Banco de Portugal Economic Studies

Volume II

BANCO DE PORTUGAL
EUROSYSTEM

3

3

volume II

Banco de Portugal
Economic Studies

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal
T +351 213 130 000 | estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSISTEMA

Lisbon, 2016 · www.bportugal.pt

BANCO DE PORTUGAL ECONOMIC STUDIES | Volume II - n.º 3 | Lisbon 2016 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt • Edition Economics and Research Department • Design Communication Directorate | Image and Graphic Design Unit | Editing and Publishing Unit • ISSN 2183-5217 (online)

Content

Editorial

July 2016

The third 2016 issue of the Banco de Portugal Economic Studies contains three diverse essays covering techniques for forecasting GDP, new results decomposing changes in Portuguese output and employment into trends and cycles, and the analysis of bank runs.

The first paper, by Dias, Pinheiro and Rua, is entitled "A bottom-up approach for forecasting GDP in a data rich environment". It provides a survey of the literature on forecasting using factor models and applies the methodological contributions to forecasting models of the authors to the forecasting of the Portuguese GDP and its components.

Factor models use data on many variables from multiple sources and from this large pool of data they extract a smaller orthogonal set of variables, its principal components, which capture the essential features of the data. Factor analysis is typically used in the statistical analysis of data in Psychology and Social Sciences but it used to be relatively rare in Economics until researchers realized this type of data reduction techniques could be useful in improving forecasting models.

The use of factor models in forecasting has led to Diffusion Index models where a variable is regressed on its lags and the principal components just mentioned. In this approach the principal components used are the same for all variables, i.e. they do not depend on the variable to be forecasted. Dias, Pinheiro and Rua have improved upon this by using a methodology that combines optimally the principal factors for each specific variable to be forecasted and does not waste information, the Targeted Diffusion Index model (TDI). In this essay they use this methodology and compare it to others in their accuracy for nowcasting (forecasting the current quarter: the present needs to be forecasted because variables such as the GDP are known only with delays and subject to revisions) and forecasting variables one quarter ahead.

The results show that the TDI outperforms other forecasting methodologies when dealing with quarterly growth rates for GDP and for its components. The paper also shows that aggregating the separate forecasts of GDP components provides better accuracy than forecasting directly GDP. However, if a component of GDP is itself disaggregated (for instance consumption divided into durables and non-durables) there seem to be no substantial gains in the forecasting accuracy of that component. The results of Dias, Pinheiro and Rua are significant in the sense that better forecasting is always helpful in providing the basis for better economic policies. However they also show that these new methods may contribute to better forecasting if we are concerned mostly with broad variables. Good policies may depend on specific details from narrowly defined areas or industries and in these cases factor models may not be the best tool.

The second essay in this issue is "Portugal: Trends, cycles, and instability in output and unemployment over 2008-2012" by J.R. Maria. Maria uses an empirically driven model, with only a few design assumptions, to estimate changes in output and unemployment gaps over the period under analysis. The motivation for the analysis could not be more salient: were large negative output changes and unemployment increases in Portugal in 2008-2012 the results of demand shocks, supply shocks or changes in trends? To what extent were these shocks or changes in trends domestic or mostly coming from outside the country? These are the sorts of questions this essay tries to answer.

The model builds on work developed in institutions such as the IMF or the ECB and is designed for a small country, Portugal, in a large monetary union, the Euro Area (EA), where causality is one-directional. Despite not being closely tied to a specific economic model, there are fundamental design assumptions that provide structure to the model. These include assuming identical long-run growth rates in the trends for output, unemployment and real interest rates in both Portugal and the Euro Area (EA). The analysis works by modeling explicitly inflation, nominal interest rates, output and unemployment gaps, meaning that all effective variables and their long-run trends are jointly estimated. The unemployment gap equation is particularly interesting because it represents a modern version of the classic Okun's law, where expected unemployment gaps are also allowed to play a role.

The results of the analysis are interesting and sometimes surprising. According to the essay, the long-run trend for the annual output growth rate in Portugal and in the EA was 1.8%. Furthermore, output in Portugal was 2% above trend at the end of 2007. From that point on there was a persistent downward movement in the output trend and a continuous upward trend in unemployment that seem to have stopped only after the end of 2012. The results of the analysis for the EA reveal high correlations in the output and unemployment gaps in Portugal and in the EA. Another interesting result is that the static relationship between the unemployment gap and the output gap is stable in the EA but changing in Portugal: the absolute effect of a one percent output gap on unemployment has been increasing: it was close to -0.3 percentage points (p.p.) in 2008 but closer to -0.6 p.p. in 2012. A surprising non-result is that changes in the real exchange rate seem to have had little or no effect on output or unemployment. As expected, the model finds recessive periods in the EA area that together with increases in risk premia have negative consequences for Portugal. Finally, one of the main conclusions of the essay is that the most important events identified were adverse movements in the trends for output and unemployment rather than just temporary shocks. If these results are confirmed by further analysis this means that only structural policies may be effective in improving Portugal's long-run growth.

The paper by Panetti, entitled "Bank runs: Theory and Evidence" reviews the literature studying bank runs and offers some novel ideas on the topic.

Bank runs are the prototypical case of financial instability. They are rare but salient economic phenomena where savers rush to withdraw their deposits, either physically forming waiting lines at the bank or, these days, by massive simultaneous attempts to use online banking services. The textbooks tell us that bank runs were less uncommon in the US before the creation of the Federal Reserve System and the introduction of deposit insurance. Since then, bank runs practically disappeared from developed countries until the runs on the US bank Countrywide Financial and the UK bank Northern Rock in 2007, two of the dramatic events in the initial stages of the Great Recession. Others have followed. For example in January 2016 the oldest ongoing bank in the world, Monte dei Paschi in Siena, Italy, suffered a bank run.

Understanding bank runs is important because it also means understanding banking systems and, more generally, financial systems. Many problematic events in financial markets are very similar to bank runs even when retail banks are not involved. One example could be the run on the investment bank Bear Stearns in 2008 or the panic driven withdrawals from money market funds. There are even interpretations of the so-called "hot money" flight in international financial crises, such as those that occurred in Asian countries in the 1990s, as being situations similar to bank runs.

The basic idea explaining bank runs starts from a view of banking as financing long term and illiquid assets with short-term and liquid liabilities. Normally the uncorrelated liquidity needs of depositors are well handled by the banks. However if many depositors want their money back at the same time, a bank may not be able to comply. Why would all depositors (or financial investors in more general settings) want their funds at the same time? Panetti´s review identifies two strands of the literature focusing on two different explanations. The first type relies on the expectations of depositors: if they expect a bank to fail, this by itself may cause it to fail. In this limit case a bank run is a self-fulfilling prophecy: for some exogenous stochastic reason (panics, sunspots. . .) depositors converge on having negative expectations regarding the ability of a bank to repay the loans. Faced with a tide of withdrawals the bank actually fails to repay, in circumstances where the fundamentals are otherwise sound and the bank has no insolvency problems. The other explanation links bank runs to situations where the fundamentals point to insolvency. In this case there may be reasons for the depositors to question the ability of a bank to repay the loans, because they expect low returns on the banks´ investments.

Panetti begins by studying expectations-driven bank runs building on developments of the seminal 1983 contribution by Diamond and Dybvig. Panetti shows that banks can make themselves immune to runs by distorting upwards their investments in liquid assets with low returns at the expense of illiquid assets with higher returns. However, that comes at the cost of lowering returns and liquidity insurance for depositors. Panetti draws the conclusion that competition in the industry leads banks to find a balance between the provision of liquidity insurance and the risk of bank runs and to be run proof in most cases even in the absence of government intervention. However in that equilibrium there is always some positive probability that bank runs will happen. Surprisingly, Panetti reveals that if the government imposes liquidity requirements with the goal of preventing any runs, the welfare costs of this policy will actually be quite small. It is also in dealing with this case, where banks need liquidity but are solvent, that central banks liquidity assistance plays a fundamental role.

The explanation for bank runs based on insolvency is reviewed next. A surprising result obtained from the basic model in this literature is that there is no room for government policies to improve on the outcomes of runs on insolvent banks. However, as the basic model is enriched with more realistic assumptions, such as incomplete markets for securities, asymmetric information for credit market participants and the existence of systemic liquidity risks, a role for policy emerges. In such models we learn that inter-bank market trades help banks avoid default but also that regulators can improve on market outcomes by imposing liquidity requirements.

The last part of Panetti´s essay examines models of bank runs with a Global Game approach that includes both the expectations and the insolvency causes for runs. The models have depositors that are imperfectly informed about the true state of the economy and the bank. The model generates insolvency bank runs when the information the depositors receive is sufficiently bad or, if the information is sufficiently good, a zero probability of a run. In between these two cases, there is, in equilibrium, some probability that there will be an expectations-driven bank run. In this model the literature shows that liquidity requirements solve the problem of expectations-driven runs but also that this policy must be complemented by central bank liquidity provision in agreement to the classic "lender of last resort" doctrine. A final note is that these policies may have the unintended consequence of increasing banks´ moral hazard. To avoid that problem it may be desirable not to fully prevent runs.

All in all, Panetti's essay examines the foundations and the instruments of public policies towards banks giving us a deeper understanding of the issues involved and of the need to balance competing goals when defining these policies.

A bottom-up approach for forecasting GDP in a data rich environment

Francisco Dias
Banco de Portugal

Maximiano Pinheiro
Banco de Portugal

António Rua
Banco de Portugal

July 2016

Abstract
In an increasingly data-rich environment, the use of factor models for forecasting purposes has gained prominence in the literature and among practitioners. In this article, we extend the work of Dias, Pinheiro and Rua (2015) by assessing the forecasting behaviour of factor models to predict several GDP components and investigate the performance of a bottom-up approach to forecast Portuguese GDP growth. We find supporting evidence of the usefulness of factor models and noteworthy forecasting gains when conducting a bottom-approach drawing on the main aggregates of GDP. (JEL: C22, C53)

Introduction

Over the past decades the stream of economic information available to policymakers increased enormously owing to the widespread development of statistical systems. Naturally, economic agents and in particular international organizations and central banks, as a general practice, follow these large information sets when assessing the ongoing economic developments and designing policy responses. In such a data rich environment, tackling an information set which can comprise hundreds of time series raises methodological challenges in terms of econometric modelling.

In particular, the use of factor models for forecasting purposes has become an increasingly widespread tool to forecast macroeconomic variables in a data rich environment. See, for example, Stock and Watson (1998, 2002a,b) and Giannone *et al.* (2008) for the United States, Marcellino *et al.* (2003) and Angelini *et al.* (2011) for the euro area, Artis *et al.* (2005) for the UK, Schumacher (2007, 2010, 2011) for Germany, Barhoumi *et al.* (2010) for France, de Winter (2011) and Reijer (2013) for the Netherlands, and for a cross-country study encompassing several European countries see Rünstler *et al.* (2009).

E-mail: fadias@bportugal.pt; mpinheiro@bportugal.pt; antonio.rua@bportugal.pt

Factor models are designed to reduce the overall size of the number of series in large datasets to a manageable scale so as to circumvent the issue of the curse of dimensionality. Basically, these models condense the informational content of large datasets in a set of few unobserved series, the common factors, which account for a sizeable fraction of the overall co-movements amongst the entire set of series. Hence, these factors capture in a parsimonious way the main features of the dataset and are therefore included as regressors in forecasting models in place of the original variables.

However, the part of the information other than the one conveyed by the small set of estimated factors is not considered in the forecasting equation. Such a procedure may disregard potentially relevant information for the variable to be forecasted or the forecast horizon under consideration.

To overcome this potential caveat Dias *et al.* (2010) proposed an alternative procedure. In particular, a targeted diffusion index was suggested that takes on board not only all the factors of the database but also their information content for the variable being forecasted and the forecast horizon. This index consists of a double weighted average of all the factors of the dataset that take into account both the explanatory power of each factor for the variable to be forecasted as well as the relative importance of the factor in capturing the co-movements of the series in the database. In an application for the United Sates, such an approach proved superior to the standard factor model in forecasting several macroeconomic variables. Furthermore this method has also been recently applied by Dias *et al.* (2015) for Portugal to forecast GDP growth providing encouraging results.

Herein, we extend the work by Dias *et al.* (2015) and assess the usefulness of the above-mentioned factor models to forecast GDP components.

Typically, two key reasons are pointed out to motivate the importance of such an assessment. First, forecasting the disaggregated components enhances a more comprehensive understanding of the aggregate. In many institutions, such as in central banks, it is crucial to have a full picture of the underlying economic developments so as to enhance the formulation of useful economic policies.

Second, there is the traditional debate on whether one should forecast an aggregate variable directly or indirectly by aggregating the forecasts of its components. In particular, the accuracy of forecasting directly the aggregate is compared with the performance obtained via a bottom-up approach. Previous empirical work along this line includes Fair and Shiller (1990) for the United States GNP, Zellner and Tobias (2000) for GDP growth in industrialized countries, Marcellino *et al.* (2003) for several euro area aggregates, Hubrich (2005) and Duarte and Rua (2007) for inflation in the euro area and Portugal, Esteves (2013) for euro area GDP, among others.

From a theoretical standpoint, it has been argued that if the data generating process is known for all components then the forecast obtained by aggregation of the disaggregated forecasts dominates in terms of forecasting performance (see, for example, Rose (1977), Tiao and Guttman (1980), Kohn (1982) and Lütkepohl (1984). However, in practice, the data generating processes are not known and instead must be specified and estimated. In such a context, the superiority of the bottom-up approach is no longer assured (see Lütkepohl, 1984). Hence, it boils down to an empirical issue.

In this article, we evaluate the relative forecasting performance of factor-augmented models to predict each individual GDP component. Thereafter, drawing on the forecasts for the GDP components, we assess if pursuing a bottom-up approach can deliver forecasting gains when predicting GDP growth.

Brief overview of factor models

Factor models were initially developed in the late seventies and early eighties by Geweke (1977), Sargent and Sims (1977) and Geweke and Singleton (1981). However the empirical applications of these models until the early nighties were confined to a handful set of variables.

Stock and Watson (1998, 2002a,b) and Forni *et al.* (2000, 2001, 2004, 2005) contributed to a large extent to this literature extending the possibility of applications of factor models to datasets comprising hundreds of series. In fact, James Stock and Mark Watson, in their seminal work, were the first to propose the use of the principal component method to estimate common factors in the presence of a large number of economic series. Due to its generality and overall simplicity, it became quite popular among forecasters. In particular, we will confine our exercise to the static factor framework. Note that the dynamic version of the factor model specification can be rewritten in an equivalent static factor formulation (see, for example, Stock and Watson, 2005). Moreover, for forecasting purposes, which is the focus of this article, such distinction is pretty much irrelevant as mentioned by Bai and Ng (2007).

In the static factor model the forecasting exercise comprises two stages. In the first step, which involves the estimation of the factors, the model specification assumes that each and every variable in the large dataset results as a sum of two components: the first which is a linear combination of a small set of latent unobserved static factors, which is common to all variables, and an idiosyncratic component which is specific to each variable. Under these circumstances it has been proposed to estimate the unobserved common factors relying on the principal components technique which is shown to provide a consistent estimator of the factor space under fairly general conditions.

The principal components are ordered according to their relative importance which reflects its ability to capture the common dynamics in the whole dataset. Typically, the few major top-ranked principal components capture a sizeable share of the comovements amongst the series in the dataset, and the relevant number of factors to be considered in the forecasting model in step two can be chosen relying on some information criteria (see, for example, Bai and Ng, 2002).

In the second step, the variable to be forecasted is projected on the set of factors obtained in the previous step (which corresponds to the so-called Diffusion Index (DI) model) and possibly on lags of the dependent variable (termed DI-AR model). In this setting, whichever variable is being forecasted the same set of factors is included in the forecasting equation, *i.e.*, the dependent variable plays no role in the selection of the factors to be considered in the regression.

Hence, the entire set of lower ranked factors are discarded, irrespectively of their information content for the variable to be forecasted and the forecasting horizon, which may lead to a loss of potentially useful information. Under these circumstances, the standard modelling procedure may prove to be limited for forecasting purposes since it does not take into account neither the specific variable to be forecasted nor the forecast horizon in the selection of the factors to be included in the forecasting equation.

To overcome this shortfall, Dias *et al.* (2010) suggested a Targeted Diffusion Index (TDI) in place of the above mentioned set of main factors in the second step of the standard factor model approach. The TDI is a convex linear combination of all the factors of the dataset, where the weights attached to each factor reflect both the share of the total variance captured by the factor and its correlation with the variable to be predicted at the relevant forecast horizon. Thus, this procedure takes into account the entire set of factors combined in a single index using both the information conveyed by the relative importance of the factors in the spirit of Stock and Watson and the information content of each factor for the specific variable and horizon to be predicted.

This approach avoids dismissing potentially relevant information contained in the dataset and tries to obtain a better match between the available data and the variable to be predicted. Dias *et al.* (2010) show that such a modelling strategy improves substantially the forecasting performance *vis-à-vis* the DI model for several US macroeconomic variables while Dias *et al.* (2015) find encouraging results for forecasting GDP growth in Portugal.

Forecasting GDP components

Preliminaries

Concerning the data, we resort to the updated dataset compiled by Dias *et al.* (2015) for the Portuguese economy which comprises 126 series. It includes both hard and soft data covering business and consumers surveys (43 series), retail sales (4 series), industrial production (7 series), turnover in industries and services (20 series), employment, hours worked and wage indices in industries and services (24 series), tourism nights spent in Portugal (3 series), car sales (3 series), cement sales, vacancies and registered unemployment (5 series), energy consumption (3 series), nominal exports and imports of goods (10 series), real effective exchange rate, Portuguese stock market index and ATM/POS series. All series are seasonally adjusted and, with the exception of survey data, taken in logarithms. As usual, the series are then differenced.

Regarding GDP and its components, the series are available from the Portuguese National Statistics Office (INE) as from the first quarter of 1995 up to the fourth quarter of 2015 on a seasonally adjusted basis and in real terms. We focus on the corresponding quarter-on-quarter rates of change and restrict the analysis to nowcasting and forecasting one-quarter ahead horizon as the gains of using factor-augmented models are relatively negligible for more distant horizons (see also Dias *et al.*, 2015).

The out-of-sample period runs from the first quarter of 2002 up to the fourth quarter of 2015. Such a long out-of-sample period allows us to put to test more thoroughly the relative performance of the competing models and allows us to consider a sub-sample analysis. In particular, we consider two sub-samples namely from 2002 Q1 up to 2007 Q4 and from 2008 Q1 up to 2015 Q4. The latter period corresponds to a challenging period by all standards as the Portuguese economy underwent marked macroeconomic changes since the latest economic and financial crisis.

As usual in this type of exercises, we consider as the benchmark a univariate autoregressive model with the lag order determined by standard BIC criteria in each round of the recursive expanding window exercise. In the case of the DI model, where the number of factors have to be determined ex-ante, we consider four factors as discussed extensively in Dias *et al.* (2015).

We present the Mean Squared Forecast Error (MSFE) for the benchmark model and the relative MSFE for the competing models which is computed as the ratio between the MSFE of the competing model to the MSFE of the benchmark. Hence, a relative MSFE lower than one means that the competing model outperforms the benchmark whereas if it is higher than one it is the opposite. Finally, to assess the statistical significance of the forecasting gains relative to the benchmark we compute the Diebold and Mariano (1995) test (in the tables *,**,*** denote statistical significance at 10, 5, and 1 per cent levels, respectively).

Results

i) Private consumption

Let us start with the analysis of the results for private consumption as a whole (see Table 1). For the entire out-of-sample period, both the DI and TDI models deliver statistically significant superior nowcasting performance. The gains *vis-à-vis* the benchmark are 26 per cent and 37 per cent, respectively. In terms of the behavior over the out-of-sample period, the gains are larger in the first part of the sample attaining 43 and 48 per cent, respectively. Although the gains are lower in the second part of the sample, they are still quite sizeable in particular in the case of the TDI model (33 per cent). Hence, the TDI model always presents the lowest relative MSFE for any sample period for nowcasting purposes.

Out-of-sample period	2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon	$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
Total						
AR model (MSFE)	0.009	0.010	0.005	0.006	0.012	0.014
DI model (Relative MSFE)	0.74*	1.03	0.57**	1.48	0.80	0.89
TDI model (Relative MSFE)	0.63**	0.73	0.52*	0.95	0.67*	0.66
Durables						
AR model (MSFE)	0.389	0.380	0.225	0.236	0.511	0.488
DI model (Relative MSFE)	0.73	0.92	0.86	1.13	0.68	0.84***
TDI model (Relative MSFE)	0.61**	0.88	0.64	1.39	0.60*	0.69**
Non-durables						
AR model (MSFE)	0.005	0.005	0.004	0.004	0.006	0.007
DI model (Relative MSFE)	0.88	1.16	0.42***	1.27	1.09	1.12
TDI model (Relative MSFE)	0.78	0.86	0.53**	0.71	0.89	0.91

TABLE 1. Private consumption forecasting results

When forecasting one quarter-ahead, as expected, the gains are lower. In particular, the DI model does not outperform the benchmark while the TDI model presents a relative gain of 27 per cent, albeit not statistically significant. In this case, the forecasting performance turns out to be better in the second part of the sample for both models with the TDI model standing out.

We now turn to the components of private consumption namely durable and non-durables. In the case of nowcasting durables consumption, the gains obtained with DI and TDI models are similar to those recorded for private consumption as a whole. While the DI model seems to perform relatively better in the second part of the sample, the TDI model presents a stable performance across the sub-samples. For the one-quarter ahead horizon, the forecasting behavior worsens and the gains are substantially reduced. Nevertheless, the gains in the second part of the sample are statistically significant.

Regarding non-durables consumption, both DI and TDI models outperform the univariate benchmark for nowcasting. However, the gains are smaller than those reported above. In terms of its behavior over time, both models present a much better performance in the first part of the sample than in the second. For the one-quarter ahead horizon, the DI model does not outperform the benchmark whereas TDI does but not by much.

Overall, the TDI model delivers the best forecasting performance, both for the aggregate and any of its components. The gains are noteworthy when nowcasting whereas for the one quarter-ahead horizon the improvements are substantially reduced. Furthermore, these findings are relatively robust across sample periods.

ii) Public consumption

In the case of public consumption expenditures, we find that both the DI and TDI models perform better when augmented with autoregressive components determined by standard information criteria (denoted as DI-AR and TDI-AR, respectively). As one can see from Table 2, the DI-AR is not able to outperform the simple univariate autoregressive model when nowcasting or forecasting one-quarter ahead. Furthermore, this holds true for any out-of-sample period.

Out-of-sample period	2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon	$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
AR model (MSFE)	0.004	0.009	0.000	0.000	0.008	0.016
DI-AR model (Relative MSFE)	1.05	1.00	1.06	1.08	1.05	1.00
TDI-AR model (Relative MSFE)	0.96	0.82	0.89	0.82	0.96	0.82

TABLE 2. Public consumption forecasting results

In contrast, the TDI-AR model always presents a relative MSFE lower than one whatever the horizon or sample period. In particular, the relative gain is higher for forecasting one-quarter ahead (namely 18 per cent) than for nowcasting purposes. However, none of these improvements appear as statistically significant.

Hence, these results seem to support the view that it is hard to improve significantly over a simple autoregressive model for public consumption (see also Esteves, 2013). One should mention that quarterly public consumption in Portugal presents a relatively smooth profile as it results typically from the quarterly distribution of annual figures.

iii) Investment

The results concerning investment are reported in Table 3. Both the DI and TDI models present noteworthy gains when nowcasting investment as a whole namely 30 and 40 per cent, respectively. Similar sizeable gains are found across sample periods with the TDI model always performing better than the DI model. For the one-quarter ahead horizon, both the DI and TDI models deliver a gain around 10 per cent. Such an improvement reflects the behavior in the first part of the sample where the gains are quite large since in the second part of the sample the performance is close to the benchmark.

Out-of-sample period	2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon	$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
Total						
AR model (MSFE)	0.212	0.186	0.102	0.093	0.295	0.256
DI model (Relative MSFE)	0.70**	0.91	0.76	0.55	0.68*	1.01
TDI model (Relative MSFE)	0.60**	0.90	0.52	0.46*	0.62**	1.02
Machinery and equipment						
AR model (MSFE)	0.853	0.996	0.190	0.193	1.351	1.598
DI model (Relative MSFE)	0.86	0.77	1.01	1.21	0.85	0.73
TDI model (Relative MSFE)	0.82	0.77	0.94	1.07	0.80	0.74
Transport equipment						
AR model (MSFE)	2.394	2.402	1.599	1.638	2.991	2.976
DI-AR model (Relative MSFE)	0.89	0.91	1.03	1.07	0.83	0.84*
TDI-AR model (Relative MSFE)	0.86	0.93	1.06	1.12	0.78*	0.86
Construction						
AR model (MSFE)	0.120	0.109	0.095	0.081	0.140	0.130
DI model (Relative MSFE)	1.02	1.02	1.00	0.72	1.03	1.17
TDI model (Relative MSFE)	0.87	1.09	0.58	0.82	1.02	1.21
Other						
AR model (MSFE)	4.614	5.227	4.890	5.493	4.406	5.027
DI model (Relative MSFE)	0.99	0.79*	0.90	0.67*	1.07	0.88
TDI-AR model (Relative MSFE)	1.02	0.86	1.13	0.78	0.93	0.93

TABLE 3. Investment forecasting results

We also consider the various components of investment namely machinery and equipment, transport material, construction and other investment. Starting with machinery and equipment, both DI and TDI models improve on the univariate autoregressive model while delivering similar gains for both nowcasting and forecasting one-quarter ahead (around 20 per cent). Moreover, the gains are basically concentrated in the second part of the out-of-sample period.

Regarding investment in transport equipment, the gains are more subdued (around 10 per cent) with both models presenting again a close relative performance for both horizons. The improvement is also more significant in the later part of the sample period.

Concerning construction, only the TDI model presents a gain for nowcasting as the DI model does not improve on the benchmark. For the one quarter-ahead horizon, none of the models deliver a gain for the out-of-sample period as whole. In terms of sub-samples, both models perform relatively better in the first part of the sample although none of the gains are statistically significant.

Finally, for other investment, there are basically no gains to report when nowcasting while we find some improvement when forecasting one quarter-ahead, reflecting to a large extent the performance in the first part of the sample.

In sum, factor models tend to perform better, in most cases, than the benchmark. However, while for investment as a whole, the gains are clearly noteworthy with the TDI model standing out, the results for investment components are relatively disappointing. This seems to suggest that factor models can be potentially more useful when forecasting larger aggregates and not so much when one is interested in very specific components. This seems natural, as factor models tend to exploit the main commonalities in the data and therefore less appropriate to pinpoint narrow components.

iv) Exports

In Table 4 we report the results for aggregate exports as whole as well as for both exports of goods and services separately. Concerning total exports, we find large gains in both DI and TDI models when nowcasting and to a lesser extent for the one-quarter ahead horizon, with the TDI model performing slightly better than the DI model. For both horizons and in the case of the two models the gains are larger in the second part of the sample although not statistically significant as the improvement seems to be concentrated in a small number of observations.

Out-of-sample period	2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon	$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
Total						
AR model (MSFE)	0.079	0.079	0.036	0.037	0.111	0.110
DI model (Relative MSFE)	0.45	0.86	0.65	0.96	0.40	0.84
TDI model (Relative MSFE)	0.42	0.80	0.68	1.00	0.36	0.76
Goods						
AR model (MSFE)	0.112	0.112	0.035	0.036	0.170	0.170
DI model (Relative MSFE)	0.48	0.95	0.85	1.24	0.42	0.90
TDI model (Relative MSFE)	0.42	0.86	0.84	1.11	0.35	0.82
Services						
AR model (MSFE)	0.087	0.081	0.139	0.131	0.048	0.043
DI model (Relative MSFE)	1.07	1.01	0.75	0.98	1.77	1.08
TDI model (Relative MSFE)	0.98	1.29	0.73	1.09	1.51	1.75

TABLE 4. Exports forecasting results

In terms of components, there are noteworthy differences between goods and services. In the case of services, factor models do not seem to outperform the autoregressive benchmark. For exports of goods, the assessment is similar to the one reported above for exports as a whole albeit the gains are slightly lower. Such evidence seems to suggest that not much is gained from considering disaggregated exports.

v) Imports

We also assessed the relative performance of factor models to forecast imports (see Table 5). For imports as a whole, we find large and statistically significant gains when nowcasting, with the TDI model performing once again better than the DI model. The results are even stronger when one focus on the second part of the out-of-sample period. For the one-quarter ahead horizon, the gains are smaller and more visible in the first part of the sample.

Out-of-sample period	2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon	h = 0	h = 1	h = 0	h = 1	h = 0	h = 1
Total						
AR model (MSFE)	0.100	0.095	0.049	0.054	0.138	0.126
DI model (Relative MSFE)	0.56*	0.78	0.83	0.64	0.48*	0.82
TDI model (Relative MSFE)	0.48**	0.77	0.65	0.69	0.43**	0.79
Goods						
AR model (MSFE)	0.126	0.114	0.053	0.052	0.180	0.160
DI model (Relative MSFE)	0.50**	0.81	0.71	0.70	0.46**	0.84
TDI model (Relative MSFE)	0.45**	0.78	0.56*	0.66	0.42**	0.81
Services						
AR model (MSFE)	0.154	0.156	0.165	0.177	0.147	0.140
DI model (Relative MSFE)	1.21	1.00	1.19	1.14	1.23	0.87
TDI model (Relative MSFE)	1.15	1.13	1.08	1.33	1.20	0.95

TABLE 5. Imports forecasting results

In terms of components, likewise exports, we also find that there are only gains when one focuses on goods since factor models do not improve on the autoregressive model in the case of services. For imports of goods, the assessment is broadly similar to the one reported for imports as a whole.

A bottom-up approach for GDP

Given the forecasts for the several components of GDP, we now turn to the question of whether forecasting GDP growth can be improved via a bottom-up approach or not. To assess if a bottom-up approach can deliver superior results than forecasting directly GDP growth, one has to pick a model to forecast it directly. Hence, we first evaluate the forecasting performance of the above considered models to forecast directly GDP growth, similarly to what has been done for each GDP component. In Table 6, we report the corresponding results.

Out-of-sample period	2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon	$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
AR model (MSFE)	0.008	0.008	0.007	0.006	0.008	0.010
DI model (Relative MSFE)	0.49**	1.00	0.42**	1.01	0.53	1.00
TDI model (Relative MSFE)	0.37***	0.73	0.23**	0.56**	0.47*	0.80

TABLE 6. GDP forecasting results

As in Dias *et al.* (2015), we find that factor models outperform the benchmark with the TDI model standing out.[1] For nowcasting, the gains achieved with the latter model are very large (63 per cent for the out-of-sample period as a whole) and statistically significant whatever the out-of-sample period considered. As expected, for the one-quarter ahead horizon the gains are smaller (27 per cent for the full out-of-sample period) and more pronounced in the first part of the sample. Hence, we will use the TDI model for forecasting directly GDP growth as the benchmark when evaluating the relative performance of the bottom-up approach for GDP.

In Table 7, we present the results for the bottom-up approach for GDP as well as for its main components. In particular, we first assess a bottom-up approach for each of the main aggregates of GDP, namely private consumption, investment, exports and imports. That is, we evaluate if forecasting each of the main GDP components directly is better than conducting the corresponding bottom-up approach. For instance, we analyze if forecasting directly private consumption is better than aggregating the forecasts of durables and non-durables using the corresponding national accounts weights. For each series we pick the overall best performing model based on the previous analysis, which we recall in the second column of Table 7 following the ordering of the discussion in the previous section.

1. These results do not correspond exactly to those reported in Dias *et al.* (2015) as quarterly national accounts have been revised by INE due to the adoption of ESA 2010 and the sample period herein considered has been extended.

Out-of-sample period		2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon		$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
	Models						
Private consumption							
Direct (MSFE)	1) TDI	0.006	0.007	0.003	0.005	0.008	0.009
Bottom-up (Relative MSFE)	2) TDI+TDI	1.03	1.01	1.14	1.05	1.00	0.99
Public consumption							
Direct (MSFE)	3) TDI-AR	0.004	0.008	0.000	0.000	0.007	0.013
Investment							
Direct (MSFE)	4) TDI	0.126	0.167	0.053	0.043	0.181	0.260
Bottom-up (Relative MSFE)	5) TDI+TDI-AR+TDI+DI	1.02	1.10	0.94	1.22	1.04	1.09
Exports							
Direct (MSFE)	6) TDI	0.033	0.063	0.024	0.036	0.040	0.083
Bottom-up (Relative MSFE)	7) TDI+AR	1.03	1.02	0.99	0.97	1.05	1.03
Imports							
Direct (MSFE)	8) TDI	0.048	0.073	0.032	0.037	0.060	0.100
Bottom-up (Relative MSFE)	9) TDI+AR	1.00	0.98	0.96	0.94**	1.01	0.99
GDP							
Direct (MSFE)	TDI	0.003	0.006	0.002	0.003	0.004	0.008
Bottom-up with main aggregates (Relative MSFE)	1)+3)+4)+6)+8)	0.79*	0.97	0.84	1.39	0.78	0.85*
Bottom-up with detailed components (Relative MSFE)	2)+3)+5)+7)+9)	0.85	1.21	0.98	2.05	0.81	0.96

TABLE 7. Forecasting via a bottom-up approach

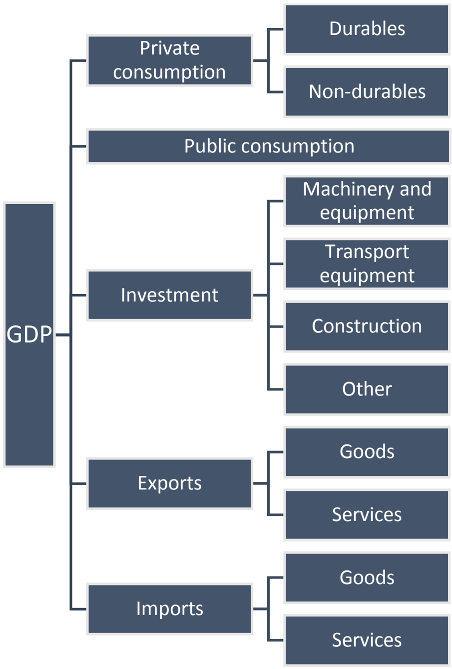


FIGURE 1: GDP disaggregation

We basically find that pursuing a bottom-up approach for each of the main GDP components does not improve the nowcasts neither the one-quarter ahead forecasts as the relative MSFE almost always exceeds one for the full out-of-sample period. This evidence seems to suggest that factor models are less useful when one focuses on relatively small components of GDP as they may be driven, to a large extent, by idiosyncratic forces. Naturally, given that factor models are designed to capture the main underlying common forces, the use of factor-augmented models tends to be less appropriate if one is interested in very narrow components of economic activity.

Regarding GDP, we consider two alternative levels of disaggregation for the bottom-up approach (see Figure 1). A level of disaggregation that considers the main aggregates of GDP (namely private consumption, public consumption, investment, exports and imports) and a higher level of disaggregation that uses more detailed GDP components (that is, consumption of durables, consumption of non-durables, public consumption, investment in machinery and equipment, transport equipment, construction, other investment, exports of goods, exports of services, imports of goods and imports of services).

We find that the bottom-up approach improves substantially the nowcasts of GDP growth.[2] In particular, the largest gains are attained when one considers the level of disaggregation that draws on the main aggregates of GDP. Such an approach delivers a statistically significant gain of 21 per cent over the best performing model for forecasting directly GDP growth. The gain is relatively stable across sub-samples.

For the one quarter-ahead horizon, the bottom-up approach that delivers the best results is once again the one that uses the main GDP components. Although it improves marginally *vis-à-vis* the direct forecast of GDP growth for the full out-of-sample period, there is a statistically significant gain of 15 per cent in the second part of the sample which is by all standards a very challenging period.

Hence, although a bottom-up approach does not seem to improve the forecasting performance in the case of the main aggregates of GDP, sizeable gains can be obtained when such an approach is pursued for GDP. This is particularly true when one considers a level of disaggregation based on the main GDP components. These results also reflect the fact that factor-augmented models are naturally more oriented to forecast broader aggregates of economic activity.[3]

Conclusions

Given the proved usefulness of factor-augmented models to forecast Portuguese GDP growth, we conducted a similar forecasting exercise to evaluate its performance in forecasting GDP components. As in the case of GDP, we find that factor models typically outperform the univariate autoregressive benchmark with the TDI model of Dias *et al.* (2010) standing out. Such an evidence reinforces the usefulness of the TDI approach in a broader context.

In particular, we find that the gains are larger when nowcasting and, as expected, decrease as one extends the forecast horizon. Moreover, we find that factor-augmented models are less useful when one forecasts relatively narrow components of GDP. In the presence of broadly based datasets, this seems a natural result as factor models tend to be more appropriate to capture the dynamics of broad GDP aggregates.

Drawing on the forecasts for the GDP components, we also assessed the forecasting behavior of the bottom-up approach *vis-à-vis* the direct approach.

2. We also considered the case where imports are modelled as a function of global demand weighted by the imported content (see Esteves *et al.*, 2013) but it did not lead to an improvement in the forecasting performance.

3. In the appendix, we also report the corresponding results considering the Mean Absolute Forecast Error (MAFE) as in Dias *et al.* (2015). The findings are qualitatively similar.

In this respect, we find significant gains when conducting a bottom-approach for GDP growth, in particular, when one considers a disaggregation level that draws on the main aggregates of GDP.

References

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011). "Short-term forecasts of euro area GDP growth." *Econometrics Journal*, 14(1), C25–C44.

Artis, M., A. Banerjee, and M. Marcellino (2005). "Factor forecasts for the UK." *Journal of Forecasting*, 24, 279–298.

Bai, J. and S. Ng (2002). "Determining the number of factors in approximate factor models." *Econometrica*, 70, 191–221.

Bai, J. and S. Ng (2007). "Determining the number of primitive shocks in factor models." *Journal of Business & Economic Statistics*, 25, 52–60.

Barhoumi, K., O. Darné, and L. Ferrara (2010). "Are disaggregate data useful for factor analysis in forecasting French GDP?" *Journal of Forecasting*, 29(1-2), 132–144.

de Winter, J. (2011). "Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models." DNB Working Papers 320, Netherlands Central Bank.

Dias, F., M. Pinheiro, and A. Rua (2010). "Forecasting using targeted diffusion indexes." *Journal of Forecasting*, 29(3), 341–352.

Dias, F., M. Pinheiro, and A. Rua (2015). "Forecasting Portuguese GDP with factor models: Pre- and post-crisis evidence." *Economic Modelling*, 44(C), 266–272.

Diebold, F. and R. Mariano (1995). "Comparing predictive accuracy." *Journal of Business & Economic Statistics*, 13, 253–263.

Duarte, C. and A. Rua (2007). "Forecasting inflation through a bottom-up approach: how bottom is bottom?" *Economic Modelling*, 24, 941–953.

Esteves, P. S. (2013). "Direct vs bottom–up approach when forecasting GDP: Reconciling literature results with institutional practice." *Economic Modelling*, 33, 416–420.

Esteves, P.S., F. Cardoso, and A. Rua (2013). "The import content of global demand in Portugal." Economic Bulletin, Autumn, 107-121, Banco de Portugal.

Fair, R. and J. Shiller (1990). "Comparing information in forecasts from econometric models." *American Economic Review*, 80, 375–389.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). "The Generalized Dynamic-Factor Model: Identification And Estimation." *The Review of Economics and Statistics*, 82(4), 540–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2001). "Coincident and Leading Indicators for the Euro Area." *Economic Journal*, 111(471), C62–85.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2004). "The generalized dynamic factor model consistency and rates." *Journal of Econometrics*, 119(2), 231–255.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting." *Journal of* *the American Statistical Association*, 100, 830–840.

Geweke, J. (1977). "The Dynamic Factor Analysis of Economic Time Series." In *Latent variables in socio-economic models*, edited by D. J. Aigner and A. S. Goldberger, chap. 19. North-Holland.

Geweke, J. and K. Singleton (1981). "Maximum likelihood 'confirmatory' factor analysis of economic time series." *International Economic Review*, 22(1), 37–54.

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

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

Kohn, R. (1982). "When is an aggregate of a time series efficiently forecast by its past?" *Journal of Econometrics*, 18, 337–349.

Lütkepohl, H. (1984). "Forecasting Contemporaneously Aggregated Vector ARMA Processes." *Journal of Business & Economic Statistics*, 2(3), 201–214.

Marcellino, M., J. Stock, and M. Watson (2003). "Macroeconomic forecasting in the euro area: country specific versus euro wide information." *European Economic Review*, 47, 1–18.

Reijer, A. (2013). "Forecasting Dutch GDP and inflation using alternative factor model specifications based on large and small datasets." *Empirical Economics*, 44(2), 435–453.

Rose, D.E. (1977). "Forecasting aggregates of independent ARIMA processes." *Journal of Econometrics*, 5, 323–345.

Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth, and C. Van Nieuwenhuyze (2009). "Short-term forecasting of GDP using large datasets: A pseudo real-time forecast evaluation exercise." *Journal of Forecasting*, 28(7), 595–611.

Sargent, T. J. and C. A. Sims (1977). "Business Cycle Modelling Without Pretending to Have too Much A-Priori Economic Theory." In *New methods in business cycle research*, edited by C. Sims *et al.* Federal Reserve Bank of Minneapolis.

Schumacher, C. (2007). "Forecasting German GDP using alternative factor models based on large datasets." *Journal of Forecasting*, 26(4), 271–302.

Schumacher, C. (2010). "Factor forecasting using international targeted predictors: The case of German GDP." *Economics Letters*, 107(2), 95–98.

Schumacher, C. (2011). "Forecasting with Factor Models Estimated on Large Datasets: A Review of the Recent Literature and Evidence for German GDP." *Journal of Economics and Statistics*, 231(1), 28–49.

Stock, J. and M. Watson (1998). "Diffusion Indexes." NBER Working Paper 6702, National Bureau of Economic Research.

Stock, J. and M. Watson (2002a). "Forecasting using principal components from a large number of predictors." *Journal of the American Statistical*

Association, 97(460), 1167–1179.

Stock, J. and M. Watson (2002b). "Macroeconomic forecasting using diffusion indices." *Journal of Business & Economic Statistics*, 20(2), 147–162.

Stock, J. and M. Watson (2005). "Implications of dynamic factor models for VAR analysis." NBER Working Paper 11467, National Bureau of Economic Research.

Tiao, G. C. and I. Guttman (1980). "Forecasting contemporal aggregates of multiple time series." *Journal of Econometrics*, 12, 219–230.

Zellner, A. and J. Tobias (2000). "A Note on aggregation, disaggregation and forecasting performance." *Journal of Forecasting*, 19(5), 457–469.

Appendix

Out-of-sample period		2002Q1-2015Q4		2002Q1-2007Q4		2008Q1-2015Q4	
Forecast horizon		$h = 0$	$h = 1$	$h = 0$	$h = 1$	$h = 0$	$h = 1$
	Models						
Private consumption							
Direct (MAFE)	1) TDI	0.55	0.64	0.38	0.57	0.69	0.69
Bottom-up (Relative MAFE)	2) TDI+TDI	1.01	1.01	1.06	1.03	0.99	0.99
Public consumption							
Direct (MAFE)	3) TDI-AR	0.37	0.52	0.07	0.16	0.59	0.78
Investment							
Direct (MAFE)	4) TDI	2.72	3.05	1.71	1.68	3.48	4.08
Bottom-up (Relative MAFE)	5) TDI+TDI-AR+TDI+DI	1.00	1.02	1.02	1.15	0.99	0.98
Exports							
Direct (MAFE)	6) TDI	1.42	1.90	1.09	1.45	1.66	2.24
Bottom-up (Relative MAFE)	7) TDI+AR	1.02	0.98	1.01	0.94	1.03	1.00
Imports							
Direct (MAFE)	8) TDI	1.56	2.09	1.28	1.46	1.78	2.57
Bottom-up (Relative MAFE)	9) TDI+AR	0.96**	0.98	0.93**	0.96***	0.98	0.99
GDP							
Direct (MAFE)	TDI	0.41	0.62	0.31	0.49	0.49	0.71
Bottom-up with main aggregates (Relative MAFE)	1)+3)+4)+6)+8)	0.87**	0.98	0.91	1.12	0.86*	0.91
Bottom-up with detailed components (Relative MAFE)	2)+3)+5)+7)+9)	0.95	1.07	1.01	1.32	0.92	0.94

TABLE 8. MAFE results for the bottom-up approach

Portugal: Trends, cycles, and instability in output and unemployment over 2008–2012

José R. Maria
Banco de Portugal

July 2016

Abstract
This article presents a trend-cycle decomposition of Portuguese Gross Domestic Product and unemployment over 2008–2012. Results show that product and labour markets were primarily marked by low frequency movements in the trend component, and less so by cyclical factors. Economic policy should therefore not neglect the structural properties of these markets, resting solely centered around standard business cycle objectives. Okun's law—the negative correlation between the output and unemployment gaps—remained empirically relevant, but not without noteworthy trend instability. All results are based on a semi-structural model with rational expectations, tailored for a small economy integrated in a credible monetary union. (JEL: C51, E32, F45)

Introduction

Portugal experienced an unstable 2008–2012 period, marked in 2011 by the request for international financial assistance, agreed with the European Union (EU), and the International Monetary Fund (IMF).

Gross Domestic Product (GDP) fell around 10% over 2008–2012, while unemployment soared, reaching 16.7% of the labour force. Behind such dramatic events are, among other reasons, (i) spillover effects from the international financial crisis, which intensified in the second half of 2008; (ii) co-movements in sovereign risk hikes across vulnerable euro area countries (Ireland, Greece, Cyprus, Italy, Spain); (iii) the need to reduce macroeconomic imbalances; and (iv) sudden stops in credit flows, which intensified financial fragmentation.

The sharp deterioration in product and labour market conditions, possibly interacting with financial factors and high credit spreads, calls for a model-based assessment of such developments: What drove such events? Was it

Acknowledgements: I am indebted to Pierre Lafourcade, who wrote a substantial part of the model's code. I thank João Amador, António Antunes, Isabel Horta Correia and Sara Serra for helpful discussions.

E-mail: jrmaria@bportugal.pt

a cyclical downturn, motivated by large and persistent negative demand shocks, partially imported, or the result of deeper structural problems? What was the relative importance of these disturbances? How to monitor and assess such events analytically? How did standard textbook's macro-modelling strategies behave under such extreme events? In particular, what happened to Okun's law (the negative correlation between output and unemployment gaps)?

This article discusses, on the one hand, the relative importance of several disturbances using a semi-structural model with rational expectations. On the other hand, it evaluates Okun's law robustness throughout the 2008–2012 period. The discussion takes into account the results of a multivariate filter named herein, for ease of reference, "*Model Q*." Key theoretical references are Carabenciov *et al.* (2013) and European System of Central Banks (2016). The current version is tailored for a small economy integrated in the credible monetary union, where the risk-free nominal interest rate is set by the monetary authority of the model—in this case the European Central Bank (ECB). It includes several innovations relative to standard approaches, namely identical long-run restrictions in both the small economy and the rest of the monetary union (identical long-run growth rates in the trend component of output; unemployment rate levels; and real interest rates). The model lacks microfoundations, although each behavioural equation is a fairly standard textbook's equation with an economic interpretation (Berg *et al.* 2006), namely a policy equation defining official interest rates' responses, an inflation equation, an output equation and a version of Okuns' law. All shocks are stochastic and orthogonal. Some are labelled demand, supply and monetary policy shocks. For simplicity, those affecting trend components are grouped under the designation of "non-cyclical disturbances." *Model Q* embeds unobserved components and is estimated with Bayesian techniques.

The main result suggests that Portuguese product and labour markets were mainly hit by low frequency developments in trends, and less so by cyclical factors. The economy was nevertheless hit by other adverse shocks, notably the two exogenous recessive periods in the euro area, and abnormal sovereign risk hikes. This outcome complements the results reported by Castro *et al.* (2014). The increase in the trend component of the unemployment rate confirms the results obtained by Centeno *et al.* (2009), although current estimates are more volatile and depict a steeper outcome.

Model Q substantiates a decrease in the level of the trend component of Portuguese output, over 2008–2012, in line with other methodologies. Okun's law remained empirically relevant, however, not without noteworthy trend instability—evaluated by (pseudo) real time estimates. It should be emphasized that the current version of the model is silent about all economic forces driving trends. They are simply given by highly flexible stochastic processes.

This article has the following structure: section 2 sketches the model focusing solely on the main equations for Portugal. Model-based decompositions of output and unemployment rates are reported in Section 3. The instability of Okun's law is evaluated in Section 4. Section 5 concludes, puts forward tentative policy implications, and possible ways to extend the model.

A two-country model for a small euro area economy

Model Q considers two regions: a small euro area economy—in this case Portugal—and the rest of the monetary union. The model mixes stringent and rigid ingredients with relatively flexible elements, although the small economy is effectively "tying its hands" with the rest of the union (an expression from Giavazzi and Pagano (1988)). A central ingredient is the assumption of a credible monetary union. This restriction implies that the nominal exchange rate is a credible institutional feature, expected to remain fixed, and that the ECB sets nominal interest rates in line with a fully credible long-run inflation target, set herein at 2.0%. Short- and medium-run inflation expectations may deviate from target, depicting high persistence, but not long-run expectations, when all shocks' impact have dissipated.

The ECB mandate in *Model Q* is translated into a policy function that only reacts to developments in euro area aggregates, an assumption that can also be found in micro-founded general equilibrium models, *e.g.* PESSOA (Almeida *et al.* 2013). The trend component of the real interest rate, which may deviate from a fixed long-run benchmark, is also determined solely by euro area data, and assumed identical in both regions.

Among the flexible elements, a special focus should be placed on all trend components of product and labour markets. In addition, short and medium-run real interest rates in the two regions may differ substantially, and persistently, due to region-specific inflation expectations, while price differentials may have long-lasting effects on real exchange rates. Nominal interest rates can drift apart due to an exogenous risk premium.

Behavioural and a-theoretical equations for Portugal

This section briefly presents the core set up for Portugal.[1] With the exception of nominal interest rates, all other variables have functional forms expressed in "gaps," *i.e.* in deviations from unobserved trends (identified with a "$\sim$"). Euro area aggregates are identified with a "$*$."

1. A comprehensive assessment of the model, including all estimation results, can be found in Maria (2016).

Okun's law associates herein unemployment gaps at quarter t, namely $u_{gap,t} = u_t - \tilde{u}_t$, to its own lead and lagged values, and to the output gap, $y_{gap,t-1} = y_{t-1} - \tilde{y}_{t-1}$. More precisely,

$$(1 + \alpha_1\alpha_2)u_{gap,t} = \alpha_1 u_{gap,t-1} + \alpha_2 u_{gap,t+1} - \alpha_3 y_{gap,t-1} + \varepsilon_{u_{gap},t}, \quad (1)$$

where u_t is the Portuguese unemployment rate, y_t is actual GDP data, and $\varepsilon_{u_{gap},t}$ is an idiosyncratic disturbance. The trend component of unemployment embodies a fixed term, u, shared by both Portugal and the euro area; $\tilde{u}_t = \rho_u u + (1-\rho_u)\tilde{u}_{t-1} + \tilde{u}_{g,t}$, where $\tilde{u}_{g,t}$ is an autoregressive process with its own disturbance $\varepsilon_{\tilde{u},t}$. The presence of lagged values captures labour market frictions, while lead values introduces more flexibility in the model by allowing for expectations to also play a role.[2]

The inflation equation of *Model Q* associates current price changes to lagged and expected inflation, the output gap, and to changes in the real exchange rate. More precisely,

$$\begin{aligned}(1 + \lambda_1\lambda_2)(\pi_t - \pi) = \lambda_1(\pi 4_{t-1} - \pi) + \lambda_2(\pi 4_{t+4} - \pi) + \\ + \lambda_3 y_{gap,t-1} + \lambda_4 \pi 4_{q,t-1} - \varepsilon_{\pi,t},\end{aligned} \quad (2)$$

where $\pi = 2.0\%$ is the long-run inflation anchor. Variables $\pi 4_t$ and $\pi 4_{q,t}$ measure year-on-year changes in consumer prices and in the real exchange rate, respectively (an increase in $\pi 4_{q,t}$ represents a real depreciation). Disturbance term $\varepsilon_{\pi,t}$ is labelled "supply shock." The associated negative sign ensures that a positive supply shock is consistent with downward inflation pressures, as in Carabenciov *et al.* (2013).

The output equation includes the real interest rate gap, $r_{gap,t} = r_t - \tilde{r}^*_t$, the foreign output gap, $y^*_{gap,t-1}$, and the real exchange rate gap, $q_{gap,t} = q_t - \tilde{q}_t$. More precisely,

$$\begin{aligned}(1 + \beta_1\beta_2)y_{gap,t} = \beta_1 y_{gap,t-1} + \beta_2 y_{gap,t+1} - \beta_3 r_{gap,t-1} + \\ + \beta_4 y^*_{gap,t-1} + \beta_5 q_{gap,t-1} + \varepsilon_{y_{gap},t},\end{aligned} \quad (3)$$

where $\varepsilon_{y_{gap},t}$ is a disturbance term henceforth labelled "domestic shock."

It should be noted that $i_t = i^*_t + \psi_t$ and i^*_t are nominal interest rates (the latter set by the ECB), where $\psi_t = \rho_i\psi_{t-1} + \varepsilon_{i,t}$ is an exogenous risk premium, $0 < \rho_i < 1$ and $\varepsilon_{i,t}$ is a i.i.d risk premium shock. The evolution of these exogenous variables over 1999Q1 and 2015Q2 are depicted in Figure 1. In addition, $r_t = i_t - \pi_{t+1}$ is the real interest rate; $\pi_{t+1} = 4(p_{t+1} - p_t)$ measures expected inflation conditional on information up to period t; $q_t = p^*_t - p_t$ is the real exchange rate, computed with Harmonized Indices of Consumer

2. A general equilibrium model where the unemployment-inflation relationship considers current, lagged, and future unemployment can be found in Ravenna and Walsh (2008).

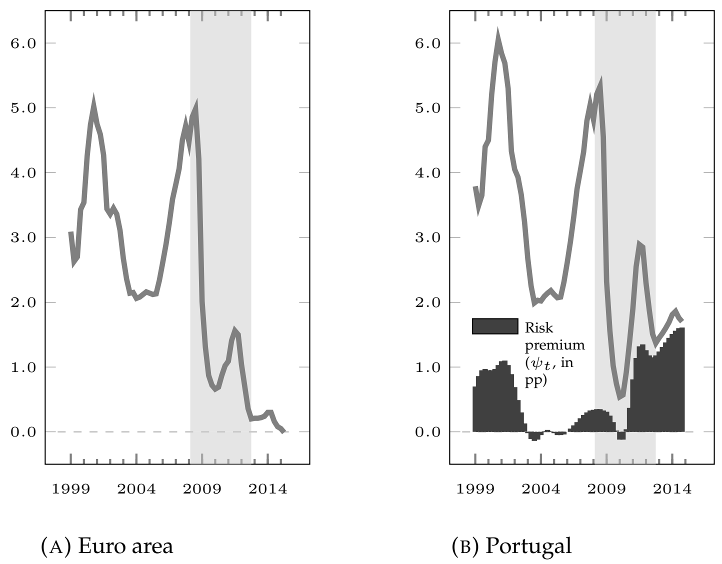


FIGURE 1: Nominal interest rates (%)

Source: Banco de Portugal, Eurostat and own calculations.

Notes: Interest rates of the euro area i_t^* are given by ECB's official interest rates. In the Portuguese case they are given by $i_t = i_t^* + \psi_t$, where ψ_t is an exogenous risk premium computed as in Castro *et al.* (2014). The shaded area identifies the 2007Q4-2012Q4 period.

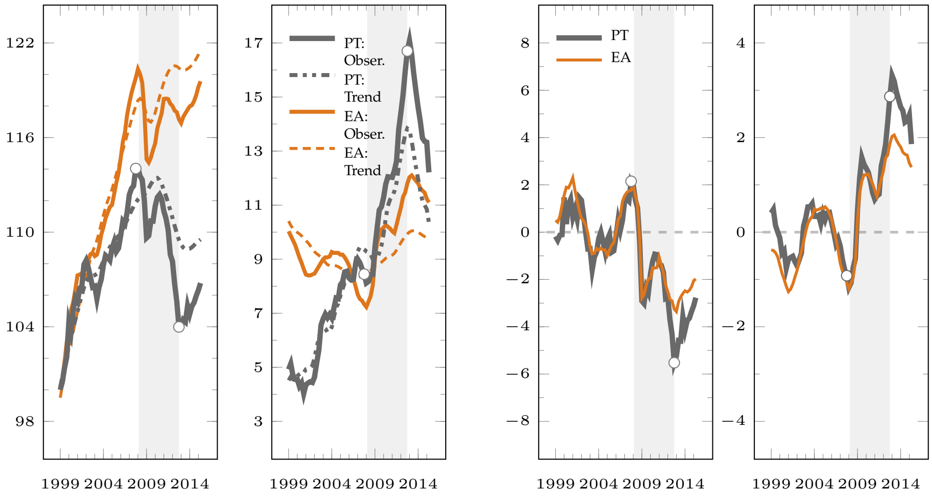
Prices for the euro area and Portugal, respectively. The real exchange rate gap $q_{gap,t}$ follows an autoregressive processes with disturbance term $\varepsilon_{q_{gap},t}$, and the trend component $\tilde{q}_t$ is modelled as a random walk with disturbance $\varepsilon_{\tilde{q},t}$. As in Carabenciov *et al.* (2013), the trend component of the real interest rate is assumed to evolve around a fixed benchmark r, namely $\tilde{r}_t^* = \rho_{\tilde{r}}^* \ r + (1 - \rho_{\tilde{r}}^*) \ \tilde{r}_{t-1}^* + \varepsilon_{\tilde{r},t}^*$. The trend component of output includes a long-run fixed term, y_g, shared by both Portugal and the euro area, $\tilde{y}_t = \ \tilde{y}_{t-1} + y_g + \tilde{y}_{g,t}$, where $\tilde{y}_{g,t}$ is an autoregressive process with disturbance $\varepsilon_{\tilde{y},t}$.

Finaly, the interest rate equation is given by

$$i_t^* = \gamma_1 i_{t-1}^* + (1-\gamma_1)\left[(\tilde{r}_t^* + \pi 4_{t+4}^*) + \gamma_2(\pi 4_{t+4}^* - \pi) + \gamma_3 y_{gap,t-1}^*\right] + \varepsilon_{i^*,t}. \quad (4)$$

This equation clarifies that the ECB sets nominal interest rates i_t^* by only reacting to developments in euro area aggregates. Changes in i_t^*, however, have a direct impact on Portuguese nominal and real interest rates.

The estimation period of *Model Q* ends in 2015Q2. Ideally, the information set should begin after the inception of the euro. However, given that 1999Q1-2015Q2 is relatively short and plagued by an unprecedented economic crisis, the information set was extended backwards until 1995Q1, which allows for 82 observations over 1995Q1-2015Q2. Results for 1995Q1-1998Q4 are ignored.



(A) GDP (B) Unemployment (C) Ouput gap (D) Unempl. gap

FIGURE 2: Observed variables and trends

Source: Banco de Portugal, Eurostat and own calculations.

Notes: Output is in logs and normalized to GDP=100 in 1999Q1. Unemployment is in percentage of the labour force. Output gaps are in percentage, and unemployment gaps in percentage points (pp). Portugal and the euro area are identified with PT and EA, respectively. The shaded area identifies the 2007Q4-2012Q4 period.

Trends and cycles

Figure 2 depicts actual and trend components of output and unemployment rates in Portugal, as well as the implied output and unemployment gaps.

Results suggest that actual output was above trend by 2007Q4, around 2%, but rapidly moved below trend as the international financial crisis gained momentum. Actual and trend levels came closer around 2011, but only briefly. This period is marked by the beginning of a persistent downward movement in both actual and trend components of output. The model flexibility can thus easily accommodate a positive long-run growth rate that is common to both regions—estimated to be around 1.8%—,with domestic unobserved short-run rates that are persistently negative. The downward movement came to an halt by 2013, and thus outside the period under analysis.

The trend component of the Portuguese unemployment rate is marked by a sharp upward movement almost over the entire sample period. It only recedes outside the period under analysis. Its behaviour is in general consistent with the view that the Portuguese labour market was not only fundamentally unprepared to cope with the crisis, featuring a worrisome institutional architecture before the crisis (Centeno *et al.* 2009). Trend levels are highly volatile, namely in comparison with the results reported by Centeno *et al.* (2009). This difference is not a surprise since the current version of *Model Q* embodies no *ex ante* restriction on the volatility behind developments in trend components, in contrast with Centeno *et al.*. The estimated volatility is only respecting an a-theoretical law of motion that, among other effects, does not have associated economic factors nor isolates undesirable impacts.[3]

In comparison with the euro area, there are signs of similarities, and signs of sharp differences. Both output and unemployment gaps reveal high synchronicity. The linear correlation coefficients between output gaps (Figure 2c) or unemployment gaps (Figure 2d) over 1999Q1-2015Q2 are close to 0.9. The Portuguese data is more volatile: the standard deviation of the unemployment and output gaps stand at 1.9 and 1.2, respectively, which compares with 1.7 and 1.0 in the euro area. The results are consistent with the view that the crisis left visible marks in both regions, although the differences are quite impressive by 2012Q4. The larger output gap in the euro area was close to 3% in absolute terms, while the Portuguese was close to 5%. Developments in trend levels in the two regions show sharper differences, although the assumed structure from which they are estimated is identical. In product markets, the first euro area recession coincides with an abrupt reduction in the trend component that does not occur in Portugal. During 2012—the second recessive period in the euro area—the euro area showed a relatively minor decrease in trend levels, while Portugal maintained a persistent decline. The differences between the two regions are also visible in the trend component of the unemployment rate, which depicts an initial downward trend in the euro area, before the crisis inception, in contrast with the Portuguese case. During 2008Q1-2012Q4, the increase registered in the euro area is much smaller than in Portugal.

Historical decompositions over 2008–2012

Table 1 quantifies the contributions of each shock to output between 2007Q4 and 2012Q4. It disaggregates actual data between domestic factors and other factors, the latter including the contribution of monetary policy shocks (ε_i^*). The sum of all contributions equals actual data. Domestic shocks include demand (stemming from $\varepsilon_{y_{gap}}$), supply (ε_π), non-cyclical (which aggregate $\varepsilon_{\tilde{u}}$, $\varepsilon_{\tilde{y}}$ and $\varepsilon_{\tilde{q}}$), and risk premium shocks (ε_i). Shocks linked to foreign factors feature a similar structure. The contributions associated with $\varepsilon_{q_{gap},t}$ and $\varepsilon_{u_{gap},t}$ are included in "Other factors: Rest".

3. It fully ignores, for instance, the series break in Labour Force Survey statistics that took place in 2011. In this year, a period when trend component estimates increase sharply, Statistics Portugal introduced a new data collection scheme (associated to the use of telephone interviews; questionnaire changes; and new field work supervision technologies).

	Portugal: Output			Euro Area: Output		
	2007Q4	2012Q4	Δ	2007Q4	2012Q4	Δ
Actual data	30.2	20.1	-10.1	28.6	26.0	-2.6
Domestic factors						
Demand ($\varepsilon_{y_{gap}}$)	0.7	-1.5	-2.2	0.0	0.0	0.0
Supply (ε_{π})	0.0	0.3	0.2	0.0	0.0	0.0
Non-Cyclical	-4.8	-16.5	-11.6	0.0	0.0	0.0
Labour market ($\varepsilon_{\tilde{u}}$)	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
Output market ($\varepsilon_{\tilde{y}}$)	*-4.8*	*-16.5*	*-11.6*	*0.0*	*0.0*	*0.0*
Rest	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
Risk premium (ε_{i})	-0.3	-1.2	-0.9	0.0	0.0	0.0
Other factors						
Foreign factors	1.7	-3.0	-4.7	6.4	-5.2	-11.6
Demand ($\varepsilon^*_{y_{gap}}$)	*1.7*	*-2.7*	*-4.4*	*1.8*	*-3.2*	*-5.0*
Supply (ε^*_{π})	*0.0*	*-0.4*	*-0.4*	*0.0*	*-0.1*	*0.0*
Non-Cyclical	*0.0*	*0.1*	*0.1*	*4.7*	*-1.9*	*-6.6*
Rest	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
Monetary Policy (ε^*_{i})	0.0	-0.1	-0.1	0.0	0.0	0.0
Rest	32.9	42.1	9.3	22.1	31.1	9.0

TABLE 1. Decomposition of output over 2007Q4-2012Q4

Source: Own calculations.

Notes: Actual data is in logs and differ from the observed by a constant. The sum of all contributions equals actual data. Real exchange rate shocks $\varepsilon_{\tilde{q}}$ are included in “Non-cyclical: Rest”, whereas $\varepsilon_{q_{gap},t}$ are in “Other factors: Rest”. The component "Other factors: Rest" also includes the growth rate y_g.

Over the period 2008–2012, the most significant domestic shock driving the fall in output is the non-cyclical shock. The contribution reached -11.6 pp. Among the remaining domestic shocks, demand played a more important role than supply shocks, although the nominal side of the economy recorded significant changes.[4] Domestic demand shocks accounted for -2.2 pp. Finally, the increase in sovereign risk premium is estimated to have subtracted output by 0.9 pp.

Results suggest that Portugal was also significantly affected by the two recessive periods in the euro area. Foreign factors amounted to -4.7 pp over 2008–2012. The importance of the negative foreign shocks is consistent with real impacts computed by Castro *et al.* (2014), following the sharp contraction in the Portuguese external demand. The negative contribution reported herein gained momentum during 2011 and lasted until late 2012.

The contribution of monetary policy shocks is virtually nil in both regions, while the aggregator “Other factors: Rest” reached 9.3 pp, influenced by the

4. In 2009, the reduction in inflation was largely unexpected in both regions. In addition, inflation expectations remained systematically below 2% over the last part of the sample period (see Maria (2016)).

	Portugal: Unemployment rate			Euro Area: Unemployment rate		
	2007Q4	2012Q4	Δ	2007Q4	2012Q4	Δ
Actual data	-1.5	6.7	8.3	-2.7	1.8	4.5
Domestic factors						
Demand ($\varepsilon_{y_{gap}}$)	0.0	0.3	0.3	0.0	0.0	0.0
Supply (ε_{π})	0.0	-0.2	-0.1	0.0	0.0	0.0
Non-Cyclical	2.3	6.2	3.9	0.0	0.0	0.0
Labour market ($\varepsilon_{\tilde{u}}$)	*2.3*	*6.2*	*3.9*	*0.0*	*0.0*	*0.0*
Output market ($\varepsilon_{\tilde{y}}$)	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
Rest	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
Risk (ε_{i})	0.1	0.8	0.6	0.0	0.0	0.0
Other factors						
Foreign factors	-1.0	1.8	2.8	-3.8	0.8	4.6
Demand ($\varepsilon^{*}_{y_{gap}}$)	*-1.0*	*1.6*	*2.6*	*-1.0*	*1.8*	*2.8*
Supply (ε^{*}_{π})	*0.0*	*0.2*	*0.2*	*0.0*	*0.0*	*0.0*
Non-Cyclical	*0.0*	*0.0*	*0.0*	*-2.8*	*-1.1*	*1.7*
Rest	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.1*
Monetary Policy (ε^{*}_{i})	0.0	0.0	0.0	0.0	0.0	0.0
Rest	-2.9	-2.2	0.7	1.2	1.0	-0.2

TABLE 2. Decomposition of the unemployment rate over 2007Q4-2012Q4

Source: Own calculations.

Notes: Actual data differ from the observed by a constant. Real exchange rate shocks $\varepsilon_{\tilde{q}}$ are included in “Non-cyclical: Rest”, whereas $\varepsilon_{q_{gap},t}$ are in “Other factors: Rest”. The component "Other factors: Rest" also includes the contribution of $\varepsilon_{u_{gap},t}$.

impact of the long-run growth rate y_g. Note also that the table’s upper-right region of zeros respects the working hypothesis that Portuguese shocks have no effect on the euro area.

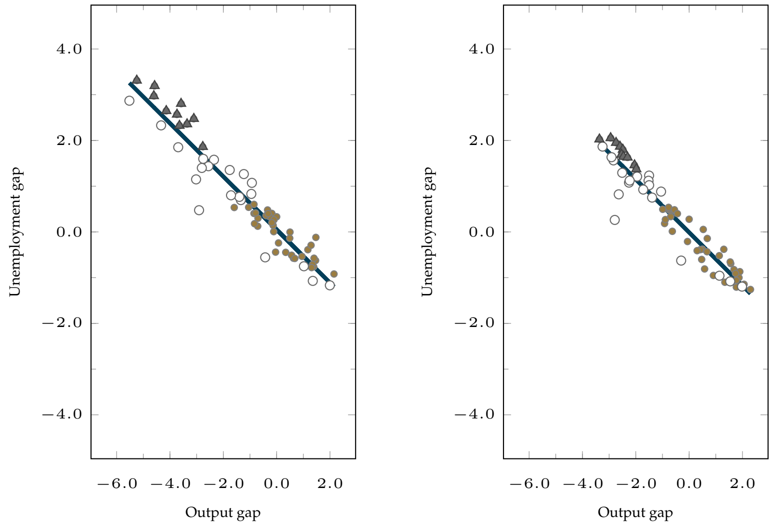
This paper fails to associate a large importance to real exchange rate shocks (included in the aggregate “Rest” of the domestic factors). Its virtually nil contribution may nevertheless suggest that the relative price of final consumption goods may not be a meaningful competitivenesses variable, and that further work is needed to create a more useful concept.

Table 2 reports the results for the unemployment rate. The outcome is qualitatively identical to that already disclosed for output, basically explained by the presence of an Okun’s law. Over the period 2008–2012, the non-cyclical shock is the most significant shock driving the upward movement in the unemployment rate.

Okun’s law over 2008–2012

This section evaluates the behaviour of Okun’s law over 2008–2012, and assesses the stability of trend components.

Figures 3a and 3b depict static representations of unemployment and output gaps. These scatter plots reorganize Figures 2c and 2d, which are



(A) Portugal

(B) Euro area

FIGURE 3: Okun's law

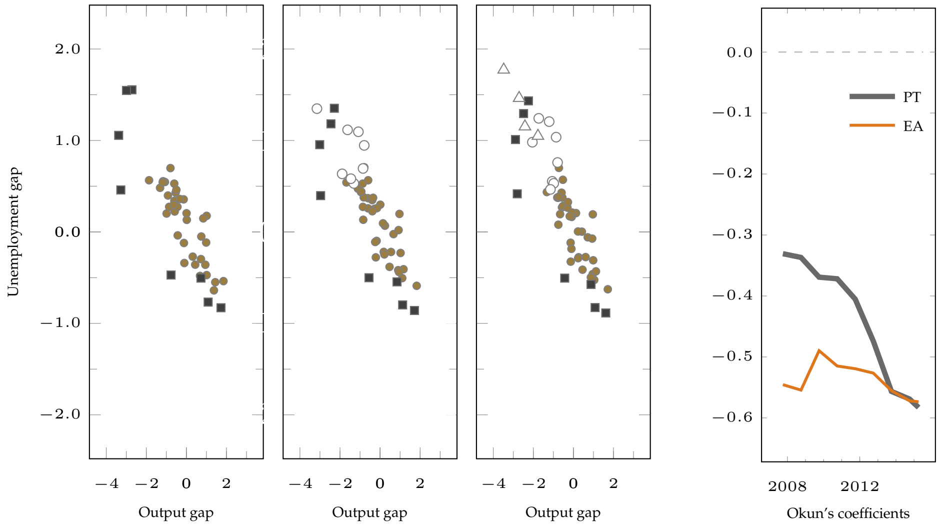
Source: Banco de Portugal, Eurostat and own calculations.

Notes: White dots cover the 2008Q1–2012Q4 period. Black triangles cover the 2013Q1-2015Q2 period.

functionally determined by the dynamic versions of Okun's law (defined in the Portuguese case by equation (1)).

Results suggest a relatively close relationship between unemployment and output gaps in both Portugal and the euro area. Over 2008–2012, the data points have basically moved from positive output gaps towards larger and larger negative output gaps in both regions (given by the white dots), with unemployment gaps depicting a mirror image. The subsequent period is evaluated by the model as a gradual movement backwards (the black triangles). These static relationships share another remarkable similarity: if the output gap increases by 1%, the unemployment gap decreases by 0.6 pp both in Portugal and in the euro area.

Figures 3a and 3b use all information up to 2015Q2, and therefore do not unveil how Okun's law changed as new data became available after 2008. Figure 4 fills this gap. Figure 4a, 4b and 4c depict recursive scatter plots where the end of each sample period is used as an identifier, namely 2009Q4, 2011Q4, and 2012Q4. Movements in the ordered pairs are identified with different symbols and colours. More precisely, squares, circles and triangles highlight how data coordinates changed as new information become available. The results reveal a close relationship between unemployment and output gaps, around a linear trend, but not without important revisions.



(A) 2009Q4 (B) 2011Q4 (C) 2012Q4 (D) 2008Q1–2015Q4

FIGURE 4: Instability of Okun's law in Portugal

Source: Banco de Portugal and own calculations.

Notes: Squares represent data points between 2008Q1 and 2009Q4; white circles between 2010Q1 and 2011Q4; and triangles between 2012Q1 and 2012Q4. Recursive estimates of Okun's coefficients cover the period 2007Q4-2015Q2.

Between 2009Q4 and 2012Q4, for instance, there is a considerable movement in data coordinates, with changes in the degree of clustering and in extreme values. Between 2009Q4 and 2012Q4, movements from positive output gaps towards larger and larger negative gaps also show instability signs, as depicted for instance by movements in the black squares.

Figure 4d plots "Okun's coefficients" using recursive estimates starting in 2007Q4. Each coefficient is defined as the negative derivative linking output and unemployment gaps. The estimates, derived from static representations of Okun's law, remained relatively stable in the euro area, around -0.55. In contrast, the Portuguese case is marked by a downward trend, suggesting a considerable movement in the static output-unemployment relationship. By the end of the sample, as expected by the results reported in Figures 3a and 3b, Portuguese and euro area coefficients coincide. This negative relationship depends among other factors on firms' decisions regarding how to adjust employment in response to temporary deviations in output, degree of job security, or social and legal constraints of firms' adjustment of employment (Blanchard 1997).

Given that observed data is invariant, the results imply that trend component estimates recorded important revisions. Uncertainties about the precise level of structural unemployment and the unemployment gap across euro area countries, using estimates from different sources (European Commission, Organisation for Economic Co-operation and Development, and IMF) are not a novelty in the empirical literature, and was highlighted for instance by European System of Central Banks (2012).

Conclusions and policy implications

This article shows that Portuguese output and unemployment over 2008–2012 are poorly assessed if unobserved trend developments are ignored. According to a semi-structural model with rational expectations, tailored for a small economy integrated in the credible monetary union—*Model Q*—, what happened in Portugal was not primarily a cyclical event, but a low frequency downward movement in the trend component of output, mirrored by an increase in the trend component of the unemployment rate.

Results confirm the desirability to achieve one of the main goals of the Economic and Financial Assistance Programme of 2011, established between the Portuguese authorities, the EU and the IMF: to remove structural impediments behind potential growth. Given that the model is silent about all economic forces driving trends, a possible way forward is to investigate causal relationships behind the estimated developments, and strengthen markets' linkages.

Results also show that the dramatic events over 2008–2012 were aggravated by the recession in the euro area, and by the higher Portuguese risk premia. Taken together, however, their importance does not outweigh all impacts coming from changes in trends. Economic policy should therefore not neglect the structural properties of these markets, resting solely centred around standard business cycle objectives.

Model Q embodies a relatively close relationship between unemployment and output gaps over all sample periods. However, there are signs of instability in trend components, making economic monitoring a difficult task.

Finally, additional ways to proceed include making the model geographically more comprehensive (*e.g.* more Member States), and structurally richer, with more information (capturing for instance financial frictions, alternative inflation measures, additional imported inflation impacts or more meaningful competitiveness variables). The analysis of the euro area is acknowledged to be incomplete. *Model Q* lacks the rest of the world economy, with prices and quantities playing their adjustment role. This is most probably an area of future work.

References

Almeida, Vanda, Gabriela Castro, Ricardo Mourinho Félix, José R. Maria, and Paulo Júlio (2013). "Inside PESSOA: A detailed description of the model." Working Paper 16, Banco de Portugal.

Berg, Andrew, Philippe Karam, and Douglas Laxton (2006). "A practical model-based approach to monetary policy analysis-overview." IMF Working Paper No. 06/81, International Monetary Fund.

Blanchard, Olivier (1997). *Macroeconomics*. Prentice Hall.

Carabenciov, Ioan, Charles Freedman, Roberto Garcia-Saltos, Ondra Kamenik, Douglas Laxton, and Peter Manchev (2013). "GPM6: The Global Projection Model with 6 Regions." IMF Working Paper No. 13/87, International Monetary Fund.

Castro, Gabriela, Ricardo Félix, Paulo Júlio, and José R. Maria (2014). "Financial frictions and shock transmission: the Portuguese economy." Economic Bulletin Autumn, Banco de Portugal.

Centeno, Mário, José R. Maria, and Álvaro Novo (2009). "Unemployment: Supply, Demand, and Institutions." In *The Portuguese economy in the context of economic, financial and monetary integration*, pp. 215–258. Banco de Portugal.

European System of Central Banks (2012). "Euro area labour markets and the crisis." Occasional Paper Series No. 138, ECB.

European System of Central Banks (2016). "Labor market modelling in the light of the financial crisis." Occasional Paper Series, *Forthcoming*, ECB.

Giavazzi, F. and M. Pagano (1988). "The advantage of tying one's hands: EMS discipline and central bank credibility." *European Economic Review*, 32, 1055–1075.

Maria, J. R. (2016). "Output and unemployment Portugal 2008-2012." Working Paper 3, Banco de Portugal.

Ravenna, Federico and Carl E. Walsh (2008). "Vacancies, unemployment, and the Phillips curve." *European Economic Review*, 52(8), 1494 – 1521.

Bank Runs: Theories and Policy Applications

Ettore Panetti
Banco de Portugal

July 2016

Abstract
In the present paper, I review the foundations of bank runs, and of the incentives of the economic agents to join them, as a base for discussing possible regulatory interventions to alleviate their effects. To this end, I study both self-fulfilling as well as fundamental runs, and propose a reconciliation of the two approaches, via the introduction of "global games". My policy conclusions highlight the role of competition and liquidity requirements to tame self-fulfilling runs. Moreover, market incompleteness and the increasing complexity of modern financial systems justify the imposition of liquidity requirements, in the presence of systemic aggregate liquidity risk.
(JEL: E21, E44, G01, G20)

Introduction

Bank runs are not only a phenomenon of the remote past:[1] in fact, they may occur whenever long-term illiquid assets are financed by short-term liquid liabilities, and the providers of short-term funds all lose confidence in the borrower's ability to repay, or are afraid that other lenders are losing their confidence. There exists a wide consensus that many U.S. money market funds have experienced runs after the bankruptcy of Lehman Brothers in 2008 and, more generally, that the financial crisis of 2007-2009 can be interpreted as a run of financial intermediaries on other financial intermediaries (Gorton and Metrick 2012). The empirical literature shows that, during that period, the U.S. endured a peak-to-trough decline in real per capita GDP of 4.8%, with a widespread impact on asset markets, housing markets, government debt and unemployment (Reinhart and Rogoff 2009, 2014). These numbers justified a massive government intervention,[2] as well

Acknowledgements: I thank Luca Deidda, Filomena Garcia, Elena Mattana, and the seminar participants at Banco de Portugal and many other institutions for their valuable comments. The opinions expressed in this article are mine, and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. Any errors and omissions are my sole responsibility.

E-mail: etpanetti@bportugal.pt

1. In the period 1825-1929, the U.S. economy experienced seven major bank runs, and twenty non-major ones (Jalil 2015). Afterwards, no episode was registered, up until 2008.

2. In 2008-2009, the U.S. Treasury invested more than $400 billion in the "Troubled Asset Relief Program", to rescue several financial and non-financial corporations hit by the financial crisis. In

as the introduction of new forms of financial regulation, in particular through the liquidity ratios of Basel III, with the explicit objective of taming the adverse effects of bank runs in the future. However, a rigorous discussion of these policies and of their effectiveness cannot prescind from an equally rigorous evaluation of the foundations of bank runs, and of the incentives of the economic agents to join them.

The aim of the present paper is to describe a theory suitable to jointly analyze these themes. To this end, I take as starting point the seminal work of Diamond and Dybvig (1983). This is the standard workhorse model for the analysis of the economics of banking, as it offers a rationale for the existence of a banking system – as a mechanism to pool risk and allocate resources in an economy hit by idiosyncratic liquidity shocks – as well as a natural framework to study bank runs. With this tool in hand, I study bank runs emerging from self-fulfilling expectations of banks' depositors, as well as from extreme fluctuations of the fundamentals of the economy. The first approach speaks to those who argue that bank runs are a consequence of illiquidity, caused by extrinsic events (like sunspots or panic attacks) completely independent from the observed state of the economy. In contrast, according to the second approach, bank runs are a consequence of insolvency, caused by fundamental shocks affecting the returns on banks' investments. To reconcile these two points of view, I conclude by introducing the "global game" approach, where runs are expectations-driven, but also explicitly depend on banks' risk profiles, and on the underlying state of the economy. For each of these approaches, I will sketch the main findings of the literature, and their policy implications.

The Diamond-Dybvig Model

I start my analysis with a description of the Diamond-Dybvig model. This framework focuses on banks engaging in liquidity and maturity transformation, through illiquid long-term loans and liquid short-term deposits, which are the main components of banks' asset and liabilities in the real world.

The economy lives for three periods, labeled $t = 0, 1, 2$, and is populated by risk-averse agents, all with an endowment $e = 1$ at date 0, and nothing afterwards. At date 1, each agent draws an idiosyncratic type θ, which is private information to herself, and takes value 0 with probability $1 - \pi$, and 1 with probability π. The idiosyncratic types affect the point in time at which each agent enjoys consumption, according to the welfare function

---

the same period, the Federal Reserve, through its liquidity facilities, extended credit to the U.S. financial system for around \$1.5 trillion.

$W(c_1, c_2, \theta) = \theta u(c_1) + (1-\theta)u(c_2)$. Clearly, those agents whose realized type is $\theta = 0$ are only willing to consume at date 2, and those whose realized type is $\theta = 1$ are only willing to consume at date 1. Thus, I interpret the types θ as "liquidity shocks" and the probability π as "liquidity risk". Moreover, I refer to the agents as late (or patient) and early (or impatient) consumers, respectively.

Being risk averse, the agents would like to insure themselves against liquidity risk. However, we make the simplifying assumptions that they are isolated, and markets are incomplete.[3] Thus, the only channel left is through the banking system. The economy is populated by a large number of banks, operating in a perfectly-competitive market with free entry. At date 0, the agents deposit their endowments, and the banks offer them a deposit contract $\{d_1, d_2\}$, stating how much they can withdraw and consume at date 1 and 2, depending on their reported types. To finance the deposit contract, the banks invest the deposits (the only liability on their balance sheets) into two assets: the first one is a storage technology (analogous to liquidity or cash) that yields 1 unit of consumption at date $t+1$ for each unit invested at any date t, and is a cheap – although not remunerative – way to roll over resources from one period to the next; the second one is a long-term asset, that yields $R > 1$ units of consumption at date 2 for each unit invested at date 0, but only $r < 1$ units at date 1. This long-term asset can be interpreted as a loan to a production unit, that takes time to mature and is partially illiquid, or can be liquidated before maturity, with a low recovery rate equal to r. Competition and free entry ensure that the banks have incentives to look after their depositors in order to attract them and survive into operation. In other words, in a competitive banking equilibrium, the banks choose a portfolio allocation between storage and loans and a deposit contract so as to maximize the expected welfare of their depositors, subject to their budget constraints.

In such an environment, Diamond and Dybvig (1983) show that the competitive banking equilibrium is equivalent to the "first best" allocation, where a benevolent social planner, who wants to maximize the expected welfare of the agents, perfectly insures them against liquidity risk. In such an equilibrium, the banks give the depositors an amount of late consumption lower than what they would get if they invested all their endowments in the long-term asset ($d_2 < R$), in exchange for an amount of early consumption higher than what they would get from mere storage ($d_1 > 1$), and the fact that the agents are risk averse implies that this transfer is welfare-improving. Moreover, such a consumption allocation satisfies the

3. The hypothesis of market incompleteness is crucial for the the results of the Diamond-Dybvig model: if the agents were allowed to trade in a complete market for state-contingent claims, banks would be redundant (Allen and Gale 2004). However, it is easy to show that the competitive banking equilibrium dominates an "autarkic" equilibrium, where the agents independently choose their portfolio allocations, and rebalance them in a secondary asset market.

incentive compatibility constraint $d_1 \leq d_2$, that ensures truth-telling when the realizations of the idiosyncratic types are private information. To sum up, a perfectly-competitive banking system with free entry, holding long-term illiquid loans financed by short-term liquid deposits, allow an efficient allocation of resources, in the presence of idiosyncratic liquidity risk.

Self-fulfilling Bank Runs

According to Diamond and Dybvig (1983), the fact that the banks offer a deposit contract equivalent to the first best makes them intrinsically fragile. To see that, assume that a bank, at date 0, commits to offer the equilibrium deposit contract $\{d_1, d_2\}$ to all withdrawers and liquidate the long-term asset to fulfill this obligation. Under this hypothesis, the economy exhibits two equilibria: one where only the early consumers withdraw at date 1, and one where also all late consumers withdraws at date 1, and store to consume at date 2. This second equilibrium may occur whenever all late consumers expect that all the other late consumers withdraw, and know that the bank does not have sufficient resources to pay d_1 to all withdrawers. In this case, the bank is subject to a "run". To see this more intuitively, notice that, if a late consumer expects no other late consumer to run, she clearly prefers to withdraw at date 2, as $d_1 \leq d_2$. However, if she expects all the other late consumers to run, a late consumer would prefer to join the run (and get $X + rY$) rather than waiting until date 2, when she gets 0, as the banks have liquidated all the long-term assets in portfolio. In other words, according to this narrative, bank runs are an exclusive consequence of depositors' self-fulfilling expectations about the behavior of the other depositors and bank illiquidity, not of fundamental shocks affecting the value of banks' assets.

Clearly, this explanation relies on the commitment of the banks to offer the equilibrium deposit contract.[4] To relax this assumption, assume that the banks, at date 0, choose the portfolio allocation $\{X, Y\}$ and deposit contract $\{d_1, d_2\}$ taking into account the strategic decision of the depositors about whether to run or not at date 1. Moreover, assume that, at date 1, the banks serve the depositors on a first-come-first-served basis (i.e. according to the so-called "sequential service constraint"). In this way, a run might affect the fraction of the depositors who are served, and the portfolio allocation between storage and the long-term asset. To see that more clearly, write the budget constraint of a bank subject to a run at date 1 as $X + rY = \delta d_1$, where δ is the fraction of depositors that can be served, given the portfolio allocation $\{X, Y\}$ and the amount of consumption d_1 stated in the deposit contract.

4. In fact, if they committed to not liquidate the long-term asset (a policy often referred to as "suspension of convertibility") the run equilibrium would not exist.

Cooper and Ross (1998) show that a run equilibrium exists if and only if δ is lower than 1, i.e. the bank is not able to serve all depositors in the case of a run. Put differently, if such a condition is satisfied, banks are illiquid, and the economy exhibits two equilibria: a run equilibrium and a no-run equilibrium. Then, the depositors coordinate a choice between the two in accordance with the realization of an extrinsic event – a "sunspot" – completely uncorrelated to the fundamentals of the economy, and happening with some exogenous probability q. Sunspots are seen as a way to account for depositors' animal spirits, panic attacks, or self-fulfilling expectations, and have been extensively employed in the literature on financial crises to model self-fulfilling runs (Peck and Shell 2003).

In turn, a bank, knowing the equilibrium selection mechanism and the probability of the sunspot q, chooses a portfolio allocation $\{X, Y\}$ and deposit contract $\{d_1, d_2\}$ at date 0 so as to maximize the expected welfare of its depositors, subject to its budget constraint. However, notice that δ, the fraction of depositors who are served at a run, also regulates the existence of the run equilibrium itself, and depends on the portfolio allocation and deposit contract. Thus, at date 0, a bank can choose them so as to rule out the run equilibrium, and be completely run-proof. More formally, a bank calculates two portfolio allocations and deposit contracts, either with possible runs (i.e. such that $\delta < 1$) or run-proof (i.e. such that $\delta \geq 1$), and then chooses in equilibrium those that maximize the expected welfare of its depositors. In the first case (possible runs), the incentives to provide more risk sharing against a run (that would increase d_1) are higher than the incentives to serve the highest possible number of depositors (that would lower d_1 so as to increase δ). Thus, a bank chooses a higher amount of storage than in a benchmark equilibrium without runs: in other words, a run generates a credit tightening. In the second case (run-proof), instead, a bank is able to provide the first-best allocation of resources if the recovery rate of the long-term asset is sufficiently high to ensure that $\delta \geq 1$; otherwise, it makes the contract run-proof by lowering d_1, i.e. by reducing risk sharing, and, in extreme cases, by also cutting credit and holding excess storage. These results highlight that, when facing the possibility of self-fulfilling runs, a bank's choice between being run-proof or not essentially boils down to finding the correct balance between (i) providing risk sharing against consumption fluctuations during a run and (ii) minimizing the probability of its occurrence.

In a calibrated dynamic general-equilibrium version of this model, Mattana and Panetti (2016) assume that the banks follow an "equal service constraint": all the depositors who withdraw get an equal share of the available resources, even in the case of a run.[5] In this environment, the banks

5. The equal service constraint resembles some contractual arrangements observed in the real world: money market mutual funds, for example, serve their depositors pro-rata. Despite equal

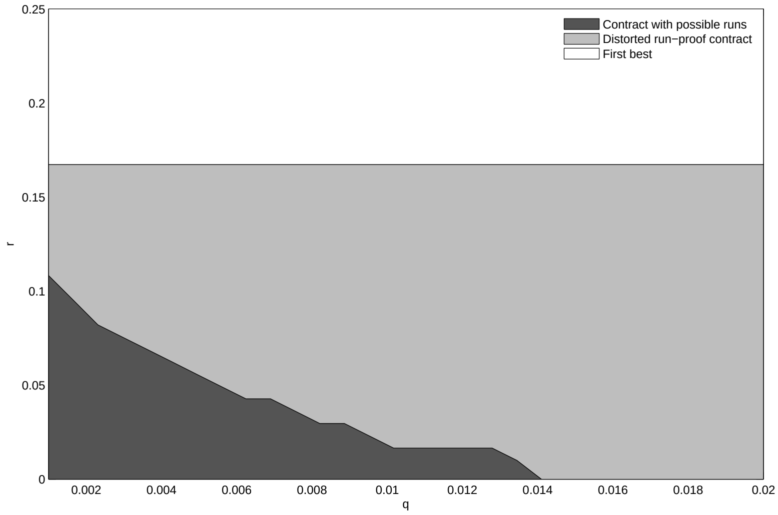


FIGURE 1: The competitive banking equilibrium, for different values of the probability of the sunspot q (on the x-axis) and recovery rate r (on the y-axis). Source: Mattana and Panetti (2016).

can offer a run-proof contract equivalent to the first-best only when the recovery rate r is above 17%, as showed in Figure 1. For values below that threshold, the banks distort the allocation of resources with respect to the first best: whenever the probability of the sunspot q and the recovery rate r are both sufficiently low (in the calculations, below 1.4% and 11%, respectively), the risk-sharing motivation dominates the objective of preventing a run, and the banks choose a contract with possible runs; above those values, instead, the opposite is true, and they choose a distorted run-proof contract.

This conclusion leads to two compelling arguments for policy. First, the message of Diamond and Dybvig (1983) is that bank runs are an inevitable consequence of liquidity and maturity transformation. Thus, government intervention, in the form of deposit insurance and central banks' liquidity assistance via the discount window (coupled with sophisticated interbank markets), is necessary to ensure that solvent banks stay liquid. The present results provide a complementary argument: in the presence of extrinsic uncertainty, that might trigger bank runs, competition and free entry in the banking system provide the correct incentives for banks to find the right balance between risk sharing and the willingness to avoid runs, even in the absence of government assistance. The second argument refers to the costs

service being technically different from sequential service, the distortions that they impose on the equilibrium portfolio allocation and deposit contract are similar.

of policy intervention: assume that a regulator wants to impose a liquidity requirement, with the objective of making the banks always run-proof ($\delta \geq 1$), irrespective of the levels of recovery rate and probability of the sunspot.[6] What would the cost of such a policy be? From what said above, this constraint would distort the competitive banking equilibrium only when the recovery rate and the probability of the sunspot are both so low that the risk sharing motivation dominates the objective of avoiding a run, as in any other case the banks is already run-proof. Thus, making the banks always run-proof would come at the cost of lower risk sharing. Quantitatively, the welfare costs are decreasing in both the recovery rate and the probability of the sunspot, and are in any case below 0.16%.[7] Arguably, this is a small number: the only comparable work (Van den Heuvel 2008) finds that the welfare costs of capital requirements are one order of magnitude higher.

Fundamental Runs

Modeling self-fulfilling bank runs is indisputably appealing, and is also corroborated by some early studies on the U.S. National Banking Era (Friedman and Schwartz 1963) as well as more recent ones of the 2007-2009 financial crisis (Foley-Fisher *et al.* 2015) and some experimental evidence (Arifovic *et al.* 2013). However, the drawback of this approach is that it relies on exogenous extrinsic uncertainty (i.e. the sunspots). Put differently, it is difficult to argue that bank runs are completely disconnected from the circumstances of the real economy. For example, Gorton (1988) argues that the bank runs of the U.S. National Banking Era could have been predicted by a leading indicator based on the level of business failures. This observation reminds us that, while bank runs are often a consequence of bank illiquidity, they might also originate from insolvency issues. These arguments have given rise to the so-called "business-cycle view", according to which bank runs are a consequence of variations in the fundamentals of the economy, that make banks unable to meet their commitments.

To see that more clearly, assume that the return on the long-term asset R (which represents the aggregate state of the economy) is a random variable that is realized at date 2, but about which all depositors get a perfectly-informative signal at date 1. Moreover, assume that the banks serve their depositors according to the equal service constraint, and are exogenously constrained to offer an "incomplete" deposit contract, in which the amount

6. Arguably, this is the only policy intervention that can be discussed in the present framework, as the equilibrium determination relies on the realization of an extrinsic event like the sunspot.

7. The welfare costs are calculated in consumption equivalents, i.e. the constant proportional increase in consumption that the "regulated" banking equilibrium would need in order to ensure the same expected welfare as the unregulated equilibrium.

of early consumption d_1 is independent of the aggregate state of the economy. Under these hypotheses, Allen and Gale (1998) show that, in a competitive banking equilibrium, the patient depositors are all happy to wait until date 2 to withdraw, whenever the signal about the aggregate state is sufficiently "good" (i.e. R is high). Conversely, when the signal is sufficiently "bad" (i.e. R is low), all late consumers attempt to withdraw at date 1, thus triggering a "fundamental run". Interestingly, the corresponding consumption allocation is equivalent to the first best, where a benevolent social planner offers a complete contract $\{d_1(R), d_2(R)\}$, fully dependent on the realization of the aggregate state of the economy R. This happens because, in this economy, it is efficient from a welfare perspective to share the resources equally among all bank depositors, whenever the aggregate state is sufficiently low, and to give a constant amount of early consumption, whenever the aggregate state is sufficiently high. In a competitive banking equilibrium, this can be achieved with an incomplete deposit contract, coupled with the possibility to have fundamental runs, during which the depositors are served according to the equal service constraint and get the same amount of consumption, irrespective of whether they are early or late consumers.

Thus, we get the rather surprising result that a competitive banking equilibrium with fundamental runs, under the hypotheses described above, is efficient. Equally surprising is the robustness of this result. In a follow-up paper, Allen and Gale (2004) study an environment where the banks face aggregate liquidity risk, and can hedge against it by buying and selling assets in a complete market for state-contingent claims at date 0, and in a secondary market at date 1. In this set-up, the banks, when exogenously constrained to offer an incomplete deposit contract, default if hit by a negative shock, and the corresponding consumption allocation is again efficient. Hence, the common conclusion of these two papers is that, in an economy with both idiosyncratic and aggregate liquidity risk, there is no justification from a welfare perspective for the introduction of financial regulation: there is no way through which a regulator can avoid bank insolvency and make some depositors better off, while keeping all others at least as well off. However, this result crucially depends on the completeness of asset markets: in fact, if markets were incomplete, liquidity regulation would allow a regulator to indirectly manipulate the equilibrium price in the secondary market, and improve welfare.

Arguably, the incompleteness of the deposit contract plays a crucial role for these results. This is a fair assumption for many reasons, such as legal arrangements, or asymmetric information between banks and depositors, or transaction costs. However, digging into the microfoundations of this incompleteness leads to some interesting considerations. Panetti (2013) studies a Diamond-Dybvig model with aggregate liquidity risk: the fraction of early consumers π that each bank faces is random. Moreover, the total fraction of early consumers in the whole banking system can be either fixed or random, implying non-systemic or systemic aggregate liquidity risk, respectively, and its distribution is known at date 0, when the banks choose the portfolio allocation and the deposit contract. Importantly, the depositors can borrow and lend among themselves in a bond market at an interest rate $\hat{R}$, without being observed by their banks. The unobservability is a plausible assumption because, in this way, the depositors can extend their investment opportunities beyond traditional banks and towards market-based "new financial intermediaries", which is a phenomenon that has been extensively observed in the recent past (Guiso *et al.* 2002). Moreover, because of this unobservability, the deposit contract becomes *endogenously* incomplete, as the ratio between late consumption and early consumption $d_2(R)/d_1(R)$ in the deposit contract does not depend on the realization of aggregate liquidity risk. Under these assumptions, whenever the economy faces non-systemic aggregate liquidity risk, interbank market trades allow the banks to avoid default. However, the competitive banking equilibrium is inefficient, because of the presence of a pecuniary externality in the bond market that makes the interest rate $\hat{R}$ too high, and that the banks do not internalize. Therefore, a regulator can indirectly lower the interest rate and increase total welfare by imposing minimum liquidity requirements, that can be either bank-specific or one-size-fits-all.

These conclusions change when the economy faces systemic aggregate liquidity risk, that prevents interbank markets to clear. To see that, assume that only two aggregate states are possible at date 1: aggregate liquidity risk can be either systemically high or systemically low, with some known probability. Under this scenario, at date 0, the banks choose a very low amount of liquidity whenever the ex-ante probability of high aggregate liquidity risk is systemically low, and default at date 1 if systemic liquidity risk is actually realized. On the contrary, banks hoard liquidity at date 0 whenever the ex-ante probability of high aggregate liquidity risk is systemically high, and at date 1 roll it over to date 2, if systemic liquidity risk is actually not realized. More interestingly, the competitive banking equilibrium is again inefficient because of the pecuniary externality on the bond market: the interest rate $\hat{R}$ is too high when the ex-ante probability of high aggregate liquidity risk is systemically low, and too low when the ex-ante probability of high aggregate liquidity risk is systemically high. Hence, a regulator can improve total welfare by imposing countercyclical liquidity requirements: a *minimum* liquidity requirement whenever the probability of high aggregate liquidity risk is systemically low, or a *maximum* liquidity requirement whenever the probability of high aggregate liquidity risk is systemically high. The indirect effect of this policy is to lower the incidence of both bank default and liquidity hoarding on the competitive banking equilibrium.

The Global Game Approach

So far, I have described two competing theories about bank runs: one based on self-fulfilling expectations leading to illiquidity, and one based on fundamental shocks leading to insolvency. However, in practice, distinguishing illiquidity from insolvency is controversial, if only because the evaluation of the solvency of a financial institution essentially depends on the assessment of its assets.[8] These considerations are particularly important in the light of government intervention: as argued above, there are cases in which a financial regulator facing insolvency should not intervene, but the common practice of central banks, based on the doctrine of the "lender of last resort", is to provide support to solvent but illiquid banks.[9] These considerations call for a theory that reconciles self-fulfilling and fundamental runs and, at the same time, provides a criterion to distinguish them and a rationalization of government intervention. This is the aim of one of the most promising branches of the literature on banking and crises, based on the "global game" approach (Morris and Shin 1998).

To show it in more detail, I slightly modify the environment of the previous section. As before, a bank offers uncontingent early consumption d_1, and the return on the long-term asset is random: it takes the value R with probability p, and 0 with probability $1 - p$, where p is a random variable uniformly distributed over the interval $[0, 1]$ and represents the aggregate state of the economy. However, differently from above, the signal that the depositors receive at date 1, about the realization of the state, is not perfectly informative, but "noisy": it takes the form $\sigma = p + e$, with e representing a small but positive idiosyncratic noise, uniformly distributed over the interval $[-\varepsilon, +\varepsilon]$. In this environment, Goldstein and Pauzner (2005) show that a fundamental run (where all late consumers withdraw at date 1) happens whenever the signal is below a certain threshold $\underline{\sigma}$, at which all late consumers are indifferent between withdrawing at date 1 or 2, irrespective of the behavior of the others. The existence of this threshold, together with an "upper dominance region", where the signal is so good that there is no run for sure, is enough to ensure the existence of an equilibrium in the intermediate region. There, absent noisy signals, the economy would exhibit two equilibria (run and no-run). However, the fact that the signals are noisy breaks the possibility for the late consumers to coordinate, in the sense that they cannot directly infer the behavior of the others from their own behavior. Thus, in the intermediate region, there

8. As an example, the New York Times reported (September 29, 2014) that, while Lehman Brothers had valued its real estate portfolio at around $50 billion in 2008, the CEO of Bank of America (who at the time was considering a bid on Lehman) "asserted that Lehman had a $66 billion hole in its balance sheet".

9. See Alves *et al.* (2015) for a recent empirical analysis of the role of the lender of last resort in Portugal.

exists a unique equilibrium, where all late consumers, after observing their own signal, create posterior beliefs about the aggregate state and the signals received by the other late consumers, and based on these decide whether to run or not.

In particular, in the intermediate region the late consumers follow the threshold strategy "run if the signal σ is below the threshold σ^*", at which they are indifferent between withdrawing at date 1 or 2 *given* their posterior beliefs. Put differently, in this intermediate region runs are self-fulfilling, i.e. based on negative expectations about the aggregate state of the economy, but not on negative fundamentals per se: banks are solvent but illiquid. More importantly, both thresholds $\underline{\sigma}$ and σ^* are endogenously determined, and positively depend on the amount of early consumption d_1 stated in the deposit contract at date 0. Thus, the banks here face again a trade-off between higher risk sharing and higher probability of a bank run: the higher the amount of risk sharing that a bank promises (i.e. the higher d_1), the higher the probability that it is not going to be able to pay the amount stated in the deposit contract, either because of bad fundamentals (high $\underline{\sigma}$) or because of bad expectations (high σ^*).

The uniqueness of the equilibrium and the endogeneity of the two thresholds allow us to fully interpret the role played by financial regulation in this environment. Intuitively, a regulator would not find convenient to intervene when the signal falls below $\underline{\sigma}$, as a fundamental crisis is efficient. However, it would intervene in the case that the signal falls between $\underline{\sigma}$ and σ^*, where illiquidity is only a consequence of bad expectations. In a framework similar to the present one, Rochet and Vives (2004) show that liquidity requirements solve the expectations problem, but might be too costly in terms of forgone bank returns. Therefore, they should be complemented by the provision of central bank liquidity, through the discount window. This conclusion supports the main prescription of the doctrine of the "lender of last resort": central banks should lend freely to solvent but illiquid banks. However, according to the classic view by Bagehot (1873), liquidity should be provided at penalty rates, and against good collateral. On those lines, Allen *et al.* (2015) analyze the case for limiting central bank liquidity interventions in a model with runs as global games. The authors show that injecting liquidity into the banks in the case of a run, in order to reduce its likelihood ex-ante, might have the unintended consequence of increasing banks' moral hazard. Thus, the optimal liquidity injection should never fully prevent runs.

Conclusions

The aim of the present paper has been to describe the foundations of bank runs, and of the incentive of the economic agents to join them, as a base to discuss government interventions to tame their adverse effects.[10] The main lesson that we can draw is threefold. First, bank runs are not an inevitable byproduct of liquidity and maturity transformation, as argued by Diamond and Dybvig (1983), and a lot can be done against them: in particular, competition in the banking system provides the right incentives for banks to avoid risky investment strategies, that might harm depositors' savings and give rise to self-fulfilling runs. This message is particularly important in our current economy, as unsecured deposits represent a large and increasing share of total bank liabilities, both in the U.S. and around the world (Peristiani and Santos 2014), and the so-called "shadow banking system" provides liquidity and maturity transformation without access to deposit insurance and central bank discount windows. Second, government intervention can make the banking system more resilient to self-fulfilling runs, either ex post, via central banks' emergency liquidity assistance, or ex ante, via liquidity requirements. However, while the former should always be partial, in order to tame banks' moral hazard, the latter should be preferred, as its costs are quantitatively small. Finally, there are many cases where government intervention against fundamental runs is not justifiable from a welfare perspective. Nevertheless, market incompleteness and the increasing complexity of modern financial systems, where "traditional" banks coexist with new market-based intermediaries, calls for a further tailoring of financial regulation, especially in the face of systemic aggregate liquidity risk.

References

Allen, Franklin, Elena Carletti, Itay Goldstein, and Agnese Leonello (2015). "Government Guarantees and Financial Stability." CEPR Discussion Paper No. 10560.

Allen, Franklin and Douglas Gale (1998). "Optimal Financial Crises." *Journal of Finance*, 53(4), 1245–1284.

Allen, Franklin and Douglas Gale (2004). "Financial Intermediaries and Markets." *Econometrica*, 72(4), 1023–1061.

Alves, Nuno, Diana Bonfim, and Carla Soares (2015). "Surviving the perfect storm: the role of the lender of last resort." mimeo.

Arifovic, Jasmina, Janet Hua Jiang, and Yiping Xu (2013). "Experimental evidence of bank runs as pure coordination failures." *Journal of Economic Dynamics and Control*, 37(12), 2446–2465.

Bagehot, Walter (1873). *Lombard Street: A Description of the Money Market*. Henry S. King & Co., London, UK.

10. Given its limited scope, this analysis abstracted from other issues of key importance for bank runs, like the role of asset markets, bank capital, or financial contagion, that are the focus of other equally important branches of the literature.

Cooper, Russell and Thomas W. Ross (1998). "Bank runs: Liquidity costs and investment distortions." *Journal of Monetary Economics*, 41(1), 27–38.

Diamond, Douglas W. and Philip H. Dybvig (1983). "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, 91(3), 401–419.

Foley-Fisher, Nathan C., Borghan Narajabad, and Stephane H. Verani (2015). "Self-fulfilling Runs: Evidence from the U.S. Life Insurance Industry." Federal Reserve Board Finance and Economics Discussion Series No. 2015-032.

Friedman, Milton and Anna J. Schwartz (1963). *A Monetary History of the United States, 1867-1960*. Princeton University Press.

Goldstein, Itay and Ady Pauzner (2005). "Demand-Deposit Contracts and the Probability of Bank Runs." *Journal of Finance*, 60(3), 1293–1327.

Gorton, Gary B. (1988). "Banking Panics and Business Cycles." *Oxford Economic Papers*, 40(4), 751–781.

Gorton, Gary B. and Andrew Metrick (2012). "Getting up to Speed on the Financial Crisis: A One-Weekend-Reader's Guide." NBER Working Paper No. 17778.

Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (eds.) (2002). *Household Portfolios*. MIT Press, Cambridge, MA.

Jalil, Andrew J. (2015). "A New History of Banking Panics in the United States, 1825–1929: Construction and Implications." *American Economic Journal: Macroeconomics*, 7(3), 295–330.

Mattana, Elena and Ettore Panetti (2016). "A Dynamic Quantitative Macroeconomic Model of Bank Runs." mimeo.

Morris, Stephen and Hyun Song Shin (1998). "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *American Economic Review*, 88(3), 587–597.

Panetti, Ettore (2013). "A Theory of Bank Illiquidity and Default with Hidden Trades." In *Essays on the Economics of Banks and Markets*, Institute for International Economic Studies Monograph Series No. 84. Stockholm University.

Peck, James and Karl Shell (2003). "Equilibrium Bank Runs." *Journal of Political Economy*, 111(1), 103–123.

Peristiani, Stavros and Joao Santos (2014). "Depositor Discipline of Risk-Taking by U.S. Banks." Liberty Street Economics, Federal Reserve Bank of New York.

Reinhart, Carmen M. and Kenneth S. Rogoff (2009). "The Aftermath of Financial Crises." NBER Working Paper No. 14656.

Reinhart, Carmen M. and Kenneth S. Rogoff (2014). "Recovery from Financial Crises: Evidence from 100 Episodes." *American Economic Review*, 104(5), 50–55.

Rochet, Jean-Charles and Xavier Vives (2004). "Coordination Failures and the Lender of Last: Was Bagehot Right After All?" *Journal of the European Economic Association*, 2(6), 1116–1147.

Van den Heuvel, Skander J. (2008). “The Welfare Costs of Bank Capital Requirements.” *Journal of Monetary Economics*, 55(2), 298–320.