0.033	4		0.040	2		0.034	4		0.045	3		0.041	1	
	2	0.041	2		0.131	2		0.035	1		0.036	1		0.036	1		0.054	4		0.043	1	
	3	0.040	2		0.143	4		0.032	2		0.034	1		0.034	2		0.055	4		0.044	1	
	Average	*0.036*	3		*0.122*	3		*0.031*	2		*0.036*	1		*0.032*	2		*0.048*	4		*0.039*	1	
Method 1 -Bridge Model																						
	0	**0.85**	1		**0.86**	1		**0.88**	2		**0.82**	2	*	**0.98**	1	*	**0.79**	5		1.04	1	*
	1	1.11	1	*	**0.76**	3	*	1.14	1		**0.88**	1	*	1.12	1		**0.59**	4		**0.95**	5	*
	2	**0.65**	2		**0.63**	4		**0.43**	2	*	**0.95**	3		**0.74**	5	*	**0.59**	3	*	**0.47**	5	*
	3	**0.77**	4	*	**0.60**	4	*	**0.72**	5	*	1.12	5		**0.57**	1	*	**0.59**	5	*	**0.81**	4	*
	Average	**0.89**	2		**0.67**	4		**0.86**	2		1.06	3		**0.86**	1		**0.66**	5		**0.86**	5	*
Method 2 - Standard PC																						
	0	**0.94**	1		**0.94**	4		1.08	1		1.07	1		1.05	1	*	**0.91**	3		**0.90**	2	*
	1	**0.69**	2		**0.86**	1		**0.71**	2		1.14	5		**0.72**	2	*	**0.66**	3		**0.82**	2	*
	2	**0.67**	2	*	**0.89**	1		**0.65**	2		1.24	2		**0.59**	2		**0.59**	1		**0.64**	2	*
	3	**0.65**	2		**0.65**	4	*	**0.67**	2	*	1.07	3		**0.79**	5	*	**0.51**	1	*	**0.75**	2	*
	Average	**0.71**	2		**0.79**	4		**0.75**	2		1.15	2		**0.78**	2		**0.64**	1		**0.77**	2	
Method 3 - Targeted PC																						
	0	**0.94**	1		**0.91**	2		1.08	1		1.07	1		1.05	1	*	**0.93**	1		**0.95**	2	*
	1	**0.96**	1		**0.83**	2		1.05	1		1.17	1	*	1.00	1		**0.67**	2		**0.87**	1	*
	2	**0.85**	3		**0.89**	2		**0.84**	3	*	1.31	5	*	**0.95**	2		**0.58**	3		**0.90**	3	
	3	**0.83**	3	*	**0.66**	4		**0.59**	2	*	1.01	1	*	**0.71**	2	*	**0.49**	5	*	**0.91**	1	
	Average	1.00	3		**0.76**	2		1.02	3		1.25	1		**0.97**	1		**0.64**	2		**0.95**	1	
Method 4 - Weigthed PC																						
	0	**0.93**	1		**0.95**	1		1.08	1		1.06	1		1.05	1	*	**0.94**	1		1.05	1	*
	1	**0.94**	1		**0.86**	1		1.02	1		1.15	1	*	1.00	1		**0.69**	1		1.10	1	
	2	1.01	1	*	**0.88**	1		1.23	1	*	1.58	1	*	1.05	1	*	**0.61**	1		1.07	1	*
	3	1.17	1	*	**0.78**	1		1.38	1	*	1.83	1	*	1.20	1		**0.49**	1	*	1.10	1	*
	Average	1.01	1		**0.81**	1		1.17	1		1.38	1		1.07	1		**0.64**	1		1.08	1	
Method 5 - Correlation Oriented PC																						
	0	**0.92**	1		**0.96**	1		1.06	1		1.02	1		1.14	1	*	**0.92**	4		**0.93**	2	
	1	**0.74**	2		**0.86**	1		**0.78**	2		1.08	1		**0.79**	2	*	**0.71**	1		**0.83**	2	*
	2	**0.81**	2		**0.86**	1		**0.97**	2	*	1.38	2		**0.93**	2	*	**0.59**	1		**0.74**	2	*
	3	**0.80**	2	*	**0.70**	3		**0.98**	2	*	1.14	3	*	**0.94**	3	*	**0.50**	1	*	**0.74**	2	*
	Average	**0.80**	2		**0.80**	3		**0.93**	2		1.14	3		**0.96**	2		**0.65**	1		**0.80**	2	
Method 6 - PLS																						
	0	1.61	4		2.00	1		1.37	2		1.26	2		1.33	2	*	1.78	1		**0.98**	4	*
	1	1.17	2		2.02	1		**0.78**	2		**0.89**	2		**0.78**	4		1.30	1	*	**0.75**	2	
	2	1.05	4	*	1.27	1		**0.86**	2	*	1.07	2		**0.82**	2		1.21	4	*	**0.81**	2	
	3	1.12	2		1.21	1	*	1.08	5	*	1.22	4	*	1.02	5		1.30	4	*	**0.81**	4	*
	Average	1.19	2		1.45	1		1.01	2		1.10	2		**0.95**	4		1.35	1		**0.85**	4	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

Table 10: Out-of-sample RMSE for y-o-y forecasts: Database of subsectors, AR terms included in all models

Out-of-sample RMSE for y-o-y forecasts: Database of Sectors, AR terms included in all models

	No. of periods ahead *h*	Overall GFCF	Regressors(a)	DM	Public GFCF	Regressors(a)	DM	Private GFCF	Regressors(a)	DM	Private Housing GFCF	Regressors(a)	DM	Private Productive GFCF	Regressors(a)	DM	Construction GFCF	Regressors(a)	DM	Overall GFCF excluding contruction	Regressors(a)	DM
Method 0 -AR Model	0	0.026	2		0.094	3		0.022	1		0.030	4		0.022	1		0.033	4		0.029	1	
	1	0.035	4		0.091	3		0.033	4		0.040	2		0.034	4		0.045	3		0.041	1	
	2	0.041	2		0.131	2		0.035	1		0.036	1		0.036	1		0.054	4		0.043	1	
	3	0.040	2		0.143	4		0.032	2		0.034	1		0.034	2		0.055	4		0.044	1	
	Average	*0.036*	3		*0.122*	3		*0.031*	2		*0.036*	1		*0.032*	2		*0.048*	4		*0.039*	1	
Method 1 -Bridge Model																						
	0	**0.90**	5	*	1.08	1	*	**0.97**	4		1.17	4		**0.93**	4		1.46	1	*	1.42	1	*
	1	**0.94**	5		1.14	3	*	**0.80**	4		1.23	2		**0.84**	5		**0.66**	3		**0.84**	2	
	2	**0.55**	1	*	**0.63**	4		**0.70**	1	*	1.33	1		**0.64**	1	*	1.24	4		**0.74**	1	*
	3	**0.61**	2	*	**0.58**	2		**0.83**	5		1.18	1		**0.81**	5		1.04	1	*	**0.78**	3	*
	Average	**0.78**	2		**0.77**	2		**0.84**	4		1.25	1		**0.80**	5		1.15	1		**0.94**	2	
Method 2 - Standard PC																						
	0	**0.92**	5		1.22	5	*	1.00	4		1.53	2		**0.89**	5		1.37	5	*	1.23	4	*
	1	**0.52**	5		**0.99**	4		**0.62**	5		1.22	4		**0.69**	5		**0.68**	5	*	**0.91**	4	
	2	**0.67**	2		**0.38**	4		**0.80**	4		1.49	2		**0.56**	1	*	**0.79**	4		**0.63**	1	
	3	**0.60**	2		**0.41**	3		**0.74**	1	*	1.39	3		**0.65**	1	*	**0.75**	2	*	**0.73**	4	
	Average	**0.72**	5		**0.65**	3		**0.79**	4		1.38	3		**0.76**	5		**0.87**	5		**0.90**	4	
Method 3 - Targeted PC																						
	0	**1.00**	4		1.20	5		**0.82**	3		1.81	3		1.01	5		**0.99**	5		1.23	4	*
	1	**0.66**	4		**0.99**	4		**0.63**	4		1.23	3		**0.69**	4		**0.77**	3		**0.85**	5	
	2	**0.59**	4		**0.56**	5		**0.68**	3		1.68	5		**0.81**	2	*	**0.74**	5		**0.78**	5	
	3	**0.63**	2	*	**0.32**	1		**0.63**	5		1.24	1	*	**0.74**	5	*	**0.73**	4		**0.69**	1	
	Average	**0.69**	4		**0.74**	4		**0.73**	3		1.52	3		**0.80**	4		**0.80**	5		**0.90**	5	
Method 4 - Weigthed PC																						
	0	1.18	1	*	1.22	1		1.14	1		1.96	1		1.27	1	*	1.40	1		1.51	1	*
	1	**0.99**	1	*	1.02	1		1.02	1	*	1.72	1	*	**0.85**	1		1.08	1		**0.91**	1	
	2	**0.98**	1	*	**0.41**	1		**0.95**	1	*	2.29	1	*	**0.56**	1	*	**0.98**	1		**0.72**	1	
	3	1.07	1	*	**0.52**	1	*	1.08	1	*	2.56	1	*	**0.67**	1		1.01	1	*	**0.94**	1	*
	Average	1.03	1		**0.69**	1		1.03	1		2.08	1		**0.78**	1		1.07	1		**0.97**	1	
Method 5 - Correlation Oriented PC																						
	0	**0.90**	5		1.08	5	*	**0.92**	5		1.87	2		1.18	3	*	1.28	4		1.08	5	*
	1	**0.53**	5		1.03	2		**0.60**	5		1.13	2		**0.66**	4		**0.81**	4	*	**0.88**	5	
	2	**0.74**	2		**0.39**	3		**0.87**	2		1.44	2		**0.79**	1		**0.72**	4		**0.59**	1	*
	3	**0.66**	2		**0.46**	3	*	**0.78**	2	*	1.48	3		**0.67**	1		**0.78**	2	*	**0.83**	5	
	Average	**0.77**	5		**0.66**	3		**0.84**	5		1.44	3		**0.84**	5		**0.85**	4		**0.90**	5	
Method 6 - PLS																						
	0	1.70	2	*	2.02	1		1.29	2		1.45	1	*	1.39	2		1.86	1		1.13	2	
	1	1.23	2	*	2.06	1	*	**0.88**	2		1.01	4		**0.88**	2		1.34	1	*	**0.75**	2	
	2	**0.98**	2	*	1.32	1		**0.84**	2		1.10	2	*	**0.86**	2		1.16	2		**0.79**	2	
	3	1.09	2		1.28	1	*	1.01	2	*	1.11	2		**0.95**	2	*	1.24	4	*	**0.69**	1	*
	Average	1.19	2		1.50	1		**0.97**	2		1.16	2		**0.97**	2		1.34	1		**0.85**	2	

Notes:

RMSE in levels in the case of Method 0 and in percentage of Method 0 in the case of the remaining methods.

Bold indicates a smaller RMSE than the one of method 0. Shading indicates the best method at each step, given that it has a smaller RMSE than method 0.

(a) Indicates the number of regressors included in the initial specification of the equation that corresponds to the lowest RMSE. In the case of Method 6, indicates the number of orthogonal components extracted from the matrix of regressors X.

(*) Indicates that the RMSE is statistically different than the one of the corresponding AR model according to the Diebold-Mariano test at a 10% significance level.

D Bridge models results

Table 11: Method 1 - Surveys included in the regressions: Database of Totals, q-o-q data, AR terms included in the models

Method 1 - Surveys included in the regressions: Database of Totals, q-o-q data, AR terms included in the model

No. of steps ahead	Overall GFCF	Public GFCF	Private GFCF	Private Housing GFCF	Private Productive GFCF	Construction GFCF	Overall GFCF excluding contruction
0	Retail Trade: Orders expectations over the next 3 months	Retail Trade: Volume of stock currently hold	Retail Trade: Business activity expectations over the next 3 months	Retail Trade: Business activity expectations over the next 3 months	Retail Trade: Business activity expectations over the next 3 months	Industry: Competitive position outside EU Retail Trade: Confidence Indicator	Industry: Factors limiting the production - Labour Industry: Assessment of stocks of finished products
1	Industry: Duration of production assured by current order-book levels	Industry: Production expectations for the months ahead	Industry: Factors limiting the production - Labour	Industry: Production expectations for the months ahead	Retail Trade: Orders expectations over the next 3 months Industry: Competitive position outside EU	Industry: Employment expectations for the months ahead	Retail Trade: Employment expectations over the next 3 months
2	Retail Trade: Volume of stock currently hold Construction: Factors Limiting the Production Shortage of material and/or equipment	Construction: Factors Limiting the Production Other factors Industry: Current level of capacity utilization	Construction: Factors Limiting the Production Shortage of material and/or equipment	Industry: Export expectations for the months ahead	Construction: Factors Limiting the Production Shortage of material and/or equipment	Industry: Assessment of export order-book levels	Retail Trade: Employment expectations over the next 3 months
3	Industry: Factors limiting the production - Equipment	Construction: Factors Limiting the Production Other factors	Services: Expectation of the demand over the next 3 months	Construction: Factors Limiting the Production Weather conditions	Industry: Competitive position inside EU	Industry: Factors limiting the production - Other Industry: Factors limiting the production - Labour Services: Expectation of the demand over the next 3 months Construction: Factors Limiting the Production None	Industry: New orders in recent months

Table 12: Method 1 - Surveys included in the regressions: Database of Sectors, q-o-q data, AR terms included in the models

Method 1 - Surveys included in the regressions: Database of Sectors, q-o-q data, AR terms included in the model

No. of steps ahead	Overall GFCF	Public GFCF	Private GFCF	Private Housing GFCF	Private Productive GFCF	Construction GFCF
0	CONS Duration of production assured by current order-book levels FOBE Factors limiting the production - Equipment	Retail Trade Volume of stock currently hold CONS Production trend observed in recent months	Retail Trade Business activity expectations over the next 3 months	INVE Production expectations for the months ahead	Retail Trade Business activity expectations over the next 3 months	INVE Employment expectations for the months ahead CONS Factors limiting the production - Equipment CONS Duration of production assured by current order-book levels FOBE Factors limiting the production - Equipment
1	FOBE Factors limiting the production - Equipment FOBE Assessment of export order-book levels INTM Competitive position inside EU	FOBE Production trend observed in recent months CNDU Selling price expectations for the months ahead CONS Assessment of export order-book levels	INTM Competitive position inside EU	FOBE Factors limiting the production - Equipment INTM Export expectations for the months ahead Public works (civil engineering) Factors limiting the production Shortage of material and/or equipment CONS Factors limiting the production - Equipment	INTM Competitive position inside EU	INVE Employment expectations for the months ahead CONS Factors limiting the production - Equipment
2	INTM Competitive position outside EU Public works (civil engineering) Factors limiting the production Shortage of material and/or equipment	Retail Trade Volume of stock currently hold INVE Assessment of stocks of finished products	INVE Factors limiting the production - Labour INTM Competitive position inside EU	Public works (civil engineering) Factors limiting the production Shortage of material and/or equipment INVE Factors limiting the production - Labour CONS Factors limiting the production - Labour	CNDU Factors limiting the production - Labour INTM Competitive position inside EU Public works (civil engineering) Factors limiting the production Shortage of material and/or equipment	INTM Competitive position outside EU INTM Competitive position inside EU
3	CDUR Competitive position domestic market	CNDU Assessment of stocks of finished products	CDUR Competitive position domestic market	CDUR Factors limiting the production - None	CNDU Factors limiting the production - Labour Residential Building Factors limiting the production Other factors	CDUR Factors limiting the production - None INTM Export expectations for the months ahead INVE Factors limiting the production - Equipment

Table 13: Method 1 - Surveys included in the regressions: Database of Totals, y-o-y data, AR terms included in the models

Method 1 - Surveys included in the regressions: Database of Totals, y-o-y data, AR terms included in the model

No. of steps ahead	Overall GFCF	Public GFCF	Private GFCF	Private Housing GFCF	Private Productive GFCF	Construction GFCF	Overall GFCF excluding contruction
0	Retail Trade: Business activity expectations over the next 3 months	Services: Evolution of the demand over the past 3 months	Retail Trade: Business activity expectations over the next 3 months	Construction: Evolution of your current overall order books Construction: Building activity development over the past 3 months	Retail Trade: Business activity (sales) development over the past 3 months	Construction: Evolution of your current overall order books Services: Confidence Indicator	Retail Trade: Business activity expectations over the next 3 months
1	Retail Trade: Business activity expectations over the next 3 months	Industry: Employment expectations for the months ahead	Retail Trade: Business activity expectations over the next 3 months	Services: Evolution of the demand over the past 3 months	Retail Trade: Business activity (sales) development over the past 3 months	Services: Evolution of the demand over the past 3 months Industry: Production expectations for the months ahead	Retail Trade: Business activity expectations over the next 3 months
2	Retail Trade: Orders expectations over the next 3 months Retail Trade: Business activity (sales) development over the past 3 months	Industry: Export expectations for the months ahead Industry: Production expectations for the months ahead	Retail Trade: Orders expectations over the next 3 months Retail Trade: Business activity (sales) development over the past 3 months	Retail Trade: Business activity expectations over the next 3 months Industry: Assessment of stocks of finished products Retail Trade: Orders expectations over the next 3 months	Retail Trade: Orders expectations over the next 3 months Retail Trade: Confidence Indicator	Industry: Employment expectations for the months ahead Industry: Export expectations for the months ahead	Construction: Factors Limiting the Production Shortage of material and/or equipment Industry: Factors limiting the production - None Retail Trade: Employment expectations over the next 3 months Construction: Evolution of your current overall order books
3	Construction: Factors Limiting the Production Weather conditions	Industry: Export expectations for the months ahead Industry: Duration of production assured by current order-book levels Industry: Current level of capacity utilization	Construction: Factors Limiting the Production Shortage of material and/or equipment Retail Trade: Volume of stock currently hold	Construction: Factors Limiting the Production Weather conditions Construction: Evolution of your current overall order books	Construction: Factors Limiting the Production Shortage of material and/or equipment	Construction: Factors Limiting the Production None Industry: Factors limiting the production - Equipment	Construction: Factors Limiting the Production Shortage of material and/or equipment Industry: Competitive position inside EU

Table 14: Method 1 - Surveys included in the regressions: Database of Sectors, y-o-y data, AR terms included in the models

Method 1 - Surveys included in the regressions: Database of Sectors, y-o-y data, AR terms included in the model

No. of steps ahead	Overall GFCF	Public GFCF	Private GFCF	Private Housing GFCF	Private Productive GFCF	Construction GFCF	Overall GFCF excluding contruction
0	Retail Trade Business activity expectations over the next 3 months Residential Building Building activity development over the past 3 months Retail Trade Confidence Indicator	Non-residential building Employment expectations over the next 3 months	Retail Trade Business activity expectations over the next 3 months INVE Confidence Indicator	INVE Production expectations for the months ahead	Retail Trade Business activity (sales) development over the past 3 months Retail Trade Confidence Indicator FOBE Duration of production assured by current order-book levels Retail Trade Business activity expectations over the next 3 months	Residential Building Building activity development over the past 3 months	Retail Trade Business activity expectations over the next 3 months
1	Retail Trade Business activity expectations over the next 3 months FOBE Assessment of export order-book levels	INTM Competitive position domestic market	FOBE Assessment of export order-book levels Retail Trade Business activity expectations over the next 3 months	Public works (civil engineering) Evolution of your current overall order books INVE Production expectations for the months ahead	Retail Trade Confidence Indicator INVE Employment expectations for the months ahead	CONS Confidence Indicator CONS New orders in recent months	Retail Trade Orders expectations over the next 3 months
2	INVE Factors limiting the production - Labour	INTM Competitive position domestic market INTM Competitive position outside EU CNDU Selling price expectations for the months ahead	CONS Factors limiting the production - Labour	CDUR Factors limiting the production - Other	INVE Factors limiting the production - Labour	CONS Assessment of export order-book levels	CONS Factors limiting the production - Labour
3	CNDU Factors limiting the production - Labour CONS Factors limiting the production - Labour	Total building Other factors	CONS Factors limiting the production - Labour CDUR Factors limiting the production - Demand	CNDU Factors limiting the production - Labour	CONS Factors limiting the production - Labour INVE Factors limiting the production - Other CDUR Factors limiting the production - Demand	INTM Export expectations for the months ahead	CONS Factors limiting the production - Labour INVE Factors limiting the production - Labour

WORKING PAPERS

2000

1/00 UNEMPLOYMENT DURATION: COMPETING AND DEFECTIVE RISKS
— *John T. Addison, Pedro Portugal*

2/00 THE ESTIMATION OF RISK PREMIUM IMPLICIT IN OIL PRICES
— *Jorge Barros Luís*

3/00 EVALUATING CORE INFLATION INDICATORS
— *Carlos Robalo Marques, Pedro Duarte Neves, Luís Morais Sarmento*

4/00 LABOR MARKETS AND KALEIDOSCOPIC COMPARATIVE ADVANTAGE
— *Daniel A. Traça*

5/00 WHY SHOULD CENTRAL BANKS AVOID THE USE OF THE UNDERLYING INFLATION INDICATOR?
— *Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

6/00 USING THE ASYMMETRIC TRIMMED MEAN AS A CORE INFLATION INDICATOR
— *Carlos Robalo Marques, João Machado Mota*

2001

1/01 THE SURVIVAL OF NEW DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

2/01 GAPS AND TRIANGLES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

3/01 A NEW REPRESENTATION FOR THE FOREIGN CURRENCY RISK PREMIUM
— *Bernardino Adão, Fátima Silva*

4/01 ENTRY MISTAKES WITH STRATEGIC PRICING
— *Bernardino Adão*

5/01 FINANCING IN THE EUROSYSTEM: FIXED VERSUS VARIABLE RATE TENDERS
— *Margarida Catalão-Lopes*

6/01 AGGREGATION, PERSISTENCE AND VOLATILITY IN A MACROMODEL
— *Karim Abadir, Gabriel Talmain*

7/01 SOME FACTS ABOUT THE CYCLICAL CONVERGENCE IN THE EURO ZONE
— *Frederico Belo*

8/01 TENURE, BUSINESS CYCLE AND THE WAGE-SETTING PROCESS
— *Leandro Arozamena, Mário Centeno*

9/01 USING THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ferreira Machado, Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

10/01 IDENTIFICATION WITH AVERAGED DATA AND IMPLICATIONS FOR HEDONIC REGRESSION STUDIES
— *José A.F. Machado, João M.C. Santos Silva*

2002

1/02 QUANTILE REGRESSION ANALYSIS OF TRANSITION DATA
— *José A.F. Machado, Pedro Portugal*

2/02 SHOULD WE DISTINGUISH BETWEEN STATIC AND DYNAMIC LONG RUN EQUILIBRIUM IN ERROR CORRECTION MODELS?
— *Susana Botas, Carlos Robalo Marques*

3/02 MODELLING TAYLOR RULE UNCERTAINTY
— *Fernando Martins, José A. F. Machado, Paulo Soares Esteves*

4/02 PATTERNS OF ENTRY, POST-ENTRY GROWTH AND SURVIVAL: A COMPARISON BETWEEN DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

5/02 BUSINESS CYCLES: CYCLICAL COMOVEMENT WITHIN THE EUROPEAN UNION IN THE PERIOD 1960-1999. A FREQUENCY DOMAIN APPROACH
— *João Valle e Azevedo*

6/02 AN “ART”, NOT A “SCIENCE”? CENTRAL BANK MANAGEMENT IN PORTUGAL UNDER THE GOLD STANDARD, 1854 -1891
— *Jaime Reis*

7/02 MERGE OR CONCENTRATE? SOME INSIGHTS FOR ANTITRUST POLICY
— *Margarida Catalão-Lopes*

8/02 DISENTANGLING THE MINIMUM WAGE PUZZLE: ANALYSIS OF WORKER ACCESSIONS AND SEPARATIONS FROM A LONGITUDINAL MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Pedro Portugal, Ana Rute Cardoso*

9/02 THE MATCH QUALITY GAINS FROM UNEMPLOYMENT INSURANCE
— *Mário Centeno*

10/02 HEDONIC PRICES INDEXES FOR NEW PASSENGER CARS IN PORTUGAL (1997-2001)
— *Hugo J. Reis, J.M.C. Santos Silva*

11/02 THE ANALYSIS OF SEASONAL RETURN ANOMALIES IN THE PORTUGUESE STOCK MARKET
— *Miguel Balbina, Nuno C. Martins*

12/02 DOES MONEY GRANGER CAUSE INFLATION IN THE EURO AREA?
— *Carlos Robalo Marques, Joaquim Pina*

13/02 INSTITUTIONS AND ECONOMIC DEVELOPMENT: HOW STRONG IS THE RELATION?
— *Tiago V.de V. Cavalcanti, Álvaro A. Novo*

2003

1/03 FOUNDING CONDITIONS AND THE SURVIVAL OF NEW FIRMS
— *P.A. Geroski, José Mata, Pedro Portugal*

2/03 THE TIMING AND PROBABILITY OF FDI: AN APPLICATION TO THE UNITED STATES MULTINATIONAL ENTERPRISES
— *José Brandão de Brito, Felipa de Mello Sampayo*

3/03 OPTIMAL FISCAL AND MONETARY POLICY: EQUIVALENCE RESULTS
— *Isabel Correia, Juan Pablo Nicolini, Pedro Teles*

2004

2/04 REAL EXCHANGE RATE AND HUMAN CAPITAL IN THE EMPIRICS OF ECONOMIC GROWTH
— *Delfim Gomes Neto*

3/04 ON THE USE OF THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ramos Maria*

4/04 OIL PRICES ASSUMPTIONS IN MACROECONOMIC FORECASTS: SHOULD WE FOLLOW FUTURES MARKET EXPECTATIONS?
— *Carlos Coimbra, Paulo Soares Esteves*

5/04 STYLISED FEATURES OF PRICE SETTING BEHAVIOUR IN PORTUGAL: 1992-2001
— *Mónica Dias, Daniel Dias, Pedro D. Neves*

6/04 A FLEXIBLE VIEW ON PRICES
— *Nuno Alves*

7/04 ON THE FISHER-KONIECZNY INDEX OF PRICE CHANGES SYNCHRONIZATION
— *D.A. Dias, C. Robalo Marques, P.D. Neves, J.M.C. Santos Silva*

8/04 INFLATION PERSISTENCE: FACTS OR ARTEFACTS?
— *Carlos Robalo Marques*

9/04 WORKERS' FLOWS AND REAL WAGE CYCLICALITY
— *Anabela Carneiro, Pedro Portugal*

10/04 MATCHING WORKERS TO JOBS IN THE FAST LANE: THE OPERATION OF FIXED-TERM CONTRACTS
— *José Varejão, Pedro Portugal*

11/04 THE LOCATIONAL DETERMINANTS OF THE U.S. MULTINATIONALS ACTIVITIES
— *José Brandão de Brito, Felipa Mello Sampayo*

12/04 KEY ELASTICITIES IN JOB SEARCH THEORY: INTERNATIONAL EVIDENCE
— *John T. Addison, Mário Centeno, Pedro Portugal*

13/04 RESERVATION WAGES, SEARCH DURATION AND ACCEPTED WAGES IN EUROPE
— *John T. Addison, Mário Centeno, Pedro Portugal*

14/04 THE MONETARY TRANSMISSION N THE US AND THE EURO AREA: COMMON FEATURES AND COMMON FRICTIONS
— *Nuno Alves*

15/04 NOMINAL WAGE INERTIA IN GENERAL EQUILIBRIUM MODELS
— *Nuno Alves*

16/04 MONETARY POLICY IN A CURRENCY UNION WITH NATIONAL PRICE ASYMMETRIES
— *Sandra Gomes*

17/04 NEOCLASSICAL INVESTMENT WITH MORAL HAZARD
— *João Ejarque*

18/04 MONETARY POLICY WITH STATE CONTINGENT INTEREST RATES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

19/04 MONETARY POLICY WITH SINGLE INSTRUMENT FEEDBACK RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

20/04 ACOUNTING FOR THE HIDDEN ECONOMY: BARRIERS TO LAGALITY AND LEGAL FAILURES
— *António R. Antunes, Tiago V. Cavalcanti*

2005

1/05 SEAM: A SMALL-SCALE EURO AREA MODEL WITH FORWARD-LOOKING ELEMENTS
— *José Brandão de Brito, Rita Duarte*

2/05 FORECASTING INFLATION THROUGH A BOTTOM-UP APPROACH: THE PORTUGUESE CASE
— *Cláudia Duarte, António Rua*

3/05 USING MEAN REVERSION AS A MEASURE OF PERSISTENCE
— *Daniel Dias, Carlos Robalo Marques*

4/05 HOUSEHOLD WEALTH IN PORTUGAL: 1980-2004
— *Fátima Cardoso, Vanda Geraldes da Cunha*

5/05 ANALYSIS OF DELINQUENT FIRMS USING MULTI-STATE TRANSITIONS
— *António Antunes*

6/05 PRICE SETTING IN THE AREA: SOME STYLIZED FACTS FROM INDIVIDUAL CONSUMER PRICE DATA
— *Emmanuel Dhyne, Luis J. Álvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lünnemann, Fabio Rumler, Jouko Vilmunen*

7/05 INTERMEDIATION COSTS, INVESTOR PROTECTION AND ECONOMIC DEVELOPMENT
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

8/05 TIME OR STATE DEPENDENT PRICE SETTING RULES? EVIDENCE FROM PORTUGUESE MICRO DATA
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

9/05 BUSINESS CYCLE AT A SECTORAL LEVEL: THE PORTUGUESE CASE
— *Hugo Reis*

10/05 THE PRICING BEHAVIOUR OF FIRMS IN THE EURO AREA: NEW SURVEY EVIDENCE
— *S. Fabiani, M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, A. Stokman*

11/05 CONSUMPTION TAXES AND REDISTRIBUTION
— *Isabel Correia*

12/05 UNIQUE EQUILIBRIUM WITH SINGLE MONETARY INSTRUMENT RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

13/05 A MACROECONOMIC STRUCTURAL MODEL FOR THE PORTUGUESE ECONOMY
— *Ricardo Mourinho Félix*

14/05 THE EFFECTS OF A GOVERNMENT EXPENDITURES SHOCK
— *Bernardino Adão, José Brandão de Brito*

15/05 MARKET INTEGRATION IN THE GOLDEN PERIPHERY – THE LISBON/LONDON EXCHANGE, 1854-1891
— *Rui Pedro Esteves, Jaime Reis, Fabiano Ferramosca*

2006

1/06 THE EFFECTS OF A TECHNOLOGY SHOCK IN THE EURO AREA
— *Nuno Alves , José Brandão de Brito , Sandra Gomes, João Sousa*

2/02 THE TRANSMISSION OF MONETARY AND TECHNOLOGY SHOCKS IN THE EURO AREA
— *Nuno Alves, José Brandão de Brito, Sandra Gomes, João Sousa*

3/06 MEASURING THE IMPORTANCE OF THE UNIFORM NONSYNCHRONIZATION HYPOTHESIS
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

4/06 THE PRICE SETTING BEHAVIOUR OF PORTUGUESE FIRMS EVIDENCE FROM SURVEY DATA
— *Fernando Martins*

5/06 STICKY PRICES IN THE EURO AREA: A SUMMARY OF NEW MICRO EVIDENCE
— *L. J. Álvarez, E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilmunen*

6/06 NOMINAL DEBT AS A BURDEN ON MONETARY POLICY
— *Javier Díaz-Giménez, Giorgia Giovannetti , Ramon Marimon, Pedro Teles*

7/06 A DISAGGREGATED FRAMEWORK FOR THE ANALYSIS OF STRUCTURAL DEVELOPMENTS IN PUBLIC FINANCES
— *Jana Kremer, Cláudia Rodrigues Braz, Teunis Brosens, Geert Langenus, Sandro Momigliano, Mikko Spolander*

8/06 IDENTIFYING ASSET PRICE BOOMS AND BUSTS WITH QUANTILE REGRESSIONS
— *José A. F. Machado, João Sousa*

9/06 EXCESS BURDEN AND THE COST OF INEFFICIENCY IN PUBLIC SERVICES PROVISION
— *António Afonso, Vítor Gaspar*

10/06 MARKET POWER, DISMISSAL THREAT AND RENT SHARING: THE ROLE OF INSIDER AND OUTSIDER FORCES IN WAGE BARGAINING
— *Anabela Carneiro, Pedro Portugal*

11/06 MEASURING EXPORT COMPETITIVENESS: REVISITING THE EFFECTIVE EXCHANGE RATE WEIGHTS FOR THE EURO AREA COUNTRIES
— *Paulo Soares Esteves, Carolina Reis*

12/06 THE IMPACT OF UNEMPLOYMENT INSURANCE GENEROSITY
ON MATCH QUALITY DISTRIBUTION
— *Mário Centeno, Alvaro A. Novo*

13/06 U.S. UNEMPLOYMENT DURATION: HAS LONG BECOME LONGER OR SHORT BECOME SHORTER?
— *José A.F. Machado, Pedro Portugal e Juliana Guimarães*

14/06 EARNINGS LOSSES OF DISPLACED WORKERS: EVIDENCE FROM A MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Anabela Carneiro, Pedro Portugal*

15/06 COMPUTING GENERAL EQUILIBRIUM MODELS WITH OCCUPATIONAL CHOICE AND FINANCIAL FRICTIONS
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

16/06 ON THE RELEVANCE OF EXCHANGE RATE REGIMES FOR STABILIZATION POLICY
— *Bernardino Adao, Isabel Correia, Pedro Teles*

17/06 AN INPUT-OUTPUT ANALYSIS: LINKAGES VS LEAKAGES
— *Hugo Reis, António Rua*

2007

1/07 RELATIVE EXPORT STRUCTURES AND VERTICAL SPECIALIZATION: A SIMPLE CROSS-COUNTRY INDEX
— *João Amador, Sónia Cabral, José Ramos Maria*

20/07 IS THE EURO AREA M3 ABANDONING US?
— *Nuno Alves, Carlos Robalo Marques, João Sousa*

21/07 DO LABOR MARKET POLICIES AFFECT EMPLOYMENT COMPOSITION? LESSONS FROM EUROPEAN COUNTRIES
— *António Antunes, Mário Centeno*

2008

1/08 THE DETERMINANTS OF PORTUGUESE BANKS' CAPITAL BUFFERS
— *Miguel Boucinha*

2/08 DO RESERVATION WAGES REALLY DECLINE? SOME INTERNATIONAL EVIDENCE ON THE DETERMINANTS OF RESERVATION WAGES
— *John T. Addison, Mário Centeno, Pedro Portugal*

3/08 UNEMPLOYMENT BENEFITS AND RESERVATION WAGES: KEY ELASTICITIES FROM A STRIPPED-DOWN JOB SEARCH APPROACH
— *John T. Addison, Mário Centeno, Pedro Portugal*

4/08 THE EFFECTS OF LOW-COST COUNTRIES ON PORTUGUESE MANUFACTURING IMPORT PRICES
— *Fátima Cardoso, Paulo Soares Esteves*

5/08 WHAT IS BEHIND THE RECENT EVOLUTION OF PORTUGUESE TERMS OF TRADE?
— *Fátima Cardoso, Paulo Soares Esteves*

6/08 EVALUATING JOB SEARCH PROGRAMS FOR OLD AND YOUNG INDIVIDUALS: HETEROGENEOUS IMPACT ON UNEMPLOYMENT DURATION
— *Luis Centeno, Mário Centeno, Álvaro A. Novo*

7/08 FORECASTING USING TARGETED DIFFUSION INDEXES
— *Francisco Dias, Maximiano Pinheiro, António Rua*

8/08 STATISTICAL ARBITRAGE WITH DEFAULT AND COLLATERAL
— *José Fajardo, Ana Lacerda*

9/08 DETERMINING THE NUMBER OF FACTORS IN APPROXIMATE FACTOR MODELS WITH GLOBAL AND GROUP-SPECIFIC FACTORS
— *Francisco Dias, Maximiano Pinheiro, António Rua*

10/08 VERTICAL SPECIALIZATION ACROSS THE WORLD: A RELATIVE MEASURE
— *João Amador, Sónia Cabral*

11/08 INTERNATIONAL FRAGMENTATION OF PRODUCTION IN THE PORTUGUESE ECONOMY: WHAT DO DIFFERENT MEASURES TELL US?
— *João Amador, Sónia Cabral*

12/08 IMPACT OF THE RECENT REFORM OF THE PORTUGUESE PUBLIC EMPLOYEES' PENSION SYSTEM
— *Maria Manuel Campos, Manuel Coutinho Pereira*

13/08 EMPIRICAL EVIDENCE ON THE BEHAVIOR AND STABILIZING ROLE OF FISCAL AND MONETARY POLICIES IN THE US
— *Manuel Coutinho Pereira*

14/08 IMPACT ON WELFARE OF COUNTRY HETEROGENEITY IN A CURRENCY UNION
— *Carla Soares*

15/08 WAGE AND PRICE DYNAMICS IN PORTUGAL
— *Carlos Robalo Marques*

16/08 IMPROVING COMPETITION IN THE NON-TRADABLE GOODS AND LABOUR MARKETS: THE PORTUGUESE CASE
— *Vanda Almeida, Gabriela Castro, Ricardo Mourinho Félix*

17/08 PRODUCT AND DESTINATION MIX IN EXPORT MARKETS
— *João Amador, Luca David Opromolla*

18/08 FORECASTING INVESTMENT: A FISHING CONTEST USING SURVEY DATA
— *José Ramos Maria, Sara Serra*