Nova School of Business and Economics

Universidade Nova de Lisboa

Thesis, presented as part of the requirements for the Degree of

Doctor of Philosophy in Economics

Essays on Banking and the Transmission of Monetary Policy

Miguel Mendonça de Sousa Boucinha

Student N. 634

A Thesis carried out under the supervision of Professor Pedro Pita Barros

May 2017

Contents

Foreword

This thesis sheds light on the behaviour of banks, the main providers of financial services for households and firms in the euro area. The analysis spans both structural and cyclical topics including the determinants of banks' capital decisions (Chapter 1), assessing the degree of competition (Chapter 2), efficiency and productivity (Chapter 3) as well as the transmission of monetary policy (Chapter 4). It therefore covers both topics associated with banks' decisions and with market equilibrium outcomes. Other important topics, including the relationship between bank lending and economic growth, are outside the scope of this work.

The first three chapters use the Portuguese banking system as a case-study, based on a unique dataset of supervisory data. The analysis starts from the early 1990s, thereby including the process of liberalization, consolidation and financial innovation. The first chapter investigates why banks hold significant capital buffers above the required regulatory minimum, despite evidence that capital is a relatively expensive source of financing when compared to deposits or bonds. The second chapter assesses how the competitive behaviour of Portuguese banking groups in the loan market has evolved through time and investigates for cross-sectional heterogeneity across different types of banks. The third chapter analyses the production technology of Portuguese banks. It explores how the changes in the banking system following the country's euro area accession reflected in developments in marginal costs, scale economies, cost efficiency and overall productivity.

The last chapter focuses on the transmission of monetary policy, and in particular on the role of the portfolio rebalancing channel in the transmission of large scale asset purchase programmes. The analysis is based on recent evidence for the euro area and investigates whether the ECB's expanded Asset Purchase Programme (APP) led investors to rebalance their bond portfolios towards riskier assets and, in the case of banks, towards lending to the real economy. Identification is based on cross-sectional information, using a unique granular dataset on security holdings for the institutional sectors of each euro area country and for the largest banking groups in the euro area.

The Determinants of Portuguese Banks' Capital Buffers*

Miguel Boucinha

Abstract

The purpose of this paper is to shed light on why Portuguese banks hold significant capital buffers above the required regulatory minimum, despite the fact that capital is a relatively expensive source of financing compared to deposits or bonds. The level of banks' capital buffers is found to be positively influenced by several broad risk measures, indicating that banks tend to hold extra capital in order to cover for increased risk. Loan loss provisions and high and stable profitability are found to be substitutes for capital buffers. Larger banks seem to hold less excess capital, which could be linked with higher portfolio diversification and technological advantages in screening and monitoring activities, but also with moral hazard associated with higher expectations of public support against the background of the "too-big-to-fail" hypothesis. A negative business cycle effect is also found, suggesting that the lending cycle may be pro-cyclical.

Keywords: Banking; Excess Capital; Risk; Panel Data

JEL: G21; G28; C23

* The author is indebted to António Antunes, João Amador, Luís Catela Nunes, Pedro Pita Barros, Nuno Ribeiro and Carlos Santos helpful comments and suggestions. All remaining errors are my responsibility. This paper was presented as the author's Masters project at Universidade Nova de Lisboa.

1. Introduction

Since the mid 1980's, there has been an increasing effort to decrease the distortionary effects of excessive regulation in the Portuguese banking sector, following the revolution (1974) and nationalization of Portuguese banks (1975). However, it was not until the early 1990's that credit ceilings were abolished and deposit rates were fully liberalized. This period has also hosted full entry liberalization and the beginning of privatizations, which lasted until 1996. Hence, during the early 1990's, the Portuguese banking sector has met greater sophistication through market segmentation, risk differentiation and increased competition, which allowed for considerable efficiency gains.

In this context of decreasing direct controls on banks' conduct, capital adequacy regulation has become relatively more important. As established in the 1988 Basel Capital Accord (Basel I), each bank is required to hold at least 8% of its Risk Weighted Assets (RWA) in capital. Most banks' solvency ratios are well above the regulatory minimum and a better understanding of the determinants of these capital buffers in one regulatory regime may shed light on the relevance and desirability of changes in the regulatory framework. On the other hand, it may make it easier to assess the factors underlying banks' solvency ratios.

The main findings of the analysis are that the magnitude of banks' capital buffers is positively influenced by several broad risk measures, indicating that banks tend to hold extra capital in order to cover for increased risk. Loan loss provisions and high and stable profitability are found to be substitutes for capital buffers. Larger banks seem to hold less excess capital, which could be linked with higher portfolio diversification and technological advantages in screening and monitoring activities, but also with moral hazard associated with higher expectations of public support against the background of the "too-big-to-fail" hypothesis. A negative business cycle effect is also found. While this may be a desirable result from a micro-prudential perspective, it also suggests that the lending cycle may be pro-cyclical.

This study is organized as follows: Section 2 discusses some hypotheses to be examined; section 3 summarizes previous empirical findings on the subject; section 4 presents the data; section 5 presents the model to be estimated; section 6 discusses estimation issues; section 7 presents the estimation results and section 8 concludes.

2. Hypotheses

The reasoning behind capital adequacy regulation is that banks should hold enough capital in order to assure that failure risk is minimal. This is important due to the negative externalities bank failures impose on their depositors and the potential for moral hazard in the behaviour of limited liability stockholders, but especially due to the possibility of generating systemic risk with severe effects to the real economy.[1]

In this context, since solvency regulation in the framework of Basel I is generally acknowledged to have poor risk sensitivity, one would expect banks with higher risk to hold higher capital buffers.[2] On the other hand, the literature suggests that capital is a relatively expensive source of financing when compared to deposits or bonds. [3] Hence, banks' capital decisions reflect the trade-off between the benefits and the costs of holding excess capital (Milne and Whalley 2001).

One of the main benefits of holding high capital buffers is lower failure costs due to a decrease in the probability of failure. On the other hand, a decrease in the capital ratio below the regulatory minimum would imply extra supervisory scrutiny which would in turn decrease bank value. Furthermore, adjustments to the capital level bare direct costs, i.e. the transaction costs of issuing and repurchasing shares, as well as indirect costs from the signals they send to markets – issuing new shares may be interpreted as a signal that shares are overvalued.[4] Dietrich and Vollmer (2004) argue that banks use excess capital as a strategic tool which provides banks with increased bargaining power when renegotiating loans. In the context of asymmetric information, strong solvency ratios may also be interpreted as a signal of the bank's low probability of failure. Hence, higher solvency ratios allow for better credit ratings, therefore decreasing the cost of financing and improving the bank's reputation. In the presence of liquidity constraints, banks may also hold excess capital in order to provide for unexpected investment opportunities. The business cycle is also likely to impact banks' capital buffers, as the value of capital as an insurance against failure should be higher during periods of increased credit risk and risk aversion.

[1] For a comprehensive discussion on the relevance of financial stability see Crockett, A. (1997).
[2] See BCBS (1999).
[3] Check Myers and Majluf (1984).
[4] Cornett and Tehranian (1994) report statistically significant negative share price reactions to announcements of equity issues in the banking industry.

According to the "too big to fail" hypothesis, larger banks feel that authorities would support them if they faced difficulties due to the important externalities and contagion risk their failure would pose. This moral hazard effect makes big banks willing to purchase less insurance against failure. On the other hand, if portfolio diversification increases with bank size and is not captured by the risk measures, larger banks are less likely to experience large drops in their capital ratios.[5] One may also consider the costs and benefits of screening and monitoring borrowers in order to better acknowledge their risk. If there are scale economies in screening and monitoring, one would expect larger banks to choose relatively more of these activities to the detriment of excess capital. As Alfon et al. (2004) remark, big banks may also be less liquidity constrained and/or have smaller costs in adjusting capital to optimum levels and thus issue comparatively more capital or debt on demand rather than hold large capital reserves.

Overall, the main hypotheses to be tested can be summarized as follows:

Hyp. 1: Banks'capital buffers are likely to be persistent;
Hyp. 2: Higher risk should be associated with higher capital buffers;
Hyp. 3: Other froms of addressing risk should be associated with lower capital buffers;
Hyp. 4: Larger banks are likely to hold lower capital buffers;
Hyp 5: The business cycle is likely to be negatively correlated with capital buffers.

3. Previous Empirical Findings

There is an extensive literature regarding banks' solvency. However, most of it regards US banks and capital ratios rather than excess capital. The literature on the determinants of European banks' capital buffers includes Stolz and Wedow (2005) for German banks, Ayuso et al. (2002a) and Ayuso et al. (2002b) for Spanish banks and Lindquist (2003) for Norwegian banks. To our knowledge there is no such work for Portuguese banks. A dynamic model is estimated in the first three papers, whereas Lindquist (2003) does not explicitly model for the persistence of the capital buffer, i.e. for adjustment costs in banks' level of capital.

[5] It is, however, important to note that Portuguese banks' capital requirements are already adjusted for diversification.

The conclusions on the effect of risk on capital buffers vary somewhat between models. Ayuso et al. (2002a) find a negative influence of non-performing loans on Spanish banks' capital buffer, which the authors expected since this is an ex-post risk measure. Ayuso et al. (2002b) also find evidence of a negative relationship between non-performing loans and excess capital. However, they also find a negative impact of the loan growth rate (which may be interpreted as a measure of banks' willingness to increase portfolio risk) and a positive impact of the share of risk free assets to total assets. These results suggest the counter-intuitive interpretation that more risky banks tend to hold less insurance against default.[6] Lindquist (2003), using a more sophisticated measure of risk but a less sophisticated modelling approach, also finds a negative risk effect which the author did not expect. Stolz and Wedow (2005) find a positive relationship between banks' liquidity and excess capital, which the authors argue may be interpreted as a positive risk effect as they proxy liquidity by banks' holdings of shares and bonds, and capital buffers may be held in order to hedge for the corresponding market risk.

Ayuso et al. (2002a) and Ayuso et al. (2002b) find a negative effect of the price of insurance, proxied by banks' Return on Equity (ROE), in capital buffers. Lindquist (2003) finds the same result using the β-coefficient for the Norwegian banking industry as a proxy for the cost of excess capital, since it is a measure of the industry level risk premium. Being an industry level variable, this proxy has the obvious shortcoming of allowing for no cross-section variation. Stolz and Wedow (2005) find a negative relationship between banks' Return on Assets (ROA) and excess capital, suggesting that banks with high returns may use profits to increase capital and therefore need to hold smaller capital buffers as insurance.

The four papers find a negative relationship between the business cycle and capital buffers. From a regulator's perspective this may seem like a reassuring result as banks seem to protect themselves with higher excess capital during downturns, when loan default rates are higher. On the other hand, banks may increase excess capital during downturns through changes in their portfolio in order to reduce the risk weighted assets on which the regulatory

[6] Crocket, A (1997) presents potential explanations for this result, most of which are based on moral hazard in banks' behaviour and the potential for regulatory arbitrage.

minimum level of capital is based.[7] Hence, from a macroeconomic perspective, this behaviour may not be quite as desirable as it is likely to amplify rather than dampen business cycles, as during downturns firms are more likely to be denied credit (at a reasonable cost) which should increase bankruptcies.

Stolz and Wedow (2005), Ayuso et al. (2002b) and Lindquist (2003) also find a negative effect of banks' size on excess capital. Moreover, Lindquist (2003) finds a significant positive effect of competitors' capital buffers which is interpreted as evidence of peer pressure as banks use solvency in order to signal the market of their credibility.[8] However, this result must be interpreted with caution, since Lindquist (2003) does not include the lagged dependent variable in the model, which in the presence of persistence in capital buffers may make inference invalid.[9]

4. The Data

Estimation in the present study is based on an unbalanced panel of yearly data from banks' financial statements reported to Banco de Portugal, Statistical Bulletins issued by Banco de Portugal and Reuters. Consolidated figures are used (except for banks that do not belong to any banking group and thus do not consolidate their data and variables for which data is only available on an individual basis) as capital requirements are imposed at the consolidated group level.[10] The dataset used for estimation covers 17 Portuguese banks from 1994 to 2004, though profitability data since 1993 was used. The choice of the period for analysis was made with the purpose of maximizing the number of observations while avoiding structural changes in the industry. Hence, by 1994 Portuguese banks had adapted to the regulatory framework developed in Basel I but by 2004 they had not yet started to adapt to the new Basel II rules. The first three observations of newly created banks were excluded in order to allow for some stabilization of their activity. Small banks specialized in investment banking were also excluded as they are likely to behave differently.

[7] Furfine (2000) develops a structural dynamic model of a banking firm, finding evidence that US banks adjust their loan portfolios to capital shocks.

[8] In fact, the hypothesis that the corresponding elasticity equals one is not rejected.

[9] See Bond (2002), for instance.

[10] For these variables, group level data is obtained by adding data for each of its members.

The capital buffer (BUF) is defined as the ratio between excess capital and the regulatory minimum. NPL1 is the ratio of non-performing loans overdue for less than one year to total loans and CREDG is the growth rate of total loans. PROV is the coverage ratio of non-performing loans by specific provisioning. ROE and ROA are each bank's Return on Equity and Return on Assets, whereas CF is banks' cash flow normalized by total assets. The variance of profits (VPROF) is the variance of banks' past ROA and is computed using profits from the previous three years to the current year in order to obtain a meaningful measure of volatility while minimising the loss of observations. For the same reason, profitability data since 1993 is used. STK is the weight of volatile income financial assets in banks' total assets and MKTD is the ratio of market debt (total liabilities deducted of deposits) to total liabilities. TIER1 is the ratio of Tier 1 to total own funds. Banks' size (SIZE) is measured by the natural logarithm of total assets. PSIG is the change in the Lisbon Stock Exchange general index and is thus constant across banks. YGAP is a simple output gap measure obtained through the application of a Hodrick-Precott filter to the real output series.[11] Since this filter is known to have a poor fit for the first and the last observation, it was applied to an output series covering more years than the sample. YGAP is defined as the ratio of output gap to potential output. MERGER is a dummy variable equal to one when a bank has been involved in a merger. Ratios and growth rates are defined in percentage points.

Descriptive statistics of the included variables are presented in Table 1. With regard to the distribution of the capital buffer in the sample, it should be noted that its value for the largest banks (4th quartile of total assets) is little above a third of the average buffer for the smallest banks in the sample (1st quartile of total assets). As seen in Figure 1, the in-sample aggregate capital buffer tends to be somewhat lower than the banking system's actual capital buffer. This difference may be explained by the fact that the banks that were eliminated from the sample – newly created banks and institutions with very specific activities – tend to have abnormally high levels of excess capital.

From the beginning of the sample to the end of the past decade, capital buffers have declined as the economy recovered from the 1993 recession, credit accelerated and

[11] The Hodrick-Prescott filter was implemented using the Stata command hprescott8.

decreased exchange rate risk due to the adoption of the Euro was anticipated. Furthermore, there was an historical excess of capital in the Portuguese banking system, and rationalization of its use increased during this period. From the turning of the millennium, as capital ratios approached the regulatory minimum and economic growth slowed down, the banking system's aggregate capital buffer initiated a slow and steady recovery.

5. The Model

Considering the small size of the available sample, which makes it impossible to estimate a comprehensive general model, the option was to estimate a parsimonious model and then test additional hypotheses on the initial specification presented below:

$$BUF_{it} = \beta_0 BUF_{it-1} + \beta_1 NPL1_{it} + \beta_2 PROV_{it} + \beta_3 CREDG_{it} + \\ + \beta_4 STK_{it} + \beta_5 SIZE_{it} + \beta_6 YGAP_t + \beta_7 MERGER_{it} + \delta + \eta_i + \varepsilon_{it} \quad (1)$$

where δ is a constant term, η_i is an unobservable variable that captures idiosyncratic features of each institution that are constant over time and $\varepsilon_{i,t}$ is a random shock

The lagged dependent variable is intended to capture capital buffers' persistence. As argued by Ayuso et al. (2002a) and Ayuso et al. (2002b), its coefficient may be interpreted as a measure of adjustment costs in capital buffers and its expected sign is thus positive. NPL1 is a credit risk measure that intends to capture the flow rather than the stock of non-performing loans, thus decreasing (but not fully eliminating) the ex-post character of this variable. Hence, β_1's sign should depend on how much lag it has. If it still measures ex-post risk, a negative sign is expected as banks where higher credit risk has materialized are expected to have lower excess capital. If, on the other hand, it is a forward looking risk measure, and since the current solvency regulation is known to have poor risk sensitivity, one would expect a positive sign as banks with higher credit risk should, *ceteris paribus*, be willing to purchase more insurance.

The coverage of non-performing loans by provisioning is expected to have a negative effect as banks that have already provisioned for more of their overdue credit should require

smaller capital buffers. Provisions are thus imperfect substitutes for capital as they are intended to cover expected rather than unexpected losses.

The high credit growth observed during the sample period may have contributed to decrease capital buffers through a direct effect if banks have not anticipated this growth. On the other hand, one would expect a positive effect if banks anticipated high credit growth and responded to it with a precautionary excess capital increase. Furthermore, since an increase in granted loans is not expected to materialize immediately in an increase in non-performing loans, controlling for credit growth may be important for a correct interpretation of NPL1.

Banks with a higher weight of stocks in their total assets are expected to hold higher capital buffers as their assets should be more volatile. As argued above, both banks' size and the output gap are expected to have a negative impact on excess capital.

A negative coefficient on $MERGER_{i,t}$ would suggest mergers consume capital, whereas a positive sign could be explained by precautionary behaviour or simply by the acquisition of a strongly capitalised bank.

5.1. Additional variables tested

a) $ROA_{i,t}/CF_{i,t}$ and $VPROF_{i,t}$

High and stable earnings are expected to decrease the level of excess capital as profits are the first line of defense against unexpected losses.

b) $PSIG_t$

Good stock market performance should increase capital buffers as banks tend to choose these times to issue new capital and the value of banks' capital should increase due to the likely increase in listed banks' share price and the increase in profits from stock holdings. Hence, the hypothesis that the stock market effect is stronger for banks with a higher weight of shares in their assets is also tested.

c) $ROE_{i,t}$

Higher cost of capital, proxied by banks' Return on Equity (ROE), is expected to have a negative impact in capital buffers.

d) $TIER1_{i,t}$

Banks with a higher ratio of Tier 1 to total own funds are expected to require smaller capital buffers as this ratio may not fall below 50%. Hence, for banks close to the minimum, a negative shock to Tier 1 capital will have a higher impact in the capital ratio as Tier 2 capital will also decrease. Furthermore, banks close to the minimum allowed ratio should have higher capital adjustment costs as supplementary capital is cheaper and faster to issue than core capital. These effects should be reflected in capital buffers as they are not considered in the definition of the regulatory minimum capital.

e) $MKTD_{i,t}$

Higher weight of market debt (total liabilities deducted of deposits) in total liabilities is expected to positively influence capital buffers as, on the one hand, banks should hedge for the increased exposure to liquidity risk and to changes in market sentiment, and on the other hand, banks with higher market debt should target higher credit ratings as the price of issued bonds depends on the issuer's rating.[12]

6. Methodology

The main advantages of using panel data are capturing both cross-section and time-series variation as well as allowing for meaningful inference using a sample with a relatively small number of cross-section observations over a short time period. Allowing for dynamics in the underlying process is relevant not only to infer on the persistence of the series but also to ensure that the estimates for other parameters are consistent.

In the estimation of dynamic models with a small number of time-series observations such as the ones described above, traditional estimation methods result in inconsistent estimates. Maximum Likelihood estimators may be inconsistent if the distribution of the

[12] The importance of rating targets had been acknowledged in Ayuso et al. (2002b) but, to our knowledge, had not been explicitly tested.

initial conditions is miss-specified.[13] In fact, in panels with a small number of time-series observations, this estimator's attractive properties depend on quite strong and un-testable assumptions. Direct Ordinary Least Square (OLS) estimation would also result in inconsistent estimates as $BUF_{i,t-1}$ would be correlated with the error term $v_{i,t} = (\eta_i + \varepsilon_{i,t})$ due to the presence of time invariant individual effects. The Within Groups estimator solves for this source of inconsistency as it eliminates the fixed effects by transforming the variables into deviations from their means:

$$BUF_{i,t-1} - \frac{1}{T-1}\left(BUF_{i,1} + ... + BUF_{i,t} + ... + BUF_{i,T-1}\right) \tag{2}$$

$$v_{i,t} - \frac{1}{T-1}\left(v_{i,2} + ... + v_{i,t-1} + ... + v_{i,T}\right) \tag{3}$$

However, the estimator will still be biased for small T as $-\frac{BUF_{i,t}}{T-1}$ in (2) will be correlated with $v_{i,t}$ in (3) and $-\frac{v_{i,t-1}}{T-1}$ in (3) will be correlated with $BUF_{i,t-1}$ in (2).[14]

Application of OLS after taking first differences of the variables would still yield inconsistent estimates as the regressor $\Delta BUF_{i,t-1} = BUF_{i,t-1} - BUF_{i,t-2}$ would be correlated with the error term $\Delta\varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}$. This problem may, however, be solved by using Two-Stage Least Squares (TSLS) with instrumental variables that are both correlated with $\Delta BUF_{i,t-1}$ and orthogonal to $\Delta\varepsilon_{i,t}$, as proposed by Anderson and Hsiao (1981).

Arellano and Bond (1991) build on this approach by developing an asymptotically efficient estimator in a General Method of Moments (GMM) framework, using an instrument matrix of the form:

[13] See Bond (2002).

[14] See Nickel (1981) for details.

$$Z_i=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ BUF_{i,1} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & BUF_{i,2} & BUF_{i,1} & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & BUF_{i,3} & BUF_{i,2} & BUF_{i,1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (4)$$

Where rows correspond to the first-differenced equations for periods t=3,4,…,T for individual i.

However, as Arellano and Bover (1995) remark, in the presence of persistence in the dependent variable, lagged levels produce weak instruments for differences. Nevertheless, assuming lagged differences are uncorrelated with the fixed effects, one may explore additional moment conditions by estimating level equations using lagged differences as instruments. This is the intuition behind System GMM proposed in Blundell and Bond (1998) and which basically consists in estimating a system of both difference and level equations using lagged levels to instrument differences and lagged differences to instrument levels. System GMM thus allows for increased efficiency, especially when the dependent variable is persistent, which is likely to be the case with capital buffers.[15]

There are one and two-step versions of this estimator. While the two-step version is asymptotically more efficient, its standard errors are known to be severely downward biased in finite samples.[16] Hence the finite sample two-step covariance matrix correction developed in Windmeijer (2005) is used.

Table 2 summarises the instruments used and the underlying assumption on the correlation of each regressor with the error term. Conservative assumptions have been made since they are relevant to the validity of the conclusions and not testable.[17]

The use of too many instruments relative to the number of cross-section observations is known to overfit endogenous variables, thus creating biased estimates. Hence, rather than using all available lags to instrument each variable, regressions were first estimated using

[15] The correlation between $BUF_{i,t}$ and $BUF_{i,t-1}$ is around 0.8.

[16] See, for example, Bond and Windmeijer (2002).

[17] Ayuso et al. (2002b), for instance, assume exogeneity of bank size.

one lag. The number of lags was then increased and the specification with the highest p-value for the Hansen J test of overidentifiyng restrictions was chosen. To further address this problem the instrument matrix Z_i in (4) was collapsed by adding into:

$$Z_i' = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ BUF_{i,1} & 0 & 0 & \cdots \\ BUF_{i,2} & BUF_{i,1} & 0 & \cdots \\ BUF_{i,3} & BUF_{i,2} & BUF_{i1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (5)$$

implying the use of one instrument for each variable and lag distance, rather than one instrument for each time period, variable and lag distance.

7. Estimation Results

Table 3 presents Blundell-Bond two-step finite sample covariance matrix corrected System GMM estimates for the model presented in section 5 and a reduced model (1.A). Estimation was carried out in Stata 8.0 using the *xtabond2* routine developed in Roodman (2005). The Hansen J test of overidentifying restrictions is not significant in any of the specifications above, which means that there is no evidence that the instruments used are invalid. AR (1) and AR (2) are the application of the autocorrelation tests developed in Arellano and Bond (1991) to check for first and second order autocorrelation in the residuals of the differenced equations. The fact that there is evidence of first order but not second order autocorrelation implies that the model is well specified in levels, as expected. Furthermore, the F-test for the null hypothesis that all coefficients equal zero is safely rejected in both models. Since the coefficient in $CREDG_{i,t}$ is not found to be statistically significant, this variable is eliminated from the regression and analysis is focused in the reduced model (1.A).[18]

Capital buffers are found to be persistent as the coefficient in $BUF_{i,t-1}$ is significant and positive, thus presenting evidence in favour of the adjustment cost hypothesis. There is statistically weak evidence of a negative relationship between $NPL1_{i,t}$ and excess capital,

[18] Even though the coefficient on the lagged dependent variable is not signifficant when credit growth is controlled for, it is relevant in models that do not include $CREDG_{i,t}$. This variable, on the other hand, has not shown to be signifficant even in specifications that do not include $BUF_{i,t-1}$.

which may suggest that this variable does not fully eliminate the ex-post character of the outstanding stock of non-performing loans or that credit risk is not a relevant determinant of banks' capital buffers. This may be because the regulatory capital ratio already adjusts enough to cover the extra credit risk (which is unlikely) or because the precautionary effect of increasing capital buffers to cover credit risk is cancelled by the moral hazard and regulatory arbitrage effects described in Crocket (1997) that actually lead more risky banks to hold less excess capital.

The negative coefficient in $PROV_{i,t}$ is consistent with the hypothesis that provisions are a substitute for capital buffers. The fact that banks with a higher weight of stocks in their total assets seem to hold higher capital buffers suggests that banks with higher exposure to market risk choose to hold more excess capital in order to cover for the excess risk not considered in the regulatory minimum requirements. As expected, a significant size effect is also found. However, no statistically significant effect of mergers is found, suggesting that mergers and acquisitions have not taken place at the expense of the overall system's capital.

Finally, as documented in the literature, a negative relationship between the output gap and capital buffers is found, which conveys that banks tend to cover the extra risk in cycle downturns with excess capital but also that the lending cycle may be pro-cyclical.

7.1. Additional Hypotheses

Since the variance of profits is a broad risk measure, the hypothesis that banks with higher and more stable profits require smaller capital buffers is tested on a model that does not include other risk measures. For increased robustness, income is measured both by ROA and by cash flow. Qualitatively, the conclusions are the same, confirming that banks with higher and less variable income do in fact tend to hold less excess capital. However, the coefficient in $ROA_{i,t}$ is statistically weaker than the one in $CF_{i,t}$. The F-test for the hypothesis that both the coefficient in profitability and the one in profit volatility are zero is significant at 10% in model (a.1) and at 1% in model (a.2).

Tests for the relevance of the growth of the PSI General stock market index were conducted on a specification which does not include the output gap as stock market

performance is strongly correlated with the business cycle.[19] Statistically significant evidence of a positive effect of stock market performance in capital buffers was found, suggesting that banks tend to choose times of good stock market performance to issue new capital or that during these periods the value of banks' capital is boosted by the likely increase in listed banks' share price and the increase in profits from stock holdings. These conclusions must, however, be interpreted with care, as the result may be spurious if changes in $PSIG_t$ do not cause changes in $BUF_{i,t-1}$ but, on the other hand, changes in the output gap cause changes both in capital buffers and in stock market performance.[20] If it is true that the positive impact of $PSIG_t$ in $BUF_{i,t-1}$ is related to impacts in profits from banks' holdings of stocks, one would expect the effect to be larger for banks with a higher weight of stocks in their total assets. The positive but not statistically significant coefficient in $PSIG*STK_{i,t}$ provides weak evidence for this hypothesis.

No significant effect of the cost of capital, proxied by ROE, was found. This may be related to the fact that the cost of capital is an adjustment cost to the capital buffer, since these costs are already taken into consideration through the inclusion of the lagged dependent variable. Furthermore, $ROE_{i,t}$ is measured in book-values whereas the true cost of capital is related to the market value of banks' return on equity.[21]

The negative but statistically weak effect of the ratio of Tier 1 to total capital provides weak evidence for the hypothesis presented in 5.1 d). Weak evidence is also found for the hypothesis that banks with a higher ratio of market to total debt hold higher capital buffers. Furthermore, this effect seems to be present only in the 94% larger banks.[22]

Given the small sample size, it is reassuring to note that the sign and significance of most coefficients remains stable across a wide range of specifications.[23] The effect of NPL1,

[19] Check the correlation matrix of the variables in the Appendix.

[20] In fact, a regression of $PSIG_t$ on a constant and $YGAP_{t-1}$ yields a significant coefficient on $YGAP_{t-1}$ whereas a regression of $YGAP_t$ on $PSIG_{t-1}$ does not, suggesting that the business cycle determines stock market performance but the opposite is not true.

[21] It is, however, impossible to compute banks' market value return on equity, as there are no estimates for most Portuguese banks' market value of equity.

[22] Evidence for this fact is once again weak as the hypothesis that both $MKTD1_{i,t}$ and $MKTD*SIZE_{i,t}$ are zero may not be rejected.

[23] In fact, estimation by Generalized Least Squares assuming Random Effects – the methodology used in Lindquist (2003) – provided the same qualitative results.

however, is not robust to the different specifications, suggesting that either banks do not adjust their capital buffers to credit risk or this variable is still a poor proxy for expected credit risk. Possible differences between domestic and foreign owned banks were also investigated through the inclusion of a dummy variable. However, this feature has not shown to be relevant.

8. Conclusions

The main purpose of this study was investigating which factors determine Portuguese banks' capital buffers, through the estimation of a dynamic panel data model.

Observed persistence in capital buffers suggests that there are relevant adjustment costs in banks' excess capital. On the other hand, high and stable profits and more conservative provisioning policies were found to be imperfect substitutes for higher capital. A statistically weak and not robust negative credit risk effect was found, suggesting that the credit risk proxy used does not fully eliminate the ex-post character of the proxies used in the literature. The intuitive and reassuring result of a positive risk effect was found for broad measures of asset risk and for the weight of market liabilities. Hence, these measures of risk seem to be relevant for the definition of the optimum capital level of banks. Rating objectives also seem to have a positive effect on capital buffers, a hypothesis which had been previously discussed in the literature but not explicitly tested. The hypothesis that larger banks tend to hold less excess capital was also confirmed.

The business cycle was found to have a negative effect implying that banks protect themselves when higher credit risk materializes, but also that their optimal choice of capital may amplify economic cycles. On the other hand, banks' capitalisation is positively related to the performance of stock markets.

These findings allow for a better understanding of the factors underlying changes in banks' capital buffers and provide a better basis for the discussion about regulatory changes in this field. These findings allow for a better understanding of the factors underlying changes in capital reserves of Portuguese banks and provide a better basis for the discussion about regulatory changes in this field. In particular, obtained results confirm the idea that banks adjust their capital reserves in response to changes in the risks they face, i.e. both

those directly resulting from changes in the macroeconomic environment throughout the cycle, and those resulting from banks' own decisions. However, it should be noted that with the analysis undertaken, it is not possible to identify the effects of the regulator's explicit and implicit intervention to lead institutions to adopt corrective measures of a prudential nature.

9. References

Alfon, I., Arguimón, I. and Bascuñana-Ambrós, P. (2005). "How Individual Capital Requirements Affect Capital Ratios in UK Banks and Building Societies". *Working Paper* No. 515, Banco de España.

Anderson, T.W. and Hsiao, C. (1981). "Estimation of Dynamic Models with Error Components". *Journal of the American Statistical Association*, No. 76, p.p. 598-606.

Arellano, M. and Bond, S. (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations". *Review of Economic Studies*, No. 58, p.p. 277-297.

Arellano, M. and Bover, O. (1995). "Another look at the Instrumental Variable Estimation of Error-Component Models." *Journal of Financial Econometrics*, No. 68, p.p. 29-52.

Ayuso, J., Pérez, D. and Saurina, J. (2002a). "Are Capital Buffers Pro-cyclical?" *Working Paper* No. 224, Banco de España.

Ayuso, J., Pérez, D. and Saurina, J. (2002b). "Los Determinantes del Excedente de Recursos Proprios de las Entidades Españolas". *Revista de Estabilidad Financiera*, Número 2, Mayo, Banco de España.

Banco de Portugal (2006). "Financial Stability Report 2006".

Basel Committe on Banking Supervision (1999). "A New Capital Adequacy Framework." *Consultative paper* issued by the BCBS, Bank for International Settlements, Basel.

Baum, C. S. (2004). "Hprescott: Stata Module to Implement Hodrick-Prescott Filter for Time-Series Data". *Statistical Software Components* S447001, Boston College Department of Economics, revised 01 Aug 2006.

Blundell, R.W. and Bond, S.R. (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models". *Journal of Econometrics*, No. 87, p.p. 115-143.

Bond, S. (2002). "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice". *Portuguese Economic Journal* 1, p.p. 141-162.

Bond, S. and Windmeijer, F. (2002), "Finite Sample Inference for GMM estimators in Linear Panel Data Models". *Cemmap Working Paper Series,* No. CWP04/02, Institute for Fiscal Studies, London.

Cornett M. and H. Tehranian (1994). "An Examination of Voluntary versus Involuntary Security Issuance by Commercial Banks: the Impact of Capital Regulations on Common Stock Returns". *Journal of Financial Economics*, Vol. 35, pp. 99-122, 1994.

Crockett, A. (1997). "The Theory and Practice of Financial Stability". *Essays in International Finance,* No. 203, April 1997.

Dietrich, D. and Vollmer, U. (2004). "Why Do Banks Hold Capital in Excess of Regulatory Requirements? A Functional Approach". *IWH Discussion Paper* No. 192.

Furfine, C. (2000). "Evidence on the Response of US Banks to Changes in Capital Requirements". *BIS Working Papers*, No. 88, June 2000.

Hodrick, R. and Prescott, E. (1997). "Post-war U.S. business cycles: An empirical investigation". *Journal of Money, Credit and Banking*, 29, 1-16.

Lindquist, K. (2003). "Banks' buffer capital: How important is risk?" *Working Paper* 2003/11, Norges Bank.

Milne, A. and E. Whalley (2001). "Bank Capital Regulation and Incentives for Risk-Taking". *mimeo*.

Myers, S. C. and Majluf, N. S. (1984). "Corporate financing and investment decisions when firms have information that investors do not have". *Journal of Financial Economics*, 13, pp. 187-221.

Nickel, S. J. (1981). "Biases in Dynamic Models with Fixed Effects". *Econometrica*, No. 49, p.p. 1417-1426.

Roodman, D. (2005). “Xtabond2: Stata Module to Extend xtabond Dynamic Panel Data Estimator.” *Statistical Software Components* S435901, Boston College Department of Economics, revised 05 Jul 2006.

Stolz, S. and Wedow, M. (2005). “Banks’ regulatory capital buffer and the business cycle: evidence for German savings and commercial banks”. *Discussion Paper Series 2: Banking and Financial Studies*, No 7/2005, Deutsche Bundesbank Research Centre.

Windmeijer, F. (2005). “A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators”. *Journal of Econometrics* No 126, p.p. 25-51.

Tables and Figures

Table 1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
$BUF_{i,t-1}$	152	46.297	37.591	1.180	208.430
$NPL1_{i,t}$	152	1.046	0.942	0.000	7.076
$PROV_{i,t}$	148	65.626	15.407	0.000	100.000
$CREDG_{i,t}$	135	19.313	28.757	-51.660	148.769
$STK_{i,t}$	152	2.902	2.526	0.009	12.153
$SIZE_{i,t}$	152	15.425	1.598	11.561	18.122
$YGAP_t$	187	0.588	2.430	-2.378	4.392
$ROA_{i,t}$	162	0.873	1.437	-0.080	10.970
$CF_{i,t}$	152	2.197	2.260	0.100	17.050
$VPROF_{i,t}$	111	0.212	0.721	0.000	4.421
$PSIG_t$	187	11.627	23.895	-20.700	65.200
$PSIG*STK_{i,t}$	152	32.514	113.030	-180.393	792.372
$ROE_{i,t}$	152	9.186	6.490	-4.154	32.791
$TIER1_{i,t}$	152	74.982	13.722	50.000	100.000
$MKTD_{i,t}$	152	42.554	24.492	11.006	100.000
$MKTD*SIZE_{i,t}$	152	642.570	326.752	162.142	1363.399

NOTE: ROA covers the period 1993-2004 while the remaining variables cover 1994-2004

Table 2. Instruments Used

Variable	Instrument	Underlying Assumption
$BUF_{i,t-1}$	$BUF_{i,t-2}$	Pre-determined
$NPL1_{i,t}$	$NPL1_{i,t-2}$	Endogenous
$PROV_{i,t}$	$PROV_{i,t-2}$	Endogenous
$CREDG_{i,t}$	$CREDG_{i,t-2}$	Endogenous
$STK_{i,t}$	$STK_{i,t-2}$	Endogenous
$SIZE_{i,t}$	$SIZE_{i,t-2}$	Endogenous
$ROA_{i,t}$	$ROA_{i,t-2}$	Endogenous
$CF_{i,t}$	$CF_{i,t-2}$	Endogenous
$VPROF_{i,t}$	$VPROF_{i,t-2}$	Endogenous
$ROE_{i,t}$	$ROE_{i,t-2}$	Endogenous
$TIER1_{i,t}$	$TIER1_{i,t-2}$	Endogenous
$MKTD_{i,t}$	$MKTD_{i,t-2}$	Endogenous
$PSIG_t$	$PSIG_t$	Exogenous
$YGAP_t$	$YGAP_t$	Exogenous
$MERGER_{i,t}$	$MERGER_{i,t}$	Exogenous

Table 3. Estimation Results

Variable	Model (1)	Model (1.A)
$BUF_{i,t-1}$	0.260 (0.98)	0.396 (1.84)*
$NPL1_{i,t}$	-11.478 (-1.64)	-8.196 (-1.43)
$PROV_{i,t}$	-1.125 (-2.07)*	-1.013 (-2.49)**
$CREDG_{i,t}$	0.249 (0.78)	
$STK_{i,t}$	11.457 (1.97)*	16.315 (2.03)*
$SIZE_{i,t}$	-20.989 (-2.69)**	-21.193 (-4.49)***
$YGAP_t$	-2.372 (-2.92)***	-2.089 (-2.73)**
$MERGER_{i,t}$	-6.440 (-0.60)	3.136 (0.39)
CONST	409.801 (2.55)**	387.258 (4.88)***
Hansen	0.695	0.705
AR (1)	0.050	0.054
AR (2)	0.958	0.652
F	0.000	0.000

NOTES: t-statistics are reported in parenthesis.
***, ** and * indicate statistical signifficance at the 1%, 5% and 10% confidence level, respectively, in a two-tailed t-test.
p-values are reported for the Hansen, AR(1), AR(2) and F tests

Table 4.1 Additional Hypotheses - Profitability, Stock Market Growth Rate

Variable	Model (a.1)	Model (a.2)	Model (b.1)	Model (b.2)
$BUF_{i,t-1}$	0.748 (11.76)***	0.675 (7.82)***	0.396 (1.92)*	0.358 (2.67)**
$NPL1_{i,t}$			-1.632 (-0.25)	-5.589 (-1.31)
$PROV_{i,t}$			-0.986 (-2.32)**	-0.640 (-1.24)
$STK_{i,t}$			14.501 (1.99)*	11.586 (1.68)
$PSIG_t$			0.215 (2.24)**	-0.667 (-0.74)
$PSIG*STK_{i,t}$				0.374 (1.08)
$ROA_{i,t}$	-3.316 (-1.17)			
$CF_{i,t}$		-3.274 (-2.14)**		
$VPROF_{i,t}$	8.614 (3.37)***	4.972 (2.10)*		
$SIZE_{i,t}$	-4.252 (-0.95)	-9.230 (-1.93)*	-15.896 (-3.24)***	-15.925 (-3.31)***
$YGAP_t$	-2.233 (-3.31)***	-2.001 (-3.29)***		
$MERGER_{i,t}$	-6.662 (-0.91)	-5.123 (-0.92)	-3.799 (-0.43)	6.127 (0.45)
CONST	79.960 (1.15)	166.078 (2.09)*	299.725 (3.26)***	284.566 (3.51)***
Hansen	0.513	0.695	0.557	0.529
AR (1)	0.023	0.020	0.046	0.061
AR (2)	0.896	0.983	0.469	0.209
F	0.000	0.000	0.001	0.001

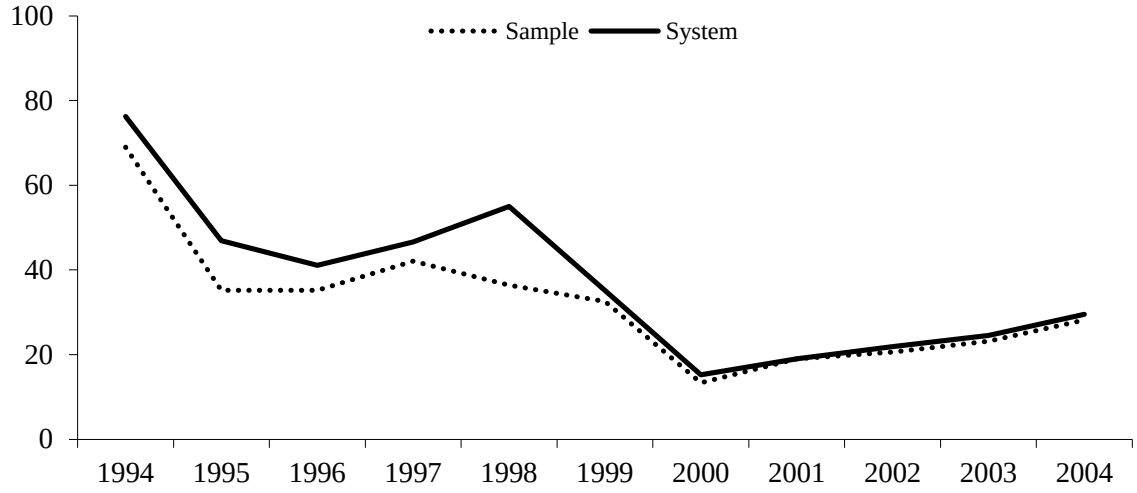
NOTES: t-statistics are reported in parenthesis.
***, ** and * indicate statistical signifficance at the 1%, 5% and 10% confidence level, respectively, in a two-tailed t-test.
p-values are reported for the Hansen, AR(1), AR(2) and F tests

Table 4.2 Additional Hypotheses - Cost of Capital, Weight of Tier 1 Capital in Total Capital, Weight of Market Liabilities in Total Liabilities

Variable	Model (c)	Model (d)	Model (e.1)	Model (e.2)
$BUF_{i,t-1}$	0.359 (2.68)**	0.377 (2.75)**	0.374 (2.59)**	0.365 (2.67)**
$NPL1_{i,t}$	-8.725 (-1.43)	-8.440 (-2.00)*	-8.295 (-2.14)**	-3.778 (-0.70)
$PROV_{i,t}$	-1.010 (-2.55)**	-1.103 (-3.34)***	-1.111 (-2.88)**	-0.983 (-3.49)***
$STK_{i,t}$	8.364 (1.89)*	15.681 (2.88)**	14.824 (2.99)***	17.850 (4.18)***
$ROE_{i,t}$	-0.119 (-0.05)			
$TIER1_{i,t}$		-0.524 (-0.44)		
$MKTD_{i,t}$			0.542 (1.04)	-4.495 (-1.18)
$MKTD*SIZE_{i,t}$				0.354 (1.33)
$SIZE_{i,t}$	-15.700 (-2.49)**	-24.077 (-2.41)**	-20.241 (-5.86)***	-36.963 (-2.75)**
$YGAP_t$	-1.650 (-2.72)**	-1.917 (-3.07)***	-1.817 (-2.81)**	-1.435 (-2.18)**
$MERGER_{i,t}$	-3.249 (-0.25)	5.693 (0.44)	5.743 (0.56)	1.635 (0.15)
CONST	324.365 (2.93)***	477.204 (1.99)*	358.870 (6.53)***	582.475 (3.17)***
Hansen	0.507	0.862	0.798	0.942
AR (1)	0.040	0.021	0.044	0.032
AR (2)	0.524	0.664	0.585	0.497
F	0.000	0.000	0.000	0.000

NOTES: t-statistics are reported in parenthesis.
***, ** and * indicate statistical signifficance at the 1%, 5% and 10% confidence level, respectively, in a two-tailed t-test.
p-values are reported for the Hansen, AR(1), AR(2) and F tests

Figure 1. Evolution of Portuguese banks' Capital Buffers (per cent)



An Assessment of Competition in the Portuguese Banking System in the 1991-2004 Period*

Miguel Boucinha Nuno Ribeiro

Abstract

The purpose of the present paper is to assess the competitive behaviour of Portuguese banking groups during the period ranging from 1991 to 2004, using the non-structural test developed by Panzar and Rosse. The main findings are that the Portuguese banking system has experienced weak competition between 1991 and 1996, underwent a period of restructuring until 2000, and from then until 2004 behaved consistently with perfect competition. Both private and, more markedly, domestic banks, seem to have competed more aggressively on occasions, and no relationship between competitive behaviour and bank size was identified.

JEL classification: G21; L13; C23

Keywords: Banking; Competition; Panzar and Rosse.

* The views expressed in this article are those of the authors and do not necessarily reflect those of the Banco de Portugal.

1. Introduction

Over the last two decades the Portuguese financial system went through major changes in the competitive environment in which financial intermediaries operate. The first liberalisation steps were taken in the mid 1980's with the opening of domestic markets to private initiative, against the background of an almost fully nationalised banking system. Until the early 1990's, banks remained tightly regulated in many dimensions of their activity. For instance, both prices and quantities in deposit and loan markets were administratively set or severely constrained, while great discretion remained in authorities' hands in what concerns banks' entry (both domestic and foreign) and branching decisions.[24] In this setting, the lack of competitive forces in the 1980's resulted directly from regulatory interference instead of stemming from market players' conduct. In fact, only in the early 1990's banks started to carry out their business in a full market environment, i.e. interest rate setting was free in all operations, while credit ceilings, a system of credit quotas defined at the bank level that was in effect during the 1980's, were abolished. In this way, we defined a priori the post-1990 period as the focus of our analysis of competitive conditions prevailing in the Portuguese banking market. This involved the identification of time series patterns in competition and the direct test of regime shift associated to participation in the euro area.

Figure 1 shows the evolution of concentration in the Portuguese banking industry according to the 3 and 5-bank concentration indices (C3 and C5, respectively, on the left scale) and to the Herfindahl-Hirshman Index (HHI, on the right scale), all derived from banks' total assets.[25] The three presented indicators show that concentration has increased over time as the deregulation period was followed by a consolidation trend across the market. From Figure 1, it is observable that over the period comprised between 1991 and 2004 there were two main consolidation waves. Until 1996, as the privatization program progressed, concentration increased almost linearly. Afterwards it remained relatively stable even though deep changes in the shareholding structure and control in some of the largest banks were observed in 2000.

[24] See Ribeiro (2007) for a brief overview of the liberalisation process.

[25] For the *k*-largest banks of a market with *n* banks, $C_k = \sum_{i=1}^{k} s_i$ and $HHI = \sum_{i=1}^{n} s_i^2$, where s_i is the market share of bank *i*. Also note that the unit of observation is the economic banking group, rather than the banking legal entity. For more details on this point, refer to Section 4.

Another relevant development to the national banking sector during the period under scrutiny was the event of privatizations and the increase in the weight of foreign banks, as the market was liberalized. In fact, if there were ten public banking groups operating in 1991, by 1996 there was only one — Caixa Geral de Depósitos. Consequently, the market share of public banks in total assets has decreased from close to 60% in 1991 to around 20% since 1996. On the other hand, the number of foreign banks has increased from 1991 to 2004 and so has their market share, which was around 5% in 1991 and slightly over 20% in 2004. The recorded increase in the market share of foreign banks was greater than that observed in the number of active foreign banks. This reflects a major acquisition carried on by a foreign bank during the above mentioned consolidation wave in 2000. In fact, similarly to most other European markets, de novo organic growth by foreign players in the Portuguese market was not particularly successful in the retail business.

According to the traditional structure-conduct-performance paradigm (Bain 1951), an increase in concentration should be linked to a decrease in competition. However, this result contradicts common wisdom and anecdotal evidence regarding the behaviour of the Portuguese banking sector during the period under analysis. This fact, which could be explained both by Baumol's contestability theory (Baumol 1982) and by the efficiency hypothesis posted by Demestz (1974), motivates the use of the non-structural test described below.

The approach taken to assess the degree of competition consists on specifying tests based on empirical reduced form revenue equations, as stated originally in Panzar and Rosse (1987). Revenues are explained by a vector of input prices and the sum of the corresponding elasticities is a statistic, the so-called H- statistic, with useful properties in the inference of competitive behaviour. However, the use of this statistic is not immune to criticism based on the assumptions underlying its use as a measure of competition in banking markets. Hence, the Panzar-Rosse methodology is a non-structural approach, as opposed to estimable specifications rooted on static oligopoly models, which establish testable relationships between market structure, direct measures of strategic behaviour and competition. Further, this methodology carries the crucial hypothesis that banks are essentially producers of a single product in the credit market, while all funding sources, including deposits, are considered inputs in banks' production function. Another controversial issue across most empirical studies is the definition of the appropriate variable to represent banks' revenue:

either interest income or total income used in levels or scaled by total assets. These issues as well as the possible ways of overcoming them in the empirical specification are discussed in detail in this paper. Particular attention is dedicated to the definition of the interest revenue variable in order to restrict the analysis to the domestic loan market, the most liable to exhibit market power in the investment side of banks' balance sheet. Similarly, the definition of the cost of funding variable takes due care of the different role of each bank as a provider of immediacy or payment services, in order to control for systematic differences in the presence of banks in those segments of funding markets where market power is more prone to emerge. Average funding cost is also adjusted to account for the presence of a bank in the interbank market both as a creditor and debtor.

Despite the shortcomings, the simplicity of this methodology explains its popularity in the study of competition in banking markets. For instance, it does not require price and quantity data on the services provided by banks, an issue that can often be problematic in the estimation of empirical structural equations of banks' behaviour, either because they are not available to researchers or due to the fluidity of these services in what concerns establishing a measure of their quantity. Another appealing property of this methodology is the fact that it allows for the inference of the interaction between input price shocks to the cost function and the revenue function, without requiring the estimation of output demand or cost functions. In addition, there is no need to worry about the appropriate relevant market in a geographic sense, as the input price to revenue relationship captures possible local market product differentiation on average, in the aggregate.

In the next section, the relationship between some common competition models and the results of the Panzar and Rosse (P-R) approach are derived. Section 3 presents a brief summary of previous empirical findings on the subject, whereas Section 4 presents the data and empirical methodology employed and results are shown in Section 5. Section 6 concludes.

2. The Panzar-Rosse approach

2.1 Competitive long-run equilibrium

To start with, let us establish the main positive result derived in Panzar-Rosse (1987), concerning the magnitude of the H-statistic in long-run perfect competition equilibrium.

From duality theory and under some regularity conditions, for some arbitrary production function $y = f(x_1, \dots, x_M)$ where x_i are M inputs, there is a function $C(y, w)$, the cost function, which results from the minimisation of the total production cost $C = \sum_{i=1}^{M} w_i x_i$ for each given output level y.

In long-run perfectly competitive equilibrium, price should equal marginal cost and free entry and exit conditions determine zero economic profit. For our purposes it translates into the two following expressions:

(1) $p_0^C - C_y(y_0^C, w) = 0$

(2) $p_0^C y_0^C - C(y_0^C, w) = 0$

where the superscript C stands for perfectly competitive equilibrium price and output levels.

Comparative statics in the neighbourhood of the competitive equilibrium can be undertaken taking the total differential of (1) and (2) and applying Cramer's Rule to the resulting system of equations. In particular,

(3) $$\frac{\partial y^C}{\partial w_i} = \frac{C_{w_i} - y^C C_{w_i y}}{y^C C_{yy}}$$

By Shephard's Lemma, the partial derivative of the cost function with respect to each input price is the conditional demand for the input itself, i.e. $C_{w_i} = x_i$ and equation (3) simplifies to $\frac{\partial y^C}{\partial w_i} = \frac{x_i^C - y^C(\partial x_i / \partial y_C)}{y^C C_{yy}}$.

Taking R_C as the total receipt $p_C y_C$ in the competitive equilibrium and using equality (2), in equilibrium, its derivative with respect to each factor price is

(4) $$\frac{\partial R_C}{\partial w_i} = \frac{C_y \left[x_i^C - y^C \frac{\partial x_i}{\partial y} \right]}{C_{yy} y} + x_i^C$$

Taking (4), multiplying by factor prices, aggregating and dividing by total receipt, the resulting figure is the so-called *H*-statistic, i.e. the sum of factor price elasticities of total revenue such as

(5) $$H = \sum_i \frac{\partial R_C}{\partial w_i}\frac{w_i}{R_C} = \frac{C_y}{C_{yy}}\left\{\frac{\sum_i w_i x_i^C - y^C \sum_i w_i (\partial x_i / \partial y^C)}{R^C}\right\} + \frac{\sum_i w_i x_i^C}{p^C y^C}$$

Using the definitions of the cost and marginal cost functions and its equilibrium levels as stated in (1) and (2), it becomes,

(6) $$H = \frac{C_y}{C_{yy}}\left\{\frac{C - y^C C_y}{R^C}\right\} + \frac{C}{R^C} = \frac{C_y}{C_{yy}}\left\{\frac{C - y^C p^C}{R^C}\right\} + \frac{C}{R^C}$$

It is straightforward from (2) to see that the first term vanishes and the second term is equal to 1.

The mechanics of what's in place in this problem can be explained as follows. Multiplying all factor prices by the same arbitrary proportion h, as cost functions are homogeneous of degree 1 in factor prices, we know that $C(y_0^C, hw) = hC(y_0^C, w)$ and $C_y(y_0^C, hw) = hC_y(y_0^C, w)$. Multiplying (1) by h it is straightforward to see that $p_1^C = hp_0^C$ and $y_1^C = y_0^C$ are solutions of the problem, with $w_1 = hw_0$. Similarly, from homogeneity of degree 1 of the cost function, the vector $\{p_1^C, y_1^C\}$ that solves condition (1), satisfies also condition (2), so that $R_1^C = C_1^C = hC_0^C \Leftrightarrow \frac{\partial R_1^C}{\partial w_i} = hx_{i,0}^C$ and $H = \sum_i \frac{\partial R_1^C}{\partial w_i}\frac{w_i}{R_1^C} = \frac{h\sum_i w_i x_{i,0}^C}{hC_0^C} = 1$.

In looser terms, it means that each bank is facing a perfectly elastic demand schedule which is tangent to the average cost curve at its minimum. Given the homogeneity of degree one in factor prices of both the marginal and average cost functions they shift in the same proportion as a simultaneous shock to all factor prices. The same happens to prices, while each firm's market equilibrium quantity remains unaffected. In this way, shifts in factor prices transmit fully to total revenue. This is a powerful result allowing for the implementation of a direct test of market players' behaviour consistent with price taking behaviour and in which no bank is earning abnormal profits.

2.2 Monopolistic behaviour in markets for imperfect substitutes

Models of monopolistic behaviour are the most plausible to consider a priori in banking as, even if we consider banks as producers of a single product very homogeneous in its intrinsic characteristics, banks differentiate among each other by means of brand advertising and/or branch location. The resulting framework is one in which banks differentiate in quality and are profit maximisers over their own specific perceived demand. For the purpose

of studying the *H*-statistic's properties we will consider both the case of a market with a pre-fixed number of banks, which is conceptually similar to perfect monopoly behaviour in each bank's captive demand, and the case of free entry and exit a la Chamberlin. In the monopoly case, market players can sustain supra-normal profits, because they do not face the threat of entry, at least in the short-run. In the latter case, market conditions attract to or drive banks away from the market, so that in the long-run equilibrium economic rents cannot be extracted.

2.2.1 Monopolistic behaviour in markets for imperfect substitutes

For simplicity, we consider a market equally shared by a fixed number of banks, in such a way that we can concentrate on some arbitrary representative bank. When choosing its output level, the representative bank ignores the effect that changes of competitors' output have on the industry prevailing price in such a way that each bank can be considered a monopolist in its particular product variety, local market or any other characteristic mapping each bank to a different brand from the consumer perspective. In addition, we consider that a differentiable inverse demand schedule $p_i = p(y_i, n)$ exists for the representative bank such that[26]

(7) $$\begin{cases} p_y = \partial p_i / \partial y_i < 0 \\ p_n = \partial p_i / \partial n < 0 \\ \partial \gamma / \partial n \geq 0 \end{cases}$$

Where the first condition consists on assuming a standard negatively sloped demand curve, the second condition is a way of stating analytically that the market each bank faces shrinks with the number of competitors and the last condition, where γ stands for the symmetric of demand elasticity, implies that each bank's perceived demand "flattens" with the number of banks. In this class of models, perceived demand elasticity can be mapped into the Lerner index, i.e. the relative price to marginal cost spread, which is a measure of market power.

[26] The following setting borrows directly from Vesala (1995).

In such a setting, the representative bank chooses the output level y^M which solves implicitly the first order condition

(8) $$R_y(y^M, n) = C_y(y^M, w)$$

Differentiating (8) relatively to w_i, applying Shephard's Lemma and rearranging terms yields:

(9) $$\frac{\partial y^M}{\partial w_i} = \frac{1}{\partial^2\pi/\partial y^2}\frac{\partial x_i}{\partial y^M}$$

Differentiating the revenue function in the neighbourhood of equilibrium output with respect to w_i yields:

(10) $$\frac{\partial R^M}{\partial w_i} = \frac{\partial R^M}{\partial y^M}\frac{\partial y^M}{\partial w_i} = \frac{\partial R^M}{\partial y^M}\frac{1}{\partial^2\pi/\partial y^2}\frac{\partial x_i}{\partial y^M}$$

The *H*-statistic is obtained by multiplying (10) by w_i and aggregating over the set of inputs:

(11) $$H = \sum_i \frac{\partial R^M}{\partial w_i}\frac{w_i}{R^M} = \frac{\partial R^M}{\partial y^M}\frac{1}{(\partial^2\pi/\partial y^2)R^M}\sum_i \frac{\partial x_i}{\partial y^M} =$$

$$= \left[p(y^M, n)\left(1 - \frac{1}{\gamma(y^M, n)}\right)\right] C_y \frac{1}{(\partial^2\pi/\partial y^2)R^M} \leq 0$$

This expression is clearly non-positive as $\partial^2\pi/\partial y^2$ is negative by the second order condition of profit maximisation and $\gamma(.)$ Is higher than one in equilibrium for a profit maximising monopolist. Hence, a negative (or null) value of the *H*-statistic is consistent with a monopoly (n=1), a perfect cartel that replicates monopoly market outcomes and markets with a predetermined number of banks behaving monopolistically. Another interesting result is that *H* is strictly decreasing with respect to the symmetric of the perceived demand elasticity, i.e. higher (less negative) values of *H* can be interpreted as higher effective market power exercised by the monopolist.

2.2.2 Long-run (Chamberlinean) monopolistic competition equilibrium

Considering the number of banks as endogenous and no longer predetermined as in the previous section, the long-run equilibrium is attained when no incentives prevail to bank entry. Positive rents attract new banks to the market shifting the representative bank's

demand schedule up to the point where each bank is operating at its economic profit break-even point.[27] Analytically it implies that in addition to the representative bank's first order condition stating that marginal revenues equate marginal costs as in (8), a zero-profit condition has to be imposed, such as:

$$R(y^*, n^*) = C(y^*, w) \tag{12}$$

where n^* is the long-run equilibrium number of symmetric banks.

Totally differentiating both conditions with respect to factor prices and the number of banks, solving for $\partial y^n / \partial w_i$, multiplying by factor prices and aggregating over all factors yields:

$$H = 1 + \frac{R_y^* \left[R_n^* R_y^* - R^* R_{yn}^* \right]}{R^* \pi_{yy}^* R_n^*} \tag{13}$$

Rewriting the numerator in terms of the inverse demand function it becomes:

$$H = 1 + \frac{R_y^* (y^*)^3 \left(p_y^*\right)^2}{R^* \pi_{yy}^* R_n^*} \frac{\partial \gamma}{\partial n} \tag{14}$$

As $R_y^* = p_y^* y + p^* = \frac{\gamma - 1}{\gamma} p^*$ and $R_n^* = y_p^* p_n^* p^* + p_n^* y^* = p_n^* (1 - \gamma) y^*$, expression (14) simplifies to:

$$H = 1 + \frac{R_y^* (y^*)^3 \left(p_y^*\right)^2}{R^* \pi_{yy}^* R_n^*} \frac{\partial \gamma}{\partial n} = 1 - \frac{1}{\gamma} \frac{p^* y^* (y^*)^2 \left(p_y^*\right)^2}{R^* y^* \pi_{yy}^* p_n^*} \frac{\partial \gamma}{\partial n} = 1 - \frac{1}{\gamma^3} \frac{(p^*)^2}{y^* \pi_{yy}^* p_n^*} \frac{\partial \gamma}{\partial n} \leq 1 \tag{15}$$

The second term is clearly positive, given the assumptions in (7). Further, all else constant, the *H*-statistic is increasing in the perceived demand elasticity and converges to 1 as it approaches infinity, replicating the perfect competition outcome derived above.

It should be borne in mind, however, that this concept of equilibrium is built up on assumptions of individual firm profit maximisation, taking all other firms' actions as constant. This assumption implies no strategic interaction of market players, arguably a too naïve description of market players' behaviour in an oligopoly (for e.g., see Kreps (1990) for a particularly sanguine sceptical discussion of this model).

[27] This concept of long-run equilibrium is usually labelled "Chamberlinian equilibrium" and borrows directly from Chamberlin (1933).

Anyway, the non-acceptance of both the null hypothesis of $H \leq 0$ and $H = 1$ has been interpreted in previous empirical applications as resulting from a market environment which is not consistent with either long-run perfectly competitive equilibrium or perfect tacit collusive behaviour. Instead, it must contain elements that are akin to both monopoly and perfect competition, representing some intermediate position, as it is the case in the monopolistic competition equilibrium (see, e.g., Henderson et al. (1980), page 193).

3. Previous empirical findings

Most studies applying the Panzar and Rosse methodology reject both the hypothesis of monopoly (or perfect cartel) behaviour and that of perfect competition. In a cross-country analysis for the EU-15 for the period between 1997 and 2003, Casu and Girardone (2006) find a value for the *H*-statistic between zero and one, thus rejecting both monopoly and perfect competition, at EU-15 level, as well as for most countries individually, including Portugal.[28] Koutsomanoli-Fillipaki and Staikouras (2004) also reject both monopoly and perfect competition, for a period ranging from 1998 to 2002. However, in this study, the *H*-statistic is only estimated for the EU-15 as a whole and not for Portugal individually. Bikker et al. (2006) present a cross-country study where competitive conditions are estimated for 101 countries between 1986 and 2005. Since this paper intends to make a methodological point, several different specifications are estimated, and the results obtained for Portugal range from monopoly, in their preferred specifications, to perfect competition in the models they consider misspecified.[29] In general, available results for Portugal are similar to those found for other countries. For a more complete summary of results obtained in applications of the Panzar and Rosse methodology, refer to Table 1 in Casu and Girardone (2006) and to Table 1 in Bikker et al. (2006).

4. Data and empirical methodology

The dataset used in this study was obtained from banks' financial statements reported to Banco de Portugal. The database comprises an unbalanced panel of yearly data for all active banks operating in Portugal from 1991 to 2004. Since detailed consolidated accounting data

[28] Perfect competition was not rejected for Finland and monopoly was not rejected for Greece. However, the authors note that inference for these two countries was based on a small number of observations.

[29] See the next Section for a discussion on this methodological point.

is available only for the most recent period, data on individual basis was used instead. However, since interest lies in comparing the behaviour of different economic units rather than legal entities, data for banks belonging to the same economic group were aggregated into one decision making unit.

All banks operating in Portugal are required to report financial statements to Banco de Portugal. However, there is a large number of small banks that mostly operate in investment banking and are thus likely to behave differently from most commercial or universal banks. Hence, in order to obtain a sample of reasonably homogenous banks, observations for institutions with less than 15 branches or 15 employees were eliminated. Banks that do not take any deposits from customers were also eliminated, and it was checked that positive values for total assets and equity were reported for all observations in the sample. The first two years of activity of new banks were also eliminated, as it seems reasonable to assume that during early stages banks may behave differently than once their activity is stabilized. 25 banking groups and a total of 197 observations survived this process. For each year, the final sample includes no less than 92% of loans granted to customers, 94% of customer deposits held and 92% of total assets of the Portuguese banking system. All money variables were deflated using the GDP deflator.

The definition of banks' outputs and inputs is by no means simple. On the one hand, the "production approach" to bank modelling regards banks as firms producing services which are related to loans and deposit accounts, thus measuring output by the number of deposit accounts serviced and the number of loans originated and input by labour and physical capital. On the other hand, according to the "intermediation approach" (Sealey and Lindley (1977)), banks' main activity is granting loans and investing in securities and other assets using funds obtained through deposits, purchased funds and equity.[30] Hence, while both approaches agree in classifying labour and physical capital as inputs, they present a conflicting view as to whether deposits should be classified as an output or as an input. Since economic theory does not give clear guidance as to which modelling approach best describes the behaviour of the banking firm, it is somewhat reassuring to note that estimated cost functions appear relatively insensitive to which approach is followed (Humphrey (1990)). On the other hand, there is empirical evidence suggesting that deposits overall behave

[30] See Freixas and Rochet (1998), pp. 77-79, on the production and intermediation approach.

primarily as inputs (examples are Gilligan and Smirlock (1984), Hughes and Mester (1993), Shaffer (1994) and Hughes, Mester and Moon (2000)). The Panzar and Rosse methodology used in this study follows the intermediation approach. Hence banks are modelled as firms that use labour, physical capital and funds in order to produce loans.

There has been considerable debate in the literature as to whether the dependent variable used to estimate empirical P-R equations should consider total or only interest revenue. On the one hand, the fact that the P-R method considers loans as banks' main output, suggests interest revenue should be used. On the other hand, the increase in the relative importance of commissions and fees in banks' total revenue should not be neglected. Hence, in the baseline specification, we choose to use interest revenue as the dependent variable and include the ratio of other income to interest income as a regressor, thus accounting for changes in income structure. Nonetheless, in order to check for the sensitivity of results, models where the dependent variable includes both interest and commission and fee income were also estimated.[31]

A third matter of discussion is whether banks' size should be controlled for. Since it seems overly simplistic to assume banks' size is uncorrelated with input prices, it is likely that the exclusion of a scaling variable could bias the estimates for the elasticities of factor prices. On the other hand, Bikker et al. (2006) point out that the use of a scaling variable (either as a regressor or by defining the dependent variable as the ratio of revenue to total assets rather than the absolute value of revenue) effectively turns the revenue equation into a price equation, and the sum of the elasticities of the output price with respect to input prices is positive by definition, and independent of the industry's degree of competition. Hence, including a scaling factor in the estimated equation could introduce a positive bias the estimate of *H*. The possible presence of "errors-in-variables" due to the approximation of input prices should, however, act in the opposite direction, as it should bias the estimated coefficients downwards, whether a scaling factor is used or not.

[31] The choice to include the ratio of other revenue to interest revenue is particularly interesting since $lnTR = \ln(IR + OR) \approx lnIR + (OR/IR)$, as put forward in Bikker et al. (2006). Note, however, that OR/IR may not be exogenous in the regression, and so we do not test whether its coefficient equals minus one (in which case the regression would be equivalent to one where the dependent variable is total income) or zero (in which case the regression would in turn be equivalent to one where the dependent variable is interest income and other income is not controlled for). Nonetheless, estimation results where OR/IR is not considered and the dependent variable is either total income (see Table 3) or interest income (not shown but available upon request) have yielded similar estimates for the *H*-statistic to those obtained with the baseline specification.

The baseline specification for the empirical reduced form revenue equation stemming from log-linear marginal revenue and cost functions is presented below[32]:

$$(16) \quad ln(Interest\ Revenue)_{it} = h_1 ln\, w_{L_{it}} + h_2 ln\, w_{K_{it}} + h_3 ln\, w_{F_{it}} + X'_{it}\beta + \delta + \eta_i + \varepsilon_{it}$$

where Panzar and Rosse's *H*-statistic is obtained by $\sum_{k=1}^{3} h_k$, X_{it} is a vector of control variables, δ is a constant term, η_i is an unobservable variable that captures idiosyncratic features of each institution that are constant over time and ε_{it} is a random shock.

The dependent variable used in the baseline specification is the natural logarithm of interest revenue obtained from loans granted to domestic clients. The option to focus the analysis only on the portion of interest revenue earned on loans rather than including all interest income is explained by the fact that banks are known to have little market power on the remaining interest earning business, such as interbank and securities activities. As such, the analytical interest lies on testing how competitive banks are in customer lending. Hence, a broader specification of the dependent variable would very likely overestimate the competitive conditions in the banking system. This is a novelty feature of this study worth emphasising as, to our knowledge, all previous studies apply this methodology to all interest revenues.

The average price of labour — w_L — is proxied by the ratio of labour costs to the number of employees, whereas the ratio of (tangible and intangible) capital expenditure to (tangible and intangible) fixed assets — w_K — proxies the cost of capital and the ratio of interest paid to interest bearing debt — w_F — measures banks' average funding cost.

As to what concerns the control variables, vector X_{it} includes the ratio of demand deposits to total deposits and that of market liabilities to total liabilities in order to account for banks' funding mix, whereas the ratio of short-term loans to total loans and of interbank assets to customer loans, on the other hand, intend to capture the asset structure. The increasing importance of banks' off-balance-sheet activity is controlled for by the inclusion of the ratio of off-balance-sheet activity to total assets. The ratio of assets to branches intends to capture different branching strategies, measuring systematic differences in banks' branch

[32] The use of the loglinear form is widespread among studies applying the P-R methodology, as it typically improves the regression's goodness of fit and may reduce simultaneity bias (De Bandt and Davies 2000). Furthermore, Molyneux et al. (1996) found that a loglinear revenue equation yielded similar results to more flexible translog equation.

density. The share of customer loans that have defaulted during each year is a credit risk measure that attempts to capture the flow rather than the stock of non-performing loans, thus decreasing the ex-post character of this variable. In turn, the ratio of equity to assets should proxy banks' risk aversion once credit risk is controlled for. The ratio of other revenue (composed of net commission and fee income) to interest revenue, as discussed above, intends to capture the increasing role of non-interest revenue in banks' income. The inclusion of asset quartile dummies in the equation is a compromise solution that intends to address the misspecification described in Bikker et al. (2006), while at the same time controlling at least partially for the correlation between banks' size and input prices.[33] Finally, dummy variables identifying when a merger has occurred, or if a bank is foreign or public are also included.[34] Descriptive statistics of the included variables are presented in Table 1.

Equation (16) was first estimated for the whole sample, including domestic and foreign as well as private and state owned banks for the period ranging from 1991 to 2004. However, as discussed in Section 1, this time period is by no means homogenous, since during the early to mid-90's the Portuguese banking system underwent a phase of privatizations, consolidation and liberalization, while preparing for euro area participation. Hence, we perform a sequential test for differences in the *H*-statistic through time by first estimating an equation using data for the first four years in the sample and checking whether the estimate for the fourth year is statistically different from that obtained for the initial period comprising the first three years. If so, a new period starting on the fourth year would be created; if not, 1994 would be pooled with 1991-1993. This process was repeated until 2004, restricting each period to comprise at least three years of data.[35]

In order to test for differences in the competitive behaviour of different types of banks, equations considering only domestic banks and only private banks were estimated. Even though it would be more informative to allow the estimate of the degree of competition to

[33] In the sample used, the simple correlation coefficient between the logarithm of assets and the price of labour, capital and funds is 0.36, -0.07 and -0.23 respectively.

[34] Only domestic public banks are classified as public, since public banks operating abroad are likely to exhibit a different behaviour than that of local public banks.

[35] One reason to maintain this restriction is connected to the small size of the available sample. Furthermore, De Bandt e Davies (2000) advocate the importance of the time series component as they find variable results for cross-section equations even with information for a larger number of banks.

vary across types of banks, this option is not feasible due to the small number of public and foreign banks in the sample. Hence, the statistical significance of the difference in competitive behaviour observed across types of banks is not tested.

As several authors have pointed out (examples are Bikker and Haaf (2002) and Hempel (2002)), small banks may have more market power in local markets, whereas larger banks are generally believed to face greater competition. Note that this argument may not apply directly to Portugal, considering the relatively small size of the national market and the fact that most regions tend to be served by at least one large bank. Nevertheless, the only truly effective way of addressing the misspecification pointed out in Bikker et al. (2006) while avoiding the introduction of other sources of bias is to analyse the competitive behaviour of similarly sized banks, thus avoiding the need to use a scaling variable. Hence, the same estimation procedure is applied to small and large banks separately, where small and large is defined according to whether a bank's total assets are above or below average total assets for each year. We divide the sample in only two groups in order to minimize the loss of degrees of freedom, and the average is chosen over the median as, due to the high concentration in the Portuguese banking market, the group of the 50% largest institutions is very heterogeneous in terms of size. Once again, the small sample size introduces limitations in that it renders unfeasible the estimation of different measures of competition through time when the sample is divided between small and large banks.

The interpretation of the *H*-statistic depends on whether or not banks are in a state of long run equilibrium. In fact, while the result that the sum of factor price elasticities of a monopolist's reduced form revenue function must be non-positive holds even in the short run, results for models of perfect and monopolistic competition depend on the assumption that the firms are observed in long-run equilibrium (see Panzar and Rosse (1987)). As is common practice in studies applying the Panzar and Rosse methodology, we use the fact that in equilibrium risk-adjusted rates of return should be equalised across banks.[36] Thus, banks' return on assets (ROA) should be uncorrelated with input prices when the market is in equilibrium. A direct test of equilibrium consists in estimating the equation for revenue with ROA as the dependent variable and performing a test to the null hypothesis $H = 0$

[36] See for example Shaffer (1982) and Molyneux et al. (1994).

(equilibrium) against $H < 0$ (disequilibrium), where H is the sum of factor price elasticities with respect to the profitability measure.

To test for the robustness of the results, some alternative specifications were estimated, including the use of total rather than interest income as the dependent variable and using alternative scaling variables, such as the natural logarithm of total assets, scaling income by total assets and not controlling for size differences at all.

In order to account for the existence of non-measurable bank specific factors that are invariant through time, the fixed-effects estimator was used, thus allowing for possible correlation between the individual effect and the explanatory variables.

5. Results

Table 2 presents fixed-effects estimation results for equation (16), as well as for the auxiliary regression used to perform the long run equilibrium test. The estimates for elasticities of interest income with respect to each of the three inputs considered prove to be positive. The estimate for the *H*-statistic, laying at 0.691, changes only marginally when statistically non-significant variables are eliminated from the regression.[37] The test for monopoly performed is a one sided test for the null hypothesis $H \leq 0$ versus the alternative $H > 0$. The former is clearly rejected in favour of the latter, thus providing compelling evidence against the hypothesis that the Portuguese banking system has operated as a monopoly or a perfect cartel on average during the period under scrutiny.[38] If, on the other hand, banks were under perfect competition, the *H*-statistic should equal one. A two sided test for this hypothesis is thus performed and, as shown in Table 2, the corresponding p-value is close to 10%, so that it is not clear whether perfect competition should be rejected or not. Another relevant result reported in Table 2 is that the application of the long-term equilibrium test described in the previous section does not allow for the rejection of the null hypothesis, therefore providing no evidence to reject the assumption that the Portuguese banking industry was in long-run equilibrium during the relevant period Hence, one may

[37] Results of this more parsimonious regression are not reported but are available from the authors upon request.

[38] As remarked in Bikker et al. (2006), even though a large number of studies applying the Panzar and Rosse methodology have used a two sided test, thus specifying the alternative hypothesis as $H \neq 0$, this is incorrect, as under monopoly the *H*-statistic may assume negative values (see Section 2.2.1).

conclude that, on average, in the period ranging from 1991 to 2004, the behaviour of Portuguese banks cannot be assessed as consistent with alternative forms of monopoly-like conduct (such as perfect cartel or monopolistic competition in a market without the threat of entry), and it is not clear whether it is consistent with perfectly competitive behaviour, or whether it is best described as stemming from a long-run monopolistic competition model with weak market power.

The control variables regarding the funding mix have not shown to be significant, whereas the negative sign on the variable which measures the maturity structure of granted loans suggests that banks for which the weight of short term loans is more important tend to earn less revenue, which is consistent with the fact that credit risk adjusted spreads tend to be lower on short than on long term loans. The same reasoning, i.e. differences in spreads, may be used for the interpretation of the result that banks with a higher ratio of interbank assets to customer loans tend to earn lower interest revenue from customer loans, whereas it seems natural that banks which are more active in the interbank and securities market relatively to the customer loan market, for a given value of total assets, earn less revenue from the latter business. Banks with more off balance sheet activity seem to earn higher interest revenues, which is possibly driven by the fact that banks with more off balance sheet activity tend to have a riskier profile.

Banks with relatively less, and possibly larger branches tend to earn higher interest income, whereas the measures of credit risk and of risk aversion have the expected signs, even if the former is not statistically significant. As expected, the coefficient on the variable which controls for the ratio of other revenue to total revenue yields a negative sign. As expected, the estimated coefficients for the dummy variables identifying the quartile of the asset distribution to which each bank belongs indicate that, all else constant, smaller banks tend to earn less revenue. As to what concerns the remaining control variables, mergers do not seem to have a significant impact on interest revenue earned whereas, *ceteris paribus*, foreign banks seem to earn more interest revenue, while the opposite result is found for public banks.

Table 3 shows a series of robustness tests in the form of alternative specifications to Model [1] which illustrate the impact in results of different choices regarding the dependent variable and the scaling variables used. As to what concerns different definitions of the dependent variable, specifications where (the natural logarithm of) total rather than interest

income is chosen as the dependent variable are presented. From comparison of the first two and of the last two lines in Table 3, one finds that considering total rather than interest revenue has virtually no impact, either in the point estimate for *H*-statistic or in the tests conducted upon it.

Another relevant robustness test involves checking the sensitivity of results to different scaling variables. From the results presented in Table 3 one finds that, as expected, specifying the dependent variable as the ratio of revenues to assets yields virtually the same results as specifications where the natural logarithm of assets is included as a regressor.[39] If, instead, one controls for the size difference of banks through the use of asset quartile dummies, the estimate for the *H*-statistic decreases slightly when interest income is used as the dependent variable. A more relevant change is the fact that, as reported in Bikker (2006), standard deviations are compressed in "price equations", which might lead to reject perfect competition too often, even if a point estimate that is closer to one is obtained. As shown in the last column of Table 3 above, specifications where no scaling variable is used yield a somewhat lower estimate for the *H*-statistic, whereas results for hypothesis tests remain unchanged relatively to specifications where total assets are considered.

Hence, considering a series of alternative specifications, it remains quite clear that, from 1991 to 2004, the Portuguese banking sector has not operated under monopoly. However there is now greater evidence towards the rejection of perfectly competitive behaviour as well, and so the hypothesis of an intermediate situation, such as that of monopolistic competition, seems more appropriate.

As briefly discussed in the first section of this study, the Portuguese banking system underwent significant changes during the sample period. Hence in order to investigate whether the process of liberalization and consolidation has had an impact in competitive conditions, the estimated *H*-statistic is allowed to vary over time without any particular functional form being imposed upon it, through the method described in the previous section. Aggregation tests to find homogenous periods were conducted on the specification presented in Model [1]. As a result, three periods were obtained: a first period of consolidation and adjustment to less restrictive regulations — 1991 to 1996; a second period of post-consolidation adjustment, which includes the beginning of euro area participation —

[39] This was expected since $ln(x/y) = \ln(x) - \ln(y)$.

1997 to 2000; and a final period of relative maturity, ending in 2004. Table 4 presents results for fixed-effects estimates of the *H*-statistic for the three periods, as well as for the whole sample, using as scaling variables the natural logarithm of total assets (line 1) and dummy variables for asset quartiles (line 2) as well as using no scaling variable (line 3).

During the first period, perfect competition is rejected in all specifications and monopoly is rejected only when the natural logarithm of total assets is used as a scaling variable. For the other specifications, which according to Bikker et al. (2006) should be more reliable, the monopoly hypothesis is not rejected and the estimate of the *H*-statistic is either negative or close to zero. Furthermore, during this first period of intense consolidation and privatization, there is no evidence to reject the hypothesis that the Portuguese banking system was operating in equilibrium. Hence, one concludes that the degree of competition was relatively low during this period. As to what concerns the next period, even if conclusions for the hypothesis tests on the H-statistic are the same, there is evidence that the system was not operating under long-run equilibrium. Hence, estimated coefficients constitute no evidence of collusive behaviour, since while the rejection of monopoly in the first specification remains valid in this context, non-rejection of $H \leq 0$ under disequilibrium no longer implies that the industry has behaved jointly as a monopoly. In the most recent period, while there is strong evidence to reject perfect monopoly (as well as perfect cartel or short run monopolistic behaviour without threat of entry) in all specifications, perfect competition is no longer rejected under any of the three specifications. Furthermore, differences in the estimate of the *H*-statistic between the intermediate and the most recent period are, both in the magnitude and in the statistical significance of the estimated coefficients, more striking than those found between the first and the second period.

Since foreign banks are likely to behave differently from domestic banks, a replication of the above results while restricting the sample to domestic banks is presented in Table 5. An increase in the estimated value of the *H*-statistic across most periods and specifications is observed.[40] In fact, even if conclusions regarding the first period remain unchanged, perfect competition is only rejected for domestic banks between 1997 and 2000 when the natural logarithm of total assets is included in the regression, which suggests domestic banks have

[40] Exceptions are the estimates for the H-statistic for the first and the last period when the natural logarithm of total assets is used as a scaling variable, which are slightly lower for domestic than for all banks.

behaved more competitively than the banking system as a whole during this period. Obtained results for the period between 2001 and 2004 when the natural logarithm of total assets is not considered explicitly are consistent with the hypothesis that domestic banks might have behaved too competitively in this period, which may be rationalized under a more complex, dynamic model, where banks aggressively fight for increased market share in order to capitalize on it with high profits in the future.[41] Furthermore, there is no evidence that domestic banks have not operated under long run equilibrium in the eve of the EMU. Restricting the estimate for the degree of competition exerted by domestic banks to be constant from 1991 to 2004, if one once again concentrates on specifications where the natural logarithm of total assets is not explicitly considered, higher values for the *H*-statistic are obtained when the sample is restricted to domestic banks, and there is no evidence to reject perfect competition. As can be seen by comparison of tables 4 and 5, when the logarithm of assets is included as a scaling variable, results for the relevant hypothesis tests do not change even though the point estimate for *H* decreases.

Table 6 shows the result of similar estimations ran on a sample which includes only private banks. Since the number of public banks in the sample is relatively small, it is not surprising to find that results are quite close to those obtained for all banks. Nevertheless, the degree of competition as inferred by the *H*-statistic is slightly higher for private banks between 1991 and 1997 as well as when the full 1991-2004 period is considered, in which case there is also less compelling evidence towards the rejection of the perfect competition hypothesis when only private banks are considered.

As to what concerns the comparison of the H-statistic between small and large banks, no robust difference was found. In fact, in addition to results not being robust to slight changes in banks' classification, equality of the estimates obtained for each group is not statistically rejected. Hence, during the period under analysis, there is no evidence that small banks have been able to exert higher market power due to a stronger presence in local markets where competition is less aggressive. The fact that this hypothesis, which has been widely stated and tested for other countries, does not apply to Portugal, should be linked to the smaller size of the national market as compared to those for which the relevant result has been obtained, since in larger countries it is common to find banks which have a strong position in the

[41] This sort of strategic behaviour may stem from the presence of switching or search costs.

region where they operate, despite having little weight in the national market as a whole. This fact mitigates the existence of fully distinct local markets in the Portuguese case.

6. Conclusions

The main conclusion to retain from this study is that on average, over the period from 1991 to 2004, Portuguese banks do not seem to have operated either under perfect competition or under perfect monopoly, but rather consistently with long-run monopolistic competition. During this period, both private and domestic banks seem to have competed more aggressively on average than the banking system as a whole, and perfect competition may not be rejected for these two types of banks.

An investigation of changes in competitive behaviour throughout the period suggests that competition was relatively weak between 1991 and 1996, even though results suggest domestic and especially private banks exhibited slightly higher competitive behaviour. An adjustment period followed between 1997 and 2000, in which behaviour consistent with long run equilibrium is rejected both for the banking system as a whole and for the group of private banks, whereas for domestic banks the hypothesis of behaviour consistent with perfectly competitive long run equilibrium is not rejected. In the more recent period, ranging from 2001 to 2004, strong competition was observed, and it is possible that domestic banks have competed more aggressively than expected in the framework of a static model with no distortions. Hence, the results suggest that the deregulation and liberalization process experienced by the Portuguese banking sector, including euro area participation, catalysed an increase in competition, particularly in what concerns the credit market.

One should, nonetheless, bear in mind the limitations of the non-structural approach employed, particularly regarding the hypotheses implicitly imposed on the underlying model for banks' behaviour. Therefore, obtained results should be compared with those derived using alternative methods in order to draw more general conclusions on the degree of competition in the Portuguese banking sector.

7. References

Bain, J. “Relation of profit rate to industry concentration: American manufacturing, 1936-1940”, *Quarterly Journal of Economics*, Vol. 65, p.p. 293-324.

Baumol, W. J. “Contestable markets: An uprising in the theory of industry structure”. *American Economic Review*, 72, 1-15, 1982.

Bikker, J. A. and Haaf, K. “Competition, concentration and their relationship: an empirical analysis of the banking industry”, *Journal of Banking and Finance*, 26, 2191-2214, 2002.

Bikker, J., Spierdijk, L. and Finnie, P. “Misspecifiation of the Panzar-Rosse Model: Assessing Competition in the Banking Industry," *DNB Working Papers* 114, Netherlands Central Bank, Research Department, 2006.

Casu, B. and Girardone, C. “Bank competition, concentration and efficiency in the single european market”. *The Manchester School Working Papers* 74 (4), 441-468, 2006.

Chamberlin, E. H. “Theory of monopolistic competition”. *Harvard University Press*, Cambridge, 1933.

De Bandt, O. and Davies, E. P. “Competition, contestability and market structure in European banking sectors on the eve of EMU”. *Journal of banking and finance*, Vol. 24, No 5, p.p. 1045-1066(22), June 2000.

Demsetz H., “Two Systems of Belief about Monopoly”. In Goldschmid H.J., Mann H.M. and J.F. Weston (eds), *Industrial Concentration: The New Learning*, 164-184, Boston: Little Brown, 1974.

Freixas, X. and Rochet, J.-C. “Microeconomics of Banking”, *MIT Press*, Cambridge, London, 1998.

Gilligan, T. and Smirlock, M. “An empirical study of joint production and scale economies in commercial banking”. *Journal of Banking and Finance*, 8, 67-77, March 1984.

Hempel, H. S. “Testing for competition among German banks”. *Discussion paper* 04/02, Economic research centre of the Deutsche Bundesbank, 2002.

Henderson, J. M. and Quandt, R. E. “Microeconomic theory: A mathematical approach”. *McGraw-Hill*, 1980.

Hughes, J. P. and Mester, L. J. “A quality and risk-adjusted cost function for banks: Evidence on the too-big-to-fail doctrine.” *Journal of Productivity Analysis*, 4, 293-315, September 1993.

Hughes, J. P., Mester, L. J. and Moon, C. “Are scale economies in banking illustrative or elusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production”. *Working Paper* 00-04, Federal Reserve Bank of Philadelphia, 2000.

Humphrey, D. B. “Why do estimates of bank scale economies differ?” *Economic Review*, Federal Reserve Bank of Richmond, 76(5), 38-50, September/October 1990.

Molyneux, P., Lloyd-Williams, D.M. and Thornton, J. “Competitive conditions in European banking”. *Journal of Banking and Finance*, Vol. 18(3), p.p. 445-459, May 1994.

Molyneux, P., Thornton, J. and Llyod-Williams, D. M. "Competition and market contestability in Japanese commercial banking". *Journal of Economics and Business*, Vol. 48(1), p.p. 33-45, February 1996.

Panzar, J. C. and Rosse, J. N. “Testing for ‘monopoly’ equilibrium.” *Journal of Industrial Economics*, 35, 443—456, 1987.

Ribeiro, N. “Debt growth: Factors, institutional issues and implications — The Portuguese case”. In Enoch, C. and Ötker-Robe, I. (eds.), “*Rapid credit growth in central and eastern Europe: Endless boom or early warning?*” International Monetary Fund, 2007.

Sealey, C.W., Jr. and Lindley, J.T. (1977), "Inputs, Outputs and the Theory of Production and Cost at Depository Financial Institutions", *Journal of Finance*, Vol. 32, pp. 1251-1266.

Shaffer, S. “A nonstructural test for competition in financial markets,” Federal Reserve Bank of Chicago. *In Proceedings of a conference on bank structure and competition*, pp. 225—243, 1982.

Shaffer, S. “A revenue restricted test of 100 large banks.” *Applied Financial Economics*, 24, 193-205, 1994.

Vesala, J. “Testing for Competition in Banking: Behavioral Evidence from Finland”, *Bank of Finland Studies*, E:1, 1995.

Koutsomanoli-Fillipaki, N. and Staikouras, C. “Competition and concentration in the new European banking landscape”. Presented in *Money Macro and Finance Research Group 36th Annual Conference*, 2004, Cass Business School.

Tables and Figures

Figure 1

Concentration in the Portuguese banking system

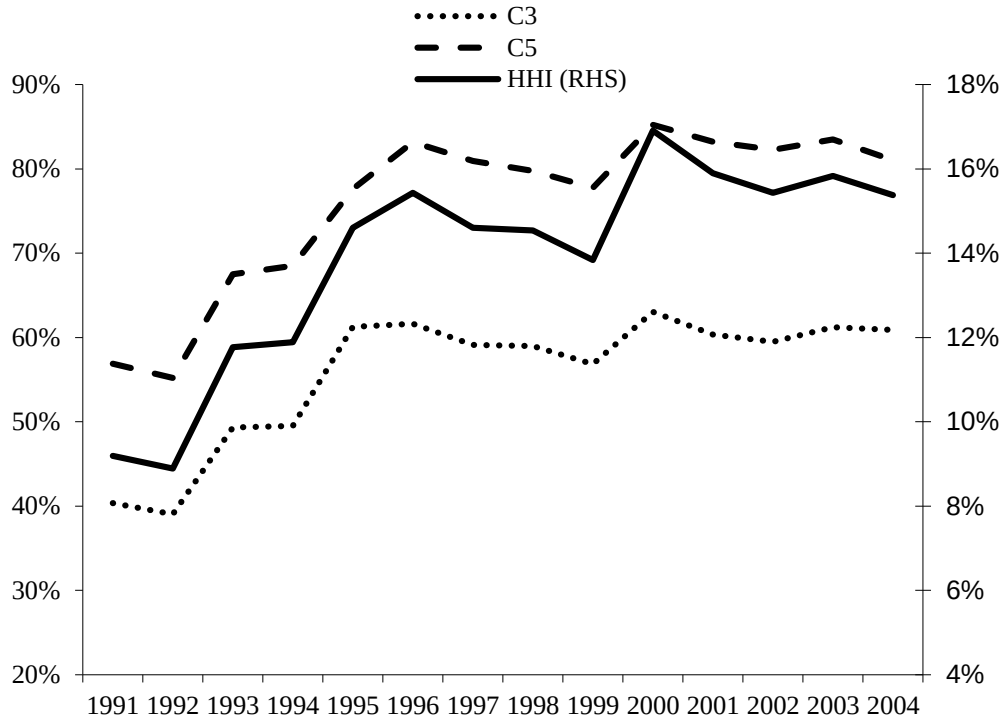


Table 1

Descriptive statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Interest Revenue	197	307.6	340	2.5	1445.6
Total Revenue	197	348.7	385	2.5	1702.3
Interest Revenue/Assets	197	4.2	2	0.4	10.6
Total Revenue/Assets	197	4.7	2	0.8	10.9
w_L (thousands of euros)	197	21.2	3.9	7.4	31.9
w_k	197	13.3	5.6	4	41.9
w_F	197	6.9	4.5	1.1	27.6
Demand Deposits/Total Deposits	197	35.6	9.5	18.7	65.2
Market Liabilities/Total Liabilities	197	44.1	14.6	13.4	88.7
Short-term Loans/Total Loans	197	46.7	20.8	9.3	93.8
Interbank Assets/Customer Loans	197	45.2	36.5	0.3	169.7
Off Balance Sheet Activity/Assets	197	190.1	2365.9	3.4	33228.2
Assets/Branches	197	27.3	10.9	4.4	58.7
NPL flow	197	1.5	1.3	0	7.6
Equity/Assets	197	6.9	2.7	1.5	29.8
Other Revenue/Total Revenue	197	14.6	12.8	1	105.6
ROA	197	0.49	0.57	-4.74	2.58
Assets	197	9295.6	11447.3	70.5	45172.8

Note: Money variables are in millions of 1991 Euros (unless otherwise stated) and ratios are in percentages.

Table 2

Baseline estimation results

	Model [1] ln(IR)		Equilibrium Test ROA	
$\ln w_L$	0.28	*0.16*	-0.69	*1.80*
$\ln w_K$	0.23	*0.08*	-0.49	*0.82*
$\ln w_F$	0.18	*0.11*	3.19	*1.91*
Demand Deposits/Total Deposits	-0.79	*0.71*	3.84	*6.51*
Market Liabilities/Total Liabilities	-0.63	*0.58*	-5.73	*6.10*
Short-term Loans/Total Loans	-1.03	*0.31*	6.62	*3.68*
Interbank Assets/Customer Loans	-0.46	*0.10*	-3.85	*1.20*
Off Balance Sheet Activity/Assets	0.00	*0.00*	-0.02	*0.01*
Assets/Branches[(a)]	0.02	*0.01*	0.15	*0.07*
NPL flow	1.58	*2.33*	-193.52	*82.80*
Equity/Assets	-8.10	*2.62*	39.28	*14.49*
Other Revenue/Total Revenue	-1.73	*0.42*	8.46	*6.06*
Aqrt(25)	-1.35	*0.21*	-0.67	*1.80*
Aqrt(50)	-1.11	*0.19*	-0.94	*1.52*
Aqrt(75)	-0.27	*0.11*	-0.54	*0.97*
M	0.02	*0.06*	0.12	*0.68*
F	0.49	*0.17*	4.84	*2.48*
P	-0.82	*0.16*	-3.42	*1.84*
δ	14.22	*0.72*	9.81	*8.19*
H-Statistic	0.69	*0.18*		
p($H \leq 0$)	0.00			
p($H=1$)	0.10			
Equilibrium test (p-value)			0.41	
R^2	0.63		0.00	
# Observations	197		197	
# Banks	25		25	

Notes: Heteroskedasticity robust standard errors are presented in italic. (a) The coefficient on this variable and the corresponding standard error are multiplied by 1000.

Table 3

Alternative dependent and scaling variables

	Scaling variable		
Dependent variable	ln *A*	*A* qrt.	None
Ln *(IR)*	0.70 0.00 0.04	0.69 0.00 0.10	0.61 0.00 0.06
Ln *(TR)*	0.71 0.00 0.06	0.71 0.00 0.12	0.60 0.00 0.07
Ln *(IR/A)*	————	————	0.71 0.00 0.06
Ln *(TR/A)*	————	————	0.71 0.00 0.07
# Obs. # Banks	197 25		

For each cell, the value in the centre is that of the H-statistic, whereas p-values for the test H≤0 (left) and H=1 (right) are presented below. ***, ** and * indicate evidence of disequilibrium at the 1%, 5% and 10% confidence level, respectively

Table 4

Competition through time for all banks

Scaling variable	1991-1996	1997-1999	2000-2004	1991-2004
ln *A*	0.41 0.00 0.00	0.27* 0.04 0.00	0.88 0.00 0.21	0.70 0.00 0.04
A qrt.	0.07 0.39 0.00	-.50* 0.87 0.00	0.97 0.00 0.90	0.69 0.00 0.10
None	-0.27 0.81 0.00	-.63* 0.90 0.00	0.93 0.00 0.81	0.61 0.00 0.06
Obs. Banks	197 25			

For each cell, the value in the centre is that of the H-statistic, whereas p-values for the tests H≤0 (left), H=1 (right) and $H_t=H_{t+1}$ (between periods) are presented below. ***, ** and * indicate evidence of disequilibrium at the 1%, 5% and 10% confidence level, respectively.

Table 5

Competition through time for domestic banks

Scaling variable	1991-1996			1997-1999			2000-2004			1991-2004	
ln *A*	0.35			0.59			0.84			0.60	
	0.00	0.00	0.09	0.00	0.01	0.20	0.00	0.36		0.00	0.04
A qrt.	0.17			1.19			1.37			0.98	
	0.27	0.00	0.01	0.00	0.64	0.64	0.00	0.05		0.00	0.92
None	0.12			0.85			1.38			0.95	
	0.36	0.01	0.06	0.02	0.70	0.19	0.00	0.07		0.00	0.84
Obs.	150										
Banks	21										

For each cell, the value in the centre is that of the H-statistic, whereas p-values for the tests $H \leq 0$ (left), $H=1$ (right) and $H_t=H_{t+1}$ (between periods) are presented below. ***, ** and * indicate evidence of disequilibrium at the 1%, 5% and 10% confidence level, respectively.

Table 6

Competition through time for privately owned banks

Scaling variable	1991-1996			1997-1999			2000-2004			1991-2004	
ln *A*	0.49			0.29**			0.88			0.75	
	0.00	0.00	0.29	0.05	0.00	0.00	0.00	0.19		0.00	0.05
A qrt.	0.39			-0.13*			0.80			0.74	
	0.14	0.09	0.26	0.60	0.03	0.08	0.00	0.48		0.00	0.30
None	0.01			-0.44*			0.88			0.72	
	0.49	0.01	0.38	0.78	0.01	0.03	0.00	0.65		0.00	0.18
Obs.	162										
Banks	18										

For each cell, the value in the centre is that of the H-statistic, whereas p-values for the tests $H \leq 0$ (left), $H=1$ (right) and $H_t=H_{t+1}$ (between periods) are presented below. ***, ** and * indicate evidence of disequilibrium at the 1%, 5% and 10% confidence level, respectively.

An assessment of Portuguese banks' efficiency and productivity towards euro area participation*

Miguel Boucinha, Nuno Ribeiro, Thomas Weyman-Jones

Abstract

This paper analyses the production technology of Portuguese banks during the 1992–2006 period through the estimation of a translog cost frontier. This period is of major interest because it covers Portugal's euro area accession and its impact on the banking system. Hence, critical factors impacting the banking system are identified against the background of increasing financial integration prior to the financial crisis that started in 2007 and later translated into strains in some European sovereign debt markets. Banks are modelled as firms which produce loans and other earning assets, choosing the cost minimizing combination of labour, capital and interest bearing debt, subject to holding a predetermined level of equity. According to the results of this study, technological progress has shifted the cost frontier downwards throughout the period under consideration, whereas the distance at which banks have operated from the frontier seems to have remained constant. Further, increases in production under scale economies have also contributed to the recorded increase in productivity.

JEL classification: G21, L13.

Keywords: Bank performance, Marginal costs, X-efficiency, Total factor productivity change.

* We are grateful to Diana Bonfim for helpful comments and suggestions provided since early stages of this work. Comments and suggestions provided by Mário Centeno, Filipa Lima, Pedro Pita Barros and seminar participants at the EWG-EPA 2010 Conference are also gratefully acknowledged. Any errors or omissions remain our own.

1. Introduction

Banks play a central role in the financial system and also in the real economy, as the recent financial crisis has vividly illustrated. Their smooth functioning allows for the intermediation of funds in the economy and provides for a wide range of financial services. In order to ensure this, banks need not only to adequately monitor their risks, but also to efficiently allocate their resources. Hence, the measurement of bank performance is a critical issue that has deserved considerable attention in the banking literature.

In this paper, we propose to analyse developments in the performance of the Portuguese banking system between 1992 and 2006, a period in which significant changes were observed, including the process of liberalization, consolidation and financial innovation.[42] These changes had a profound impact on the market's structure and on banks' technology and, through the analysis of a cost function, we assess how they affected banks' marginal costs and total productivity, which we decompose into the effect of returns to scale, cost efficiency change and technological progress.[43] In this way, we can not only quantify total factor productivity growth, but also identify if changes in productivity were driven by moving to a different point in the cost function, by moving closer to the cost frontier or by shifts in the frontier itself.

The analysis does not include the period of the financial crisis which emerged in 2007, since during this period banks started to experience marked increases in the prices of market funding and, more importantly, some quantitative constraints, which became progressively stronger until 2010. The presence of quantitative funding constraints is not consistent with the model defined below, since the cost function is derived solving a maximization problem in which the amount of funding is considered variable in the short-term. Furthermore, as the crisis intensified, some banks resorted in large scale to ECB lending to fund their activity. These loans are provided at very low cost, but they are supposed to be used only if necessary, and not to fund banks' normal activity. Hence, while some banks made an effort to increase funding from other sources at higher cost (mainly through customer deposits) others took advantage of the low cost funding from the ECB in a search for yield. Against this background, the banks which opted for funding at higher cost clearly did not behave as

[42] See Ribeiro (2007) and Antão et al. (2009) for a brief overview of the liberalization process.

[43] Cost inefficiency is commonly referred to in the literature as X-inefficiency.

profit maximizers in the strict sense of our model which thus proves inadequate to represent banks behaviour during this period. Lastly, as the sovereign crisis intensified, banks were faced with strong exogenous constraints stemming not only from stricter capital requirements, but also from the point of view of deleveraging their balance sheets, namely through the definition of a trajectory to lower the loan to deposit ratio in the medium term. These constraints implied that, during this period, many of the variables of interest in the model do not correspond to standard optimizing behaviour in the context of our model.

The modelling framework adopted allows for the identification, through the envelope theorem, of the estimated shadow return on equity capital prior to the global financial crisis and both before and after euro area accession. The importance of this estimated shadow return arises from the fact that equity or owner's capital is a major risk-absorbing capacity of the banking system. In other words, changes in the asset valuation of the banking firms impact directly on their equity capital, and the ensuing change in banks' leverage, in conjunction with an increased reliance on short term funding, has been identified as an extremely important indicator of the risk- absorbing capacity of the European banking systems.

Previous empirical results on the efficiency of Portuguese banks include the work by Mendes and Rebelo (1999), (2000), Pinho (2001), Canhoto and Dermine (2003), Lima (2008) and Lima and Pinho (2008). The majority of these studies estimated translog cost functions using Stochastic frontier Analysis (SFA), while in the one by Canhoto and Dermine (2003) a non-parametric frontier was estimated using Data Envelopment Analysis (DEA) and Mendes and Rebelo (2000) employ both methodologies.

Even though the above mentioned studies use different empirical and theoretical approaches to the modelling of banks' activity and cover different time periods (starting in 1987 and ending in 2004), all but the one by Mendes and Rebelo (1999) found that the productivity of Portuguese banks has increased. However, as expected given the differences in the approaches, they do not agree on the levels of X-inefficiency. The identification of shifts in best practices and changes in the distance at which banks operate from the efficient frontier also varies across studies according to the methodology employed. In fact, some studies do not allow for the distinction of the two effects, since the frontier is assumed to be constant over time, so that all productivity changes are attributed to changes in cost efficiency. Further, Pinho (2001) and Mendes and Rebelo (2000) found that state-owned

banks tend to perform worse on average whereas Canhoto and Dermine (2003) found that banks which were created after 1984 and foreign banks perform better than older banks which operated under the previously prevalent tightly regulated market conditions, including state-owned banks. Mendes and Rebelo (2000) and Lima (2008) also found that mergers contributed to increase banks' performance.

The remainder of the paper proceeds as follows. Section 2 presents the methodology and the data used in order to estimate banks' cost function and productivity. Section 3 presents the empirical results and is divided into 6 subsections, comprising the discussion of the estimates for Portuguese banks' marginal costs, the shadow cost of equity, scale efficiency, cost efficiency, technological progress and total factor productivity growth. Section 4 presents the concluding remarks.

2. Methodology and data

The modelling of banks' production has been the subject of considerable debate in the literature, essentially due to the controversy regarding the classification of customer deposits as inputs or as outputs. On the one hand, the production approach to bank modelling regards banks as firms producing services which are related to loans and deposit accounts, thus identifying as outputs the number of deposit accounts serviced and the number of loans originated and as inputs labour and physical capital. On the other hand, according to the inter- mediation approach (Sealey and Lindley (1977), banks' main activity is granting loans and investing in securities and other assets using funds obtained through deposits, purchased funds and equity.[44] There are sensible theoretical arguments supporting both approaches, and there is not a clear preference for either of them in empirical applications.[45] However, as remarked in Hughes et al. (2001) the inclusion of deposits both as inputs and as outputs would yield misleading results. In this case, the cost function would include both the level of deposits (since deposits are an output) and the price of deposits, whereas the definition of costs would include deposit related interest expenses (since deposits are an input). The

[44] See Freixas and Rochet (1998), pp. 77–79, on the production and intermediation approach.

[45] The fact that the production approach identifies as outputs the number of loans originated and deposit accounts constitutes an additional complication since this data is often unavailable. Studies which follow the production approach usually circumvent this issue by proxying the number of loans and deposit accounts by their value.

argument is that the optimal choice of one input – deposits – would not be influenced by the price of this input, since its quantity is held fixed.

There are two main reasons why holding deposits is an attractive activity for banks. On the one hand, as suggested by the production approach, deposits generate commission income and are a product which adds value in itself, as the general public does not have access to the same investment opportunities as banks. On the other hand, they are a relatively low cost and stable source of funding. Either way, a considerable part of banks' resources is dedicated to the origination and management of deposits. However, even though on a smaller scale, the acquisition and management of any input carries costs.

Hence, we follow Hughes and Mester (1993) in reasoning that deposits should be modelled as inputs (outputs) if the elasticity of total costs deducted of interest paid on deposits to the level of deposits is negative (positive). According to the result of this test (shown in Table 1), we choose to model deposits as inputs. This specification has the additional advantage of allowing for a more comprehensive definition of banks' costs, since otherwise the definition of costs would totally ignore funding costs, and so the measurement of efficiency would be limited to operational costs. Such an analysis could yield misleading results as some banks may be willing to bear higher operational costs (with employees and equipment) in order to optimize their funding structure, thus attaining lower funding costs. A similar argument motivates the inclusion of equity as a fixed input since, as remarked in Hughes et al. (2001), otherwise banks which find relatively more funding in equity and less in debt would spuriously appear to be more efficient. The fact that equity is treated as a quasi-fixed rather than a variable input is justified by regulatory and rating/reputation constraints to the choice of the optimal level of equity. Further, the costs associated to common equity issues lead banks to raise capital in relatively large tranches. As a consequence, current levels of capital need not only suffice to cover risks currently incurred, but should also accommodate future growth of assets. As such, banks may have a higher level of equity than that yielded by the individual static maximization problem.

Banks are thus assumed to minimize labour, funding and capital related costs ($w_L L + w_F F + w_K K$) subject to the production of a predetermined amount of loans ($\bar{y}_1$) and other earning assets ($\bar{y}_2$) and to the maintenance of a given level of equity (e_0):

$$C(y_1, y_2, w_L, w_F, w_k, e) = \min_{L,F,K}(w_L L + w_F F + w_K K)$$

$$s.t. \qquad (1)$$

$$F(x, e) \geq \bar{y}$$

$$e \geq e_0$$

where the variables are defined as:

$$C \equiv \sum_k w_k x_k (2)$$

$$y_1 : Total\ Loans$$
$$y_2 : Other\ Earning\ Assets$$
$$w_L : Price\ of\ Labour$$
$$w_F : Price\ of\ Funding \qquad (2)$$
$$w_K : Price\ of\ Capital$$
$$L : Labour$$
$$F : Funding$$
$$K : Capital$$
$$e : Equity$$

The price of funding is computed as the ratio between the flow of interest paid and the stock of interest bearing liabilities and the price of labour is defined as the ratio between labour costs and the number of employees, whereas the price of capital was proxied by the ratio between the sum of depreciation and general administrative costs (excluding labour) and the stock of tangible and intangible assets.

It is well known that banks, as is true with other firms, either due to agency problems or due to differences in managerial ability, do not strictly behave as profit maximizers, and some banks are closer to optimal behaviour than others. Furthermore, as usual in empirical applications, the performance of each bank is also affected by random factors, and the variables used in the estimations may be subject to measurement error. Hence, in order to analyse the cost efficiency of Portuguese banks since the early nineties, the cost function stemming from (1) is estimated using SFA models. The main equation to be estimated in the model may be expressed as:

$$lnC_{i,t} = lnC\left(t_t, y_{r,i,t}, w_{w,i,t}, e_{i,t}\right) + \left(v_{i,t} + u_{i,t}\right) \quad (3)$$

where $C\left(t_t, y_{r,i,t}, w_{w,i,t}, e_{i,t}\right)$ represents the estimated cost frontier and $C_{i,t}$ are banks' actual costs, so that a banks' observed costs are bounded below by the sum of the estimated cost frontier and a random error $\left(v_{i,t}\right)$ which is assumed to follow an i.i.d. $N(0, \sigma_v^2)$ distribution and accounts for measurement error of the level of costs, as well as for the effect of other random uncontrollable shocks. The sum of $lnC\left(t_t, y_{r,i,t}, w_{w,i,t}, e_{i,t}\right)$ and $v_{i,t}$ constitutes the stochastic frontier, and $u_{i,t}$ is a non-negative random variable which measures cost inefficiency as the difference between realized cost and the stochastic cost frontier. There are several models established in the literature which make different assumptions about the distribution of $u_{i,t}$. Battese and Coelli (1995) assume that the $u_{i,t}$ are independently distributed as truncations at zero of the $N\left(m_{i,t}, \sigma_u^2\right)$ distribution, where $m_{i,t} = z_{i,t}\delta$ and, in turn, $z_{i,t}$ is a vector of firm specific and time varying variables and δ is a vector of unknown coefficients to be estimated. This specification has the advantage of allowing for an interpretation of how some selected variables (those included in $z_{i,t}$) are related with estimated cost efficiency. The variables included in $z_{i,t}$ were the ratio of non-performing loans outstanding for less than one year to granted loans (NPL) as a measure of credit risk, banks' credit market share (Msc) as a measure of relative size, ROE and ROA as profitability measures, the equity to assets ratio (Cap. Ratio) and a measure of liquidity defined as the ratio of volatile assets to volatile liabilities (Liq. Ratio).

An alternative model, proposed in Battese and Coelli (1992) defines $u_{i,t}$ as follows:

$$u_{i,t} = u_i \exp\left(-\eta(t - T)\right) \quad (4)$$

where $u_{i,t}$ are assumed to be independently distributed as truncations at zero of the $N(\mu, \sigma_u^2)$ distribution and μ and η are parameters to be estimated. In this specification, inefficiency is firm specific and is allowed to vary through time even though, unlike in the model proposed in Battese and Coelli (1995), the ranking of firms remains constant through time. If η is not found to be statistically significant, it can be constrained to zero, so as to maximize the degrees of freedom by estimating no more parameters than needed. The cost efficiency of firm *i* at time *t* is:

$$CE_{i,t} = C\left(t_t, y_{r,i,t}, w_{w,i,t}, e_{i,t}\right)/C_{i,t} \in (0,1] \quad (5)$$

A fully efficient bank's actual cost is on the cost frontier, so that its efficiency is 100%, whereas an x% efficient bank's actual cost is above the frontier, so that it could theoretically produce the same output with only x% of its actual cost.

In order to provide a good approximation to the true cost function while preserving the available degrees of freedom and avoiding multicollinearity problems, the choice of the functional form in which the cost function is specified should obtain a balance between flexibility and parsimony. While the Cobb-Douglas specification is acknowledged to be too restrictive, the translog functional form provides a flexible local approximation and the Fourier functional form provides a global approximation. Berger and Mester (1997) found the difference between the two latter functional forms to be statistically but not economically relevant. Hence, and given the relatively small number of observations in our sample, the cost function is estimated using a translog functional form, which can be written as:

$$\begin{aligned} lnC_{i,t} = {} & \gamma_0 + \gamma_t t_t + \tfrac{1}{2}\gamma_{tt} t_t^2 + \textstyle\sum_r \gamma_{t,r} t_t lny_{r,i,t} + \sum_k \gamma_{t,k} t_t lnw_{k,i,t} + \sum_r \gamma_r lny_{r,i,t} + \\ & + \textstyle\sum_k \gamma_k lnw_{k,i,t} + \tfrac{1}{2}\left(\sum_r \sum_s \gamma_{r,s} lny_{r,i,t} lny_{s,i,t} + \sum_k \sum_l \gamma_{k,l} lnw_{k,i,t} lnw_{l,i,t}\right) + \\ & + \textstyle\sum_k \sum_r \gamma_{k,r} lnw_{k,i,t} lny_{r,i,t} + \gamma_e e_{i,t} + \tfrac{1}{2}\gamma_{ee} e_{i,t}^2 + \\ & + \textstyle\sum_k \gamma_{e,k} e_{i,t} lnw_{k,i,t} + \sum_r \gamma_{e,r} lne_{i,t} lny_{r,i,t} + v_{i,t} + u_{i,t} \end{aligned} \quad (6)$$

where the usual theoretical restrictions stemming from duality theory (i.e. symmetry and linear homogeneity in prices) are imposed:

$$\begin{aligned} & \gamma_{k,l} = \gamma_{l,k}, \forall\, k, l \\ & \sum_k \gamma_k = 1 \\ & \sum_k \gamma_{k,r} = 0, \forall\, r \\ & \sum_l \gamma_{k,l} = 0, \forall\, k, l \end{aligned} \quad (7)$$

$$\sum_k \gamma_{e,k} = 0$$

$$\sum_k \gamma_{t,k} = 0$$

In practice, symmetry is implicitly imposed in the specification of the estimated equation whereas homogeneity is obtained by normalizing input prices and total cost by w_k. The data are expressed as deviations from the overall sample mean, so that the first order coefficients correspond to the elasticities evaluated at the sample mean. The associated cost share equations implied by Shepherd's lemma:

$$S_{k,i,t} \equiv \frac{w_{k,i,t}x_{k,i,t}}{\sum_k w_{k,i,t}x_{k,i,t}} = \frac{\partial lnC_{i,t}}{\partial lnw_{k,i,t}} \quad (8)$$

were not imposed since they hold only under the assumption that no allocative inefficiency exists. Hence, our measure of X-inefficiency comprises both technical and allocative inefficiency.

The dataset used in this study was obtained from banks' financial statements reported to Banco de Portugal. The database comprises an unbalanced panel of yearly data for all banks operating in Portugal from 1991 to 2006.[46,47] Total loans were adjusted for securitization, essentially since the originating bank is generally still responsible for servicing securitized loans. Hence, if this correction were not to be undertaken, the cost efficiency of banks involved in securitization operations would be underestimated.[48] Non-per- forming loans are not included in the definition of output since they are essentially a non-income producing item in banks' balance sheet. Hence, this procedure accounts for different levels of credit risk in banks' loan portfolios and implicitly corrects for the differences in the level and in the

[46] As argued in the Introduction, the covered period includes the extremely important preparation for and pre-crisis membership of the euro area and it excludes data from 2007 due to the perturbations related to the financial crisis.

[47] Similarly to other papers, such as Fiorentino et al. (2006), Fries and Taci (2005) and Lang and Welzel (1996), the cost function was estimated using nominal data. Nonetheless, as a robustness check, the model was also estimated using deflated data and the results (not shown but available upon request) were broadly unchanged.

[48] Securitization in Portugal began in 1997 and grew rapidly in the following years, accounting for around 6 % of aggregate loans outstanding in 2004. Furthermore, considerable heterogeneity was present among banks, with a particular bank presenting a share of securitized loans as high as 34 % on a non-consolidated basis in 2004.

quality of banks' screening and monitoring activities of their borrowers' creditworthiness. Note, however, that this argument implicitly assumes that banks' loan portfolios are homogeneous. In practice, each bank may target different loan segments which have different levels of credit risk associated. Even though some effort was made in the selection of a sample of banks with a relatively similar activity in order to mitigate this problem, a full solution would involve defining a separate output for different categories of loans. In our case, this was not viable since it would imply a large increase in the number of parameters to be estimated, which could not be accommodated within the size of our sample. Data on individual rather than consolidated basis was used since it is available with greater detail for most of the sample period. Since banks which belong to the same financial group often have independent brands and activity structures, the analysis relies on bank level data. Since banks grew quite rapidly during the period under consideration, stock variables are defined as period averages so that meaningful values are between a flow and a stock variable.

All banks operating in Portugal are required to report financial statements to Banco de Portugal. However, in order to ensure that the analysis focuses on banks which operate with a similar technology, so that it is legitimate to include them in the same cost function, only universal banks with a retail branch network were included in the sample. Further, newly created banks were included only from their third year of operation, in order to avoid biases associated with short-term misalignments between setup costs and output.

After applying these filters, a sample of 25 banks, comprising a total of 280 bank-year pairs was obtained. For each year, the sample covers at least 77% of total loans, 80% of total assets and 87% of total deposits in the Por-]]tuguese banking system. Further, the market share of the five largest banks, when measured in terms of total assets in the sample, increased from around 57% in 1992 to over 80% in 2006, in similar fashion to what was observed in the whole banking system.

3. Empirical analysis

This section presents the main results of the analysis. All results are based on the estimation of Eq. 6, imposing the restrictions in 7, and are summarized in Table 2. As stated in Sect. 2, since the data are expressed as deviations from the overall sample mean, one can easily assess relevant elasticities evaluated at the mean by directly analysing single parameters. Hence, in order to clarify the exposition, the cross terms which do not have a

direct interpretation are not shown in Table 2.[49] A preliminary analysis of estimation results shows that the elasticity of cost with respect to each of the input prices is positive. Furthermore, the input price to which costs react most is the price of funding, which is not surprising considering that funding costs constitute the highest share of total costs. The sum of the parameters on the two outputs is close to one, indicating close to constant returns to scale at the sample mean. The fact that the parameter on the interaction term between the two outputs is negative indicates that there are scope economies in the joint production of loans and other earning assets. There is statistically significant cost reducing technological progress at the sample mean and banks with higher levels of equity tend to have lower costs with other inputs. The estimate for γ indicates the percentage of the total error's variance which is accounted for by cost inefficiency rather than by the classical random error and the fact that it is statistically significant in every specification provides evidence that the estimation of the cost function as a frontier is appropriate.

The first column of the table provides the results for models where estimated inefficiency is defined as a function of a series of bank characteristics. Banks with a higher ratio between the flow of non-performing loans and total loans granted show up to be more cost inefficient. This result suggests that loan delinquency works as a proxy for manager skill, i.e. managers who are poor at monitoring and screening loans are also poor at controlling costs. Banks with a larger market share in loans show up to be more cost efficient, which could indicate that larger banks are able to attract more competent managers. This hypothesis could in turn reflect larger salaries and perks than smaller banks able to pay to their top management as well as the prestige that comes from leading the largest banks. More capitalized banks show up to be less cost efficient, as do those with a higher liquidity ratio. This suggests that banks with less aggressive liquidity management also tend to be less cost efficient. One should nonetheless note that banks that choose to hold lower liquidity buffers will be subject to higher funding liquidity risk, which may be underpriced in wholesale markets in times of smooth market functioning, such as the one that prevailed during the sample period. The positive coefficient on the time trend indicates that, for the same value of loan delinquency, relative size and liquidity and capital ratios, banks' cost efficiency would have decreased through time.

[49] The full estimation results are available from the authors upon request.

In the third column of Table 2, the estimate for bank specific cost inefficiency is defined as a function of time, as shown in Eq. 4. However, since the estimate for η was not found to be statistically significant, the next column of the table presents estimation results in which it is restricted to zero in order to avoid the loss of degrees of freedom due to the estimation of redundant parameters. This is the specification used for the analysis whose results are presented in detail in the following subsections, the first of which presents results concerning estimated marginal costs for each bank as well as their behaviour over time. The next sub-section discusses the estimates for banks' shadow cost of equity capital. The third subsection discusses results concerning scale economies. The following subsections present results concerning cost efficiency and technological progress. In the last subsection changes in total factor productivity are quantified and decomposed in order to assess whether they were driven mainly by changes in the optimum technology, by the technology of each bank approaching the best practices, or simply by banks moving to a different point in the same cost function.

3.1 Marginal Costs

Using the estimated parameters for the cost function, marginal cost estimates for the production of each output may be obtained by:

$$mc_{r,i,t} \equiv \frac{\partial C_{i,t}}{\partial y_{r,i,t}} = \frac{C_{i,t}}{y_{r,i,t}} \frac{\partial ln C_{i,t}}{\partial ln y_{r,i,t}} \tag{9}$$

Note that Eq. 9 yields bank specific marginal cost estimates for both the production of bank loans and of other earning assets. Hence, the time-series presented in Table 3 were constructed by aggregating the individual estimates, using each bank's market share in loans as weights. Since funding costs constitute a major share of banks' variable costs and interest rates have decreased markedly during the period under analysis, the fact that the marginal cost estimates have decreased sharply over time is not surprising (Fig. 1). Nonetheless, an interesting question is whether real resource marginal costs also decreased through time. A proxy for banks' non-financial marginal cost is obtained by deducing the estimated marginal cost for each bank of the corresponding price of funding. As shown in columns 5 and 6 of Table 3, this measure also presents a decreasing trend, indicating that, despite contributing to the profile observed in marginal costs through time, the behaviour of interest rates alone is not enough to explain it. It should be mentioned that during the period

under analysis there was a change in the structure of banks' loan portfolio, with an increase in the share of loans to households as opposed to a decrease in the weight of loans to the public sector. This structural change should have contributed to an increase in the marginal cost of total loans. As such, the significant reduction in the estimated operational marginal cost of loans was not driven by changes in the composition of the loan portfolio.

As illustrated in Fig. 2 and documented in Table 3, the marginal cost of loans has generally been higher than that of other earning assets, indicating that it is more resource consuming to provide an additional loan than it is to invest in securities, which should be related with the screening and monitoring costs involved in granting loans. However, this difference has become less relevant through time. In order to understand this development, one should keep inmind that the output which is defined as other earning assets includes quite heterogeneous products. Furthermore, during the sample period there have been changes to the composition of this output. In fact, whereas during the early 1990's banks had significant resources invested in government bonds and deposits with the central bank, with the liberalization of the banking system and financial innovation banks started to invest in more sophisticated assets which, due to their greater complexity, require the use of more resources.

Furthermore, using data on banks' loan related interest income and stock of outstanding loans, one may compute an implicit interest rate on loans, as shown in column 7 of Table 3. Deducing the marginal cost from this interest rate, a measure of banks' price cost margin is obtained. According to the results shown in the last column of Table 3 and in Fig. 3, this measure has decreased through time, most notably during the liberalisation period and in the run up to euro area accession. This is consistent with the result found in Boucinha and Ribeiro (2009), according to which competition in the banking system has increased during the period under scrutiny.[50]

[50] The measure of implicit interest rate used is computed based on interest income and loan stocks which do not include non-performing loans. Hence, it is a proxy for the interest rate that banks charge their costumers, which should be higher than the average interest rate that they actually receive due to loan delinquency. Hence, the decrease in non-performing loans observed throughout the sample period should also have contributed to the observed decrease in banks' price-cost margin. Nonetheless, constructing a measure of interest rate which is a lower bound for the one that banks actually receive, since it includes non-performing loans but not the interest on these loans, the decreasing pattern found for the margin on loans is still present. Hence, this behaviour was not solely driven by the decrease in loan delinquency observed throughout the sample period.

3.2 Shadow cost of equity

Since the estimated cost function includes the level of equity as a fixed input, it allows for the computation of a measure of the shadow cost of equity capital as:

$$w_k^* \equiv \frac{\partial C_{i,t}}{\partial e_{i,t}} = \frac{C_{i,t}}{e_{i,t}} \frac{\partial lnC_{i,t}}{\partial lne_{i,t}} \tag{10}$$

The rationale underlying the computation of the shadow cost of equity is to provide a measure of how much banks are willing to pay for equity, since it indicates the amount that they would save in other costs as a result of an increase in the level of equity.[51]

As shown in Fig. 4 the time series obtained by aggregating the estimates for the shadow cost of equity is strongly correlated with market interest rates and with banks' weighted average cost of deposit and market debt funding. This result is consistent with shareholder capital being a source of funding in itself, so that funding costs are the ones which are most affected by the level of equity.

Even though it is in general higher that the short-term money market interest rate, the obtained measure of the shadow cost of equity, presented in Table 4, is lower than what is generally acknowledged to be a reasonable value for the actual price of equity. This result is not surprising and supports our choice of modelling equity capital as a quasi-fixed rather than a variable input, since it suggests that the regulatory and reputation constraints to the level of equity are in fact relevant, so that banks hold a higher level of equity capital than the one which would solve their static unconstrained optimization problem.

With the purpose of investigating what drives differences in banks' shadow cost of equity, this variable was regressed upon a set of bank specific variables, including each bank's capital ratio and return on equity and dummy variables which identify public banks, branches of credit institutions whose head office is in foreign countries and relatively large banks (the 60 % larger banks in each year).[52] In order to avoid simultaneity issues concerning the shadow cost of equity and banks' capital ratio and return on equity, the lag rather than the contemporary value of these variables is used. Since the dummy variable

[51] One must bear in mind the limitations of the model employed, by operating under the framework of a static optimization model estimated using non-consolidated accounting data.

[52] The fact that the dependent variable of this regression is itself an estimate means that the standard errors of this regression are not valid, since they do not account for the variance of the dependent variable.

which identifies branches of credit institutions whose head office is in foreign countries is time invariant, identification of the coefficient on this variable is not possible in a regression which includes bank specific fixed-effects. Hence, both fixed-effects and random-effects regressions are shown.

The results of these regressions, shown in Table 5, show a positive relationship between banks' capital ratio and the corresponding shadow cost of equity, but the coefficient has a high standard deviation. More profitable banks are found to have a higher shadow cost of equity, which may reflect higher risk incurred by the bank. Possibly reflecting lower perceived risk, state owned banks tend to have a lower shadow cost of equity. Branches of credit institutions whose head office is in foreign countries generally represent a relatively small portion of their banking group's assets, so that their activity hardly influences the group's credit rating and they often resort directly to the head office in order to obtain funding. Hence, it is not surprising to find that these banks tend to have a lower shadow cost of equity on average. Conversely, larger banks, which tend to be more transparent and whose equity is more likely to be traded in public markets, tend to have a higher shadow cost of equity.

3.3 Scale economies

The assessment of scale economies has been the subject of extensive discussion in the literature. Even though there are many theoretical arguments supporting their existence and they are typically invoked by bank managers as a motivation for mergers, empirical studies often fail to find them in the data. The identification of scale economies has relevant implications since it allows for inference on the adequacy of the market structure from a technological point of view.

This section assesses the presence of scale economies since the liberalization of the Portuguese banking system. A measure of scale economies is typically obtained as:

$$SE_{i,t} \equiv \sum_r \frac{\partial lnC_{i,t}}{\partial lny_{r,i,t}} \qquad (11)$$

An elasticity of cost with respect to total loans smaller (larger) than one is obtained in the presence of scale economies (diseconomies). As shown in Table 6, scale diseconomies (as defined above) were found during the early 1990's so that, all else equal, an increase in banks' size implied a more than proportional increase in costs. In the more recent period, the

estimate for the scale parameter is slightly below one, albeit not statistically different from one, indicating virtually constant returns to scale. One should, nonetheless, keep in mind that the elasticity computed according to Eq. 11 is a measure of short-run or constrained scale economies, since the level of equity is held fixed. Furthermore, since the definition of cost employed does not include the cost of equity, the measure of scale economies presented above is actually a measure of cash flow cost economies. This measure is likely to overestimate the true scale parameter, since the fact that the level of equity is held fixed implies that any increase in output must be totally financed by interest bearing debt, so that the cost of debt is forced to increase more than would be realistic.

A measure of scale economies which allows for the level of capital to change in response to changes in output could be obtained by estimating a cost function where equity is treated similarly to the other inputs. However, as men tioned above, we do not think that this would be an optimal solution, as there are important constraints to the choice of banks' level of equity capital. Furthermore, even in the more recent period, only a small number of Portuguese banks are listed in the stock exchange market, so that it is not straightforward to obtain estimates for the cost of equity.

Alternatively, as outlined in Hughes et al. (2001), cit- ing an original proposal by Hughes (1999), one can compute a measure of economic scale economies assuming that the observed level of equity capital minimizes economic cost at the shadow price of equity, since it then holds that:

$$C(t, y_r, w_k, w_e^*) = C(t, y_r, w_k, e) + w_e^* e \tag{12}$$

From the expression above, one can compute a measure of economic scale economies as:

$$ESE_{i,t} \equiv \sum_r \frac{\partial C(t, y_r, w_k, w_e^*)}{\partial y_r} \frac{y_r}{C(t, y_r, w_k, w_e^*)} \tag{13}$$

Since the level of equity capital e minimizes economic cost, the constrained marginal cost equals the long run marginal cost:

$$\frac{\partial C(t, y_r, w_k, w_e^*)}{\partial y_r} = \frac{\partial C(t, y_r, w_k, e)}{\partial y_r} \tag{14}$$

From this result and the definition of the shadow cost of equity in equation 10, expression 13 may be written as:

$$ESE_{i,t} \equiv \sum_r \frac{\partial C(t, y_r, w_k, e)}{\partial y_r} \frac{y_r}{C(t, y_r, w_k, e) - \frac{\partial C}{\partial e} e} \tag{15}$$

or

$$ESE_{i,t} \equiv \frac{\sum_r \frac{\partial lnC_{i,t}}{\partial lny_{r,i,t}}}{1 - \frac{\partial lnC_{i,t}}{\partial lne_{i,t}}} \tag{16}$$

An aggregate time-series of the estimates for scale economies obtained through the aggregation of the individual estimates yielded by the computation of Eq. 16 is presented in Table 6. While this measure presents the same decreasing profile as the constrained measure, its level is considerably lower at each year. Hence, accounting for the fact that banks' level of capital is allowed to vary according to changes in banks' output, one finds statistically significant scale economies for the full period under scrutiny, which suggests that the concentration process observed in the Portuguese banking system was at least partly driven by the opportunity to increase productive efficiency.

Estimated scale economies show up to be stronger at the end of the sample when compared to the early 1990's. This result is likely to be linked with the changes to banks' technology brought about by technological progress. In fact, the increasing automation of services should have allowed for a decrease in banks' variable costs at the expense of a more significant initial investment in technology, such as storage and processing of information and communication facilities. These technological developments in turn allowed for the setup of a dense ATM net- work and of other remote-delivery outlets such as websites, with the corresponding savings in costs associated with the need for less employees and branches. Another factor possibly contributing to the higher scale economies found in the more recent period was the increasing internationalization of banking activity brought about by technological progress, financial innovation and increasing economic integration among EU members. In fact, the expansion of the relevant market for banks' activity beyond national

borders brought about new growth opportunities while, to some extent, exposed them to increased competition from non-resident banks. Moreover, even the largest banks in the Portuguese financial system are relatively small when compared to their international counterparts.

3.4 Cost efficiency

Table 7 presents the obtained estimates for the cost efficiency of Portuguese banks between 1992 and 2006. As indicated above, results are based on a specification which does not include determinants of inefficiency. Further, η — the parameter for the change in cost efficiency through time — was not found to be statistically significant, and so was constrained to zero. Hence, the distance at which each bank stands from the cost frontier representing best practices does not seem to have changed during the period under scrutiny.[53] The aggregate estimate for inefficiency lies close to 84%, suggesting that Portuguese banks could theoretically have produced the same output while incurring only 94 % of their actual costs. Some heterogeneity across banks was found, with estimated efficiency scores ranging from a minimum of 78% to a maximum of 96%.

3.5 Technological progress

The estimated cost function also includes a time trend as a translog term, which allows for the computation of both Hicksian neutral and non neutral technological progress. Total cost reducing technological progress, i.e., shifts to the frontier brought about by the adoption of more efficient production techniques, is obtained by $-\frac{\partial lnC}{\partial lnt}$.[54] As shown in Table 8, technological progress was identified for the full period under consideration. These developments should be regarded in the context of global financial inte- gration, which catalysed the swift adoption of more effi- cient technology. The estimate for technological progress found for 2006 should be interpreted as indicating that, in this year, Portuguese

[53] The fact that the aggregate value of the cost efficiency estimate is not constant even though each bank's efficiency estimate is time invariant is motivated by a composition effect. In fact, due to changes in banks' market shares, the weights used in aggregation (the value of granted loans) are not constant and, due to mergers/acquisitions and to the emergence of new banks, estimation relies on an unbalanced panel of data.

[54] Note that this measure underestimates technological progress when the quality/variety of products increases through time.

banks operating according to the industry's best practices could produce the same output as in the previous year incurring 2.2% lower total costs.

3.6 Total factor productivity growth

In this section the parameters of the estimated cost function are used to compute a measure of total factor productivity change (*TFPC*) which may be decomposed into the effect of cost efficiency change (*EC*), of technological progress (*TC*) and of returns to scale (*RTS*) (see Bauer (1990) for details):

$$TFPC = EC + TC + RTS \tag{17}$$

or

$$ln\left(\frac{TFP_{i,t}}{TFP_{i,t-1}}\right) = ln\left(\frac{CE_{i,t}}{CE_{i,t-1}}\right) + \frac{1}{2}\left(-\frac{\partial lnC_{i,t}}{\partial t} - \frac{\partial lnC_{i,t-1}}{\partial t}\right) + \frac{1}{2}\sum_{r}\left[\left(\varepsilon_{r,i,t}\frac{1-ESE_{i,t}}{ESE_{i.t}} + \varepsilon_{r,i,t-1}\frac{1-ESE_{i,t-1}}{ESE_{i.t-1}}\right) ln\left(\frac{y_{r,i,t}}{y_{r,i,t-1}}\right)\right] \tag{18}$$

where e_r is the elasticity of cost with respect to output r and each term of the decomposition has an interesting inter- pretation. In fact, according to the expression above, total factor productivity change comprises catching-up to the cost frontier (cost efficiency change), shifts in the frontier itself over time (technical progress) and shifts along the frontier (returns to scale component). The effect of returns to scale represents the pure impact on total costs stemming from changes in output after allowing for input requirements and it is positive if a bank with increasing (decreasing) returns to scale increases (decreases) its production.

It should be taken into account that Eq. 18 is presented as proposed in Bauer (1990), with the necessary changes to account for the inclusion of equity in the estimated cost function. As such, the concept of economic scale econo- mies (ESE) is used instead of the classical measure of scale economies presented in Eq. 11.

The results for total factor productivity change in the Portuguese banking sector during the period under consideration are summarized in Table 9 and in Fig. 5. The most striking result is that total factor productivity change has been mainly driven by technological progress, which became stronger throughout the sample period. Scale efficiency change also made a positive contribution towards total factor productivity growth, especially in the years when the largest mergers occurred. Cost efficiency remained virtually constant throughout the period. Combining the three effects, one finds that there was consistent total factor productivity growth observed during the period under consideration. The value of total factor productivity growth was close to 3% during the 1990's, it was boosted by gains in returns to scale during the large mergers observed around the turn of the century, and later returned to a value close to that observed earlier.

4. Concluding remarks

This paper analyses the production technology of Portuguese banks during the 1992–2006 period through the estimation of a translog cost frontier. Banks are modelled as firms which produce loans and other earning assets, choosing the cost minimizing combination of labour, cap- ital and interest bearing debt, subject to holding a given level of equity.

Several different specifications were tested for the distribution of estimated inefficiency. Banks with higher credit risk and with more idle liquidity were found to be more cost inefficient, possibly reflecting the fact that these variables are in a way proxies for manager quality/ sophistication. Banks with higher capital ratios were also identified as being less cost-efficient. On the other hand, relatively larger banks were found to be more cost efficient, which could indicate, that larger banks are able to attract more competent managers. The more detailed analysis whose results are briefly summarised below was carried out based on the estimation results of a simpler model where no determinants of inefficiency were included.

Portuguese banks' marginal costs in the production of loans and other earning assets were found to follow to a large extent the decline in nominal interest rates observed throughout the period under consideration. Still, a significant part of the decrease in total marginal costs is explained by a reduction in the real resource marginal cost. The estimate

for this measure corresponding to lending activity was in general higher than that for other earning assets, even though the difference became less relevant in the more recent period, in which the value of the real resource marginal costs ranged between 1 and 1.3 %.

Banks' capital structure was accounted for in the analysis by including equity as a quasi-fixed input in the cost function. This procedure allowed for the computation of estimates for banks' shadow cost of equity, which should be interpreted as a lower bound to banks' true willingness to pay for equity capital. Hence, it is not surprising to find that they are lower than levels compatible with usually accepted equity risk premia. Furthermore, the estimated shadow cost of equity follows quite closely the developments in market interest rates.

On average, Portuguese banks were found to operate with a cost inefficiency level around 17%, indicating that they could theoretically produce the same output incurring only 83% of their actual cost. The magnitude of cost reducing technological progress was found to be slightly above 2% throughout the period. Accounting for banks' capital structure, significant scale economies were found, especially in the more recent years. Further, the results point to the existence of economies of scope in the joint production of loans and other earning assets.

Against the background of the liberalization and privatization of the banking system and of increasing financial innovation, the cost frontier representing best practices has shifted downwards over time. The distance between banks' actual costs and the cost frontier, on the other hand, has not changed significantly. Since banks with increasing returns to scale increased their production, there was a move along the cost function which also contributed to an increase in productivity.

Combining these results, estimates for total factor productivity change were computed, amounting to 3.2 % each year on average, which results in a total productivity increase close to 58 % between 1992 and 2006.

5. References

Antão P, Boucinha M, Farinha L, Lacerda A, Leal AC, Ribeiro N (2009), Financial integration, financial structures and the decisions of households and firms, *in:* The Portuguese economy in the context of economic, financial and monetary integration. Economics and Research Department, *Banco de Portugal*, pp 415–545.

Battese GE, Coelli TJ (1992), Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis* 3:153–159.

Battese GE, Coelli TJ (1995), A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20:325–332.

Bauer PW (1990), Decomposing TFP growth in the presence of cost inefficiency, nonconstant returns to scale, and technological progress. *Journal of Productivity Analysis* 1:287–299

Berger AN, Mester L (1997), Inside the black box: what explains differences in the efficiencies of financial institutions. *Journal of Banking and Finance* 21:895–947

Boucinha M, Ribeiro N (2009) An assessment of competition in the Portuguese banking system in the 1991–2004 period. *Working paper* 1/2009, Banco de Portugal

Canhoto A, Dermine J (2003), A note on banking efficiency in Portugal, new vs. old banks. *Journal of Banking and Finance* 27:2087–2098.

Fiorentino E, Karmann A, Koetter M (2006), The cost efficiency of German banks: a comparison of SFA and DEA. *Discussion paper series 2: Banking and Financial Studies* 10/2006, Deutsche Bundesbank

Freixas X, Rochet JC (1998), Microeconomics of banking. *MIT Press*, Cambridge

Fries S, Taci A (2005), Cost efficiency of banks in transition: evidence from 289 banks in 15 post-communist countries. *Journal of Banking and Finance* 25:55–81

Hughes JP (1999), Incorporating risk into the analysis of production. *Atlantic Econonic Journal* 27(1):1–23.

Hughes JP, Mester LJ (1993). Accounting for the demand for financial capital and risk-taking in bank cost functions. *Working paper* 93–17, Federal Reserve Bank of Philadelphia.

Hughes JP, Mester LJ, Moon CG (2001), Are scale economies in banking elusive or illusive? evidence obtained by incorporating capital structure and risk-taking into models of bank production. *Journal of Banking and Finance* 25(12):2169–2208.

Lang G, Welzel P (1996), Efficiency and technical progress in banking: empirical results for a panel of German cooperative banks. *Journal of Banking and Finance* 20:1003–1023.

Lima F (2008), Product differentiation and the measurement of cost efficiency in banking: the case of Portuguese banks, *in*: Soares J, Pina J, Catalão-Lopes M (eds.) New developments in financial modelling. *Cambridge Scholars Publishing*, Cambridge, pp 34–76.

Lima F, Pinho PS (2008), Financial disintermediation and the measurement of efficiency in banking: the case of Portuguese banks. *International Journal of Banking, Accounting and Finance* 1(2):133–148.

Mendes V, Rebelo J (1999), Productive efficiency, technological change and productivity in Portuguese banking. *Applied Financial Economics* 9:513–521.

Mendes V, Rebelo J (2000), The effect of bank M&As on efficiency: the Portuguese experience, *in:* Abreu M, Mendes V (eds) What financial system for year 2000? *Principia Academic Press*, Lisbon, pp 287–293

Pinho PS (2001), Using accounting data to measure efficiency in banking: an application to Portugal. *Applied Financial Economics* 11:527–538.

Ribeiro N (2007), Debt growth: factors, institutional issues and implications — the Portuguese case, *in:* Enoch C, Ötker Robe I (eds) Rapid credit growth in central and eastern Europe: endless boom or early warning? *Palgrave MacMillan and IMF*, Basingstoke, pp 296–311

Sealey C, Lindley J (1977), Inputs, outputs and a theory of production and cost at depository financial institutions, *Journal of Finance* 32(4): 1251–1266.

Tables and Figures

Table 1

Elasticity of banks' costs with respect to the level of deposits

Year	Estimate at weighted sample mean	p-value (H0: elasticity of deposits=0)
1992	-0.7583	*0.00*
1993	-0.6612	*0.00*
1994	-0.5908	*0.00*
1995	-0.5932	*0.00*
1996	-0.4905	*0.00*
1997	-0.4592	*0.00*
1998	-0.3727	*0.00*
1999	-0.2988	*0.00*
2000	-0.3906	*0.00*
2001	-0.3993	*0.00*
2002	-0.2788	*0.00*
2003	-0.1780	*0.00*
2004	-0.0905	*0.14*
2005	-0.1318	*0.03*
2006	-0.1780	*0.01*
Overall	-0.4211	*0.00*

Notes: This test is performed by estimating a cost function which excludes interest paid on deposits from the definition of total cost (the dependent variable) and from the computation of the price of funding and includes the level of deposits as a fixed input. If the elasticity of costs with other inputs with respect to the level of deposits is negative (positive), then deposits behave as inputs (outputs) and should be modelled as such. Total loans adjusted for securitization are used as weights in the computation of means.

Table 2

Cost frontier estimation results

	(1)	(2)	(3)
$\ln(W_L)$	0.1293	0.1381	0.1385
	0.02	*0.02*	*0.02*
$\ln(W_F)$	0.5966	0.6034	0.6028
	0.02	*0.02*	*0.02*
$\ln(y_1)$	0.6314	0.5768	0.5764
	0.01	*0.02*	*0.01*
$\ln(y_2)$	0.4562	0.4716	0.4718
	0.01	*0.01*	*0.01*
$\ln(y_1)*\ln(y_2)$	-0.2992	-0.2458	-0.2454
	0.02	*0.02*	*0.02*
t	-0.0300	-0.0188	-0.0186
	0.00	*0.00*	*0.00*
$\ln(e)$	-0.0694	-0.0638	-0.0635
	0.02	*0.02*	*0.02*
NPL	3.2562		
	1.32		
Msc	-2.1762		
	1.01		
Cap Ratio	1.5870		
	0.65		
Liq. Ratio	*0.0170*		
	0.01		
t	*0.0477*		
	0.02		
μ		0.1650	0.1625
		0.09	*0.08*
η		-0.0017	
		0.02	
γ	0.9680	0.6130	0.6077
	0.03	*0.11*	*0.09*

Notes: Standard errors are reported in italics. In the cost function specification, the constant and most cross terms were omitted as they have no direct interpretation. The complete estimation results are available from the authors upon request.

Table 3

Marginal cost estimates at the weighted sample mean (per cent)

Year	Marginal cost of loans (1)	Marginal cost of other earning assets (2)	Short-term money market interest rate (3)	Implicit price of funding (4)	Real resource marginal cost of loans (5)	Real resource marginal cost of other earning assets (6)	Implicit interest rate on loans (7)	Margin on loans (8)
1992	15.08	13.13	16.72	10.74	4.34	2.39	17.15	2.07
1993	12.79	11.13	13.17	8.94	3.85	2.19	15.45	2.66
1994	10.46	9.00	11.23	7.01	3.45	1.99	13.08	2.62
1995	10.21	9.03	9.79	7.09	3.11	1.94	12.26	2.05
1996	8.47	7.71	7.27	5.71	2.76	2.00	10.72	2.26
1997	7.16	6.65	5.61	4.65	2.51	2.00	9.24	2.08
1998	5.81	5.43	4.23	3.59	2.22	1.84	7.44	1.63
1999	4.72	4.30	2.96	2.65	2.07	1.66	5.78	1.06
2000	5.15	4.60	4.39	3.28	1.87	1.32	6.13	0.98
2001	5.07	4.64	4.26	3.47	1.60	1.17	6.33	1.26
2002	4.26	4.09	3.32	2.76	1.49	1.32	4.87	0.61
2003	3.54	3.50	2.33	2.28	1.26	1.22	4.56	1.02
2004	3.28	3.25	2.11	2.00	1.27	1.24	4.21	0.93
2005	3.42	3.36	2.18	2.07	1.35	1.30	4.05	0.63
2006	3.72	3.85	3.08	2.61	1.12	1.25	4.76	1.04

Note: Total loans adjusted for securitization are used as weights in the computation of means.

Table 4

The shadow cost of equity (per cent)

Year	Shadow cost of equity (full sample)	Short-term money market interest rate	Long-term government bond interest rate	Equity/Assets ratio	Implicit price of funding
	(1)	(3)	(4)	(5)	(6)
1992	20.28	16.72	11.86	7.56	10.74
1993	15.60	13.17	10.33	7.41	8.94
1994	11.33	11.23	10.48	6.97	7.01
1995	9.84	9.79	11.47	6.54	7.09
1996	7.42	7.27	8.56	6.29	5.71
1997	6.22	5.61	6.36	6.10	4.65
1998	5.88	4.23	4.88	6.41	3.59
1999	4.89	2.96	4.78	6.50	2.65
2000	6.19	4.39	5.60	6.04	3.28
2001	6.98	4.26	5.16	5.85	3.47
2002	5.55	3.32	5.01	6.13	2.76
2003	3.75	2.33	4.18	6.36	2.28
2004	3.34	2.11	4.14	6.33	2.00
2005	2.88	2.18	3.44	5.69	2.07
2006	1.86	3.08	3.92	5.52	2.61

Note: Total loans adjusted for securitization are used as weights in the computation of means.

Table 5

Determinants of the shadow cost of equity

	Random-Effects model	Fixed-Effects model
	(1)	(2)
$(Equity/Assets)_{t-1}$	0.14	0.08
	0.12	*0.12*
ROE_{t-1}	0.06	0.05
	0.03	*0.03*
State-owned bank (dummy)	-0.03	-0.03
	0.01	*0.02*
Branch of credit institution whose head office is in foreign countries (dummy)	-0.08	
	0.03	
60% largest banks (dummy)	0.02	0.03
	0.01	*0.01*

Notes: Standard errors are reported in italics. Time dummies and a constant were included in the regressions.

Table 6

Scale economies

Year	*Scale Economies (SE)*	p-value (*H0* : *SE* =1)	*Economic Scale Economies (ESE)*	p-value (*H0* : *ESE* =1)
1992	1.0658	*0.05*	0.9478	*0.00*
1993	1.0481	*0.10*	0.9431	*0.00*
1994	1.0227	*0.40*	0.9361	*0.00*
1995	1.0056	*0.83*	0.9353	*0.00*
1996	0.9946	*0.82*	0.9341	*0.00*
1997	0.9896	*0.65*	0.9347	*0.00*
1998	0.9872	*0.60*	0.9220	*0.00*
1999	0.9897	*0.73*	0.9217	*0.00*
2000	0.9806	*0.52*	0.9103	*0.00*
2001	0.9791	*0.55*	0.9057	*0.00*
2002	0.9798	*0.56*	0.9065	*0.00*
2003	0.9640	*0.29*	0.9029	*0.00*
2004	0.9476	*0.16*	0.8900	*0.00*
2005	0.9415	*0.11*	0.8975	*0.00*
2006	0.9364	*0.12*	0.9049	*0.00*
1992-2004	0.9712	*0.33*	0.9099	*0.00*

Notes: Total loans adjusted for securitization are used as weights in the computation of means. SE denotes scale economies as defined in Equation (11) and ESE refers to economic scale economies as defined in Equation (16).

Table 7

Cost efficiency (per cent)

Year	X-Efficiency
1992	84.30
1993	84.50
1994	84.46
1995	84.37
1996	84.28
1997	84.21
1998	83.50
1999	83.46
2000	83.08
2001	82.88
2002	82.94
2003	82.99
2004	83.18
2005	83.10
2006	83.17
1992-2006	83.31

Note: Total loans adjusted for securitization are used as weights in the computation of means.

Table 8

Technological progress (per cent)

Year	Technological progress (*eti*) (1)	p-value (*H0*: *eti* =0) (2)
1992	*-2.45*	*0.02*
1993	*-2.38*	*0.01*
1994	*-2.21*	*0.01*
1995	*-2.09*	*0.01*
1996	*-2.05*	*0.00*
1997	*-2.21*	*0.00*
1998	*-2.04*	*0.00*
1999	*-2.34*	*0.00*
2000	*-2.61*	*0.00*
2001	*-2.46*	*0.00*
2002	*-2.38*	*0.00*
2003	*-2.24*	*0.00*
2004	*-1.84*	*0.01*
2005	*-2.22*	*0.00*
2006	*-2.16*	*0.00*

Note: Total loans adjusted for securitization are used as weights in the computation of means.

Table 9

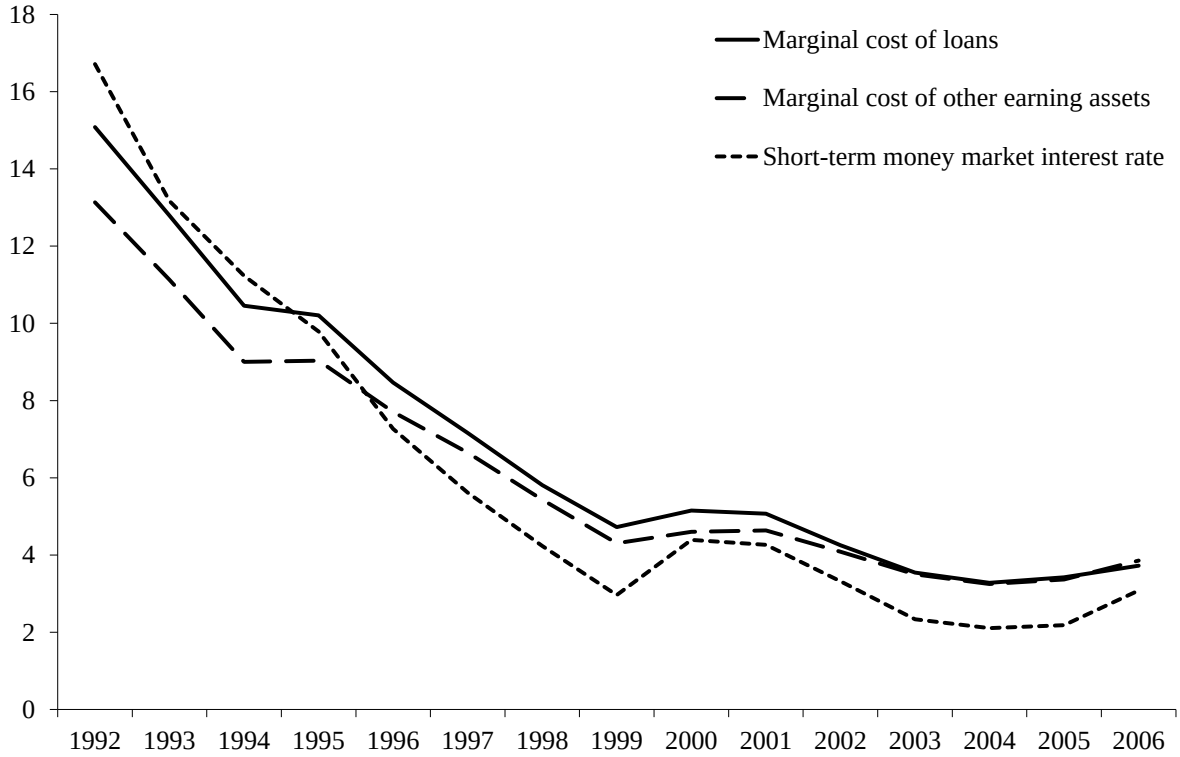
Total factor productivity growth (per cent

Year	*Scale Efficiency Change* (1)	*Technical Efficiency* (2)	*Technological Change* (3)	*Total Factor Productivity Change* (4)
1993	0.67	0.24	-2.41	3.32
1994	0.64	-0.05	-2.30	2.89
1995	0.54	-0.10	-2.15	2.59
1996	0.35	-0.11	-2.07	2.31
1997	0.51	-0.08	-2.13	2.55
1998	1.74	-0.85	-2.12	3.02
1999	1.22	-0.05	-2.19	3.36
2000	4.42	-0.46	-2.48	6.43
2001	2.22	-0.24	-2.53	4.51
2002	0.31	0.07	-2.42	2.80
2003	0.65	0.07	-2.31	3.02
2004	1.74	0.23	-2.04	4.01
2005	0.97	-0.10	-2.03	2.90
2006	0.49	0.08	-2.19	2.77

Note: Total loans adjusted for securitization are used as weights in the computation of means.

Figure 1

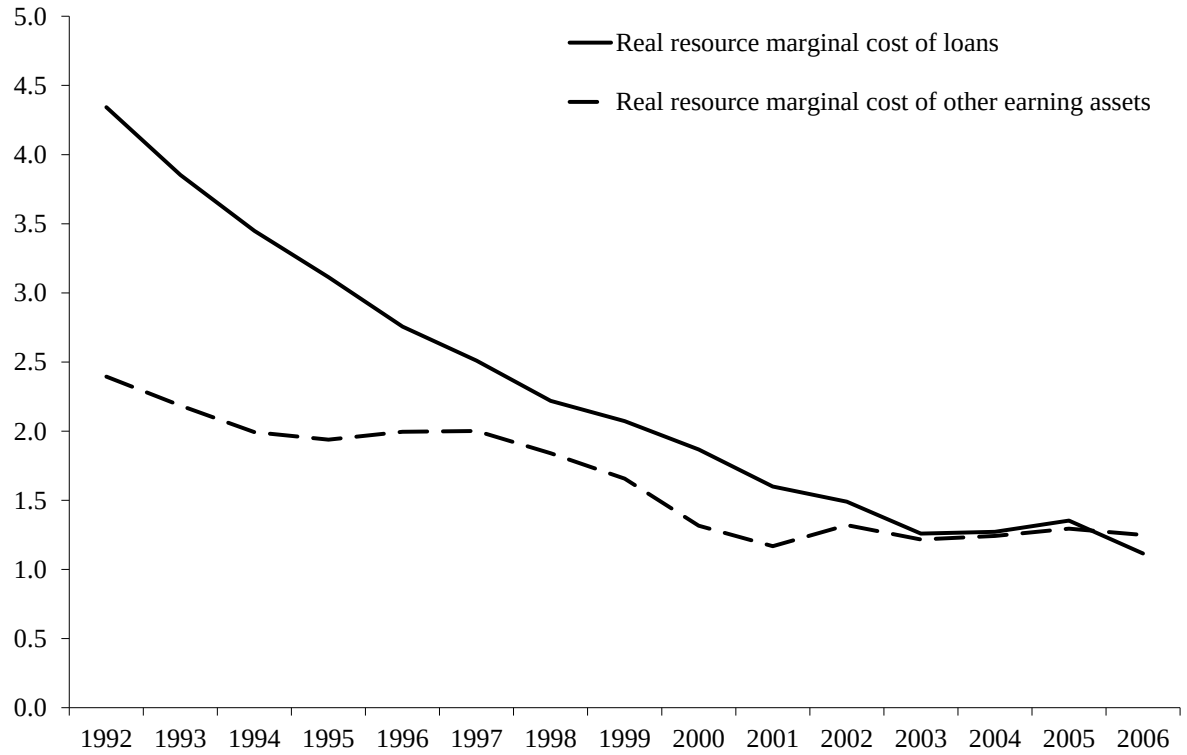
Marginal costs (per cent)



Note: Total loans adjusted for securitization are used as weights in the computation of means.

Figure 2

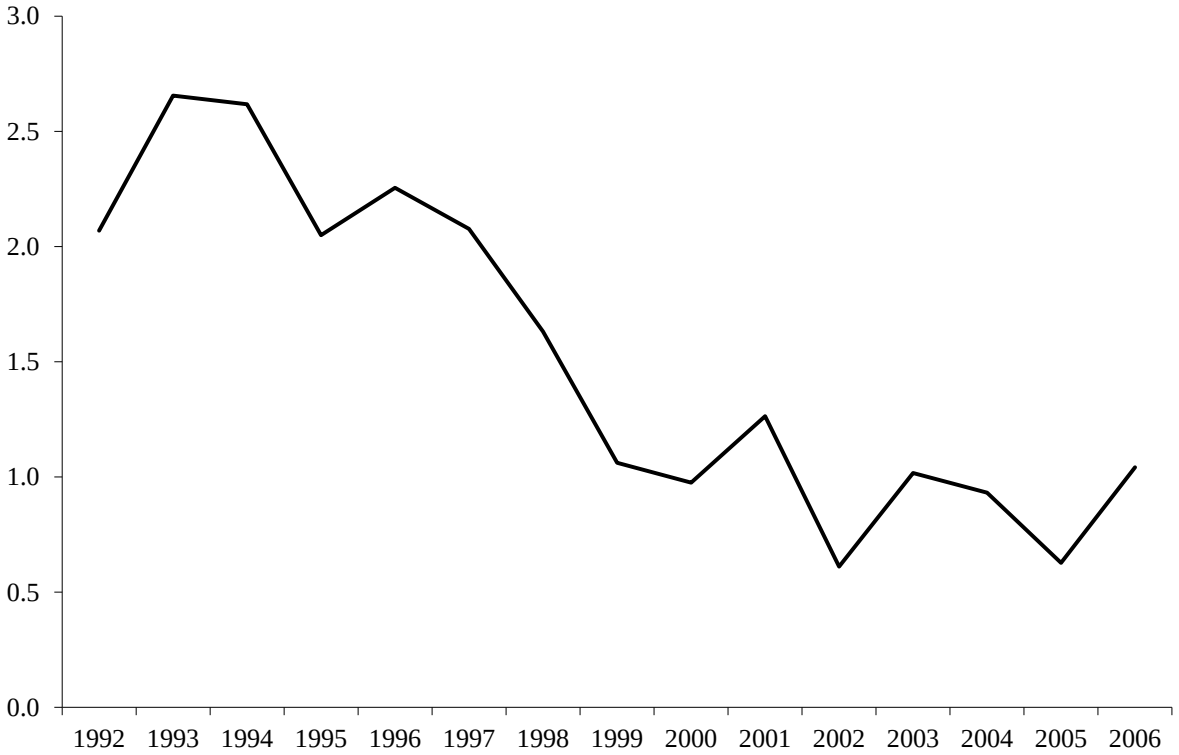
Real resource marginal costs (per cent)



Note: Total loans adjusted for securitization are used as weights in the computation of means.

Figure 3

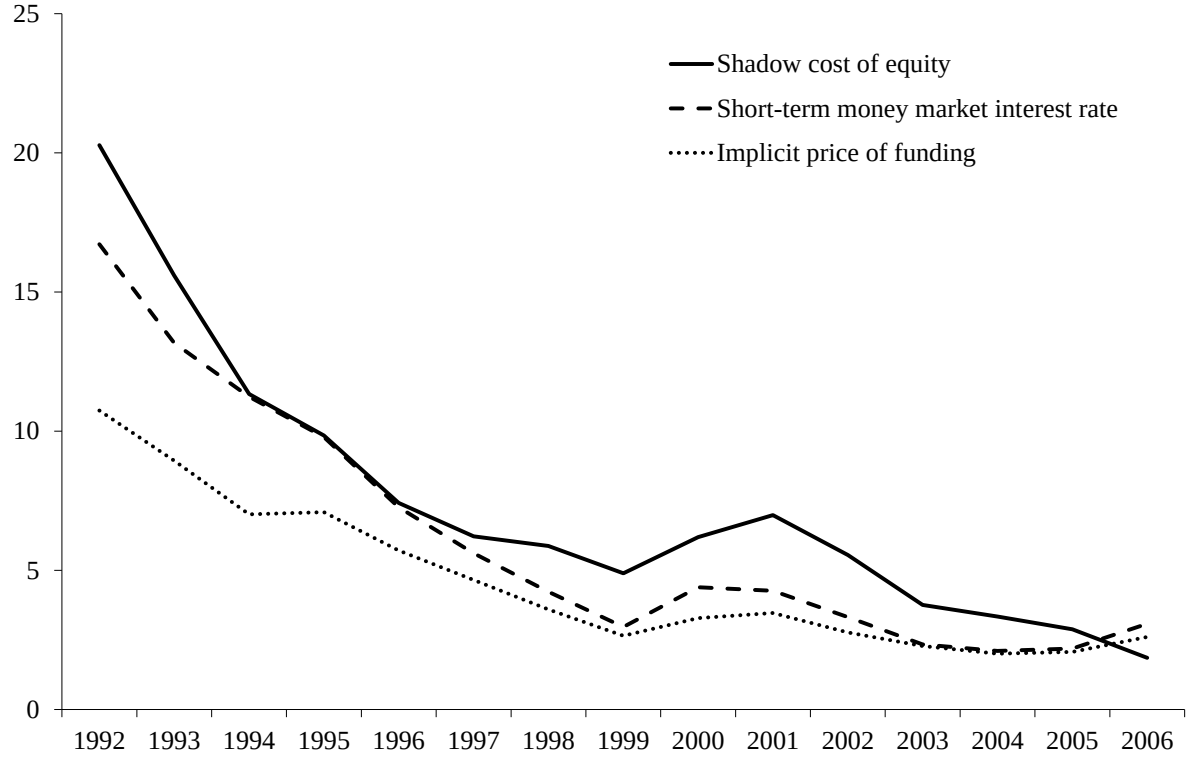
Margin on loans (per cent)



Note: Total loans adjusted for securitization are used as weights in the computation of means.

Figure 4

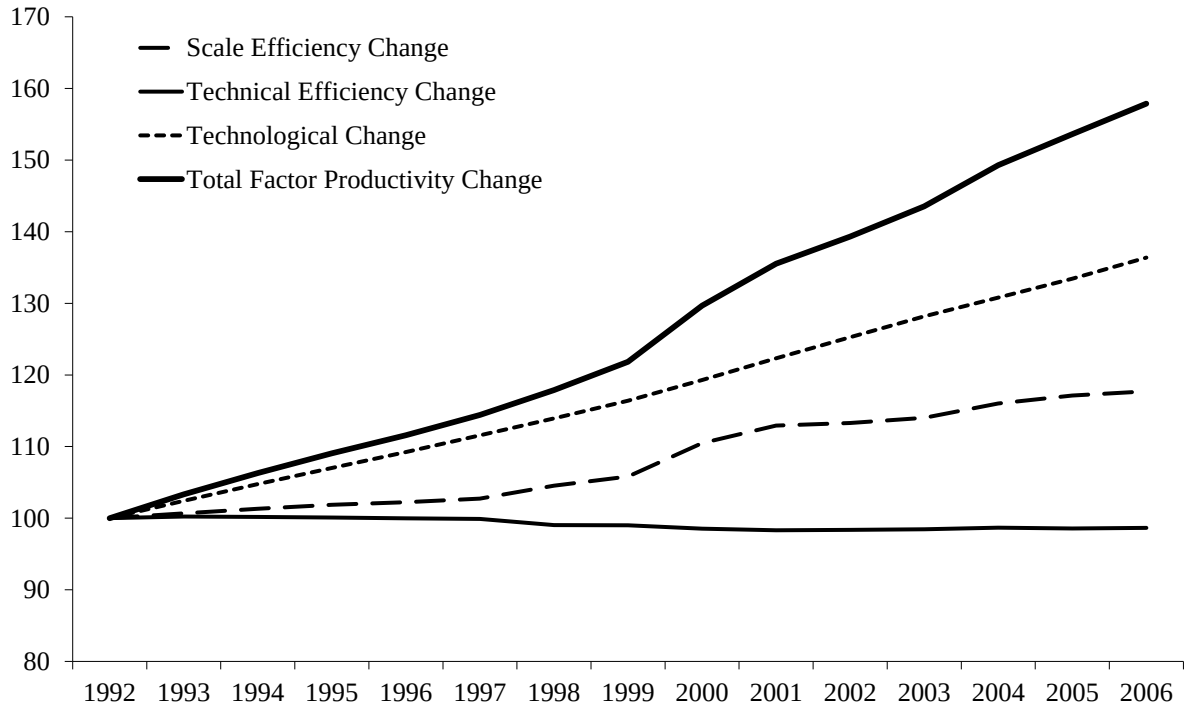
Shadow cost of equity (per cent)



Note: Total loans adjusted for securitization are used as weights in the computation of means.

Figure 5

Total factor productivity growth (Index, 1992=100)



Note: Total loans adjusted for securitization are used as weights in the computation of means.

Portfolio Rebalancing and the Transmission of Large-Scale Asset Programs: Evidence from the Euro Area

Ugo Albertazzi, Bo Becker, Miguel Boucinha*

Abstract

One of the main channels of transmission of large-scale asset programs is the so-called portfolio rebalancing channel, whereby, in a context of low yields, investors have incentives to shift their investments towards assets with higher expected returns. Using granular, security-by-security information on the composition of financial portfolios of all aggregate institutional sectors in the euro area countries, we document how asset allocation evolved around the announcement of ECB asset purchase programme (APP). In order to explore the role played by APP in tilting asset allocation, we exploit cross-sectional heterogeneity in the impact of the programme on the valuation of the financial portfolio held by each sector in Mach 2014, well before the introduction of asset purchases. We then zoom in on each of the euro area's largest banks in order to investigate the programme's impact on bank lending. Interestingly, our findings suggest that portfolio rebalancing manifested in a distinct form across euro area countries. In more vulnerable countries, where macroeconomic unbalances and relatively high risk premia remain, APP was mostly reflected into a rebalancing towards riskier securities. In less vulnerable countries, where constraints on loan demand and supply are less significant, the rebalancing was observed mostly in terms of bank loans.

JEL classification: E44, E51, G21.

Keywords: quantitative easing, unconventional monetary policy, portfolio rebalancing, search for yield.

* Respectively, Bank of Italy (Directorate General Economic Research and Statistics), Stockholm School of Economics and CEPR, and the European Central Bank (Directorate General Economics). We wish to thank Sohei Kaihatsu, Burçin Kısacıkoˇ, Enrico Sette, Stefano Siviero, Andrea Tiseno, for their comments as well as participants at the conferences and workshorps held at the Bank of Portugal (January 2016), Bank of Italy (March and October 2016), IFABS conference in Barcelona (June 2016) and European Central Bank (October 2016). Any remaining errors are our responsibility. The views expressed are our own and do not necessarily reflect those of the Eurosystem.

1. Introduction

The crisis triggered by the collapse of Lehman Brothers in September 2008, and the accompanying recession provoked the development of a new set of monetary policy tools. Central banks in all main developed countries reacted to the crisis by cutting official rates and adopting a wide range of unconventional measures. A key such measure is asset purchase programs, whereby the central bank aims at lowering long-term yields through purchases of bonds. These programs were seen as a necessary monetary policy tool to provide stimulus once policy rates approached their effective lower bound. Early programs include the US QE1, QE2 and QE3 programs undertaken by the Federal Reserve starting in 2008 and similar policies initiated by the Bank of England in early 2009.

In this paper, we examine the impact of the European Central Bank's (ECB) Expanded Asset Purchase Program (APP). This program was implemented later than the US and UK programs, against the backdrop of a very prolonged economic downturn in the Euro area which coincided with historically low inflation. The ECB announced the APP, on 22 January 2015 and the implementation started in March of 2015.[55]

A key question about asset purchase programs is whether they work, in the sense of generating a positive impact on macro-economic developments and, if so, through which channel. A direct impact from reduced interest rates is unlikely to be important for asset purchase programs (see Stein 2012).[56] Another possible channel is signaling (whereby asset purchases serve as a commitment device for relatively high future inflation targets). Krishnamurthy and Vissing-Jorgensen (2013) reject this channel for the US. Instead, these authors suggest that the US QE programs have worked largely through a narrower channel, increasing the price of the specific assets purchased under the program, with possible spillovers depending on institutional features and economic conditions.

[55] See https://www.ecb.europa.eu/explainers/tell-me-more/html/asset-purchase.en.html.

[56] This is because it is argued that these programs reduce (real) long-term interest rates mainly by compressing the term-premium incorporated in yields, rather than the expected levels of future short-term rate. Under these conditions, firms can finance themselves at a cheaper rate by issuing longer-term securities, but the (opportunity) cost of investing in the marginal project does not diminish as its return has to be confronted with the expected return achieved by investing in a sequence of short-term securities, which remains unchanged. In a few words, *LSAPs are likely to elicit a financing response on the part of firms, as opposed to a change in their capital spending plans* (Stein, 2012).

A related channel, widely emphasized in the policy debate, is the portfolio rebalancing channel: asset purchase programs exert pressure on the supply of credit to the riskier segments of the economy, typically those suffering the most from credit supply restrictions during downturns as, by reducing yields on safe long-term securities, investors have incentives to shift their investments towards assets with higher expected returns, thus taking on more risk. This search-for yield is argued to represent an important channel of transmission of purchase programs, if not the main one, as it implies that the monetary stimulus is passed-through onto sectors which, unlike issuers of securities which are eligible for the central bank purchases, cannot directly benefit of the program. Indeed, portfolio rebalancing is deemed to be able to benefit even SMEs, which typically do not issue securities on financial markets, by stimulating banks' supply of loans to this sector. According to a different view, portfolio rebalancing is instead a perverse byproduct of asset purchase programs as it implies an increased risk taking that may sow the seeds for future crises.[57] It is therefore crucial to document not only if portfolio rebalancing takes place, but also for which types of investors and assets.

This study aims at exploring the relevance of portfolio rebalancing for the transmission of the APP by exploiting granular information on the composition of security portfolios for all aggregate holding sectors in euro area countries. The announcement and introduction of the APP was associated with a positive impact on financial markets overall. Long-term yields have declined sharply over the period when the debate on a possible purchase program by the ECB has intensified. At the end of 2013 10–year benchmark government bond yields in the euro area started to decline sharply and kept doing so until the end of March 2015, right after purchases actually began. The prominent role played by APP in explaining such massive decline in yields has been demonstrated in analyses based on an event study methodology (Altavilla et al., 2015).

The core of our analysis consists in an examination of whether sectors that experienced higher gains rebalanced toward riskier assets, compared to holding sectors with smaller

[57] Portfolio rebalancing can be seen as an instance of the risk-taking channel of monetary policy, as recently documented based on granular data in Jiménez et al. (2014), applied to the specific case of asset purchase programs.

gains. We essentially take the initial impact on financial prices as a given, and ask to what extent it had a secondary effect on asset allocations across different sectors and countries over the period 2014Q1 and 2015Q2, the first data point after the decline in yields induced by expectations of the APP. In each euro–area country, we consider the securities held by the following institutional sectors: banks, money-market funds, insurance corporations, pension funds, other financial corporations, non-financial corporations, households (including non-profit institutions serving households), general governments and rest of the world.[58]

Our identification strategy relies on two assumptions. One is that, across holding sectors, incentives for rebalancing are commensurate to the changes in the value of the portfolio. If, prior to the yield decline phase, a sector was holding securities whose yields has diminished little, then such sector is supposed to be have lower incentives to search-for yield and rebalance towards riskier (higher-yield) securities. This means that we can exploit cross-sectional heterogeneity in the exposure to the APP shock, as measured by the valuation gains experienced against the background of the announcement and introduction of APP. The change in the value of the financial portfolios held in Mach 2014 by each institutional sector in a given country (hereinafter, holding sector) varied substantially. For example, the 25th percentile was a gain of around 2% and the 75th percentile around 4%.

The second assumption is that we can exploit the granularity of the dataset to address one tricky endogeneity issue. If we observed that holding sectors experiencing higher valuation gains exhibit a sharper rebalancing towards riskier securities, this would be consistent with an increase in the financing needs of riskier issuers, whereby such increased credit demand has been met by sectors that typically invest in risky securities and as such were more exposed to the APP shock (that is, experienced larger re-valuations of their portfolio). This would imply an increase in credit demand by some issuers, rather than portfolio rebalancing which is a notion involving an increase in credit supply for risky borrowers or issuers. Crucially, the availability of security-by-security information allows for the comparison of investment patterns in the same security across different sectors, effectively controlling for the credit demand channel. We present results for all securities in the dataset but we give a special focus on newly issued debt securities. This is because, by

[58] Details on data used are available at https://www.ecb.europa.eu/stats/money/shs/html/index.en.html.

construction, as long as the monetary policy stimulus succeeds in inducing rebalancing for the average investor in the economy, this can be accommodated only via an increased issuance of riskier securities. Clearly, for outstanding securities, rebalancing towards riskier securities by some investors needs to be accommodated by portfolio rebalancing in the opposite direction by some other investors.

It is also crucial to emphasize that purchases under the APP are subject to strict rules concerning their cross-country allocation based on euro-area national central banks' individual shares in ECB's capital. This rules out that the ECB was targeting securities in specific countries and the related possible endogeneity issues.

We will be using yields or spreads as risk indicators (similar to the method used by Becker and Ivashina 2015 to examine cyclical variation in the risk appetite in US insurance portfolios), but we will also look at specific dimensions of risk, such as the rating and the residual maturity.

Throughout our analysis we will also assess portfolio rebalancing across vulnerable and less vulnerable countries.[59] As will be shown, different patterns will be documented between these two groups of countries. Interest in these geographical patterns is warranted by the fact that, in the context of financial fragmentation that emerged with the euro area sovereign debt crisis, different conditions for credit and lending supply have been observed in the two areas.

The analysis of rebalancing in debt security portfolios allows for an assessment of effects on the supply of credit in the form of securities. However, in the euro area, bank lending constitutes the key source of financing for most firms – especially SMEs. We estimate the effect of the APP on bank lending by exploiting a detailed dataset for each of the largest twenty five banking groups in the euro area.

For each bank we compute the exposure to the APP shock in line with what is done for sectoral holdings, that is by considering the increase in the value of financial portfolios held at before the announcement of the APP. We then measure whether this is related to the amount of loans subsequently extended to the real economy, controlling for possible heterogeneity in demand conditions faced by lenders operating in different countries.

[59] Throughout the paper, the term "vulnerable countries" refers to Cyprus, Greece, Ireland, Italy, Portugal, Slovenia and Spain. The remaining euro area countries are referred to as "less vulnerable countries".

The results of our analysis, focusing on newly issued securities, show no statistically significant relationship between portfolio rebalancing patterns across sectors and the exposure to the APP shock for the euro area as a whole. A relationship can however be documented when focusing on more vulnerable economies only, in particular in what concerns corporate bonds held and credit (but not maturity or currency) risk. For what concerns lending activity, banks more exposed to the APP displayed larger reductions in the interest rates applied on new loans to households and, in less vulnerable countries, higher growth of credit extended to non-financial corporations. One possible explanation for our distinct findings across country groups is that in non-vulnerable countries spreads were already so compressed to begin with that, against a background of persisting home-bias, engaging in search for yield would require an unfeasibly large change in portfolio composition. Relative returns on different types of assets would then favour rebalancing towards lending activity.

The rest of the paper is organized as follows. Section 2 contains an overview of the relevant literature. Section 3 provides a brief description of the novel dataset used. Section 4 presents the econometric exercises. Section 5 summarizes the results and concludes.

2. The related literature

Several recent papers attempt to assess the effects of asset purchase programs. One group of papers aims at empirically documenting the impact on asset prices and bond yields. They rely on granular and high-frequency data to identify the response of market prices for individual securities around announcements of asset purchase programs by central banks. Overall, this strand of literature argues that asset purchase programs increase asset prices and diminish bond-yields.[60]

Other papers use bank-level information to investigate the presence of a bank lending channel of asset purchase programs by testing whether banks that end up receiving most of the liquidity injected with central bank purchases of long-term bonds have disproportionately

[60] These papers include, among others, Krishnamurthy and Vissing-Jorgensen (2013) on FED's QE, Joyce and Tong (2012) on Bank of England's program, Krishnamurthy et al (2014) on the ECB's OMT and SMP, Altavilla et al (2015) on APP. A related but different approach is in Wright (2012) who estimates a VAR with daily data where the identification is derived from the assumption that monetary policy shocks have high variance on days of FOMC meetings. He also finds an impact of monetary policy shocks on governments and corporate bonds, although only a transitory one.

increased their loan supply. Butt et al (2014), looking at UK's experience, do not find significant effects; Kandrac and Schlusche (2016), instead, find evidence of an operational bank-lending channel for the US.

Other authors attempt to assess the effects of asset purchase programs on real macro-economic variables using VAR or DSGE models. These papers look at different episodes and countries so that results are not always comparable. Nonetheless, the broad message they convey is that there is a significant impact of asset purchase programs on the real economy.[61] Since the identification in these studies comes from aggregate time series variation, the precision with which specific causal mechanisms can be pinpointed is generally weaker. To the extent that programs are introduced at non-random times, the results may be confounded.

Peydrò et al (2016) exploit granular bank-level data on individual security and borrower exposures of Italian banks. Their objective is to study how banks' investment decisions are influenced by the monetary policy rate or by the adoption of unconventional monetary policy measures, captured by the size of the central bank balance-sheet. They conclude that unconventional monetary policy measures do not induce risk-taking in the composition of security portfolios nor on lending supply. Compared to Peydrò et al (2016), whose sample period ends in 2013, we use a dataset covering a more limited time span but which covers the APP. We also analyze security-by-security holdings of all main institutional sectors, not just banks, and cover all euro area countries. Another difference is related to the indicator adopted to capture unconventional monetary policy (size of central bank balance-sheet) which cannot reflect what occurs in anticipation of the actual implementation of such policy measures. This approach, adopted in several other papers (e.g., Gambacorta et al., 2014), is not suitable for our purposes given our focus on the ECB's asset purchase program. As mentioned, APP started to reflect on the size of the central bank-balance sheet only gradually, starting in March 2015, while the price impact on long-term yields took place

[61] Baumeister and Benati (2012) use a Bayesian time-varying parameter structural VAR for a sample of advanced economies and argue that a compression in the long-term yield spread exerts a powerful effect on both output growth and inflation. Following a broadly similar approach, Kapetanios et al. (2012) studies the first round of QE in UK and suggest that QE may have had a peak effect on the level of real GDP of around 1.5 and 1.25 p.p. on real GDP level and CPI inflation respectively. Chen (2014) finds that the sole LSAPs interventions in the US had an insignificant effect on the macro-economy. She finds instead a strong effectiveness of the policy involving an extended period of near-zero interest rates, either on output or on inflation, depending on whether perfect foresight rational expectations are incorporated into the model or not.

entirely before that date.[62] We also exploit bank-level information on security holdings for a sample of large banking groups in the euro area but, unlike Peydrò et al (2016), we cannot match this information with loan-level data. Nonetheless, we will be able to provide an assessment of whether portfolio rebalancing has benefitted loan-supply by integrating our dataset with bank-level information on the amount and the cost of the credit extended to different sectors, including firms and households.[63]

Another paper close to ours is Koijen et al (2016). Based on a dataset similar to ours, the authors describe the evolution of portfolio composition across institutional sectors in the euro area and build a methodology to assess its impact on yields, different from more standard event study approaches based on high frequency data. Our focus, instead, is on how monetary policy has been affecting the composition of security portfolios.

3. The data and descriptive evidence

The security holding statistics (SHS) dataset contains granular information at individual ISIN level on securities held in the euro area. In more detail, it includes holdings by residents of each euro area and some other EU countries, as collected by the corresponding country, and holdings by non-euro area residents in custody in the euro area.[64] The dataset has a widespread coverage, at close to 90 per cent of the universe of debt securities reported in the national accounts. The SHS dataset includes also granular information on the portfolio of securities held by each of the 25 largest euro area banks. This dataset is matched with bank-level information on stocks and flows of loans granted to the non-financial private sector and on the corresponding interest rates, so as to investigate the impact of the monetary policy shock on bank lending to the real economy.

The first form of rebalancing that can be observed from raw data is across types of securities held (Figure 1). A clear rebalancing of portfolios towards equity instruments was

[62] Implicit in our choice is the idea that portfolio rebalancing is a special case of search for yield. From this perspective, what really matters is the impact of the announced measures on yields and this, in principle, could originate even without actual purchases, as shown by the notorious episode where ECB President Draghi committed to do "whatever it takes" to fulfill his keep the central bank's mandate.

[63] A similar dataset is exploited in Abbassi et al (2015) who uses granular bank-level data from Germany on individual security and borrower exposures and establishes that trading experienced banks are more likely to cut loan supply to exploit investment opportunities in financial markets.

[64] In principle the dataset also includes holdings by non-financial residents of each euro area country in custody in other euro area countries and therefore collected by these other euro area countries; however, due to data quality concerns and in order to avoid possible double-counting, this component is excluded.

observed for OFI and, to a smaller extent, private sector non-euro area investors in 2015 Q2, compared to 2014Q1. This was to a large extent driven by a higher valuation of the outstanding equity portfolio and not by new (equity) finance provided by less risk averse investors, the ultimate goal of monetary policy. Once holding amounts are adjusted for valuation effects (not shown), a visible rebalancing towards equity was observed only for OFIs, though these holdings represent a negligible share of the overall portfolio of securities. It remains therefore to be assessed whether portfolio rebalancing benefitted the supply of new credit in the form of bonds.

Table 1 shows some descriptive statistics, focusing on the sample of debt securities issued in the two quarters considered. Large differences are observable in the holding amounts, across securities and holding sectors, reflecting heterogeneity in the size of issuances and of holding sector portfolios. Portfolio valuation, m_h, is the investor specific measure of APP shock intensity and is defined as the change in the value of securities held by each sector in 2014Q1, before the anticipation of the APP. This measure displays significant variation both across institutional sectors and countries. Concerning holding sectors, the impact was particularly significant for insurance corporations and pension funds and for other financial intermediaries, reflecting the long duration of the securities held by these classes of investors. Looking at countries, a noteworthy pattern is that the stronger valuation effects are discernible in non-vulnerable countries. This finding, which may come as a surprise, is explained by the higher share of equity instruments and of investment fund and money market fund participation units in these countries, against a background in which the value of these assets was more affected by the APP than that of debt securities. Maturities are similar across groups of countries, but show considerable dispersion across individual countries and holders. Yields and spreads are higher in more vulnerable countries, as expected.

Table 2 reports similar statistics, for the two periods separately. Some increase in the average maturity and in the share of non euro-denominated bond holdings is observable between the two periods. Furthermore, not only average yields but also spreads decline, which would not be consistent with increased risk taking. However, one needs to take into account that purchase programs may possibly exert a downward pressure on expected future

short-term rates and on unit risk premium (for both term and credit risk). This pricing impact may hide a rebalancing towards relatively higher yield securities.[65]

4. Econometric evidence

4.1 The empirical framework

The objective of this section is to explore the role played by monetary policy in shaping the risk appetite of euro area investors. The empirical strategy exploits heterogeneity in the exposure to the monetary policy shock in the cross section of investors, measured by the impact of the APP on the valuation of the portfolio of securities held at 2014 Q1.

The approach used to implement this strategy essentially consists in the estimation of a regression equation with the following baseline specification:

$$h_{i,h,t} = \left(\beta_0 m_i + \beta'_0 r_{it} + \beta_0'' m_h r_{i,t}\right) + \left(\beta_1 m_h T_t + \beta_1' T_t r_{i,t} + \beta_1'' m_h T_t r_{i,t}\right) + \\ + \gamma T_t + a_{i,t} + b_{h,t} + \varepsilon_{i,h,t} \qquad (1)$$

The variable $h_{i,h,t}$ is the (log) amount of holdings of security with ISIN *i* by holding sector *h* (e.g. French investment funds), in the two periods considered (*t* is either 2014 Q1 or 2015 Q2). m_h is the intensity of the monetary policy shock specific to holding sector *h* and is defined as $m_h = w'_h e$, where w_h is a vector defining the composition at 2014 Q1 of the financial portfolio for investor h and e is the vector of the actual variations in the price of each security over the period observed. T_t is a dummy variable identifying the post-announcement period, 2015 Q2. r_{it} is the yield-to-maturity of security *i* at time *t*. A positive estimate for the coefficient β_1'' would indicate that between the two periods investors more exposed to the monetary policy shock rebalanced their portfolio towards riskier securities more intensely than other holding sectors.

Although expectations and implementation of APP are plausibly the most important drivers of financial asset prices in the period under examination, prices of securities may have also changed for other reasons unrelated to APP. This is not problematic for our approach. First, the notion of search-for-yield refers to a reduction of yields and does not

[65] It should be pointed out that we are considering the security residual maturity as a (rather crude) proxy for financial duration for which no information is available.

require the yield to diminish for a specific reason or factor, such as monetary policy. Second, the distinction between actual changes in prices and changes related to monetary policy matters for the quantification of the its effects but not for the estimation of the coefficient the coefficient β_1'', at least as long as m_h can be considered an exogenous regressor. As the portfolio composition over which m_h is computed is the one prevailing before market yields declined, it can by definition not be influenced by expectations of quantitative easing (which would lead to reverse causality).

As discussed, *on aggregate*, portfolio rebalancing can occur only if there is an additional supply of risky securities. Given that our objective is to assess the transmission of monetary policy on credit we focus on newly issued securities. For each of the two dates considered (end 2014 Q1, end 2015 Q2), newly securities are defined as those issued in the preceding 4-quarters. This is done to smooth out possible seasonality effects and to avoid capturing developments specific to a given quarter. This also means that any mechanical relationship between changes in valuations and changes in portfolio composition, which would generate spurious correlations, is avoided.[66]

Exploring the granularity of our dataset, we will conduct our estimates by also including different sets of fixed effects. These are crucial to be able to control for possible unobservable characteristics of the securities or of the holding sector which may blur the results. In particular, we can perfectly control for developments in credit risk or financing needs that are associated to a given security or issuer.[67] Indeed, starting from Kwhaia and Mian (2008), a recent and growing empirical literature in banking exploits loan-level datasets and the fact that borrowers concomitantly borrow from multiple lenders, to run estimations including (time specific) borrower fixed effects. Introducing fixed effects for each security (in each period) allows us to isolate from credit developments everything

[66] The focus on newly issued securities is useful also to overcome the problems caused by sluggishness in portfolio rebalancing. For example, one may plausibly think that retail investors do not optimally adjust their portfolio holdings in real time, but only in discrete time. Therefore, when the portfolio composition changes because of heterogeneous changes in value across the securities held, it might take time before a household makes new transactions to re-optimize the portfolio. Acquisitions of newly issued securities require, by definition, that some transaction is conducted, so sluggishness for such sectors is supposedly less relevant.

[67] Clearly when fixed effects are included in the estimations some terms of equation (1) will mechanically be dropped. Also, given that by definition a security is newly issued in one of the two periods only, introducing time-invariant security fixed effects (a_i) would be equivalent to having period-specific security fixed effects ($a_{i,t}$). We keep this notation as will also show, among the extensions, some estimations conducted on both seasoned and newly issued securities.

which is explained by specific instrument or borrower characteristics, irrespectively of whether these are time varying, time invariant, observable or not. Therefore, the introduction of these fixed effects is the most effective control for credit risk and demand conditions.

Similarly, with holding sector (time-varying) fixed effects, we can effectively control for everything that is specific to a given class of investors and has an impact on the overall size of its portfolio. This is important given that different investor categories may structurally invest in securities involving different levels of risk. As m_h may vary only across different holding sectors, all regressions are estimated by clustering errors at the level of *h*.

Search for yield is investigated using the yield to maturity as an encompassing measure of risk (r_{jt}), as well as more specific components such as credit risk (spread), maturity risk and exchange rate risk.

4.2 Results for the baseline model

Table 3 shows the estimation of model 1 for the sample of newly issued securities and for different specifications characterized by different types of controls and different sub-samples. Looking at the first three columns, referring to the estimation for the whole sample, it turns out that irrespectively of the specification adopted, the triple interaction is never positive and statistically significant.

As mentioned, heterogeneity could be expected to be significant across investors residing in different countries. An obvious breakdown is the one between vulnerable and less vulnerable euro area economies. During the sovereign debt crisis financial conditions in countries more directly involved have significantly diverged from those of other countries. In early 2014, when long-term yields started their declining trend in anticipation of the adoption of APP, differences were still sizeable, though much less so than at the peak of the sovereign debt crisis.[68]

[68] For instance, the spread between the yield on domestic 10-year sovereign bonds and the corresponding German figure was about 2 p.p. in Italy and Spain, 3 p.p. in Portugal. Spreads on sovereign may be though as a lower bound for spreads on corporates, so financial market fragmentation was still sizeable. Sovereign spreads started diminishing thereafter and reached minimum levels in March 2015 when they stabilized at smaller but still non negligible levels (1 p.p. Italy and Spain, 2 in Portugal).

These differences may have important implications for portfolio rebalancing, although a priori it is not clear in which direction they should affect the intensity of rebalancing. Opposite hypotheses can be made. On the one hand, one may conjecture that given the already higher level of risk in vulnerable countries, domestic investors would be less inclined to take on additional risks. At the same time, interest rates in less vulnerable economies were so low and possibly squeezed toward their lower bound (10-year Bund in March 2015 was 0.2 percent; yields on many shorter term bonds were negative) that in order to search-for-yield, in a context where most of the securities offer return rates close to nil, one would need to distort the portfolio composition to an extent that would be too costly or even impossible (e.g. constrained by investment policies). Of course, this reasoning implies some fragmentation in financial markets, such that investors in other countries are reluctant to invest in the countries which were more affected by the sovereign crisis. These two hypotheses have opposite implications on whether one should expect more rebalancing in one area or in the other. To gain some insight on this we explore the international coverage of our dataset to investigate the behavior of investors by focusing on the two groups separately.

Column 4 of Table 3 shows the OLS specification for the subsample of holding sectors residing in vulnerable countries. The coefficient for the triple interaction term is now positive and statistically significant, suggesting that in these countries monetary policy has brought about some rebalancing towards risky assets. Results (not shown) for the subsample of less vulnerable economies confirm that rebalancing is limited to vulnerable countries.[69] As mentioned, this can be interpreted as a sign that in a context of diminishing returns and fragmented financial markets, risk balancing is easier in vulnerable economies, where securities paying non trivial yields are available.

The robustness of the result to the introduction of fixed effects for each pair period-holding sector (column 5) suggests that it is not driven by an increase in the size of the portfolio of some sectors which are specialized in investing in more risky securities but rather a genuine tilt in asset allocation (it should be noted that these sectors would also likely exhibit larger values for m_h). More generally, this implies that the result is robust once we

[69] Results for less vulnerable countries are similar to those obtained for the whole sample, both in terms of sign and statistical significance.

control for any kind of factors affecting the entire portfolio of each holding sector considered.

The coefficient on the triple interaction term remains positive and statistically significant also when introducing (time-varying) security fixed-effects together with time-varying holding-sector fixed effects (column 6). This suggests that the rebalancing observed is not exclusively originated by a stronger than usual issuance of risky securities, something we may label confidence or credit-demand effect, but it is at least partly induced by an intensified desire of (high m_i) investors to increase their holdings of such securities. The particularly low spreads on risky securities prevailing in this period also corroborate this credit supply-side view.[70]

The documented effects are sizable. As shown in Table 4, based on the coefficients of the OLS model, for a sector with a median shock (m_i=2.42%) the semi-elasticity of the amount of holdings to the level of yield (the percentage change of the amount of holdings of a security when its yield increases by one p.p.) increases in the post period by 10 p.p.. [71] In contrast, for a sector almost not exposed to the APP shock (m_h=0.46%, 10th percentile of the distribution of m_h), such semi-elasticity is negative in 2014 Q1 and even slightly diminishes in 2015 Q2, possibly reflecting the generalized reduction of spreads (for a given increase in yield in post the increase in underlying risk, say the probability of default, is larger than in the pre-announcement).

As discussed above, one interpretation of finding evidence of rebalancing only in more vulnerable economies is that investors residing in countries where long term yield are squeezed to very low levels may find additional constraints to rebalance to riskier portfolios as this would require investing in other economies, which may be problematic in a context where home bias is still persistent. In other words, being at the zero-lower bound may be a constraint for the rebalancing rather than a factor exacerbating risk taking.

[70] This specification determines a reduction in the number of securities as some of them are held by one sector only (this is typical for Germany).

[71] Note that the OLS is the only specification where an estimate of the level of the coefficient of semi-elasticity, which summarise asset allocation, can be derived for the different sectors and in the different periods. In all the following specifications, where we also introduce fixed effects at the holding sector or at the security level, one loses information on the level and can just focus on cross-sectional differences. Accordingly, for those specifications we will be commenting only on the sign of the triple interaction term.

4.3 Extensions and robustness

Table 5 repeats the same type of regressions considering only government bonds or other securities, respectively. As can be seen, much of the rebalancing documented in Table 3 takes place within the category of securities issued by the private sector.

This is relevant as it suggests that the monetary-policy induced increase in risk appetite has benefitted mainly the supply of credit to the real economy, which is in line with the notion of portfolio rebalancing as a transmission channel of asset purchase programs. At the same time, it should be emphasized that only large corporates can issue securities on the market and these firms tend to be constrained in their access to credit.[72]

The level of the yield is a summary measure of the risk involved in investing in a given security. It may thus subsume different components, namely credit risk, maturity risk or currency risk.[73] This is explored for more vulnerable economies in Table 6 where the variable r_{jt} is replaced by three alternative measures of risk: the spread between the yield paid by the security and the risk-free rate of a corresponding maturity; the maturity of the security (in months); a dummy for non-euro denominated securities. The specification is modified so as to include, for each of these risk measures, all possible double- and triple-interaction terms.

As shown in Table 6, the results of this exercise suggest that most of the rebalancing is driven by increasing investments in securities involving higher credit risk (the only triple interaction term with a positive sign and statistically significant is that for the spread). This holds across all specifications, irrespectively of the type of fixed effects included.

One interpretation for the lack of amplified risk taking in terms of maturity is that investing in long-term assets is a relatively costly way to search for yield precisely because the term structure has flattened (to increase the yield by one p.p. one needs to lengthen the maturity by a much bigger amount compared to normal times). Absence of rebalancing towards non-euro denominated securities may signal a residual persistent fragmentation of

[72] Note also that this analysis neglects possible rebalancing taking place between these two categories of securities.

[73] Investing in non-euro currency involves some currency mismatch for resident investors that typically have liabilities denominated in euro. Information on the extent to which investors hedge against this type of risk is not available.

financial markets that is restraining the allocation of euro area investors. Note that these regressions exclude foreign investors who are responsible, together with investment funds, for much of the increase in the share of non-euro denominated securities (by definition, for them it is not clear whether investing in non-euro represents an increase or a decline in the currency mismatch).

While our focus is primarily on new issuances, we also conduct estimations on the entire sample also including seasoned securities. The main purpose of this exercise is not to assess the transmission of APP to the real economy, but rather to hint at its implications for financial stability, as the overall risk to which investors are exposed obviously needs to be measured on the entire portfolio. As shown in Table 7, when controlling for both sets of fixed effects, no visible APP-related rebalancing is detected, not even for more vulnerable economies (the coefficient for the triple interaction term in columns 3 and 6 is not significant). These results suggest that the rebalancing observed in newly issued securities was not large enough to modify the overall risk profile of the portfolios of securities held. This assessment may, of course, change over time if rebalancing continues in a context of persisting low rates.

A potential concern for identification arises if the behavior of holders was already different before the APP and, in particular, if holders that came to be more affected by the programme were already rebalancing towards riskier securities before the policy started to be anticipated by the markets. In this case one would expect to find a positive triple interaction before the policy announcement. Tables A1 and A2 in the Appendix show that this is not the case: there is no positive relationship between changes in the portfolio allocation of different sectors in the period from 2013Q4 to 2014Q1 and the extent to which these sectors were then affected by monetary policy, not even in vulnerable countries.[74]

4.4 Portfolio rebalancing in the extensive margin

The regression set up described in equation (1) is not suitable to explore the extent to which APP-related portfolio rebalancing has involved the extensive margin, that is, investments in assets issued by issuers toward which the investors were not already exposed prior to APP announcements.

[74] The choice of 2013Q4 as the starting period is driven data availability.

In order to do so, it is necessary to take into account that the dataset does not include observations for triples *i*,*h*,*t* (security, holder, period) for which the amount of holdings is nil (irrespectively of the fact that we are taking log-amounts).

To account for non-reported nil holdings one observation with a nil holding amount is added to the dataset for each pair security-holding sector that is absent from the dataset (and this for each time period). In order to keep the number of observations manageable, such "rectangularisation" of the dataset is based on security categories, or pseudo-securities, instead of actual individual securities. We defined about 2,300 categories distinguished by different combinations of issuer sector, issuer country, maturity, coupon type, nominal currency and rating.

We then define a dummy-variable identifying new holdings, i.e. security categories held in positive amount in 2015 Q2 but not in 2014 Q1. We then drop observations for 2014 Q1 and estimate a linear probability model for the new holding dummy. We estimate different specifications allowing the model to incorporate pseudo-security fixed effects and holding sector fixed effects. As the time dimension is lost, the emphasis is now on the coefficient for the term of interaction between the security yield r_i and the holding sector portfolio valuation m_h.

The results are displayed in Table 8, looking at investors in more vulnerable countries and showing that, irrespectively of the specification adopted, the coefficient for r_i*m_h is never significant. Therefore, we do not find evidence of APP-related portfolio rebalancing leading to investments in new security categories but only within such categories, possibly reflecting the presence of some constraints on the investment strategies that investors may follow.

For robustness purposes we conduct the analysis on the intensive margin in the rectangularised dataset. Results, displayed in Table 9, confirm the presence of portfolio rebalancing.

4.5 Portfolio rebalancing and lending supply for individual banks

This section intends to shed some light on the direct link between monetary policy and euro area banks' lending activity. It relies on SHS data collected for the 25 largest euro area banking groups in order to obtain a bank-level measure of the intensity of the monetary

policy shock (m_h), defined as described in Section 4.1. We then investigate the impact of this measure on quantities and prices of loans granted to the non-financial private sector. Although the small number of banks represents a constraint for the econometric exercise we conduct, the dataset is relevant in terms of coverage as it includes a large share of the euro area banking system, at around 70% of total assets.

One observation in the dataset used for these regressions is a pair *b-s*, where *b* stands for a given bank and *s* for a given borrowing sector (households and non-financial corporations). The dependent variable is the yearly growth rate of loans extended by bank *b* at the end of 2015 Q1. The regressions are estimated by also including a set of country fixed effects, as controls for (country-wide) credit demand and risk. Standard errors are clustered at the bank level. Table 10 shows that there is a statistically significant relation between the monetary policy shock and the growth rate of loans to non-financial private sector (column 1).[75] The results in the second column of the table show that this effect is not statistically different for loans to households and to non-financial corporations (NFC), even though the negative coefficient suggests that the effect is somewhat weaker for loans to the latter sector.[76] In what concerns geographical patterns, the relationship is found to be significant only for banks headquartered in less vulnerable countries. These results are robust to the inclusion of bank-specific control variables such as regulatory capital ratios, asset quality, CDS and rating (not shown). As such, it should not be driven by differences in the balance sheet strength or market perceptions of the risk of banks across jurisdictions.

We also run similar estimations where the dependent variable is the interest rates applied on new loans extended in the four quarters to 2015 Q1. No relationship is found for the non-financial private sector as a whole (Table 11, column 1). However, this masks underlying differences across sectors. High m_h banks are found to decrease the interest rates

[75] For what concerns the security portfolios of banks, we also estimated equation (1) with bank-group data and could not find any significant evidence of rebalancing, neither in vulnerable nor in non-vulnerable economies (not shown). This reassures about the fact that the results shown above for sector-by-sector holdings are not affected by possible (reverse causality) endogeneity issues that could arise if the ECB targeted the securities held by banks whose investment portfolio is made preeminently of government bonds, in response to anticipations of a rebalancing of their portfolios. Focusing on the case of Italy, Affinito et al. (2016) analyze banks' purchases of domestic government bonds in the years 2007-2013. They conclude that investing in sovereign bonds represented, for Italian lenders, a way to support their own balance sheet conditions at a time of increasing credit and liquidity risk.

[76] A test for the join-significance of both coefficients in column 2 allows for the rejection of the hypothesis that the effect for NFC is equal to zero.

applied on loans to households by comparatively more. This is not the case for loans to NFC whose interest rate displays, counter-intuitively, a positive relation with m_h,, the more so the larger the size of the loans The effect of the monetary policy measure on lending rates is not found to depend on whether the bank is headquartered in a more vulnerable economy or not.[77]

The fact that the monetary policy shock is found to be associated with higher growth of loans to NFC but not with a comparatively stronger decline in interest rates is consistent with the presence of some rebalancing within this borrowing sector towards riskier borrowers.

5. Conclusions and policy implications

In this paper we empirically study whether the APP has induced portfolio-rebalancing. As mentioned above, the portfolio rebalancing channel has attracted a lot of attention in the public debate, despite the fact that its actual relevance is largely unknown.

This channel exerts its effects by inducing an increase in risk appetite. While this may conceivably pose financial stability risks, whether an increase in risk appetite is desirable or not depends on whether the current level of risk taking is below optimum or not. One of the main conclusions of our analysis is that the APP-related portfolio rebalancing is statistically significant only for asset holders residing in more vulnerable countries, where credit conditions are still comparatively tight or, in other words, risk taking is still sub-optimal.

Regarding the transmission of the APP to the real economy via the portfolio rebalancing channel, we obtain a mixed picture. First, we show that the APP-related increase in risk-taking in vulnerable economies has affected securities issued by corporates (as opposed to sovereigns) and has resulted in more credit risk-taking (as opposed to maturity or currency risk-taking). However, when looking at lending volumes granted by banks, we obtain evidence of effects limited to non-vulnerable countries.

One possible explanation of these geographical patterns is that in non-vulnerable countries spreads were already so compressed to begin with, that in order to reach a given increase in the average yield of a given portfolio a dramatic change in its composition would

[77] In a similar vein to the exercise described in the last paragraph of Section 4.3, Tables A3 and A4 show that the there is no positive relationship between the lending behavior of different banks in the period from 2013Q4 to 2014Q1 and the extent to which these banks were then affected by monetary policy.

be needed.[78] This could also explain why in these economies some rebalancing of banks towards real-sector loans (where presumably spreads are still positive) is detectable. One possible explanation for the fact that in vulnerable economies rebalancing has concerned financial securities but not loans to the real economy is that some constraints have limited the expansion in the supply of bank loans. These could be related to regulatory or supervisory activity. Overall, our results do not support the claim that APP poses risks to financial stability while, at the same time, they are consistent with the presence of exogenous constraints limiting its pass-through to the real economy.

[78] To exemplify, if the yield curve is perfectly flat, then even an arbitrary large increase in average duration does not help in raising the average yield. In other words, when comes to search-for yield, both income and substitution effects are at play; when spreads are very much compressed, as it is the case in non-stressed countries during LSAPs, substitution effects may actually dominate.

6. References

Abbassi, P., R. Iyer, J. L. Peydro, and F. R. Tous (2015). Securities Trading by Banks and Credit Supply: Micro-Evidence from the Crisis. *Journal of Financial Economics*, forthcoming.

Affinito, M., G. Albareto and R. Santioni (2016). Purchases of sovereign debt securities by Italian banks during the crisis: the role of balance-sheet conditions. *Questioni di Economia e Finanza* (Occasional Papers), No. 330.

Altavilla, C., G. Carboni, and R. Motto (2015). Asset purchase programmes and financial markets: lessons from the euro area. No. 1864. European Central Bank, 2015.

Becker, B., and V. Ivashina. (2015). Reaching for yield in the bond market. *Journal of inance*, 70(5),1863-1902.

Butt, N. R. Churm, M. McMahon, A. Morotz, and J. Schanz (2014). QE and the bank lending channel in the United Kingdom. *Bank of England*, Working Paper No. 511.

Baumeister, C. and L. Benati (2010) Unconventional monetary policy and the great recession-Estimating the impact of a compression in the yield spread at the zero lower bound. No. 1258. European Central Bank, 2010.

Chen, H. (2014). Assessing the Effects of Large Scale Asset Purchases in a Zero-Interest-Rate Environment through the Lens of DSGE and VAR Models. Board of Governors of the Federal Reserve System, mimeo. De Nicolo, G., Dell'Ariccia G., Laeven L., and Valencia, F. (2010). *Monetary Policy and Risk Taking* (No. 2010/09). IMF Staff Position Note.

Gambacorta, L., B. Hofmann, and G. Peersman (2014). The Effectiveness of Unconventional Monetary Policy at the Zero Lower Bound: A Cross-Country Analysis. *Journal of Money, Credit and Banking*, 46.4: 615-642.

Jiménez, G., S. Ongena, J. L. Peydro, and J. Saurina (2014). Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk-Taking? *Econometrica, 82*(2), 463-505.

Joyce, M., M. Tong (2012). QE and the Gilt Market: a Disaggregated Analysis. *The Economic Journal*, 122 (November): F348-F384.

Kandrac, J., B. Schlusche (2016). Quantitative easing and bank risk taking: evidence from lending. Review of Financial Studies. *Forthcoming.*

Kapetanios, G., et al. (2012) Assessing the Economy-wide Effects of Quantitative Easing. *The Economic Journal*, 122.564 (2012): F316-F347.

Khwaja, A. I., and A. Mian (2008). Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market. *American Economic Review*, 1413-1442.

Koijen, R., F. Koulischer, B. Nguyen and M. Yogo (2016) Quantitative easing in the euro area: the dynamics of risk exposures and the impact on asset prices. *Banque de France, Document de Travail, No. 601*.

Krishnamurthy, A., S. Nagel, and A. Vissing-Jorgensen (2014). *ECB policies involving government bond purchases: Impact and channels*. Mimeo Northwestern University, University of Michigan, and University of California Berkeley.

Krishnamurthy, A., and A. Vissing-Jorgensen (2013). The ins and outs of LSAPs. *Proceedings-Economic Policy Symposium-Jackson Hole*. Federal Reserve Bank of Kansas City.

Peydró, J. L., A. Polo, and E. Sette (2016). Bank of Italy, *mimeo*.

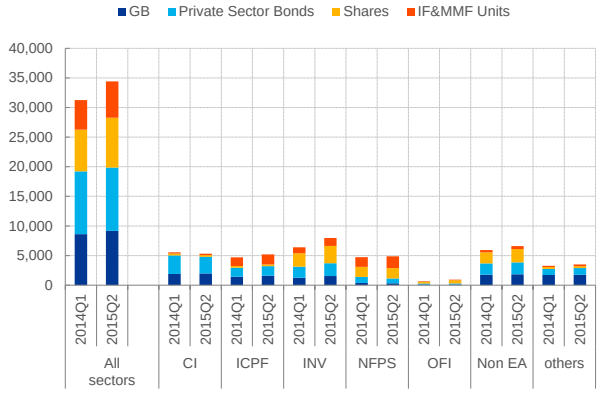
Stein, J. (2012). Remarks at The Third Boston University/Boston Fed Conference on Macro-Finance Linkages. Boston, Massachusetts, November 30, 2012.

Wright, J. H. (2012). What does Monetary Policy do to Long-term Interest Rates at the Zero Lower Bound?. *The Economic Journal*, *122*(564), F447-F466.

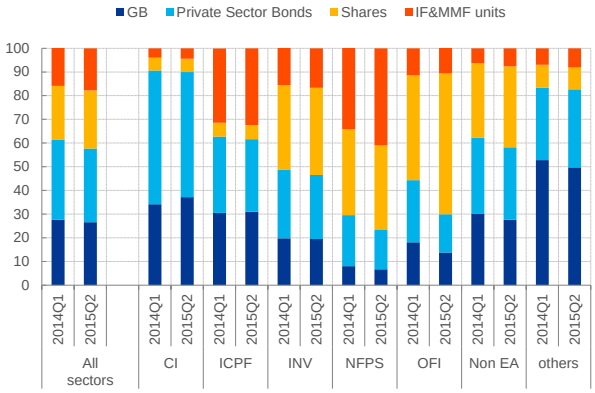
Tables and Figures

Figure 1
Portfolio rebalancing between 2014Q1 and 2015Q2 across types of instruments

a) Outstanding amounts (*billions of euro*)



b) Shares (*percentage points)*



Notes: The chart shows the investment in each type of instrument by holding sector based on market values. Excluding non-euro-area residents third-party holdings (non-euro area residents holdings reported by euro area NCBs) and non-euro area securities held by non-euro area residents. The category OFI does not include FVCs (financial vehicles); non EA does not include holdings of non-resident central banks and general government.

Table 1
Descriptive statistics for newly issued securities

	Mean	Std. Dev.	P25	P50	P75	N. Obs
Full sample						
Holding amount	19.62	157.31	0.20	1.08	6.35	235423
Log (Holding Amount)	0.12	2.55	-1.55	0.11	1.88	232626
Portfolio valuation (mh)	4.12	2.06	3.46	3.89	4.86	235423
Yield-to-maturity	2.96	2.55	1.05	2.60	4.13	235423
$Spread_{it}$	2.53	2.46	0.69	2.07	3.61	228721
$Maturity_{it}$	80.64	72.66	36.00	59.00	96.00	228721
$NonEur_{it}$	0.40	0.49	0.00	0.00	1.00	228721
Vulnerable countries						
Holding amount	22.86	217.56	0.26	1.52	6.99	50140
Log (Holding Amount)	0.31	2.40	-1.31	0.44	1.95	49869
Portfolio valuation (mh)	2.69	1.87	2.26	2.42	3.70	50140
Yield-to-maturity	3.20	2.40	1.67	3.15	3.91	50140
$Spread_{it}$	2.67	2.30	1.07	2.68	3.38	49193
$Maturity_{it}$	86.92	81.13	37.00	59.00	111.00	49193
$NonEur_{it}$	0.34	0.47	0.00	0.00	1.00	49193
Less vulnerable countries						
Holding amount	18.74	136.49	0.20	1.00	6.10	185283
Log (Holding Amount)	0.07	2.59	-1.59	0.03	1.85	182757
Portfolio valuation (mh)	4.51	1.94	3.70	4.53	5.66	185283
Yield-to-maturity	2.89	2.58	0.88	2.42	4.21	185283
$Spread_{it}$	2.49	2.50	0.59	1.91	3.71	179528
$Maturity_{it}$	78.92	70.06	36.00	59.00	95.00	179528
$NonEur_{it}$	0.41	0.49	0.00	0.00	1.00	179528

Notes: Data for 2014Q1 and 2015Q2. Holding amount in EUR millions. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. Yield-to-maturity in percent. Spread$_{it}$ *is the difference at time t between the yield-to-maturity of security* i *and the risk-free benchmark rate of a corresponding maturity, in percent.* $Maturity_{it}$ *is the residual maturity of security i* at time *t, in months.* $NonEur_{it}$ *is a dummy for securities denominated in currenciues other than the euro. mh is the change in valuation of the portfolio of sector h in 2014 Q1, in percent.*

Table 2

Descriptive statistics for newly issued securities before and after the shock

	Mean		Weighted mean		N. Obs.	
	Pre	Post	Pre	Post	Pre	Post
Full sample						
Yield-to-maturity	3.23	2.71	2.34	1.80	112159	123264
Spread $_{it}$	2.65	2.42	1.63	1.41	108880	119841
Maturity $_{it}$	79	83	93	98	108880	119841
NonEur $_{it}$	0.38	0.42	0.17	0.23	108880	119841
Vulnerable countries						
Yield-to-maturity	3.42	2.96	2.55	1.94	25514	24626
Spread $_{it}$	2.75	2.58	2.09	1.66	24983	24210
Maturity $_{it}$	81	93	82	102	24983	24210
NonEur $_{it}$	0.32	0.36	0.11	0.14	24983	24210
Less vulnerable countries						
Yield-to-maturity	3.17	2.65	2.27	1.76	86645	98638
Spread $_{it}$	2.62	2.38	1.49	1.33	83897	95631
Maturity $_{it}$	78	80	97	96	83897	95631
NonEur $_{it}$	0.39	0.43	0.11	0.14	83897	95631

Notes: Data for 2014Q1 and 2015Q2. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The t erm "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. Yield-to-maturity in percent. Spreadit is the difference at time t between the yield-to-maturity of security i *and the risk-free benchmark rate of a corresponding maturity, in percent. Maturity* $_{it}$ *is the residual maturity of security i* at time *t, in months. NonEur* $_{it}$ *is a dummy for securities denominated in currencies other than the euro. mh is the change in valuation of the portfolio of sector h in 2014 Q1, in percent.*

Table 3

Baseline estimation: newly issued securities

	Full sample			Investors in vulnerable countries		
	(1)	(2)	(3)	(4)	(5)	(6)
yield-to-maturity (r_{it})	-0.0596	-0.0551*		-0.0968*	-0.0617**	
	(-1.26)	(-1.72)		(-1.80)	(-2.44)	
portfolio valuation (m_h)	-0.122*			0.0915		
	(-1.85)			(1.12)		
post-APP period dummy (T_t)	0.114			0.594		
	(0.46)			(1.59)		
$r_{it} * m_h$	-0.0200	-0.0195	0.0171	0.0155	0.00118	0.0487***
	(-0.95)	(-1.54)	(1.30)	(0.80)	(0.09)	(2.70)
$r_{it} * T_t$	-0.00852	-0.0778		-0.274**	-0.319**	
	(-0.07)	(-0.82)		(-2.47)	(-2.61)	
$m_h * T_t$	-0.0368			-0.0445		
	(-0.78)			(-0.63)		
$r_{it} * m_h * T_t$	-0.00620	0.00718	-0.00175	0.0528**	0.0708**	0.0469*
	(-0.20)	(0.32)	(-0.35)	(2.31)	(2.37)	(1.92)
holder*time f.e.	No	Yes	Yes	No	Yes	Yes
security f.e.	No	No	Yes	No	No	Yes
N	232626	232618	182580	49869	49865	39450
R^2	0.051	0.320	0.558	0.030	0.244	0.635

Notes : Dependent variable is log of the amounts of security i held by sector h (a given institutional sector in a given country), in period t . Data for 2014Q1 and 2015Q2. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. r_{it} is the yield-to-maturity of the corresponding securitiy, in percent. m_h is the change in valuation of the portfolio of sector h in 2014 Q1, in percent. T_t is a dummy for the period 2015 Q2. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.
* p<0.10 ** p<0.05 *** p<0.01

Table 4

Semi-elasticity of the amount of security holdings to the yield-to-maturity in vulnerable countries

	m_h	2014 Q1	2015 Q2
p10	(0.46)	-9.0	-9.3
p25	(2.26)	-6.2	3.0
p50	(2.42)	-5.9	4.1
p75	(3.70)	-3.9	12.8
P90	(4.71)	-2.4	19.7

Notes : Percentage variation of holdings for a one p.p. change in the yield-to-maturity, conditional on the time period and on the portfolio valuation m_h. Based on column 4 of Table 3. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia.

Table 5

Investors in vulnerable countries; holdings of newly issued sovereign and corporate bonds

	Sovereign Bonds			Corporate bonds		
	(1)	(2)	(3)	(4)	(5)	(6)
yield-to-maturity (r_{it})	0.0289	-0.148*		-0.0829*	-0.0489	
	(0.25)	(-1.76)		(-1.78)	(-1.63)	
portfolio valuation (m_h)	0.0937			0.0962		
	(1.58)			(1.01)		
post-APP period dummy (T_t)	0.269*			0.620		
	(1.83)			(1.46)		
$r_{it}*m_h$	-0.0418	0.000525	0.0314	0.0175	0.00323	0.0518***
	(-1.20)	(0.03)	(1.45)	(1.01)	(0.24)	(2.98)
$r_{it}*T_t$	-0.113	-0.219*		-0.276**	-0.309**	
	(-1.63)	(-1.72)		(-2.24)	(-2.36)	
m_h*T_t	0.00333			-0.0510		
	(0.08)			(-0.61)		
$r_{it}*m_h*T_t$	0.0259	0.0524	0.00982	0.0535**	0.0689**	0.0525*
	(1.35)	(1.58)	(0.46)	(2.07)	(2.11)	(1.79)
holder*time f.e.	No	Yes	Yes	No	Yes	Yes
security f.e.	No	No	Yes	No	No	Yes
N	4382	4368	3904	45487	45482	35532
R^2	0.015	0.206	0.567	0.031	0.258	0.648

Notes : Dependent variable is log of the amounts of security *i* held by sector *h (* a given institutional sector in a given country), in period *t* . Data for 2014Q1 and 2015Q2. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. r_{it} is the yield-to-maturity of the corresponding securitiy, in percent. m_h is the change in valuation of the portfolio of sector *h* in 2014 Q1, in percent. T_t is a dummy for the period 2015 Q2. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.
* p<0.10 ** p<0.05 *** p<0.01

Table 6

Investment in newly issued securities of holders resident in vulnerable countries; individual risk factors

	(1)		(2)		(3)	
portfolio valuation (m_h)	0.177	(1.61)				
post-APP period dummy (T_t)	0.452	(1.46)				
Spread $_{it}$	-0.132*	(-1.98)	-0.0330	(-1.17)		
Maturity $_{it}$	0.00391	(0.99)	0.00232	(0.70)		
NonEur $_{it}$	-1.005**	(-2.22)	-1.437***	(-5.35)		
*m_h *T_t*	-0.0326	(-0.61)				
*Spread $_{it}$ *m_h*	0.0359*	(1.68)	0.00739	(0.63)	0.0230	(1.15)
*Maturity $_{it}$ *m_h*	-0.00111	(-1.23)	-0.000928	(-1.24)	-0.0000211	(-0.07)
*NonEur $_{it}$ *m_h*	0.0432	(0.40)	0.0111	(0.20)	0.0525	(0.48)
*Spread $_{it}$ *Tt*	-0.262**	(-2.34)	-0.256*	(-1.91)		
*Maturity $_{it}$ *Tt*	-0.000737	(-0.75)	0.0000207	(0.03)		
*NonEurit*Tt*	0.384*	(1.82)	0.673***	(2.84)		
*Spread $_{it}$ *m_h *Tt*	0.0529**	(2.31)	0.0571*	(1.87)	0.0435*	(1.83)
*Maturity $_{it}$ *m_h *Tt*	0.000179	(0.72)	0.0000614	(0.41)	-0.0000783	(-0.58)
*NonEur $_{it}$ *m_h *Tt*	-0.0551	(-0.84)	-0.110*	(-1.86)	-0.109**	(-2.16)
holder*time f.e.	No		Yes		Yes	
security f.e.	No		No		Yes	
N	50374		50370		40209	
R^2	0.058		0.286		0.626	

Notes: Dep. variable is log of the amounts of security i held by sector h (a given institutional sector in a given country), in period t. Data for 2014Q1 and 2015Q2. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. Spreadit is the difference at time t between the yield-to-maturity of security i *and the risk-free benchmark rate of a corresponding maturity, in percent. Maturity $_{it}$ is the residual maturity of security i* at time *t, in months. NonEur $_{it}$ is a dummy for securities denominated in currenciues other than the euro. mh is the change in valuation of the portfolio of sector h in 2014 Q1, in percent. Tt is a dummy for the period 2015 Q2. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.*
** p<0.10 ** p<0.05 *** p<0.01*

Table 7
Estimations on full portfolios (including newly issued and seasoned securities)

	Full sample			Investors in vulnerable countries		
	(1)	(2)	(3)	(4)	(5)	(6)
yield-to-maturity (r_{it})	-0.0733**	-0.0695***		0.0167	-0.00149	
	(-2.58)	(-2.77)		(0.38)	(-0.11)	
portfolio valuation (m_h)	-0.0802			0.0556		
	(-1.31)			(0.77)		
post-APP period dummy (T_t)	0.184			0.307*		
	(1.55)			(1.80)		
$r_{it} * m_h$	-0.0192**	-0.0219***	0.0139	-0.0409***	-0.0305***	0.0406***
	(-2.01)	(-2.74)	(1.27)	(-2.82)	(-6.87)	(3.94)
$r_{it} * T_t$	-0.0966*	-0.124***		-0.149***	-0.151**	-456.7
	(-1.77)	(-2.83)		(-2.72)	(-2.39)	(-0.00)
$m_h * T_t$	-0.0326			-0.0115		
	(-1.41)			(-0.33)		
$r_{it} * m_h * T_t$	0.0146	0.0213**	0.000476	0.0297**	0.0326*	-0.00772
	(1.24)	(2.16)	(0.12)	(2.29)	(1.96)	(-1.60)
holder*time f.e.	No	Yes	Yes	No	Yes	Yes
security*time f.e.	No	No	Yes	No	No	Yes
N	957680	957677	800033	249374	249372	190264
R^2	0.037	0.226	0.509	0.020	0.182	0.590

Notes : Dependent variable is log of the amounts of security i held by sector h *(* a given institutional sector in a given country), in period t . Data for 2014Q1 and 2015Q2. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. r_{it} is the yield-to-maturity of the corresponding securitiy, in percent. m_h is the change in valuation of the portfolio of sector h in 2014 Q1, in percent. T_t is a dummy for the period 2015 Q2. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.
* p<0.10 ** p<0.05 *** p<0.01

Table 8
Investors in vulnerable countries; extensive margin

	(1)	(2)	(3)	(4)
yield-to-maturity (r_{it})	0.00886*		0.0105***	
	(2.40)		(3.17)	
portfolio valuation (m_h)	-0.00176	-0.00354		
	(-0.44)	(0.74)		
$r_{it} * m_h$	-0.00101	-0.00141	-0.000412	-0.0000575
	(-0.76)	(-0.95)	(-0.41)	(-0.06)
pseudo-security f.e.	No	Yes	No	Yes
holder f.e.	No	No	Yes	Yes
N	15179	14956	15179	14956
R^2	0.002	0.326	0.074	0.44

Notes : The sample is restricted to securities held in 2015Q2. The dependent variable identifies new holdings, i.e. conditional on being held in 2015Q2, securities which were not also held in 2014Q1, for each sector *h (* a given institutional sector in a given country). The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. The dataset is rectangularised in order to account for the fact that non-reported holdings actually represent zero holdings. In order to keep the number of observations manegable, securities are grouped into around 2300 categories according to issuer sector, issuer country, maturity, coupon type, nominal currency and rating. r_{it} is the yield-to-maturity of the corresponding securitiy, in percent. m_h is the change in valuation of the portfolio of sector *h* in 2014 Q1, in percent. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.
* p<0.10 ** p<0.05 *** p<0.01

Table 9

Investors in vulnerable countries; intensive margin (*rectangularised* dataset)

	Full sample				Investors in vulnerable countries			
	(1)	(2)	(3)	(5)	(5)	(6)	(7)	(8)
yield-to-maturity (r_{it})	-0.138***	-0.130***			-0.0955***	-0.0701***		
	(-6.13)	(-6.96)			(-4.46)	(-4.30)		
portfolio valuation (m_h)	0.0503*				0.0944***			
	-1.68				-4.21			
post-APP period dummy (T_t)	0.118***				0.11			
	-3.01				-1.6			
$r_{it}*m_h$	-0.00275	-0.0048	-0.00501	-0.00328	0.0006	-0.00640*	-0.00658	-0.00574
	(-0.74)	(-1.57)	(-1.50)	(-1.20)	-0.09	(-1.70)	(-1.53)	(-0.83)
$r_{it}*T_t$	-0.0496***	-0.0263**			-0.0365**	-0.0151		
	(-4.68)	(-2.56)			(-2.35)	(-1.31)		
m_h*T_t	-0.012				-0.0386*			
	(-1.38)				(-1.71)			
$r_{it}*m_h*T_t$	0.00414*	0.00153	0.0013	0.00510**	0.0117**	0.00397	0.003	0.00789**
	(1.91)	(0.75)	(0.57)	(2.40)	(2.47)	(1.37)	(0.94)	(2.58)
holder*time f.e.	No	Yes	Yes	Yes	No	Yes	Yes	Yes
pseudo-security*time f.e.	No	No	Yes	Yes	No	No	Yes	Yes
holder*issuer f.e.	No	No	No	Yes	No	No	No	Yes
N	103402	103400	102957	74294	30817	30816	30340	21094
R^2	0.022	0.334	0.525	0.93	0.041	0.314	0.545	0.928

Notes: Dependent variable is log of the amounts of security i held by sector h (a given institutional sector in a given country), in period t. Data for 2014Q1 and 2015Q2. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. The dataset is rectangularised in order to account for the fact that non-reported holdings actually represent zero holdings. In order to keep the number of observations manegable, securities are grouped into around 2300 categories according to issuer sector, issuer country, maturity, coupon type, nominal currency and rating. rit is the yield-to-maturity of the corresponding securitiy, in percent. mh is the change in valuation of the portfolio of sector h in 2014 Q1, in percent. Tt is a dummy for the period 2015 Q2. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.
** p<0.10 ** p<0.05 *** p<0.01*

Table 10

Portfolio valuation and credit growth

	(1)	(2)	(3)	(4)
portfolio valuation (m_h)	1.633**	2.335**	2.797***	3.527***
	(2.75)	(2.68)	(4.03)	(3.57)
*m_h*Loans to Non Financial Corporations*		-1.405		-1.460
		(-1.04)		(-0.92)
*m_h*Vulnerable countries*			-3.262***	-3.429***
			(-3.64)	(-3.72)
*m_h*L_{NFC}*Vulnerable countries*				0.335
				(0.17)
sector f.e.	Yes	Yes	Yes	Yes
country f.e.	Yes	Yes	Yes	Yes
N	50	50	50	50
R^2	0.402	0.422	0.463	0.483

Notes : Dependent variable is y-o-y growth of loans to households and to non-financial corporations granted by bank h in 2015Q2. The interaction term *Loans to non-financial corporations* is a dummy-variable identifying observations for this sector (so that households becomes the baseline). m_h is the change in valuation between 2015Q2 and 2014Q1 of the portfolio held by bank *h* in 2014 Q1, in percent. In all specifications errors are clustered at the bank level. t-statistics in parentheses.

* p<0.10 ** p<0.05 *** p<0.01

Table 11

Portfolio valuation and interest rates on loans

	(1)	(2)	(3)	(4)	(5)	(6)
portfolio valuation (m_h)	0.034	-0.250*	-0.250*	0.016	-0.271***	-0.271**
	(0.72)	(-1.77)	(-1.75)	(0.40)	(-2.81)	(-2.75)
*m_h*Loans to Non Financial Corporations*		0.378**			0.383***	
		(2.46)			(3.13)	
*m_h*Loans to NFC up to € 0.25 million*			0.324*			0.320**
			(1.92)			(2.35)
*m_h*Loans to NFC above € 0.25 and up to € 1 million*			0.378**			0.413***
			(2.28)			(2.95)
*m_h*Loans to NFC above € 1 million*			0.433***			0.414***
			(3.02)			(3.70)
*m_h*Vulnerable countries*				0.05	0.071	0.071
				(0.44)	(0.24)	(0.23)
*m_h*Vulnerable countries*L_{NFC}*					-0.027	
					(-0.09)	
*m_h*Vulnerable countries*L_{NFC} up to € 0.25 million*						0.021
						(0.05)
*m_h*Vulnerable countries*L_{NFC} above € 0.25 and up to € 1 million*						-0.217
						(-0.63)
*m_h*Vulnerable countries*L_{NFC} above to € 1 million*						0.115
						(0.44)
sector f.e.	Yes	Yes	Yes	Yes	Yes	Yes
country f.e.	Yes	Yes	Yes	Yes	Yes	Yes
N	100	100	100	100	100	100
R^2	0.315	0.455	0.463	0.317	0.457	0.483

Notes : Dependent variable is the change in the interest rates on new loans applied by bank h between 2015Q2 and 2014Q1. m_h is the change in valuation between 2015Q2 and 2014Q1 of the portfolio held by bank *h* in 2014 Q1, in percent. In all specifications errors are clustered at the bank level. t-statistics in parentheses.

* p<0.10 ** p<0.05 *** p<0.01

Appendix

Table A1

Baseline estimation on newly issued securities before the APP

	Full sample			Investors in vulnerable countries		
	(1)	(2)	(3)	(4)	(5)	(6)
yield-to-maturity (r_{it})	-0.0504	-0.0647**		-0.0998*	-0.0928***	
	(-1.21)	(-2.29)		(-1.99)	(-4.15)	
portfolio valuation (m_h)	-0.120*			0.0506		
	(-1.87)			-0.69		
post-APP period dummy (T_t)	0.0452			0.000143		
	-0.69			0		
$r_{it}*m_h$	-0.0134	-0.0108	0.0122	0.021	0.0175*	0.0281**
	(-0.83)	(-1.03)	-1.12	-1.09	-1.84	-2.29
$r_{it}*T_t$	0.00554	0.0221**		0.0164	0.0389***	
	-0.38	-2.27		-0.94	-3.49	
m_h*T_t	0.00452			0.0466*		
	-0.31			-1.71		
$r_{it}*m_h*T_t$	-0.00692	-0.00847***	0.00208	-0.00744	-0.0159***	0.0118
	(-1.52)	(-3.14)	-0.56	(-1.28)	(-4.83)	-1.09
holder*time f.e.	No	Yes	Yes	No	Yes	Yes
security f.e.	No	No	Yes	No	No	Yes
N	216898	216887	172160	49980	49975	39843
R^2	0.034	0.288	0.546	0.011	0.235	0.642

Notes : Dependent variable is log of the amounts of security *i* held by sector *h (* a given institutional sector in a given country), in period *t* . Data for 2013Q4 and 2014Q1. Only holdings of newly issued securities, defined as those issued in the preceeding 4 quarters. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. r_{it} is the yield-to-maturity of the corresponding securitiy, in percent. m_h is the change between 2014 Q1 and 2015 Q2 in valuation of the portfolio held by sector h in 2014 Q1, in percent. T_t is a dummy for the period 2014 Q1. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.

* p<0.10 ** p<0.05 *** p<0.01

Table A2

Estimations on full portfolios (including newly issued and seasoned securities) before the APP

	Full sample			Investors in vulnerable countries		
	(1)	(2)	(3)	(4)	(5)	(6)
yield-to-maturity (r_{it})	-0.0559**	-0.0700***		0.0175	-0.0239	
	(-2.32)	(-3.16)		-0.46	(-1.37)	
portfolio valuation (m_h)	-0.0713			0.0184		
	(-1.24)			-0.3		
post-APP period dummy (T_t)	0.0104			-0.0343		
	-0.36			(-0.74)		
$r_{it} * m_h$	-0.0197**	-0.0174**	0.0103	-0.0384***	-0.0195***	0.0210*
	(-2.50)	(-2.58)	-1.08	(-2.93)	(-3.53)	-1.8
$r_{it} * T_t$	0.00241	0.0163***		0.00261	0.0220**	
	-0.31	-2.93		-0.18	-2.31	
$m_h * T_t$	0.000872			0.0343*		
	-0.14			-1.99		
$r_{it} * m_h * T_t$	-0.00135	-0.00481***	0.001	0.00116	-0.00607**	0.0021
	(-0.53)	(-3.16)	-0.46	-0.25	(-2.13)	-0.53
holder*time f.e.	No	Yes	Yes	No	Yes	Yes
security f.e.	No	No	Yes	No	No	Yes
N	894714	894709	745334	243120	243117	183738
R^2	0.029	0.214	0.507	0.015	0.181	0.592

Notes : Dependent variable is log of the amounts of security *i* held by sector *h (* a given institutional sector in a given country), in period *t* . Data for 2013Q4 and 2014Q1. The term "vulnerable countries" refers to Ireland, Greece, Spain, Italy, Cyprus, Portugal and Slovenia. r_{it} is the yield-to-maturity of the corresponding securitiy, in percent. m_h is the change between 2014 Q1 and 2015 Q2 in valuation of the portfolio held by sector h in 2014 Q1, in percent. T_t is a dummy for the period 2014 Q1. In all specifications errors are clustered at the holding-sector level. t-statistics in parentheses.

* p<0.10 ** p<0.05 *** p<0.01

Table A3

Portfolio valuation and credit growth before the APP

	(1)	(2)	(3)	(4)
portfolio valuation (m_h)	1.479	1.766	1.537	1.866
	(1.32)	(1.15)	(0.88)	(0.87)
*m_h*Loans to Non Financial Corporations*		-0.573		-0.659
		(-0.42)		(-0.53)
*m_h*Vulnerable countries*			-0.163	-0.422
			(-0.09)	(-0.18)
*m_h*L_{NFC}*Vulnerable countries*				0.519
				(0.19)
sector f.e.	Yes	Yes	Yes	Yes
country f.e.	Yes	Yes	Yes	Yes
N	50	50	50	50
R^2	0.270	0.272	0.270	0.272

Notes : Dependent variable is y-o-y growth of loans to households and to non-financial corporations granted by bank h in 2014Q1. The interaction term *Loans to non-financial corporations* is a dummy-variable identifying observations for this sector (so that households becomes the baseline). m_h is the change in valuation between 2015Q2 and 2014Q1 of the portfolio held by bank *h* in 2014 Q1, in percent. In all specifications errors are clustered at the bank level. t-statistics in parentheses.
* p<0.10 ** p<0.05 *** p<0.01

Table A4

Portfolio valuation and interest rates on loans before the APP

	(1)	(2)	(3)	(4)	(5)	(6)
portfolio valuation (m_h)	-0.0217	-0.0631	-0.0631	-0.0366	-0.0647	-0.0647
	(-0.48)	(-0.53)	(-0.52)	(-0.53)	(-0.45)	(-0.44)
*m_h*Loans to Non Financial Corporations*		0.0552			0.0374	
		(0.50)			(0.33)	
*m_h*Loans to NFC up to € 0.25 million*			0.0903			0.0850
			(0.86)			(0.77)
*m_h*Loans to NFC above € 0.25 and up to € 1 million*			0.0807			0.0698
			(0.66)			(0.55)
*m_h*Loans to NFC above € 1 million*			-0.00535			-0.0425
			(-0.05)			(-0.35)
*m_h*Vulnerable countries*				0.0418	-0.0393	-0.0393
				(0.54)	(-0.34)	(-0.34)
*m_h*Vulnerable countries*L_{NFC}*					0.108	
					(0.79)	
*m_h*Vulnerable countries*L_{NFC} up to € 0.25 million*						0.0321
						(0.31)
*m_h*Vulnerable countries*L_{NFC} above € 0.25 and up to € 1 million*						0.0663
						(0.54)
*m_h*Vulnerable countries*L_{NFC} above to € 1 million*						0.226
						(0.69)
sector f.e.	Yes	Yes	Yes	Yes	Yes	Yes
country f.e.	Yes	Yes	Yes	Yes	Yes	Yes
N	100	100	100	100	100	100
R^2	0.118	0.121	0.127	0.119	0.124	0.138

Notes : Dependent variable is the change in the interest rates on new loans applied by bank h between 2014Q1 and 2013Q4. m_h is the change in valuation between 2015Q2 and 2014Q1 of the portfolio held by bank *h* in 2014 Q1, in percent. In all specifications errors are clustered at the bank level. t-statistics in parentheses.

* p<0.10 ** p<0.05 *** p<0.01

Epilogue

This dissertation provides some insights on key aspects of bank behaviour and, more generally, provision of credit to the real economy.

The main findings of the first chapter are that the voluntary capital buffers held by banks in excess of the regulatory minimum are positively influenced by several broad risk measures, whereas provisions and high and stable profitability are found to be substitutes for capital buffers. Larger banks seem to hold less excess capital and capital buffers tend to be lower in times of stronger economic growth.

The analysis in the second chapter concludes that the Portuguese banking system has experienced weak competition between 1991 and 1996, underwent a period of restructuring until 2000, and from then until 2004 behaved consistently with perfect competition. Both privately owned and, more markedly, domestic banks, seem to have competed more aggressively on occasions.

The third chapter concludes that technological progress has shifted Portuguese banks' cost frontier downwards between 1992 and 2006, whereas the distance at which banks have operated from the frontier seems to have remained constant. Increases in production under scale economies have also contributed to the recorded increase in productivity.

The results of the fourth chapter suggest that portfolio rebalancing has been an active channel of transmission of the ECB's expanded Asset Purchase Programme (APP), though with important differences across countries. In countries which were more directly affected by the recent financial crisis there is evidence that the APP led to a rebalancing of portfolios towards riskier securities, in particular impacting securities issued by corporates (as opposed to sovereigns) and resulting in increased credit risk-taking (as opposed to maturity or currency risk-taking). However, evidence of effects on bank lending is limited to other countries. One possible explanation for these geographical patterns is that in less vulnerable countries spreads were already so compressed that reaching a significant increase in the average yield of a given portfolio would an unfeasibly large change in its composition. This could also explain why in these economies some rebalancing of banks towards real-sector loans (where presumably spreads are still positive) is found.