City Research Online

City, University of London Institutional Repository

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Asset encumbrance, size distribution and liquidity provision: Three essays on banking

Enrique Benito Ruiz

Supervisor:

Albert Banal-Estanol

(City, University of London)

A thesis submitted for the degree of Doctor of Philosophy in Economics

Department of Economics

City, University of London

March 2018

This page is intentionally left blank.

Table of contents

This page is intentionally left blank.

List of tables

This page is intentionally left blank.

List of figures

This page is intentionally left blank.

Declaration

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

I declare that the first paper included in the main body of the thesis, 'Asset encumbrance and bank risk: First evidence from public disclosures in Europe', is co-authored with my PhD supervisor, Dr. Albert Banal-Estanol and Dmitry Kametshin.

This page is intentionally left blank.

Acknowledgements

A mis padres

Abstract

This thesis presents three papers in the field of banking.

The first paper considers 'asset encumbrance' which refers to the existence of bank balance sheet assets being subject to arrangements that restrict the bank's ability to freely transfer or realise them. Asset encumbrance has recently become a much discussed subject and policymakers have been actively addressing what some consider to be excessive levels of asset encumbrance. Despite its importance, the phenomenon of asset encumbrance remains poorly understood. I build a novel dataset of asset encumbrance metrics based on information provided in the banks' public disclosures for the very first time throughout 2015. The study then provides descriptive evidence of asset encumbrance levels by country, credit quality, and business model using different encumbrance metrics. The empirical results point to the existence of an association between CDS premia and asset encumbrance that is negative, not positive. That is, on average encumbrance is perceived to be beneficial. Still, certain bank-level variables play a mediating role in this relationship. For banks that have high exposures to the central bank, high leverage ratio, and/or are located in southern Europe, asset encumbrance is less beneficial and could even be detrimental in absolute terms.

The second paper investigates the size distribution of the whole population of Spanish commercial, savings and cooperative banks from a dynamic perspective over the 1970-2006 period. To investigate the evolution of the size distribution, the study determines whether the data is in line with the Law of Proportionate Effect (LPE) using panel unit root tests. A key finding is that the size-growth relationship is not stable over time but changes depending on the competitive environment of banks (liberalization, deregulation and integration). When Spanish banking was highly regulated we find that smaller banks grew faster than their larger counterparts. In recent years, however, we find that larger banks grow at the same rate or faster than smaller banks, a result that lend towards LPE acceptance. Thus, the study corroborates the conditioned nature of the size-growth relationship and the size distribution of banks, as emphasized by studies of the US banking system.

Finally, the third paper investigates, from a theoretical perspective, the roles of banks and markets when both are active, there is limited participation in markets, and there exists liquidity and technology risk in the economy. In a model where banks and markets co-exist and banks are subject to runs, we show that the levels of aggregate risk and limited participation jointly determine the superiority of the mixed (market and bank deposits) economy over pure equity contracts. The study finds that if aggregate risk exceeds a certain threshold then markets may perform better than banks even for low or null levels of market participation, and it is shown that markets may perform better than banks the lower the market participation under some circumstances. The results imply that the level of bank risk taking cannot be considered in isolation, but in conjunction with the availability and access of banking and non-banking options.

Chapter 1: Introduction

The financial turmoil that began in 2007 developed into an unprecedented crisis, evidencing the vulnerabilities of financial institutions across the world and leading to a profound transformation of the banking industry. This thesis presents an investigation of three areas of study arising from such crisis and tries to shed light on related research gaps. The main body of the thesis is developed in chapters 2, 3 and 4, each one presenting a paper.

The second chapter contains the first paper which investigates asset encumbrance. The financial crisis led to an increased perception of counterparty credit risk. Unsecured wholesale funding diminished in size, whereas secured funding and, more generally, the use of assets pledged as collateral in transactions, known as asset encumbrance, rose significantly (IMF, 2013).

Increasing the proportion of secured funding, which carries lower rollover risks and is generally cheaper than equivalent unsecured funding, could translate into an increased capacity of debt repayment and a lower probability of default. But as highlighted by an incipient theoretical literature, higher asset encumbrance would result in a reduction of the assets that become available to unsecured creditors under insolvency, an effect coined as "structural subordination" (Bank of England, 2012; CGFS, 2013; Houben and Slingenberg, 2013; IMF, 2013; Juks, 2012; Le Leslé, 2012). The same seniority that secured creditors enjoy means that as more secured debt is issued, balance sheet shocks are asymmetrically concentrated on unsecured creditors, exacerbating the possibility of a run of unsecured creditors (Anhert et al., 2016; Matta and Perotti, 2015) and potentially resulting in increased bank funding costs.

Policymakers are acting decisively in order to address what some consider to be excessive levels of asset encumbrance. Some jurisdictions have introduced limits on the level of encumbrance (Australia, New Zealand) or ceilings on the amount of secured funding or covered bonds (Canada, US), while others have incorporated encumbrance levels in deposit insurance premiums (Canada). Several authors have proposed linking capital requirements to the banks' asset encumbrance levels or establishing further limits to asset encumbrance as a back-stop (Helberg and Lindset, 2014; IMF, 2013; Juks, 2012). As part

of the Basel III regulatory package, the Net Stable Funding Ratio (NSFR), an additional minimum liquidity requirement of the LCR will be introduced in 2018. The NSFR heavily penalises asset encumbrance by requiring substantial amounts of stable funding to finance encumbered assets. In Europe, regulatory reporting and disclosure requirements have been introduced and all institutions are required to incorporate asset encumbrance within their risk management frameworks. The Dutch National Bank has even committed to 'keeping encumbrance to a minimum' (De Nederlandsche Bank, 2016).

Despite the importance of asset encumbrance, the phenomenon remains poorly understood. There is not even a consensus as to how asset encumbrance should be measured, and there is limited knowledge of how asset encumbrance varies across countries or bank business models. Surprisingly, the relationship between bank risk and asset encumbrance remains unexplored empirically.

The study investigates asset encumbrance across three facets. First, it defines asset encumbrance, describes how assets become encumbered and reviews the sources of asset encumbrance. Second, it provides descriptive evidence of asset encumbrance levels by country, bank credit quality, size and business model using different encumbrance metrics. To do so, a novel dataset is built using information provided in the asset encumbrance disclosures published for the first time throughout 2015 by European banks, following a set of harmonised definitions provided by the EBA (2014). Finally, it investigates the association between bank risk and the levels of asset encumbrance empirically. In line with recent studies, the relationship between bank balance sheet ratios, based on capital adequacy, liquidity, asset quality and earnings potential (CAMEL) indicators, typically used in supervisory rating systems to classify a bank's overall condition, and implied five-year CDS spreads is estimated. It is then considered the extent to which asset encumbrance contributes to the explanatory power of such models.

The findings show that banks with higher encumbrance levels present lower CDS spreads across all three metrics of asset encumbrance considered – i.e. bank risk seems to be negatively associated with asset encumbrance. In addition, ratios measuring asset encumbrance or asset quality provide more valuable information on bank risk than capital and liquidity ratios. This result is consistent with recent literature pointing to limited reliance by markets on capital and liquidity ratios to account for overall bank risk (Ötker-

Robe and Podpiera, 2010; Chiaramonte and Casu, 2013; Hasan et al., 2015; Kanagaretnam et al., 2016).

CAMEL and other bank-level variables seem to play a mediating role in the relationship between asset encumbrance and bank risk. For banks with a high reliance on central bank funding and high levels of liquid assets or with a high leverage ratio and high levels of impaired loans, or for banks located in Southern Europe (GIIPS), asset encumbrance is less beneficial and could even be detrimental in absolute terms. Banks with high levels of loan loss provisions and Nordic banks, on the contrary, could further benefit from increasing their levels of asset encumbrance. These findings imply that regulators need to be cautious when assessing asset encumbrance levels and leaping to across-the-board conclusions about its effects.

The second paper, in the third chapter, investigates the size distribution of financial institutions. The European banking industry has experienced significant changes during the last decades. The liberalization and deregulation processes during the 1980's and the 1990's, the technological progress, the introduction of a single currency and the consolidation process experienced by the industry both before and after the financial crisis all have led to a progressive decline in the number of institutions and increased concentration of banking assets among the largest banks. After the financial crisis, the concentration of the banking industry has accelerated even more significantly, leading to a progressive decline in the number of institutions and increased concentration of banking assets among the largest banks. According a study from the European Central Bank (ECB), the number of credit institutions in the euro area declining to 5,475 at the end of 2015, from 6,767 at the end of 2008 (ECB, 2016).

Although these changes have a significant impact on the size distribution and dynamics of the banking industry, very few studies have examined the dynamic nature of the size distribution in banking. This is true even though the size distribution of banks is of fundamental interest to bankers and policy makers given size is a key determinant of bank's risk taking (Demsetz and Strahan, 1997), credit availability (Peek and Rosengren, 1996), lending relationships (Stein 2002; Berger et al., 2005), or lending specialization (Delgado et al., 2007). The existing empirical evidence on the size distribution banks is also not exempted from limitations, with the vast majority of the existing studies covering

a relatively small proportion of the banking system or very short periods of time. Sample censoring and selection bias is also a common issue with these type of studies (Sutton, 1997).

The paper investigates the size distribution of the whole population of Spanish banks from a dynamic perspective over the 1970-2006 period. The time period considered offers a unique opportunity to study the evolution of the size distribution from a period when Spanish banking was strongly regulated (with administratively fixed interest rates, compulsory investment coefficients, foreign entry restrictions and high asymmetry among different ownership forms), to a period in which competitive conditions press towards efficient decisions and savings banks outperform commercial banks in competitive credit markets.

The study focus on the Law of Proportionate Effect (LPE). The approach used is based on Tschoegl (1983), who investigates the relationship between size and growth rate of large international banks by testing three propositions that all together make up a 'strong form' of the LPE: i) the growth rate of each bank is independent of its size; ii) there is no persistence in bank's growth in two consecutive periods; and iii) the variability of growth rates is independent of the banks' size.

Evidence points to smaller banks growing faster than large banks when the banking system was highly protected. As deregulation and integration processes occur, however, larger banks tend to grow at a faster pace. In the final years of our sample period, however, we find some evidence that larger banks tend to grow at higher rates. Thus, the study corroborates the conditional nature of the size-growth relationship and the size distribution.

The fourth chapter includes the third paper which is a theoretical study looking at the liquidity provisioning role of financial markets and intermediaries when both are active, there is limited participation in markets, and there exists aggregate technology risk in the economy.

Comparing the role of market-based and bank-based economies is especially relevant in the current setting. It is well known that the US economy experienced a rapid recovery from the financial crisis compared to most European countries despite the initial shock

being similar in size and nature and several commentators have pointed to an overreliance on bank intermediaries in Europe as a key reason for such striking difference.

The role of market-based and financially intermediated systems as liquidity providers and the related underlying mechanisms that make banks vulnerable to a loss of depositors' confidence have been examined extensively in the theoretical literature. The existing studies however, do not consider the existence of bank runs, or the potential impact of market participation when there is limited participation in markets and aggregate risk in the economy, thus providing an incomplete picture when undertaking welfare comparisons.

The model setup is similar in spirit to Diamond (1997), with differences in the manner the different technologies and information structures around them are modelled. In this new setting, there are two technologies available, a short-term safe one and a long-term risky one, and we classify agent types according to three different groups whose aggregate size is known at date 0, but not its composition. Therefore, ex-ante people do not know when they need to consume or whether they are able to participate in the market. Those who will not participate in the market are a subset of the patient agents. Agents learn their type at date 1, when a market opens for trade. Information about the future returns of the risky technology is, initially, only available to participating agents.

The performance of a market-based economy is investigated first. Participating individuals will trade short-term for long-term claims at date 1 whereas non-participating individuals will reinvest any holdings of short-term claims, resulting in inefficient re-investment of liquid assets and overall underinvestment in long-term risky assets (and consequently lower prices of long-term claims) compared to a full participation situation. A bank offering a deposit contract however can aggregate wealth to avoid inefficient reinvestment and provide risk-sharing against the uncertainty in depositors' liquidity needs at the same time. In addition the bank can trade claims in the market at date 1, increasing liquidity provision and normalising market prices.

The study shows that bank runs may occur if a subset of non-participating agents receive information about the future return of the underlying assets. Depending on the number of agents that receive information and decide to withdraw early, how small is the return, and

the amount of excess liquidity that the bank holds at date 1, the bank may be forced to undertake fire sales of long-term assets, depressing the market price and reducing overall consumption. The negative impact on welfare of bank runs may however, not be sufficient to tilt the scale in markets favour. I argue that the level of the underlying aggregate technology risk and of limited market participation jointly determine the superiority of the mixed economy over pure equity contracts.

The observation that deposit contracts that are subject to bank runs tend to be better for financing low-risk assets is consistent with previous literature. The results have important policy implications. Recent regulatory reforms have focused on constraining the risk undertaken by, for example, imposing new capital and liquidity requirements. These findings, however, imply that such efforts should be accompanied by further action aimed at improving market access and participation.

Chapter 2: Asset encumbrance and bank risk: First evidence from public disclosures in Europe

2.1 Introduction

As of June 2011, Dexia, a Franco-Belgian bank, reported a strong Tier 1 Capital Ratio of 11.4%.[1] Out of the 91 institutions analysed in the European Banking Authority (EBA) stress tests, Dexia came joint 12th, with a forecast Core Tier 1 capital ratio of 10.4% under the adverse stress scenario.[2] From a liquidity standpoint, the bank had built up a buffer of €88bn in liquid securities, had decreased short-term funding needs by €47bn and its short-term ratings had been reaffirmed as investment grade by the main credit rating agencies. But just three months later, in October 2011, Dexia was partly nationalised by the Belgian and French governments. Several commentators highlighted the high levels of "encumbered" assets as the key factor precipitating its move into government arms.[3,4]

Asset encumbrance refers to the existence of financial bank balance sheet assets being subject to arrangements that restrict the bank's ability to freely transfer or realise them. Bank assets become encumbered when these are used as collateral to raise funding, for example in repurchase agreements (repos) or in other collateralised transactions such as asset-backed securitisations, covered bonds, or derivatives.[5] In the particular case of Dexia, more than €66bn of its €88bn buffer securities were encumbered through different secured funding arrangements, particularly with the European Central Bank (ECB), and were therefore unavailable for obtaining emergency funding.

Policymakers are acting decisively in order to address what some consider to be excessive levels of asset encumbrance. Some jurisdictions have introduced limits on the level of

---

[1] See Dexia 2Q & 1H 2011 Results and Business Highlights Presentation, 4 August 2011.

[2] The Core Tier 1 ratio represents the ratio of very high quality capital (shareholders' capital and reserves) to risk-weighted assets (RWA). The Tier 1 capital ratio includes, in addition to Core Tier 1 capital, other perpetual capital resources such as subordinated debt instruments with conversion features and is also expressed as a fraction of RWA.

[3] See e.g. Financial Times, 'Bank collateral drying up in rush for security', October 2011.

[4] More recently, in June 2017, Banco Popular was put into resolution by the European Single Supervisory Mechanism (SSM) and was acquired by Banco Santander for a symbolic amount of €1. Yet, as of year-end 2016, the Spanish bank Banco Popular reported a Tier 1 capital ratio of 12.3% and had passed the EBA stress tests undertaken in 2016 with a solid margin. However, nearly 40% of its total balance sheet assets were encumbered as of December 2016.

[5] Collateralisation is a common method of mitigating counterparty credit risk in derivative markets through the provisioning of margin.

encumbrance (Australia, New Zealand) or ceilings on the amount of secured funding or covered bonds (Canada, US), while others have incorporated encumbrance levels in deposit insurance premiums (Canada). Several authors have proposed linking capital requirements to the banks' asset encumbrance levels or establishing further limits to asset encumbrance as a back-stop (Helberg and Lindset, 2014; IMF, 2013; Juks, 2012). As part of the Basel III regulatory package, the Net Stable Funding Ratio (NSFR), an additional minimum liquidity requirement of the LCR will be introduced in 2018. The NSFR heavily penalises asset encumbrance by requiring substantial amounts of stable funding to finance encumbered assets. In Europe, regulatory reporting and disclosure requirements have been introduced and all institutions are required to incorporate asset encumbrance within their risk management frameworks. The Dutch National Bank has even committed to *'keeping encumbrance to a minimum'* (De Nederlandsche Bank, 2016).

Despite the importance of asset encumbrance, the phenomenon remains poorly understood. There is not even a consensus as to how asset encumbrance should be measured, and there is limited knowledge of how asset encumbrance varies across countries or bank business models. The relationship between bank risk and the overall levels of asset encumbrance remains unexplored empirically. As noted by Juks (2012), measuring the relevance of the overall asset encumbrance for a bank is far from easy due to data limitations and that asset encumbrance varies across banks and sources of asset encumbrance.[6]

As highlighted by an incipient theoretical literature and policy papers, higher asset encumbrance could result in a reduction of the assets that become available to unsecured creditors under insolvency, an effect coined as "structural subordination" (Bank of England, 2012; CGFS, 2013; Houben and Slingenberg, 2013; IMF, 2013; Juks, 2012; Le Leslé, 2012). This may be priced in by unsecured creditors, potentially increasing overall funding costs. On the other hand, increasing the proportion of secured funding, which carries lower rollover risks and is generally cheaper than equivalent unsecured funding,

---

[6] Overall asset encumbrance refers to encumbrance arising from multiple sources of asset encumbrance, namely the liabilities or obligations that give rise to encumbered assets. A typical bank will have encumbered assets from several sources and we discuss some of the most common sources of asset encumbrance in section 2.2.

could translate into an increased capacity of debt repayment and a lower probability of default.

This chapter tries to shed some light on these issues in three steps. First, we define asset encumbrance, describe how assets become encumbered and review the sources of asset encumbrance. Second, we provide descriptive evidence of asset encumbrance levels by country, bank credit quality, size and business model using different encumbrance metrics. To do so, we build a novel dataset using information provided in the asset encumbrance disclosures published for the first time throughout 2015 by European banks, following a set of harmonised definitions provided by the EBA (2014). Finally, we investigate the association between bank risk and the levels of asset encumbrance empirically. In line with recent studies, we estimate the relationship between bank balance sheet ratios, based on capital adequacy, liquidity, asset quality and earnings potential (CAMEL) indicators, typically used in supervisory rating systems to classify a bank's overall condition, and implied five-year CDS spreads. We then consider the extent to which asset encumbrance contributes to the explanatory power of such models. Our analysis provides, to the best of our knowledge, the first empirical investigation of the relationship between banks' overall asset encumbrance and bank risk.

Our findings show that banks with higher encumbrance levels present lower CDS spreads across all three metrics of asset encumbrance considered – i.e. bank risk seems to be negatively associated with asset encumbrance. In addition, and consistently with the demise of Dexia, we find that ratios measuring asset encumbrance or asset quality provide more valuable information on bank risk than capital and liquidity ratios. This result is consistent with recent literature pointing to limited reliance by markets on capital and liquidity ratios to account for overall bank risk (Ötker-Robe and Podpiera, 2010; Chiaramonte and Casu, 2013; Hasan et al., 2015; Kanagaretnam et al., 2016).

We find that CAMEL and other bank-level variables play a mediating role in the relationship between asset encumbrance and bank risk. For banks with a high reliance on central bank funding and high levels of liquid assets, such as Dexia, or with a high leverage ratio and high levels of impaired loans, such as Banco Popular, or for banks located in Southern Europe (GIIPS), asset encumbrance is less beneficial and could even

be detrimental in absolute terms.[7] Banks with high levels of loan loss provisions and Nordic banks, on the contrary, could further benefit from increasing their levels of asset encumbrance. Banks in Nordic countries rely to a large extent on covered bonds, which are perceived as a very safe investment and source of funding. These findings imply that regulators need to be cautious when assessing asset encumbrance levels and leaping to across-the-board conclusions about its effects.

The remainder of this chapter is structured as follows. Section 2.2 defines asset encumbrance and explains its sources. Section 2.3 explores the benefits and risks of asset encumbrance from the perspective of both secured and unsecured creditors. Section 2.4 presents the methodology used and the data. Section 2.5 presents the results. Section 2.6 concludes.

2.2 Asset encumbrance: Definition and sources

In this section we define asset encumbrance and describe how assets become encumbered. We also review the most common sources of asset encumbrance (i.e. the liabilities or obligations that give rise to encumbered assets).

2.2.1 Defining asset encumbrance

European regulations define encumbered assets as "*assets pledged or subject to any form of arrangement to secure, collateralize or credit enhance any transaction from which it cannot be freely withdrawn"*.[8] Assets that are not encumbered are referred to as 'unencumbered'. The Basel Committee on Banking Supervision (BCBS) defines unencumbered assets as those assets which are "free of legal, regulatory, contractual or other restrictions on the ability of the bank to liquidate, sell, transfer, or assign the asset".[9]

To clarify the definition of encumbrance, let us consider a bank (Bank A) whose assets include loans and a portfolio of securities (government or corporate bonds, equities, etc.), financed via equity capital, retail deposits and unsecured wholesale funding, as shown on

---

[7] In the demise of Banco Popular (see footnote 4), the bank had high levels of asset encumbrance and of impaired loans. As of December 2016, almost 15% of Popular's loan portfolio was non-performing compared to a European average of 5.1% (EBA 2017). Its Basel III Leverage Ratio was also high (5.31% compared to a weighted average for European banks of 5.2% as per EBA, 2017).

[8] See European Commission (2015).

[9] See BCBS (2012).

the left-hand side of figure 2.1. As show on the right-hand side of the same figure, Bank A could obtain additional funding from a counterparty, let us say Bank B, by entering into a secured financing transaction.

Figure 2.1. Bank A encumbers assets by obtaining secured funding from Bank B
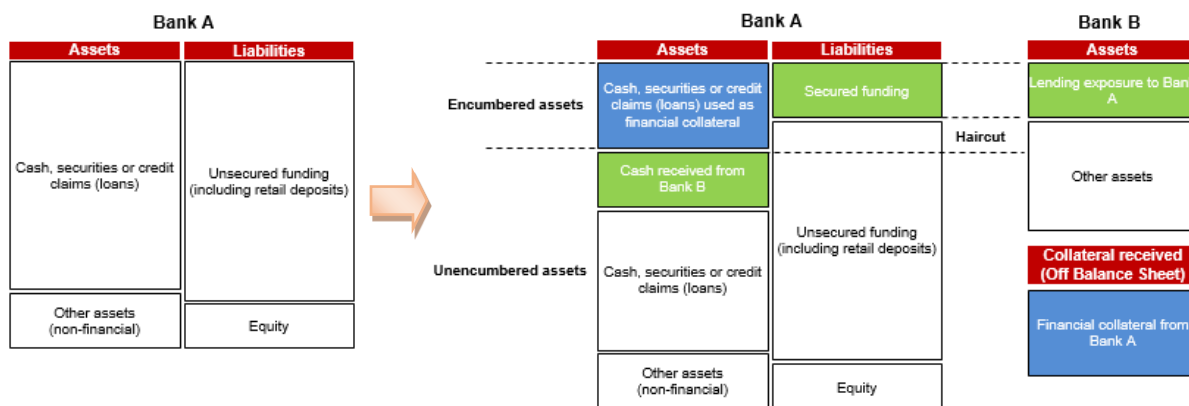
Under such an arrangement Bank A provides collateral to Bank B in order to mitigate the risk of failing to keep up with interest repayments or repaying the borrowings. In exchange, Bank A benefits from cheaper funding when compared to an equivalent unsecured transaction.[10] The arrangement imposes restrictions on Bank A in its ability to sell, transfer or dispose of the collateral provided during the term of the transaction. Bank A would consider such assets encumbered.

Figure 2.1 represents the securities provided as collateral as recorded or *recognised* in Bank A's balance sheet rather than being *transferred* to Bank B's balance sheet. Collateral obtained by Bank B is therefore represented in an off-balance sheet (OBS) rather than an on-balance sheet, and is known as 'OBS collateral' or simply 'collateral received'. The assumption that the collateral remains recognised from Bank A's balance sheet is a necessary condition for being considered an encumbered asset of Bank A. If the assets used as collateral were derecognised by Bank A then they would be recognised by Bank B and they would not be encumbered for Bank A.

In practice, the recognition or derecognition of collateral provided depends on the contractual terms of the transaction as well as its accounting treatment. Derecognition

---

[10] In addition, the arrangement may provide for savings in regulatory capital requirements to Bank B as well as lower regulatory liquidity requirements to Bank A and Bank B.

cannot occur unless the securities are 'transferred' to the counterparty. This can be achieved by using 'title transfer' arrangements, whereby full ownership of the collateral is passed on to the counterparty during the term of the transaction.[11] Collateral can also be provided under 'security interest' arrangements, which do not transfer ownership but concede rights to the counterparty to obtain full ownership of the collateral under some pre-determined event, such as failure to repay.[12] The use of one technique over the other depends on market practice. Collateral provided in secured financing transactions such as repurchase agreements (i.e. repo), is typically provided by way of title transfer whereas collateral used as a margin for OTC derivatives can be provided using both methods.[13]

The transfer of title over collateral, however, is not a sufficient condition for derecognition to occur, with the actual outcome depending on the applicable accounting treatment. Under International Financial Reporting Standards (IFRS), IAS 39 applies a set of tests to assess whether (i) the risks and rewards and (ii) control over the asset have been transferred.[14] If the risks and rewards have not been transferred, or in other words, if the collateral provided continues to be exposed to the risks of ownership of the assets such as loss in market value and/or the benefits that they generate such as dividends, then the collateral would remain recognised on its balance even if a transfer of assets has occurred. But even if the risks and rewards had been transferred, further 'control' tests are undertaken to understand which entity controls the asset. If the collateral provider could direct how the benefits of that asset are realised, then the collateral would not be derecognised either.

As illustrated in figure 2.1, the value of securities that Bank A posted as collateral is higher than the value of the borrowings. This practice is known as overcollateralisation and is intended to mitigate the risk of the collateral falling in value during the term of the transaction. It is usually undertaken by means of a 'haircut' or 'margin ratio'.[15] Collateral agreements often require a frequent (sometimes daily) marked-to-market valuation of the

---

[11] Under title transfer, Bank B would have to return the collateral (or equivalent securities) to Bank A when the original transaction matures.

[12] Security interest arrangements are also known as collateral 'pledges'.

[13] Under English Law the collateral for OTC derivatives is typically provided by way of title transfer, whereas under New York Law collateral is typically provided under security interest.

[14] The treatment under US GAAP (ASC 860) differs from IFRS since the focus is on whether the transferor has surrendered control over a financial asset.

[15] The agreed haircut or margin ratio determines the percentage by which the market value of a security is reduced for the purpose of calculating the amount of collateral being provided.

collateral and requests to top up the value of collateral, known as collateral calls, may be triggered if its market value falls below certain pre-determined threshold amounts.

Even in the case in which the collateral received is not reflected in its balance sheet, Bank B could re-use some or all of the collateral received from Bank A to obtain financing from a third party (let us say, Bank C). As illustrated in figure 2.2, this re-use of collateral by Bank B would result in the encumbrance of OBS collateral. As such, encumbrance can affect both on-balance sheet assets as well as OBS collateral. The practice of providing collateral that has been previously received is known as collateral re-use or re-hypothecation. It is common practice and may result in long 'collateral chains'.[16]

Figure 2.2. Collateral received and re-used
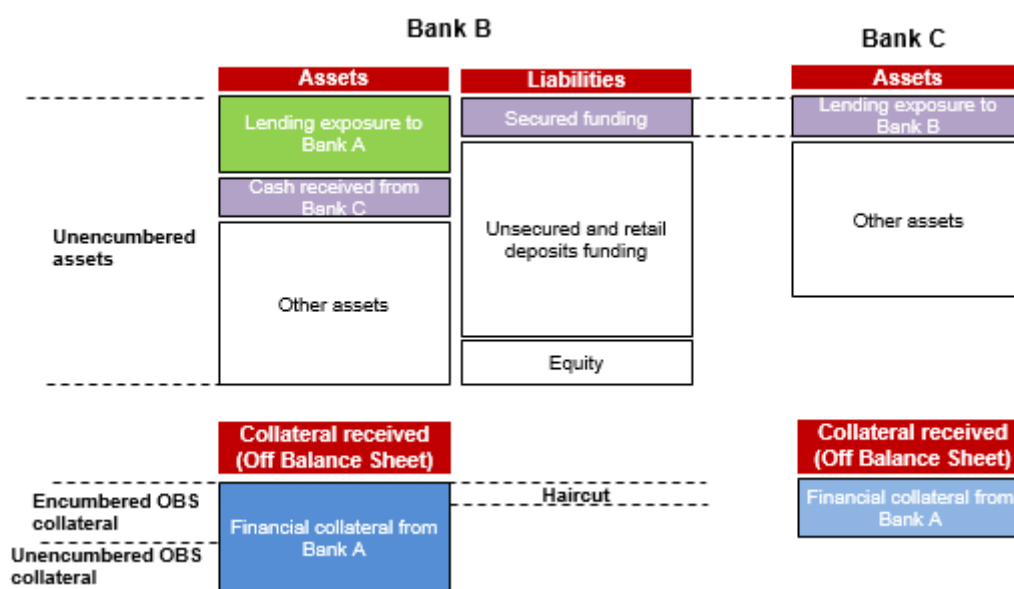
2.2.2 Sources of asset encumbrance

The liabilities or obligations that give rise to encumbered assets are known as 'sources of asset encumbrance' or 'matching liabilities'. The typical bank will have encumbered assets from several sources but the simplest institutions may rely only on a single source

---

[16] The terms re-hypothecation and re-use are often used interchangeably and we will do so here. In practice there are legal distinctions between them that may be relevant in a different context. Recent studies have analysed the concept of re-hypothecation and 'collateral velocity'. Analytical work includes Adrian and Shin (2010) and Singh (2010). More recent work has focussed on liquidity mismatches and the role of collateral in intermediation chains. Brunnermeier et al. (2014) introduced the Liquidity Mismatch Index (LMI) which compares the market liquidity of assets and the funding liquidity of liabilities, thus capturing the length of collateral intermediation chains.

or may present no encumbered assets at all. We now discuss some of the most common sources of asset encumbrance.[17]

2.2.2.1 Secured financing transactions

Secured financing transactions encompass myriad transactions involving the temporary provision of securities to borrow cash or other securities. Common types include repurchase agreements (repos), buy/sell backs or securities borrowing and lending. Collateral in repo is provided under title transfer but it remains recognised in the balance sheet of the collateral provider's (i.e. the repo seller) since the risks and rewards of the collateral are retained.[18] Thus the repo collateral would be encumbered for the collateral provider.[19]

Encumbered assets in repo are predominantly government bonds, followed by corporate bonds and covered bonds. Asset-backed securities and equities are also used as collateral. Most of the funding provided by central banks is transacted through repo. Like Dexia, many European banks were, and some still are, heavily reliant on repo financing from the ECB.

2.2.2.2. Asset-backed securities (ABS) and mortgage-backed securities (MBS)

Another potential source of asset encumbrance is securitisations. These entail ABS and MBS bonds or notes being issued and receivables, which may include retail or commercial mortgages in MBS, or credit card debt or other loans in ABS, being used as collateral.

---

[17] In addition to the sources covered in section 2.2, transactions that may result in encumbered assets include collateral swaps, also known as collateral upgrade transactions, where collateral of a different quality is exchanged. Collateralised guarantees rely on securities to secure an existing or future liability. Other arrangements, such as factoring —which include the transfer of trade receivables to an institution— may result in similar encumbrance to securitisations.

[18] If this was not the case, banks could artificially reduce its overall leverage by derecognising collateral in repurchase agreements. This treatment was exploited by Lehman Brothers' under the well-known 'Repo 105' scheme, characterised by the New York Attorney General Andrew Cuomo as a 'massive accounting fraud' and leading to a review by the accounting standard settlers of the accounting treatment of repo transactions.

[19] The International Capital Markets Association (ICMA) whose members include almost 530 market participants, has argued that asset encumbrance arising from repo transactions is, in contrast with regulatory definitions, limited to the amount of overcollateralization provided (ICMA, 2013). Their argument is based on the premise that the collateral provider receives cash which is unencumbered and therefore the net impact on the recovery of unsecured creditors should the bank was to become insolvent would be limited to such overcollateralization. We consider the potential impact on asset encumbrance on unsecured creditors in section 2.3 below and in section 2.4.1 show that a specific measure of asset encumbrance can be used to capture this phenomenon.

A traditional 'two-step' securitisation involves the initial transfer of the receivables of the originating bank to a Special Purpose Vehicle (SPV) and the sale of the ABS or MBS to investors. The overall securitisation structure is intended to make sure that there is a true sale of receivables to the SPV and that the SPS is 'bankruptcy remote'. Accounting standards however, may require that the SPV is consolidated into the 'sponsoring' bank's balance sheet, including all of its assets and liabilities, even the receivables.[20] If the underlying receivables were consolidated, this would result in the recognition of such receivables on the sponsor's balance sheet. However, tests to assess whether the assets meet the criteria for accounting derecognition, as discussed earlier, shall still be undertaken. If derecognition criteria are not met the receivables would be encumbered. This is often the case since it is common for the sponsoring bank to keep an active role in the securitisation, for example, by servicing the assets or providing support by retaining certain tranches to absorb first losses and potential risks in relation to timings in the collection of the receivables.

ABS or MBS can be used as collateral to raise funding with counterparties and central banks. Thus, a common practice across some banks, especially during the Eurozone crisis, is the retention of their self-issued ABS or MBS rather than its sale to investors.[21] If notes are retained, they would not be encumbered. But if the notes are used to raise fresh funding, for example, from the central bank via repo, the receivables would become encumbered as it occurs in securities' financing transactions.

Figure 2.3 (left-hand side) illustrates how securitised receivables can be encumbered (highlighted in green) by collateralising ABSs that are either (i) sold to investors or (ii) used as repo collateral to obtain funding from another counterparty.

---

[20] The consolidation models under IFRS and GAAP are relatively similar and are based on the criteria of entity control over the SPV.

[21] The acceptance of securitised notes as collateral in the ECB facilities led to an important increase in retention levels during the Eurozone crisis, with overall retention as a proportion of total gross issuance increasing from 26% in the first half of 2007 to 42% in the first half of 2012 (IMF, 2013).

Figure 2.3. Encumbered and unencumbered assets from securitisations (LHS) and derivative transactions (RHS)
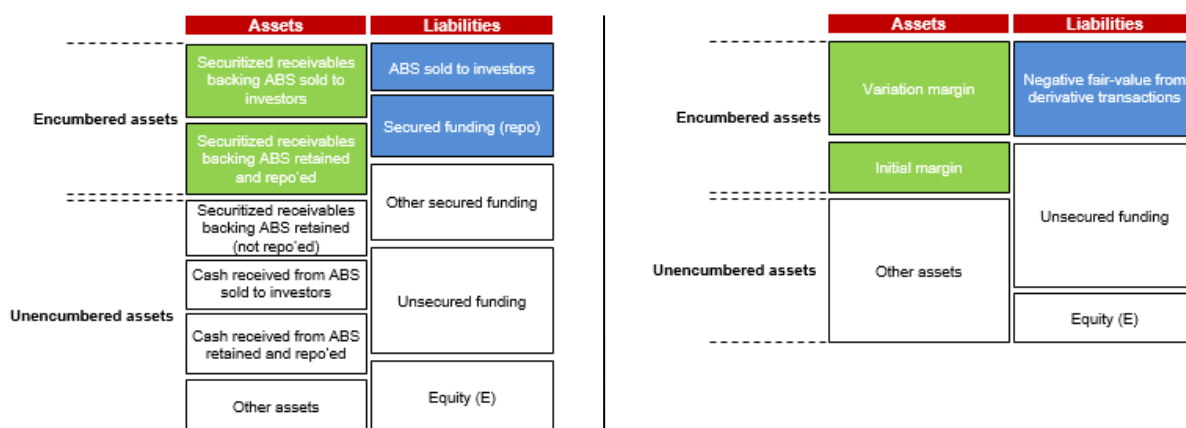
2.2.2.3. Covered bonds

Covered bonds are similar to MBS but the mortgages used as collateral always remain recognised on the consolidated balance sheet of the issuing entity and thus always generate encumbrance. The issuer and the investors have dual recourse to the collateral. This feature, together with the existence of overcollateralisation requirements and the dynamic replenishment of non-performing loans in the collateral pool imply that these instruments are perceived as being very safe. There is indeed no known default on covered bonds since their inception.

The use of covered bonds as collateral has significantly increased in recent times. For many banks in peripheral European countries (GIIPS) funding collateralised by retained covered bonds became the main source of long-term funding during the Eurozone sovereign crisis, as their access to unsecured markets was partially or fully closed (van Rixtel and Gasperini, 2013).

2.2.2.4. Derivatives

Derivatives also generate encumbrance, as collateralisation has become a key method of mitigating counterparty credit risk in derivative markets, both on over-the-counter (OTC) and exchange-traded (ETD) derivatives. Collateralisation occurs because of the provisioning of the margin, in two different forms. A variation margin is posted during the course of the transaction to cover adverse changes in value (i.e. a negative mark-to-market value). Initial margin (also known as an independent amount) is posted at the

beginning of a transaction to cover potential future adverse changes in the value of the contract, and is recalculated on a regular basis.

The margin provided is subject to restrictions and therefore constitutes encumbered assets. This is illustrated in figure 3 (right-hand side).[22] The margin can be provided in the form of cash or securities and it is common to provide re-hypothecation rights to the counterparty. According to the latest ISDA Margin Survey, for non-cleared OTC derivatives cash represents 76.6% of the collateral provided, followed by government bonds (13.4%) and other securities (10.1%), including US municipal bonds, government agency/government-sponsored enterprises (GSEs), and equities (ISDA, 2015).

2.3 Asset encumbrance: Risks and benefits

The potential negative impact of asset encumbrance on unsecured creditors has been the focus of much discussion recently. As highlighted by an incipient theoretical literature, higher asset encumbrance would result in a reduction of the assets that become available to unsecured creditors under insolvency, an effect coined as "structural subordination" (Bank of England, 2012; CGFS, 2013; Houben and Slingenberg, 2013; IMF, 2013; Juks, 2012; Le Leslé, 2012). The same seniority that secured creditors enjoy means that as more secured debt is issued, balance sheet shocks are asymmetrically concentrated on unsecured creditors, exacerbating the possibility of a run of unsecured creditors (Ahnert et al., 2016; Matta and Perotti, 2015). The resulting shifting of risks depends on the magnitude of the haircuts being applied since the required overcollateralisation reduces the amount of collateral available for unsecured funding (Eisenbach et al., 2014).

If unsecured creditors reflect the risk of structural subordination into required returns, this could result in higher overall funding costs to institutions. As stated by Dr Joachim Nigel, a former member of the executive board of the Deutsche Bundesbank in a speech at the 2013 European Supervisor Education Conference on the future of European financial supervision: "Higher asset encumbrance has an impact on unsecured bank creditors. The

---

[22] The figure assumes that the variation margin is not offset against the derivative liability (i.e. the negative fair value from the derivative) therefore becoming encumbered. Some contracts allow for such an offsetting of the variation margin. The outstanding exposure between the counterparties is settled and the terms of the derivative contracts are reset so that the fair value is zero, leading to no encumbered assets due to an exchange of the variation margin.

more bank assets are used for secured funding, the less remain to secure investors in unsecured instruments in the case of insolvency. They will price in a risk premium for this form of bank funding."

In addition, a higher amount of encumbered assets may reduce a bank's headroom to obtain funding under a stressed market environment. This could in turn trigger investor concerns about the bank's viability, as shown in Dexia's demise. During economic downturns, falling collateral values and higher haircuts result in higher overcollateralisation levels, requiring more assets to be pledged to raise a given level of funding and increasing asset encumbrance (Bank of England, 2012). This latter effect would in turn magnify the impact of asset encumbrance on unsecured creditors via structural subordination. There is evidence that during the Eurozone crisis, not only the funding costs of banks increased significantly alongside the increases in asset encumbrance, but larger over-collateralisation levels also shrank the pools of unencumbered assets, further reducing banks' ability to raise secured and unsecured funding (CGFS, 2011; CGFS, 2013; ECB, 2012).[23]

However, as it is shown in the data, higher asset encumbrance may also bring in benefits for unsecured creditors. Clearly, secured creditors benefit from the safety that collateral provides. This is reflected in lower funding costs than equivalent unsecured funding. Secured funding also carries a lower rollover risk. There is indeed evidence that repo funding was rolled-over during the financial crisis, up to the eve of default (Gorton and Metrick, 2012; Krishnamurthy et al., 2014). As a result, higher collateralisation could lead to a lower probability of default and increased capacity of debt repayment, which would also benefit unsecured creditors.

In addition, higher collateralisation could provide a reduction in the cost of settling creditors' conflicts in case of resolution or bankruptcy (Hardy, 2014). Claimants holding collateral do not have to enter the contest for residual assets and, despite increasing levels

---

[23] Recent literature has also analysed the system-wide implications of increased asset encumbrance levels and the potential for increased susceptibility to procyclical swings in the underlying value of the collateral assets (Gai et al., 2013; Haldane 2012; Krishnamurti, 2010; Perotti, 2010). More generally, the amplification role of haircut shocks in generating procyclicality has been broadly considered in the literature (see e.g. Adrian and Shin, 2010; Brunnermeier and Pedersen, 2009; Geanakoplos, 2010; Gorton and Metrick, 2012).

of structural subordination, the remaining claimants have less to fight over, thus reducing bankruptcy costs. This would also benefit unsecured creditors.

In sum, asset encumbrance carries risks of subordinating unsecured creditors, increasing funding costs and reducing a bank's headroom to obtain funding under stressed conditions. However, it may also bring benefits to both secured and unsecured creditors, which could lead to a decreased probability of default and decreased overall funding costs.

2.4 Data and methodology

Our analysis has two parts. First, we provide a descriptive analysis of asset encumbrance levels by country, bank credit quality, size and business model. We then assess the extent to which bank CDS premia are associated with asset encumbrance through a series of multivariate regressions.

2.4.1. Measuring asset encumbrance

There is currently no consensus as to how overall asset encumbrance shall be measured and different measures have been proposed. We focus on the three key ratios being used by policymakers. The computation of each ratio is illustrated in figure 2.4.

The 'asset encumbrance ratios' (AERs) capture the amount of encumbered assets as a proportion of total assets. There are two variations:

- The ratio of *encumbered assets to total assets,* which captures the overall proportion of balance sheet assets that have been encumbered. This ratio has been used by the Bank of England and the European Systemic Risk Board (ESRB) to undertake analysis of the UK and European banking sectors respectively (BoE, 2012; ESRB, 2013). We denote it as AER1.
- The ratio of e*ncumbered assets and other collateral received and re-used to total assets and total collateral received,* which captures the overall proportion of encumbered balance sheet assets as well as off-balance sheet collateral. This ratio is used by the EBA to undertake their risk assessment of the European banking system and to apply more comprehensive regulatory reporting requirements (EBA, 2016). We denote this ratio as AER2.

The third ratio focusses instead on unencumbered assets:

- The ratio of *unencumbered assets to unsecured liabilities* (UAUL), which captures the proportion of assets which are not subject to collateral agreements as a proportion of unsecured creditor's claims and provides an indication of the amount of structural subordination of unsecured creditors. According to a report from the Bank of International Settlements' Committee on the Global Financial System (CGFS, 2013), the UAUL ratio is the most appropriate measure of asset encumbrance.

As opposed to AER1 and AER2, UAUL is a measure of how many assets are available to unsecured creditors under insolvency, and should therefore capture the structural subordination of unsecured creditors more directly than AER1 and AER2. Since UAUL is measured relative to unsecured funding, this ratio would be unable to capture low levels of unencumbered assets relative to the total assets of banks that rely heavily on capital or secured funding. As opposed to AER2, AER1 and UAUL do not capture encumbrance arising from off-balance sheet activities.

Figure 2.4. Asset encumbrance metrics

$$AER1 = EA / TA$$

$$AER2 = (EA + OBR) / (TA + OCR)$$

$$UAUL = UA / UF$$
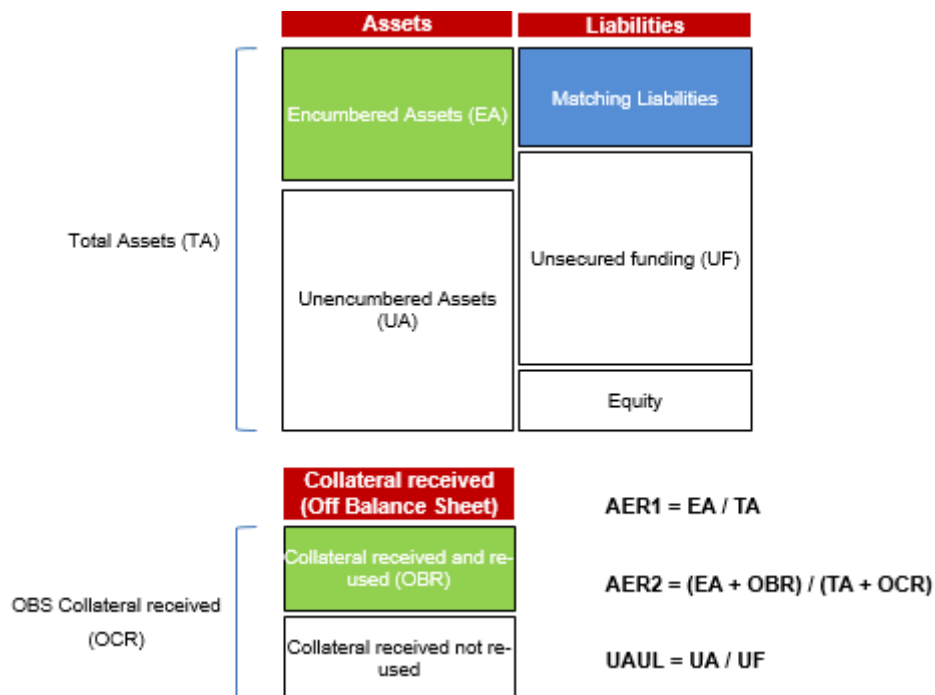
Computing asset encumbrance measures at the bank level is not straightforward since accounting data provides limited information to infer the amount of banks' encumbered assets, unencumbered assets and matching liabilities. Accounting statements are accompanied by disclosures which try to shed light on the amount of assets that are collateralising transactions but, as noted by the EBA: "existing disclosures in International Financial Reporting Standards (IFRS) may convey certain situations of encumbrance but fail to provide a comprehensive view on the phenomenon" (EBA, 2014). For this reason, the EBA introduced new guidelines in 2014 proposing the requirement to disclose asset encumbrance reporting templates. EBA guidelines do not constitute a regulatory requirement and, although most did, not all of the European institutions disclosed such information.

We extract data from the risk disclosures of banks, including information on encumbered assets, unencumbered assets, off-balance sheet collateral received and available for encumbrance, OBS collateral received and re-used and matching liabilities as of year-end 2014. We complement the disclosure data with data on total assets and equity extracted from Bankscope to compute the asset encumbrance ratios, AER1 and AER2 considered for each institution. For UAUL, we use a slightly modified version which we denote as AUAUL (Adjusted UAUL), calculated as $AUAUL_i=(Max(UAUL)-UAUL_i))/(Max(UAUL)-Min(UAUL))$ where $UAUL_i$ is bank's i ratio of unencumbered assets to unsecured liabilities. This adjustment facilitates comparisons with AER1 and AER2 by ensuring that higher encumbrance is associated with a higher AUAUL and that its values fall between 0 and 1.

2.4.2. Measuring bank risk

Our main dependent variable in the multivariate regressions is a measure of bank risk represented by banks' CDS spreads as of year-end 2015. CDS spreads are widely considered to be a good indicator of bank risk. As noted by Iannotta et al. (2013), the use of a market-based risk measures such as CDSs as opposed to other measures of risk such as accounting ratios, can help to mitigate endogeneity issues. In addition, CDS spreads are a direct measure of the risk to senior unsecured creditors and therefore can be used to estimate for the amount of structural subordination and overall bank unsecured funding costs (see e.g. Babihuga and Spaltro, 2014; Bank of England, 2014).

We use implied rather than market-based spreads in our study. For over 5,800 banks, Fitch Solutions determines the implied spreads on a daily basis using a proprietary model. The model employed is based on a 'hybrid' approach where multiple data inputs and benchmarking methodologies are used to calculate CDS spreads. CDS pricing curves are grouped, aggregated and smoothed before being used as proxies to entities without CDS curves but with similar characteristics. Inputs include banks' financial fundamental information and Distance-to-Default (DtD) information derived from the equity market. For benchmarking purposes entities are group by region, sector, credit rating, and other market derived information such as issued bond spreads over the risk-free rate adjusted for CDS-Bond basis and maturity mismatches, and equity-derived implied ratings computed using an equity probability of default type model such as KMV. Finally, adjustments for debt subornation level, currency, and restructuring differences are applied.

Since only the largest global institutions are involved in CDS issuance and not all banks disclosure asset encumbrance information and other fundamental information necessary to compute our additional explanatory variables, using implied CDS spreads ensures a minimum sample size is achieved.

In line with the existing literature, we focus on five-year senior spreads since these contracts account for 85% of the market and are highly liquid. Data is provided by Fitch Solutions and extracted from Bankscope.

2.4.3. Other variables

Explanatory variables include, in addition to the asset encumbrance measures, CAMEL and control variables. We follow Chiaramonte and Casu (2013) to select the following CAMEL variables:

Capital Adequacy:
- The *Tier 1 capital ratio*, which represents the ratio of high-quality capital (shareholders' capital, reserves and other perpetual capital resources such as subordinated debt), divided by risk-weighted assets (RWA).
- The *leverage ratio*, which is calculated as the fraction of common equity to total assets and reflects the level of indebtedness of a firm.

Liquidity:

- The *net loans to deposits and short-term funding ratio*, which is a measure of structural liquidity. A lower value of the ratio means the bank relies to a greater extent on more stable deposit funding, as opposed to wholesale funding, to finance its loan book.

- The *liquid assets to total assets ratio*, which measures the amount of liquid assets that the bank holds and that could be converted into cash to withstand a liquidity stress event.

Quality of assets:

- The *ratio of loan-loss reserve to gross loans*, which measures the quality of the loan portfolio by indicating the proportion of reserves for losses relative to the banks' loan portfolio.

- The *ratio of unreserved impaired loans to equity*, which is another indicator of the quality of the loan portfolio but expressed relative to common equity. It is also known as the 'capital impairment ratio'.

Earnings potential:

- The *return on equity ratio (ROE)*, which measures the bank's income-producing ability as reflected by its net income relative to the bank's common equity.

- The *return on assets ratio (ROA)*, which is an indicator of the return on a firm's investments and is calculated by dividing the bank's net income over its total assets.

Control variables include bank size (measured by the natural logarithm of total assets), central bank exposure to total assets and off-balance sheet exposure to total assets. We include dummy variables to differentiate the business model of the institution using three categories: 'Commercial banks and Bank holding companies (BHC)', 'cooperative and savings banks' and 'other banks'. We also include a dummy variable to identify which banks are investment grade. We use implied ratings in order to avoid compromising the sample size, in a similar fashion to CDS spreads.[24]

Country-specific dummies are included in all models to help to control for factors affecting CDS premia at the country level, including regulatory particularities. To account for the potential correlation of the errors among the banks belonging to the same business

---

[24] Implied ratings are provided by Fitch Solutions and derived from proprietary fundamental data. These provide a forward-looking assessment of the stand-alone financial strength of a bank and are categorised according to a 10-point rating scale from A to F where A denotes the maximum creditworthiness, with four interim scores (A/B, B/C, C/D and D/E).

category in a given country, we apply country-business model clustering in all our regression models.

Our final data sample includes institutions with total assets above €1bn for which CDS spreads, asset encumbrance, CAMEL and control variables are available, resulting in 367 banks.

2.4.4. Model specifications

To construct our model specifications, we follow recent studies which estimate the linear relationship between CAMEL indicators and CDS spreads (see e.g. Chiaramonte and Casu, 2013), and consider the extent to which asset encumbrance contributes to the explanatory power of these models.

Our baseline model specification is as follows:

$$CDS_i = \propto + \beta_1 AE_i + \beta_2 CAMEL_i + \beta_3 Control_i + Z_i + \varepsilon_i$$

where CDS is the natural log of the CDS spread for bank i at year-end 2015; AE is the asset encumbrance measure for bank i at year-end 2014; CAMEL represents the set of eight CAMEL variables for bank i at year-end 2014; CONTROL represents the control variables for bank i at year-end 2014; and Z corresponds to country dummies. All explanatory variables are measured at *t -1* to mitigate endogeneity concerns.

In order to assess whether the relationship between encumbrance and CDS spreads varies for different types of banks, we also look at additional specifications by interacting the asset encumbrance ratios with the CAMEL indicators and some of the control variables, using the following model:

$$CDS_i = \propto + \beta_1 AE_i + \beta_2 CAMEL_i + \beta_3 Control_i + \beta_4 AE_i * CAMEL_i + \beta_5 AE_i * Control_i + Z_i + \varepsilon_i$$

2.5 Results

In this section we present the results of our analysis. We first provide a descriptive analysis of asset encumbrance followed by the regression results.

2.5.1. Descriptive analysis and summary statistics

Table 2.1 presents the summary statistics of the variables of study. The mean values of AER1, AER2 and AUAUL are 0.13, 0.14 and 0.60 respectively. Note that there is a wide disparity across banks in our sample. AER1 and AER2 present standard deviations of 0.11 and 0.12 respectively. Although the standard deviation of AUAUL is lower (0.08), the mean and the original standard deviation of UAUL are 1.06 and 0.15.

Table 2.1. Summary statistics of asset encumbrance metrics

	Mean	Median	Std.	Min.	Max.
AER1	0.13	0.09	0.11	0.00	0.68
AER2	0.14	0.10	0.12	0.00	0.70
AUAUL	0.60	0.59	0.08	0.00	1.00

The number of observations for all variables is 367.

Table 2.2 presents the correlation matrix of encumbrance ratios. AER1 and AER2 present a high correlation of 0.974 which is expected given their similar construction with the only difference being the inclusion of off-balance sheet collateral in AER2. The correlation coefficients of AUAUL with AER1 and AER2 are 0.388 and 0.363 respectively.

Table 2.2. Correlation matrix of asset encumbrance metrics

	AER1	**AER2**	**AUAUL**
AER1	1		
AER2	0.974^{***}	1	
AUAUL	0.388^{***}	0.363^{***}	1

***: Significant at 1% level.

Given the above results, we hereafter focus on AER1 and AUAUL. Figure 2.5 presents the histograms for both ratios. Whereas AER1 shows high skewness towards the lower values (between 0 and 0.2), the AUAUL distribution is much more symmetric.

Figure 2.5. Histograms of AER1 and AUAUL
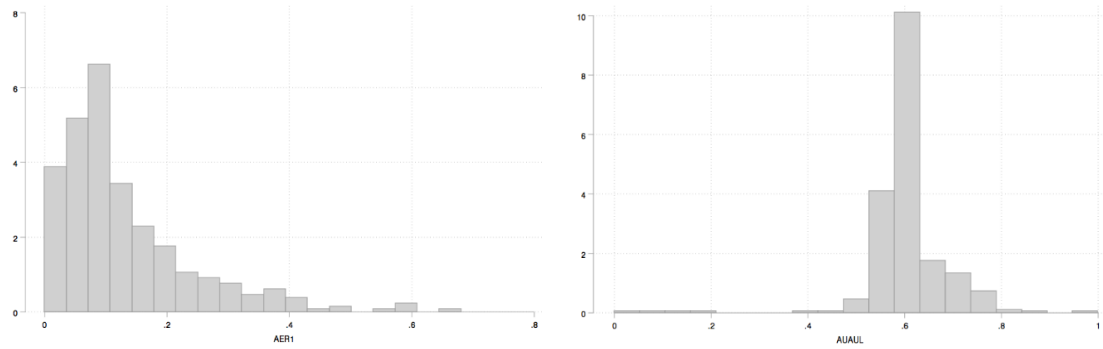
Figure 2.6 shows the mean ratio levels of AER1 (blue, left scale) and AUAUL (green, right scale) for those countries with more than one observation. The countries are shown in four groups corresponding to GIIPS, Nordic countries, core countries including Austria, Belgium, Germany, France, UK, Luxembourg and the Netherlands, and other European countries such as the Eastern European countries and Malta. Results show a wide disparity in mean encumbrance levels across countries. All the GIIPS countries present higher mean encumbrance ratios than the sample average, as do Nordic countries such as Denmark and Sweden. Denmark, in particular, presents the highest mean ratio of all countries in the sample. Nordic countries have a long tradition of covered bond issuance, which may help explain these results. Of the remaining countries, Belgium, France and the UK present higher mean encumbrance levels than the overall sample average. Belgium, Malta, Netherlands and the UK present higher mean values of AUAUL than the sample mean. Luxembourg and some of the countries classified as 'other' such as Bulgaria, Poland and Malta present the lowest values of AER1 but also the largest differences between AUAUL and AER1.

Figure 2.6. Mean AER1 (left scale) and AUAUL (right scale) by country
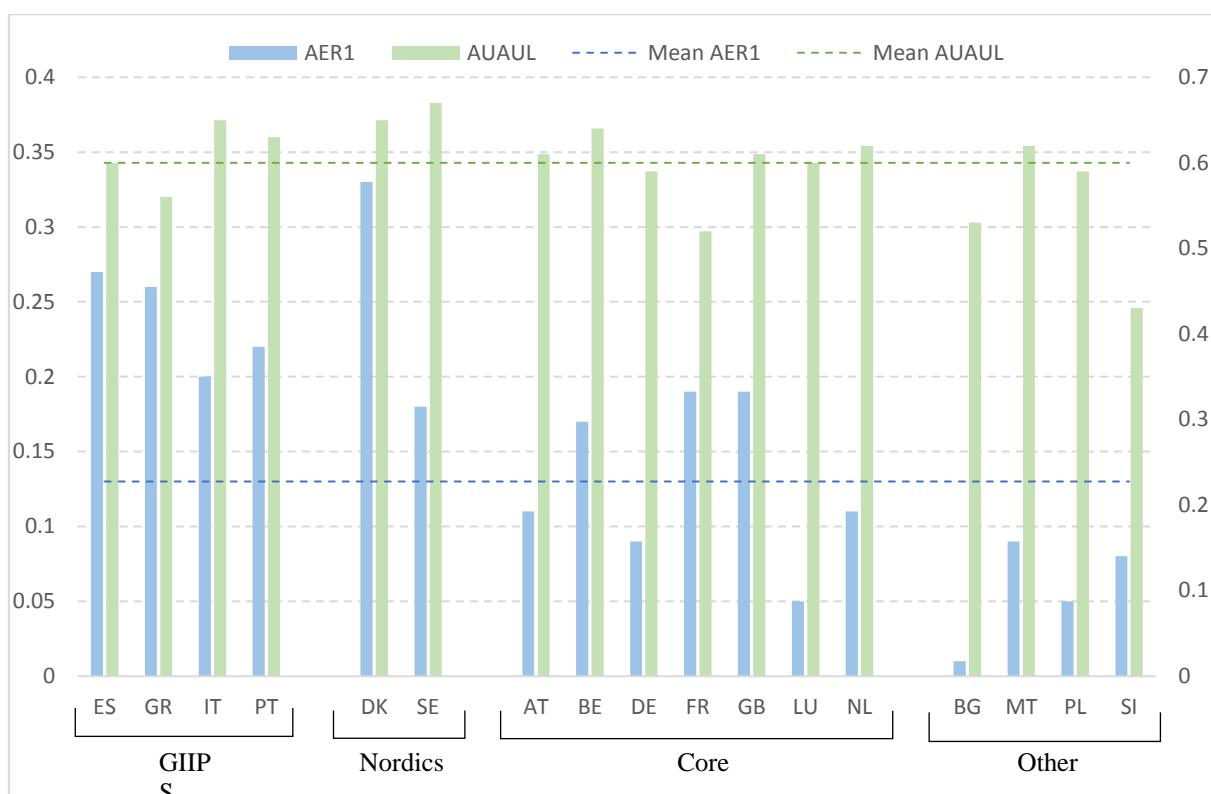
Figure 2.7 shows the mean levels of the two asset encumbrance ratios across rating categories. Banks within the most extreme categories, A/B and E/F, present the lowest mean AER1 and AER2 ratios of all categories. For AUAUL, it is banks in categories D/E and E/F that present the lowest mean values.

As shown in figure 2.8, mean encumbrance levels tend to increase with bank size, measured in terms of total assets, across all ratios. Since securitisations involve substantial costs, mostly of a fixed nature, these should be particularly costly to issue for smaller banks (Affinito and Tagliaferri, 2010; Carbó-Valverde et al., 2012; Panetta and Pozzolo, 2010).

Figure 2.9 shows the mean ratio levels by type of institution. We distinguish between 'commercial banks and bank holding companies (BHC)', 'cooperative banks', savings banks' and 'other banks', including mortgage banks and pure investment banks. Savings banks show the lowest levels for both AER1 and AUAUL. Institutions classified as 'other' show relatively high values of AER1 but not AUAUL. Cooperative banks show the highest average level of asset encumbrance when measured by AUAUL.

Figure 2.7. Mean AER1 (left scale) and AUAUL (right scale) by implied credit rating category
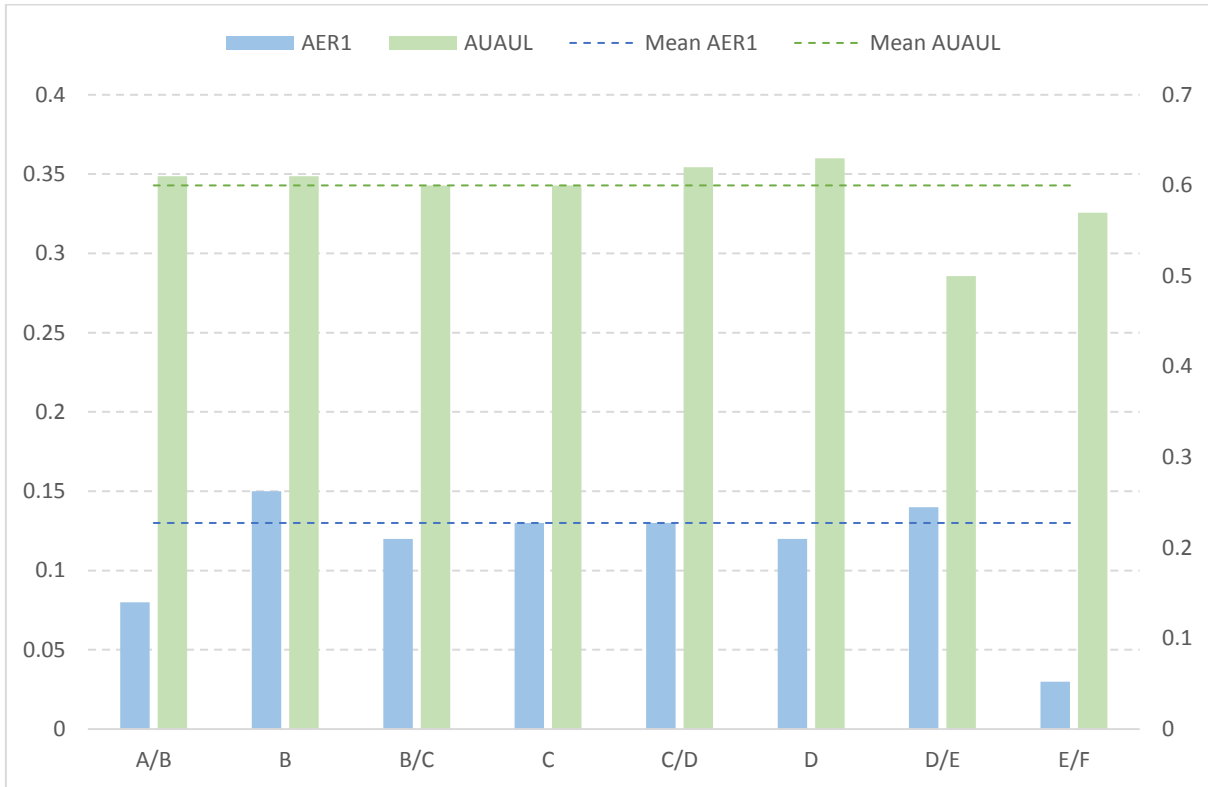
Figure 2.8. Mean AER1 (left scale) and AUAUL (right scale) by size

Figure 2.9. Mean AER1 (left scale) and AUAUL (right scale) by business type
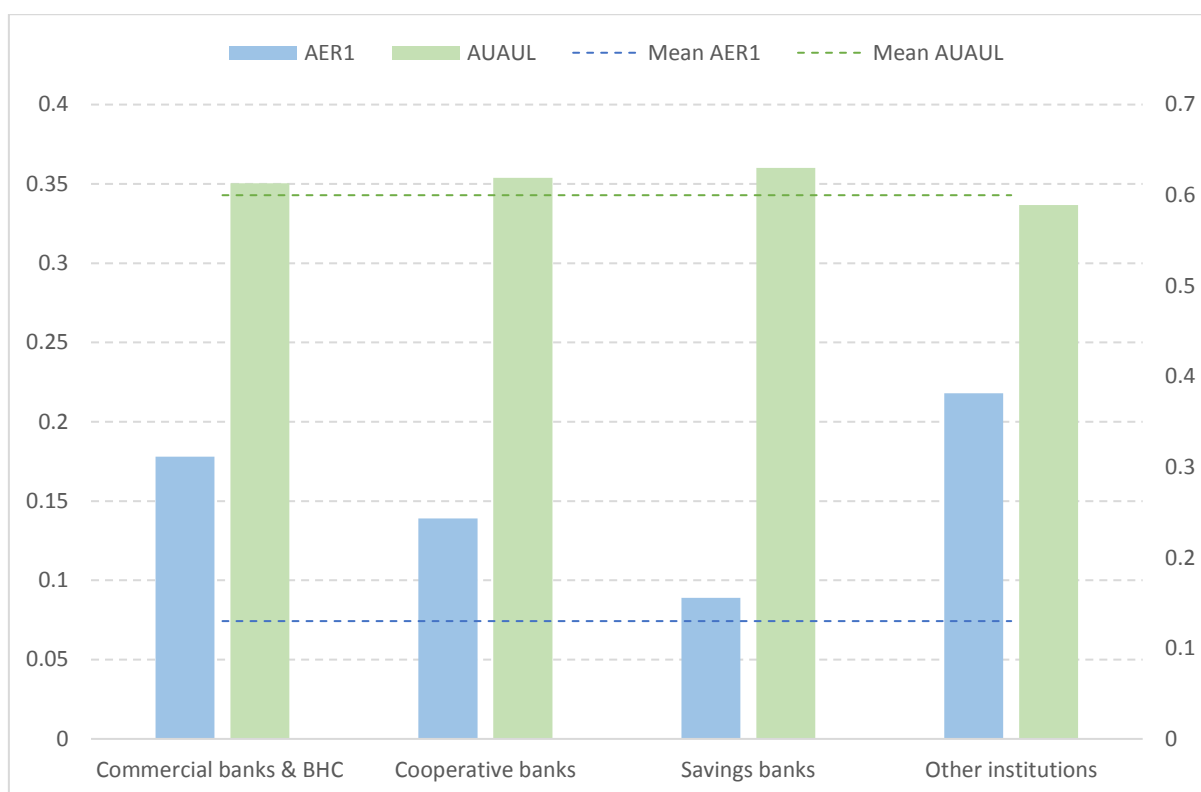
2.5.2 Regression analysis

Our regression analysis consists of a series of multivariate regressions run with the natural logarithm of a CDS spread as the dependent variable. Our first set of regressions (baseline results) includes asset encumbrance ratios, CAMEL variables, bank-specific controls and country dummies as explanatory variables. In our second set of regressions (mediating effects) we also include interactions of asset encumbrance with CAMEL and control variables.

Table 2.3 presents the summary statistics of the variables of study. The average value of the CDS spread variable is 5.14, corresponding to 171 basis points. The median value is 5.17, corresponding to 176 basis points. In terms of bank CAMELS indicators, we find that, on average, a sample bank has a tier 1 ratio of 0.15, a leverage ratio of 0.08, net loans to deposits and a short-term funding ratio of 0.77, a liquid assets to total assets ratio of 0.18. The average ratio of the loan-loss reserve to gross loans is 0.04, and the ratio of unreserved impaired loans to equity is 0.33. The average ROA and ROE are nearly 0. From all the control variables, central bank exposure presents the lowest standard deviation of 0.01.

Table 2.3. Summary statistics, variables of study

	Mean	Median	Std.	Min.	Max.
Dependent variable					
CDS Spread	5.14	5.17	0.33	4.56	6.56
Asset Encumbrance					
AER1	0.13	0.09	0.11	0.00	0.68
AER2	0.14	0.10	0.12	0.00	0.70
AUAUL	0.60	0.59	0.08	0.00	1.00
CAMEL variables					
Tier 1 ratio	0.15	0.14	0.04	0.05	0.35
Leverage ratio (Equity / TA)	0.08	0.08	0.02	0.01	0.18
Net loans to deposits and ST funding	0.77	0.75	0.24	0.05	2.05
Liquid assets / TA	0.18	0.11	0.23	0.01	2.98
Loan loss reserve / gross loans	0.04	0.02	0.04	0.00	0.27
Unreserved impaired loans / equity	0.33	0.14	0.49	0.00	4.63
ROA	0.00	0.00	0.01	-0.04	0.03
ROE	0.01	0.02	0.10	-1.08	0.32
Control variables					
Central bank exposure / TA	0.01	0.01	0.01	0.00	0.12
Off-balance sheet exposure / TA	0.08	0.06	0.10	0.00	1.66
Investment grade	0.66	1.00	0.47	0.00	1.00
BHC and commercial banks	0.19	0.00	0.39	0.00	1.00
Saving and cooperative banks	0.75	1.00	0.43	0.00	1.00
Size	6.75	6.41	0.84	6.00	9.32
GIIPS	0.28	0.00	0.45	0.00	1.00
Nordics	0.02	0.00	0.15	0.00	1.00

This table reports the summary statistics of the variables used in regressions. The number of observations for all variables is 367. 'CDS Spread' refers to the logarithm of the CDS Spread (in basis points). 'TA' refers to total assets. 'Size' refers to the logarithm of total assets (in thousand Euros).

2.5.2.1. Baseline results

Table 2.4 reports the results of the baseline regressions. We control for fixed country effects in all regressions and continuous variables are demeaned for ease of interpretation. Model 1 includes CAMEL variables, bank-specific controls and country dummies as explanatory variables but excludes asset encumbrance.

Table 2.4. Baseline results

	Dependent variable: 5-year implied CDS spreads			
	(1)	(2)	(3)	(4)
Asset Encumbrance				
AER1		-0.177***		
		(0.06)		
AER2			-0.131*	
			(0.07)	
AUAUL				-0.267***
				(0.10)
CAMEL variables				
Capital: Tier 1 Capital Ratio	1.013	1.058	1.046	1.103
	(0.94)	(0.93)	(0.92)	(0.92)
Capital: Leverage Ratio (Equity / TA)	-0.907	-1.164	-1.102	-1.195
	(1.50)	(1.42)	(1.40)	(1.52)
Liquidity: Net Loans / Deposits and Short-Term Funding	-0.057	-0.049	-0.053	-0.055
	(0.04)	(0.04)	(0.04)	(0.04)
Liquidity: Liquid Assets / TA	-0.104	-0.105	-0.103	-0.108*
	(0.07)	(0.07)	(0.07)	(0.06)
Asset quality: Loan Loss Reserve / Gross Loans	2.025***	1.927***	1.959***	2.085***
	(0.65)	(0.61)	(0.64)	(0.62)
Asset quality: Unreserved Impaired Loans / Equity	-0.066*	-0.060*	-0.062*	-0.063*
	(0.03)	(0.03)	(0.03)	(0.03)
Profitability: ROA (Net Income / TA)	-16.454***	-16.398***	-16.369***	-17.019***
	(4.77)	(4.92)	(4.91)	(4.78)
Profitability: ROE (Net Income / Equity)	0.513***	0.538***	0.528***	0.555***
	(0.17)	(0.18)	(0.18)	(0.17)
Control variables				
Central Bank Exposure / TA	1.240**	1.150**	1.171**	1.190**
	(0.50)	(0.56)	(0.55)	(0.53)
Off-Balance Sheet Exposures / TA	0.056	0.077	0.073	0.072
	(0.08)	(0.09)	(0.09)	(0.08)
Investment grade	-0.467***	-0.466***	-0.466***	-0.464***
	(0.02)	(0.02)	(0.02)	(0.02)
BHC and commercial banks	0.050	0.035	0.039	0.050
	(0.06)	(0.05)	(0.05)	(0.05)
Saving and cooperative banks	0.006	-0.007	-0.002	0.002
	(0.05)	(0.05)	(0.05)	(0.05)
Size	-0.152***	-0.149***	-0.149***	-0.157***
	(0.02)	(0.02)	(0.02)	(0.02)
Country FE	YES	YES	YES	YES
Observations	367	367	367	367
R-squared	0.79	0.79	0.79	0.79

All regressions are estimated using OLS. A constant is included, but its coefficient is left unreported. Country-specific dummies are included in all models and country-business model clustering is applied in all models. ***: Significant at 1% level; **: Significant at 5% level; *: Significant at 10% level.

Models 2-4 include the three asset encumbrance measures as explanatory variables. A negative and significant association between banks' implied CDS spreads and asset encumbrance emerges across all models. Thus, our initial evidence suggests a net positive perception of unsecured creditors towards asset encumbrance. As suggested in the

theoretical discussion, higher collateralisation could lead to a lower probability of default and an increased capacity of debt repayment. Higher collateralisation could also reduce bankruptcy costs and increase value (Hardy, 2014).

While the coefficients for AER1 and AUAUL are highly significant, AER2 is significant only at the 10% level. An increase in AER2 is also associated with a lower decrease in CDS spreads when compared to AER1 and AUAUL. In contrast to AER1, AER2 reflects the encumbrance of OBS collateral. This finding could point to a more negative perception of on encumbrance of off-balance sheet collateral compared to on-balance assets. High levels of encumbered OBS collateral are characteristic of investment banks which engage in matched book trading, the activity of carrying large volumes of repos and reverse repos, effectively re-using collateral received to finance repo liabilities.

In contrast to asset encumbrance ratios, the coefficients for capital and liquidity ratios turn out to be insignificant in all models. Variables such as asset encumbrance or asset quality seem to provide more valuable information on bank risk than capital and liquidity ratios. These results are consistent with recent literature pointing to a limited market reliance on capital and liquidity ratios to account for overall bank risk. Ötker-Robe and Podpiera (2010) find no significance in capital and liquidity ratios over the period 2004–2008 using a sample of 29 European Large Complex Financial Institutions (LCFI). Chiaramonte and Casu (2013), using a sample of 57 mostly European banks, find no statistical significance for Tier 1 and leverage and a limited statistical significance of liquidity ratios. Hasan et al. (2015) also find no statistically significant relation with the Tier 1 capital or liquidity ratios using a sample of 161 global banks in 23 countries. Kanagaretnam et al. (2016) find in a sample of 27 U.S. Bank Holding Companies (BHC) that the capital ratio is not significantly related to CDS spreads.

The coefficients for the ratio of loan-loss reserve to gross loans ratio are all positive and highly significant. The higher this ratio, the lower the quality of the loan portfolio; and therefore an increase in loan loss reserves should lead to an increase in CDS spreads. This result is consistent with Hasan et al. (2015) who also find a positive relationship between CDS spreads and the loan loss provision ratio.

The coefficients for the ratio of unreserved impaired loans to equity are all negative and significant on a 10% level. Banks with a higher value of this ratio exhibit higher impairments that have not been provisioned. Thus, this result implies that investors are not excessively concerned with such impairments. Chiaramonte and Casu (2013) also obtain this inverse relationship.

We observe opposing signs on the effects of ROA and ROE. The coefficients on ROA are all negative and highly significant. A negative sign for ROA could point to investors perceiving banks with a lower level of operating income relative to a level of investment as riskier. The coefficients for ROE are positive and highly significant, pointing to increased perceived default risk in institutions with higher profitability relative to their capital base. This result is somewhat surprising. Given the subdued profitability in traditional lending businesses in Europe, this finding could point to concerns by markets with banks that engage in highly profitable activities such as trade finance, invoice discounting or securities lending, with a comparatively low capital base.

The coefficient on the ratio of central bank exposure to total assets turns out to be positive and significant, implying that reliance on central bank funding is positively associated with bank risk. Not surprisingly, credit quality is also strongly associated with lower CDS spreads. Negative and highly significant coefficients are obtained across all models. The coefficient for size turns out to be negative, pointing out to a size advantage. The ratio of off-balance sheet items to total assets and business model variables are not statistically significant.

2.5.2.2. Mediating effects

Our second set of regressions explores the relationship between CDS spreads and key variables, including the interactions of asset encumbrance metrics with CAMEL, control variables, GIIPs and Nordic countries dummies. The results are presented in table 2.5. All models include the individual (non-interacted) CAMEL and control variables but for clarity these are not shown since the coefficients are very much in line with those presented in table 2.4.

Table 2.5. Results, mediating effects.

	Dependent variable: 5-year implied CDS spreads		
	(1)	(2)	(3)
Asset Encumbrance (AE)			
AER1	-0.712***		
	(0.19)		
AER2		-0.605***	
		(0.21)	
AUAUL			-0.552
			(0.48)
CAMEL variables (interactions)			
AE * Tier 1 capital ratio	-5.500	-5.182	3.596
	(4.97)	(4.83)	(2.95)
AE * Leverage Ratio	15.409**	15.180**	2.751
	(6.03)	(6.42)	(4.81)
AE * Net loans / deposits and ST Funding	0.522***	0.554***	0.172
	(0.16)	(0.17)	(0.34)
AE * Liquid assets / TA	0.845**	0.978**	-0.358
	(0.38)	(0.41)	(0.51)
AE * Loan Loss Reserves / Gross Loans	-12.749***	-11.524***	-8.465***
	(3.66)	(3.49)	(3.15)
AE * Unreserved impaired loans / equity	0.650**	0.595**	0.553
	(0.29)	(0.30)	(0.57)
AE * ROA	15.780	28.157	-52.962
	(29.44)	(36.22)	(67.71)
AE * ROE	-0.430	-1.059	2.658
	(1.41)	(1.83)	(4.32)
Control variables (interactions)			
AE * Central Bank Exposure / TA	10.914**	8.900*	9.887
	(4.16)	(4.65)	(10.08)
AE * Off-Balance Sheet Exposures / TA	-0.012	-0.119	3.054
	(0.82)	(0.81)	(2.15)
AE * Investment grade	0.348*	0.351*	0.317
	(0.20)	(0.20)	(0.19)
AE * BHC and commercial banks	0.048	0.146	0.404
	(0.19)	(0.23)	(0.46)
AE * Saving and cooperative banks	0.234	0.262	0.539
	(0.18)	(0.19)	(0.32)
AE * Size	0.125	0.073	-0.061
	(0.16)	(0.16)	(0.29)
AE * GIIPS	0.671***	0.466**	0.068
	(0.17)	(0.22)	(0.41)
AE * Nordics	-0.466**	-0.712**	-1.236***
	(0.21)	(0.27)	(0.41)
CAMEL and control variables	YES	YES	YES
Country FE	YES	YES	YES
Observations	367	367	367
R-squared	0.81	0.81	0.80

All regressions are estimated using OLS. A constant is included, but its coefficient is left unreported. Country-specific dummies are included in all models and country-business model clustering is applied in all models. ***: Significant at a 1% level; **: Significant at a 5% level; *: Significant at a 10% level.

We first discuss models 1 and 2 together as they yield very similar results. The stand-alone coefficients of asset encumbrance ratios (AER1 an AER2) are negative and significant. Several coefficients of the interacted CAMEL and control variables are significant, pointing to the existence of mediating effects in the relationship between asset encumbrance and CDS spreads. We first discuss the results for the interactions with control variables followed by CAMEL variables.

The coefficients for the interaction of asset encumbrance with the GIIPS and Nordic country dummies are significant but present opposite signs. GIIPS and Nordic countries present, on average, the highest levels of asset encumbrance in our sample. For GIIPS, the positive sign of the interaction term may reflect the negative perceptions of investors towards banks with high asset encumbrance levels in these countries. The coefficient, however, is not large enough to offset the negative relationship between asset encumbrance and the bank risk arising from the main effect. For Nordic countries, the negative coefficient would reflect a positive perception towards asset encumbrance arising from the issuance of covered bonds, considered very safe investments.

A positive and significant coefficient is obtained for the interaction with the ratio of central bank exposure to total assets. High asset encumbrance levels in banks with high amounts of central bank funding, as in Dexia's case, are negatively perceived by investors. The positive effect of the interaction term, for banks with high levels of central bank exposure (the maximum of which is 0.12 in our sample), offsets the negative effect of the stand-alone asset encumbrance coefficient, thus making higher levels of encumbered assets detrimental in absolute terms. The overall effect in AER1 would be -10+10.914*0.12=-0.869 whereas for AER2 it would be -10+8.9*0.12=-8.932.

A negative and significant coefficient is found for the interaction with the ratio of loan loss reserves to gross loans. Although higher loan loss reserves may point to a lower quality of the loan portfolio, excess reserves may signal a lower probability of incurring unexpected losses in the future and may therefore be perceived positively by markets. A positive and significant coefficient, however, is found for the interaction with the unreserved impaired loans to equity ratio. This could point to concerns by investors in banks with large amounts of encumbered assets that lack the reserves to deal with future

loan defaults. This finding is in line with the demise of Banco Popular, which presented high levels of asset encumbrance and simultaneously impaired loans.

The coefficients for the interactions of asset encumbrance with the Tier 1 capital and leverage ratios have conflictive signs, negative and positive, although the former turns out to be not significant. The leverage ratio is a non-risk-based measure of capital adequacy. A high value of the leverage ratio accompanied by larger amounts of encumbered assets could point to increasing risk in the loan portfolio which in turn would point to higher overall bank risk.[25] Similarly to central bank exposure, the positive effect of the interaction term offsets the negative effect of the stand-alone asset encumbrance coefficient, thus making higher levels of encumbered assets detrimental in absolute terms. In this case, the overall effects in AER1 and AER2 are -10+15.409*0.18=-7.226 and 10+15.18*0.18=-7.26 respectively.

The coefficients of the interacted liquidity variables are both positive and significant. A high number of net loans to deposit ratios points to a higher reliance on wholesale funding, which would make high levels of encumbered assets particularly detrimental. A positive sign is also found for the interaction with the liquid assets to total assets ratio. If investors are unable to identify the specific assets that are being encumbered, high levels of liquid assets accompanied by higher levels of encumbered assets could point to those liquid assets being encumbered, as in Dexia's case. The interactions of profitability variables (ROE and ROA) with asset encumbrance ratios are not significant.

Model 3 presents the results for AUAUL. While the stand-alone coefficient of AUAUL ratio turns out to be not significant, the coefficients corresponding to the interaction with the loan loss reserves to gross loans ratio and the Nordics dummy are both significant and of a negative sign. Consistent with the results of models 1 and 2, banks with high levels of asset encumbrance and with high levels of loan loss provisions, or based in Nordic countries, could benefit from increasing their levels of asset encumbrance. The effects of the remaining interacted variables are less significant but almost all, including central bank exposure and GIIPS, conserve the same sign found for models 1 and 2.

---

[25] As noted in footnote 6, Banco Popular had a relative high value of the Basel III leverage ratio (5.31%) compared to a weighted average for European banks of 5.2% as of year-end 2016.

2.6 Conclusion

Asset encumbrance has been a much discussed subject in recent literature and policymakers have been actively addressing what some regulators consider to be excessive levels of asset encumbrance. Still, the question of whether asset encumbrance is as perverse as it is portrayed arises. The risks of asset encumbrance because of the structural subordination of unsecured creditors, or because of the reduction of a bank's capacity to obtain funding, may end up being a concern for a subset of banks only. Other banks, on the other hand, could be signalling their overall 'health' by, and thus benefit from, increasing asset encumbrance levels (issuing covered bonds to private investors, for instance).

Our descriptive analysis shows a wide disparity in mean encumbrance levels across countries with southern European (GIIPS) and Nordic countries presenting higher mean encumbrance ratios than the sample average. Banks within the most and least creditworthy-rating categories present the lowest mean AER1 and AER2 ratios. For AUAUL, however, it is only banks in the least creditworthy categories that present the lowest mean values. Mean encumbrance levels tend to increase with bank size, measured in terms of total assets, across all ratios which could be explained by the substantial costs of securitisation issuance for smaller banks. By type of institution, saving banks show the lowest levels for both AER1 and AUAUL whereas cooperative banks show the highest average level of asset encumbrance when measured by AUAUL.

Our empirical analysis provides, to the best of our knowledge, the first investigation of the relationship between asset encumbrance and bank risk. We show that asset encumbrance is, on average, negatively associated with bank risk across different asset encumbrance measures. We also find that ratios measuring asset encumbrance or asset quality provide more valuable information on bank risk than capital and liquidity ratios, which is consistent with the recent literature pointing to a limited reliance by markets on capital and liquidity ratios to account for overall bank risk.

We also show that certain bank-level variables play a mediating role in the relationship between asset encumbrance and bank risk: for banks that have a high exposure to the central bank, high levels of unreserved impaired loans, high leverage ratio and/or located

in southern Europe, larger amounts of encumbered assets and encumbered OBS collateral are less beneficial and could even be detrimental in absolute terms. Banks with high levels of loan loss provisions and/or based in Nordic countries, in contrast, benefit from increased levels of asset encumbrance. Banks in Nordic countries rely to a large extent on covered bonds, which are perceived as a very safe investment and source of funding. These results suggest that regulators need to be cautious before leaping to all-encompassing conclusions when assessing the effects of asset encumbrance levels.

Chapter 3: Size, growth and bank dynamics

3.1 Introduction

The European banking industry has experienced significant changes during the last decades. The liberalization and deregulation processes during the 1980's and the 1990's, the technological progress, the introduction of a single currency and the consolidation process experienced by the industry both before and after the financial crisis all have led to a progressive decline in the number of institutions and increased concentration of banking assets among the largest banks. Over the period 1998-2005, the number of banks in the European Union fell from approximately 9200 to slightly more than 7100 whereas the average share in total assets of the 5 largest institutions however, rose from 50.3% to 54.2% (ECB 2002; 2006). Market concentration continued this upward trend during the financial crisis with the number of credit institutions in the euro area declining to 5,475 at the end of 2015, from 6,767 at the end of 2008 (ECB, 2016).

Structural, regulatory and technological changes have a significant impact on the size distribution and dynamics of the banking industry. However, few studies have examined the dynamic nature of the size distribution in banking. This is true even though the size distribution of banks is of fundamental interest to bankers and policy makers given size is a key determinant of bank's risk taking (Demsetz and Strahan, 1997), credit availability (Peek and Rosengren, 1996), lending relationships (Stein 2002; Berger et al., 2005), or lending specialization (Delgado et al., 2007).

In this chapter, we investigate the size distribution of Spanish banks from a dynamic perspective over the 1970-2006 period. The case study is particularly relevant because over this period, Spain had significant diversity in bank size and ownership forms including commercial for profit banks, not-for-profit savings banks, and credit cooperatives, with no clear market dominance of one form over the others. The Spanish banking industry was also no exception to the European deregulation and consolidation processes (Caminal et al. 1993; Salas and Saurina, 2003). The result has been an intense merger activity during the last three decades involving banks of all sizes and ownership forms. Until the financial crisis, ownership and size diversity increased over time with

savings banks and credit cooperatives representing half of the market share of Spanish retail banking by 2006, compared with less than one third in 1970. The time period considered offers a unique opportunity to study the evolution of the size distribution from a period when Spanish banking was strongly regulated (with administratively fixed interest rates, compulsory investment coefficients, foreign entry restrictions and high asymmetry among different ownership forms), to a period in which competitive conditions press towards efficient decisions and savings banks outperform commercial banks in competitive credit markets.

To investigate the evolution of the size distribution, we focus on the Law of Proportionate Effect (LPE). According to the LPE, firm growth and industry concentration are driven by unsystematic and random factors. This hypothesis, which was originally devised by Gibrat (1931)[26], implies that firm size and growth are unrelated, so that small and large firms share the same chance of growing at any given rate in any particular period. If the LPE holds for a particular industry, then concentration is expected to increase over time and the empirical size distribution will tend to become highly skewed[27]. Our approach is based on Tschoegl (1983), who investigates the relationship between size and growth rate of large international banks by testing three propositions that all together make up a 'strong form' of the LPE: i) the growth rate of each bank is independent of its size; ii) there is no persistence in bank's growth in two consecutive periods; and iii) the variability of growth rates is independent of the banks' size.

Only a few studies have examined the size-growth relationship in the banking industry and the existing evidence is not exempted from limitations. First, the vast majority of the existing studies estimate cross-sectional regressions of growth over some period as dependent variable, and initial size and (in some cases) persistence in growth rates as explanatory variables. This procedure often fails to exploit all the available information whenever data is available in a time series format. In addition, Goddard et al. (2002a) show that cross-sectional parameter estimates are biased towards LPE acceptance and the test suffers from a loss of power if there are heterogeneous individual firm effects. As a result, a few recent studies have incorporated panel data-based tests to investigate the size-growth relationship in banking. Second, the existing studies only consider banks that

---

[26] The LPE is also commonly known as Gibrat's Law.
[27] Sutton (1997) provides a simple proof of the argument using the lognormal distribution.

survive during the entire estimation period, excluding institutions that failed, merged or were acquired. If banks need to achieve a minimum size to fully exploit economies of scale, small banks may prioritize growth in order to attain this scale. Therefore, average growth among small banks may tend to be higher than among large banks. In addition, this growth prioritization may make small banks to exit with higher probability than their larger counterparts if they fail to attain the minimum size in a short period of time. In fact, several studies show that exit rates tend to decline with size in different industries[28]. Thus, this procedure may generate a sample censoring bias which may have important implications for the testing results[29]. Third, previous studies cover a relatively small proportion of the banking system, so their samples are unlikely to be perfectly representative of the population with respect of relevant characteristics such as size or ownership form. Finally, existing studies usually cover short periods of time, so they are unable to investigate how the size distribution evolves over time through the liberalization, deregulation and integration processes.

This chapter attempts to fill the previous gaps by investigating the size-growth relationship using an unbalanced panel comprised by the whole population of Spanish commercial, savings and cooperative banks. We test Tschoegl's three propositions over the period 1970-2006 and over different sub-periods corresponding to different states in the deregulation and integration processes. The size-growth relationship is estimated using the panel unit root test devised by Breitung and Meyer (1994) and an Ordinary Least Squares (OLS) test. Using Monte Carlo simulations, Bond et al. (2005) show that these two tests are the most powerful for testing unit roots in micro panels across a set of alternative tests. Since our panel includes institutions that exit over the estimation periods, the potential sample selection bias is eliminated.

We find evidence that when the banking system was highly protected, smaller banks grew faster than large banks. However, as deregulation and integration processes occur, larger banks tend to grow at a faster pace. In recent years, however, we find some evidence that larger banks tend to grow at higher rates. Thus, our study corroborates the conditional nature of the size-growth relationship and the size distribution.

---

[28] See Hall (1987), Evans (1987a; 1987b), Dunne et al. (1989), Dunne and Hughes (1994) and Fariñas and Moreno (2000).

[29] See Sutton (1997) for an exhaustive explanation of the sample censoring bias.

The rest of this chapter is organized as follows. Section 3.2 briefly summarizes the existing literature and states the main propositions to be tested. Section 3.3 contains an overview of the transformation of the Spanish banking system and the deregulation process it has undergone. Section 3.4 explains the data used and the econometric methodology. Section 3.5 presents the empirical results. Finally, section 3.6 concludes.

3.2 Literature overview

The majority of explanations on increasing levels of concentration and bank scale rely on systematic differences across institutions that lead banks obtaining advantages over others. These advantages may include economies of scale and scope, efficiency gains attained through size or the adoption of entry-deterring strategies (Berger et al., 1993; Berger et al., 1999).

A plausible alternative explanation rests on the assumption that industry concentration may tend to emerge over time naturally, as a result of random growth. Systematic factors do not play a differentiating role since they are distributed randomly over the population of firms. This idea is embodied in the Law of Proportionate Effect (LPE) and was originally devised by Gibrat (1931). According to the LPE, firm size is unrelated to growth, implying that small and large firms share the same chance of growing at any given rate in any particular period. The factors that influence firm's growth, such as growth of demand, managerial talent, innovation or the organizational structure are distributed across firms in a manner which is essentially random. Over time this stochastic process will generate a size distribution of firms which exhibits a positive skew, with a few large firms, rather more medium-sized firms, and a large number of small firms. Previous research has found that this process accords well with the actual size distribution of firms observed in many industries. Gibrat (1931) investigates the implications if each firm's growth in any year is determined randomly, and is therefore independent of its size and its growth in previous years. He shows that this non-relationship between growth and firm size has important consequences for changes in concentration and market structure over time. If the LPE holds for a particular industry, the size distribution will tend to become skewed, increasing concentration over time.

The following stochastic model results appropriate to describe observed growth or decline of firms as the result of the chance product of a large numbers of factors acting independently of each other:

$$S_{it} = S_{it-1}^{\beta} \exp(\mu_{it}) \tag{3.1}$$

where S_{it} is the size of firm i at time t, β is the parameter of the size effect, and μ_{it} is firm i's draw from the common distribution of growth rates, and is hypothesized to be distributed $\mu_{it} \sim N(\alpha_i + \delta_t, \sigma^2)$, and hence $\mu_{it} = \alpha_i + \delta_t + \varepsilon_{it}$.

In his study of size and growth among the world's largest banks, Tschoegl (1983) identifies three testable propositions which all together make up a 'strong form' of the LPE and constitute the hypothesis that all the banks' growth rates follow a random walk and are identically distributed.

P.1. The growth rate of each bank is independent of its size.

If P1 holds, the size distribution will become highly skewed and concentration will increase over time. This result implies β=1, and is strictly due to the workings of chance, requiring no assumptions about monopoly advantages, economies of scale or managerial effectiveness. Diseconomies of scale or political costs to growth in heavily regulated periods may cause smaller banks to grow faster on average than large banks resulting in β<1. Over time the size of all banks would tend to converge towards some value that might be bank specific if there is heterogeneity in α_i. In liberalisation periods, however, efficiency advantages arising from scale and scope economies and geographical growth orientation in the larger banks could cause large banks to grow quicker than smaller banks resulting in β>1.

P.2. There is no persistence in bank's growth in two consecutive periods.

P2 can be restated as $\text{cov}(\varepsilon_{it}, \varepsilon_{it-1}) = 0$. Whenever P2 does not hold, a positively serially correlated growth indicates that advantages acquired in one period carry over to the next. On the contrary, consistent negatively serially correlated growth implies the existence of

some process that systematically reverses fortune. Persistent above average growth could be found in banking due to technological improvements or high switching costs.

P.3. The variability of growth rates is independent of the banks' size.

P3 corresponds to $E(\varepsilon_{it}^2) = \sigma_t^2$. Low variability in growth rates may be explained by diversification or scale advantages that reduce uncertainty. Heteroscedastic growth rates, however, imply that these advantages are contingent on bank's size. If size diversification advantages play a role, for example due to geographical diversification, one could expect smaller banks to experience more variability in growth rates than larger banks. As Tschoegl (1983) points out, P3 has important implications for public policy and the safety net. According to Merton (1977), a deposit insurer effectively issues a put option, whose social cost depends on the variance rate per unit of time for the logarithmic changes in the value of assets. If banks reduce (increase) the variability of their growth rates one would expect that this would reduce (increase) their risk other things being equal.

If the above three propositions hold, taking logs and rearranging equation (3.1) becomes

$$\ln S_{it} - \ln S_{it-1} = \alpha_i + \delta_t + \varepsilon_{it} \tag{3.2}$$

implying that logarithmic bank growth rates follow a random walk with drift.

There is extensive research looking at the size-growth relationship but evidence for the banking industry is scarce. Alhadeff and Alhadeff (1964) compare the growth of the largest 200 US commercial banks over the period 1930-1960. The authors find that smaller banks tend to grow faster than their larger counterparts. However, they also show that top banks that survived throughout the sample period grew faster than the system as a whole and attributed this to mergers among the largest banks. Rhoades and Yeats (1974) and Yeats et al. (1975) analyse the size and growth relationship for different sized US banks over the periods 1960-1971 and 1960-1963 respectively. They find that larger banks tend to grow more slowly than small banks. Tschoegl (1983) investigates the size-growth relationship for a sample of the 100 larger international banks in each year. His analysis suggests that size and growth are unrelated, that growth rates variability declines

with size and that there is a positive but insignificant relationship between growth rates in subsequent periods. More recently, Wilson and Williams (2000) investigate the relationship between size and growth for a sample of European banks from France, Germany, Italy and the UK over the period 1990–1996. Their sample comprises about 100 commercial, savings and cooperative banks per country, so they cover a very small proportion of the banking industry for each country. The authors find no relationship between bank size and growth for France, Germany and the UK but small banks seem grow faster than large banks in Italy. Smaller banks are also found to have experienced more variability in growth rates than larger banks. This suggests that diversification advantages may make large banks less susceptible to abnormal high fluctuations in growth. Janicki and Prescott (2006), in a study similar in spirit to ours, consider the LPE and study the dynamics of the bank size distribution using a sample of US commercial banks using data from 1960 to 2005. Their findings are broadly supportive of the LPE for the whole sample. Interestingly, when the sample is broken into different size categories, they find evidence for acceptance of the LPE for the 1960s and 1970, before the deregulation and liberalization process, but not for the 1990s and the 2000-2005 periods, where largest banks seem to grow faster than the small banks.

All the previous studies estimate cross-sectional regressions of growth over some period as dependent variable and initial size and (in some cases) persistence in growth rates as explanatory variables. However, this procedure often fails to exploit all the available information whenever data is available in a time series format. Goddard et al. (2002a) also show that cross-sectional parameter estimates are biased towards LPE acceptance and the test suffers from a loss of power if there are heterogeneous individual firm effects. As a result, recent studies have incorporated panel data-based tests to investigate the size-growth relationship. Goddard et al. (2002b) investigate the growth of US credit unions during the 1990s using cross-section and panel techniques. They find that small credit unions grew slower and tended to have more variable growth than their larger counterparts. They also find negative persistence in growth rates. This may suggest that credit unions curtail growth until they retain sufficient earnings to satisfy capital requirements. The authors also conduct multivariate analyses and show that several structural and operational variables may also help to explain future growth. Goddard et al. (2004) study growth and profitability in a sample of 583 commercial, savings and cooperative banks operating in France, Germany, Italy, Spain and the UK over the period

1992 to 1998. Their results show no evidence of any relationship between size and growth, but they find a positive persistence in growth rates. They also find evidence that profit is a good predictor of future growth.

3.3 Deregulation and consolidation in Spanish banking

During the last 30 years, the European countries have implemented several regulatory changes that have transformed the European banking industry. Several efforts have been put forward to reach the necessary harmonization level required for the establishment of a unique and competitive European market of financial services. The deregulation processes of the 70s and 80s, the liberalization of capital flows, the Second Banking Directive in the early 90s, and the establishment of the Euro in several Member States, definitely contributed to boost this process. The Financial Services Action Plan (FSAP), created in 1999 dismantled most legal barriers to provide a favourable environment for the development of a Single European Financial Market, leading to the introduction of the Markets in Financial Instruments Directive (MIFID) in 2004.

Spain has not been an exception to the deregulation and consolidation processes (Caminal et al., 1993; Salas and Saurina, 2003). As a result, the Spanish banking system has undergone through a deep transformation during the last four decades. Overall, deregulation has transformed both the behaviour of banks (liberalization of interest rates, fees and commissions…) and the industry structure and distribution (entrance of new competitors, removal of artificial barriers between commercial, savings and cooperative banks…). A timeline of events in the Spanish banking sector is provided in table 3.1.

In the 1970s, the situation corresponded to a strongly regulated oligopoly, with administratively fixed interest rates, compulsory investment coefficients and entry restrictions to foreign banks. The high risk concentration of banks as well as an investment policy not guided by efficiency and profitability, together with the economic crisis of the mid-70s, made banks to face serious troubles. During this period, the relative importance of public banking increased due to the financing needs of local governments and households after the crisis. Two consequences for the industry can be remarked from this episode: First, banks decreased their stake in industrial portfolios and increased their participation in government securities. Second, banks transferred the cost of the crisis to

final clients by increasing the spread between asset and liability rates. Despite all these problems, this period was essential for the transformation of the Spanish banking industry in the following decades.

Table 3.1. Timeline of events in the Spanish banking sector

	Period			
	1970-1979	**1980-1989**	**1990-1999**	**2000-2010**
Regulatory developments	• Administratively fixed interest rates. • Compulsory investment coefficients. • Entry restrictions to foreign banks. • Commercial banks are allowed to open branches nationwide.	• Risk-based capital requirements are introduced. • Investment coefficients for commercial and savings banks are equalised. • Saving banks are allowed to open branches nationwide.	• Liquidity rules are liberalised. • Capital requirements are adapted to EU rules.	• European integration. • Introduction of the Euro. • Introduction of MiFID and passporting of services across the EU. • Set up of the Fund for Orderly Restructuring of the Banking Sector (FROB) after the crisis.
Industry trends	• Economic crisis of the mid-70s makes banks to increase asset-liability spreads. • Banks decrease their stake in industrial portfolios and increase purchase government debt.	• Expansion of commercial banks.	• Intense merger activity across commercial banks and progressive consolidation among savings and small to medium size commercial banks. • Expansion of savings banks in deposits and geographically.	• Gradual entrance of foreign banks. • Progressive consolidation accelerated.

During the 1980s, a series of important deregulation measures were introduced. In 1985, capital requirements as a function of asset risk are enacted, savings banks are free to open branches in their traditional regions, and the investment coefficients for commercial and savings banks are equalized. In 1987, interest rates and controls on fees are liberalized. In 1989, a calendar to phase out the investment coefficient is settled and savings banks

are allowed to open branches nationwide, ten years later than commercial banks. In 1990, liquidity rules were liberalized. In 1992, capital adequacy requirements were adapted to the EU rules.

All these policy measures led to a highly liberalised situation in the early 1990s and a more competitive arena. Consequently, the industry suffered a very intense merger activity in the following years aimed to increase size for competing in the European markets and preserve market power. Until 1989 there were some punctual peaks in M&A over the 70s. However, the strong consolidation process starts in the early 90s, when the liberalization process is finished. This trend can be characterized by four facts (Gual, 1993). First, M&A increase the concentration of the industry, although the strong growth of small and medium-sized institutions and the entrance of new competitors tend to reduce it. Second, this process only involves Spanish banks aiming to improve efficiency through national scale economies. Political forces can limit the entrance of foreign institutions to keep under control a sector which is considered strategic. Third, the consolidation rationalizes the sector structure, leading to the closure of several branches that became redundant due to increased competition. Finally, during the deregulation process the main target entities were small to medium size banks, and large M&A operations were avoided.

A significant expansion of savings banks occurred in the 1990s. The deregulation measures introduced in the 80s, led Spanish banks to a significant interest in diversifying their risk, achieve scale economies and defend their regional market shares (Gual, 1993). The result is an accelerated M&A process in the early 1990s, basically involving savings banks that operated in the same geographical locations. As a result, the number of saving banks operating in Spain drop from 78 in 1989 to 54 in 1992. In the following years, the amount of deposits held by savings banks raised considerably and exceeded the deposits of commercial. On the lending side, however, the differences between commercial and savings banks vanished later, over the 2000-2006 period. By the end of the century, public banking in Spain had virtually disappeared. In 2000 there is another peak which corresponds to an important period of consolidation in Spanish commercial banking led by the strong economic period and the liquidity holdings that Spanish banks accumulated in previous years. In 1999, the two largest Spanish commercial banks (BSCH and BBVA) are created as a result of two national mega mergers.

From 2000 onwards, consolidation has been prevalent. Prior to the financial crisis, over the 2000-2006 period, the European integration process and introduction of the Euro governed consolidation. Even though market integration in Spain took place mostly indirectly through interbank flows and cross-border investments, a gradual entrance of foreign banks through the opening of subsidiaries and branches benefiting from the MiFID passport occurred. The strong regional presence and significant market share of savings and cooperative banks however made it difficult for foreign banks to establish in Spain. M&A activity remained high during the whole period. After 2008 and given the severity of the crisis, a major program of financial sector reform was launched with the intervention of national and supranational authorities. The Fund for Orderly Restructuring of the Banking Sector (FROB) was set up and several mergers were encouraged through the provision of financial assistance and bail-outs. The FROB financed eight mergers and the aid granted, totaling 11,559 million euros, was lost in its entirety (Maudos and Vives, 2016). A key structural change was the restructuring of savings banks into for-profit commercial banks in 2012.

3.4 Data and methodology

3.4.1. Data

To investigate the dynamics of the bank size distribution, we use a dataset which includes individual bank level data drawn from the confidential balance sheet statements declared annually (in December) to the Banco de España by the whole population of Spanish banks over the period 1970-2006. Reliable data for cooperative banks is available only since 1980, when they started to be legally obliged to declare their statements. Thus, we conduct two different analyses. First, we study the evolution of the size-growth relationship over the 1970-2006 period for commercial and savings banks excluding cooperative banks. Then, we conduct the analysis for commercial, savings and cooperative banks over 1980-2006. This approach allows us to analyse the influence of cooperative banks on the evolution of size distribution through the liberalization process and exploit all the available information. Our final dataset is an unbalanced panel comprising all the institutions that enter and exit over the different estimation periods[30].

---

[30] Although our unbalanced panel includes banks that exit, the exit event is not included 'per se' in the estimations. Thus, our analysis should be interpreted as conditional on the survival of banks in the different estimation periods.

We use three measures of size: total assets less contra-accounts (acceptances, letters of credit, securities held on behalf of customers…), total loans and total deposits. All data is in nominal terms and is converted to Euros. Table 3.2 shows the summary statistics for assets, loans and deposits nominal size measures for the years 1970, 1980, 1990, 2000 and 2006. The total population of commercial banks increases during the first two decades, but decreases over the 90s and 2000s. The population of savings and cooperative banks, however, decreases along the years considered. As can be expected, nominal size figures increase considerably along time. The average size of savings banks is much higher than that of commercial banks during the last two decades, while the contrary occurs in the 70s and 80s. Average size of cooperatives is much lower.

Table 3.3 reports summary data on growth for the three size measures considered. These data refers to banks that survived the end-point of each sub-period considered. The higher average growth rates are reached over the 1970s and 1980s in assets and deposits and over the 1970s and 1990s in loans. Commercial and savings banks show the highest and the lowest standard deviations in growth rates respectively.

3.4.2. Methodology

Taking the log of equation 3.1 yields the following stochastic growth model for observations of bank size and growth:

$$s_{it} - s_{it-1} = \alpha_i + \delta_t + (\beta - 1)s_{it-1} + \varepsilon_{it} \qquad (3.3)$$

where s_{it} is the log size of bank i at time t. α_i and δ_i allow for individual bank and time effects respectively. The parameter β determines the relationship between size and annual growth. ε_{it} is a random disturbance, normal and IID with $E(\varepsilon_{it}) = 0$ and $var(\varepsilon_{it}) = \sigma_\varepsilon^2 > 0$.

Chesher (1979) shows that if a lagged growth term is not included in equation (3.3) then the estimate of β will be inconsistent. Thus, we modify the equation to take into account the possibility of first-order serial correlation:

$$s_{it} - s_{it-1} = \alpha_i + \delta_t + (\beta - 1)s_{it-1} + \rho\varepsilon_{it-1} + u_{it} \qquad (3.4)$$

Table 3.2. Descriptive Statistics. Nominal Size.

	1970			1980			1990			2000			2006		
	N	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.
Assets															
Total	*199*	71	182	*361*	300	908	*331*	1387	4025	*277*	4053	17338	*272*	8750	31088
Commercial Banks	*111*	86	225	*125*	602	1411	*158*	1891	5013	*139*	4811	22777	*142*	9251	37391
Savings Banks	*88*	51	100	*81*	365	570	*66*	1788	2361	*48*	7154	10593	*47*	17137	24151
Credit Unions	-	-	-	*155*	22	45	*107*	446	2879	*90*	1656	7806	*84*	4232	20486
Loans															
Total	*199*	36	106	*361*	154	502	*331*	632	1746	*277*	2245	8371	*272*	5875	19120
Commercial Banks	*111*	51	138	*125*	340	800	*158*	918	2302	*139*	2549	10804	*142*	5624	21760
Savings Banks	*88*	19	29	*81*	143	216	*66*	913	1412	*48*	5091	8131	*47*	15224	24483
Credit Unions	-	-	-	*155*	11	20	*107*	48	117	*90*	335	780	*84*	1078	2812
Deposits															
Total	*199*	51	139	*361*	204	653	*331*	836	2469	*277*	2248	8512	*272*	4613	15699
Commercial Banks	*111*	59	172	*125*	379	1016	*158*	984	2914	*139*	2219	10612	*142*	3946	17899
Savings Banks	*88*	40	75	*81*	293	440	*66*	1688	2945	*48*	5881	8658	*47*	12950	19163
Credit Unions	-	-	-	*155*	16	31	*107*	91	203	*90*	415	966	*84*	1074	2680

Figures in millions of euros.

Table 3.3. Descriptive Statistics. Growth.

	1970-1979			1980-1989			1990-1999			2000-2006		
	N	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.
Assets												
Total	*201*	1.982	1.007	*337*	1.222	1.221	*285*	0.922	1.132	*272*	0.806	1.266
Commercial Banks	*119*	2.099	1.098	*149*	0.871	1.692	*142*	0.871	1.570	*142*	0.898	1.726
Savings Banks	*82*	1.813	0. 835	*78*	1.541	0.372	*51*	0.964	0.300	*47*	0.807	0.175
Cooperative Banks	-	-	-	*110*	1.470	0.556	*92*	0.979	0.351	*83*	0.646	0.340
Loans												
Total	*201*	2.109	0.991	*337*	1.084	1.410	*285*	1.161	1.635	*272*	0.935	2.046
Commercial Banks	*119*	2.302	1.184	*149*	0.570	1.849	*142*	0.960	2.303	*142*	0.946	2.901
Savings Banks	*82*	1.834	0.511	*78*	1.575	0.568	*51*	1.253	0.248	*47*	1.013	0.184
Cooperative Banks	-	-	-	*110*	1.403	0.838	*92*	1.406	0.462	*83*	0.869	0.256
Deposits												
Total	*201*	1.280	1.387	*337*	1.228	2.384	*285*	0.521	2.736	*272*	0.948	2.914
Commercial Banks	*119*	2.036	1.359	*149*	0.810	3.402	*142*	0.079	3.797	*142*	1.248	3.983
Savings Banks	*82*	0.184	0.050	*78*	1.619	1.210	*51*	0.919	0.229	*47*	0.748	0.297
Cooperative Banks	-	-	-	*110*	1.518	0.569	*92*	0.998	0.355	*83*	0.565	1.047

Figures are based on annual logarithmic growth rates calculated for banks that are operative the last year of each period.

The majority of previous empirical studies estimate 3.4 using a cross-section approach to test the LPE. However, this procedure requires the assumption that specific bank effects are homogeneous ($\alpha_i = \alpha$). As Goddard et al. (2002a) point out, there is little basis in theory to anticipate homogeneity, and several reasons to expect heterogeneity a priori. Different banks may differ in the level of importance to size in their objective functions. If long run average costs are flat or inverted U-shaped, institutions with different objective functions can operate at different scales with similar average costs, and thus bank sizes may mean-revert towards different equilibrium values.

Breitung and Meyer (1994) and Goddard et al. (2002a) show that if there are heterogeneous bank effects but equation 3.4 is estimated assuming homogeneity in α_i, the resulting estimator of β is upward biased and inconsistent, and the test suffers from a dramatic loss of power. Intuitively, if log size is mean-reverting but the individual bank means are widely dispersed, the size distribution at *t*-1 conveys little information about which banks are above or below their own means at *t*-1, and therefore about which are expected to experience above or below average growth between *t*-1 and *t*. Accordingly, the cross sectional estimator will be close to unity. However, in this case it would be wrong to interpret this finding as supportive of the LPE, because the individual log bank sizes are stationary and mean-reverting, and there is no tendency for industry concentration to increase over time.

Recently, some papers have incorporated panel data techniques to investigate the size-growth relationship using panel unit root tests. This approach allows us to exploit all the available information from the dataset, and account for the possibility of heterogeneous bank effects.

For the purposes of panel estimation, we use the modified specification:

$$s_{it} - s_{it-1} = \alpha_i(1-\rho) + \delta_t + (\beta-1)s_{it-1} + \rho(s_{it-1} - s_{it-2}) + \eta_{it} \qquad (3.5)$$

where $\eta_{it} = \varepsilon_{it} + \rho(1-\beta)s_{it-2}$.

Equation 3.5 allows us to test the three propositions outlined in section 3.2. The analysis of the relationship between size and growth (P1) consists of testing the null hypothesis of $H_0 : \beta - 1 = 0$, with the alternative that $H_1 : \beta - 1 \neq 0$. As noted by Goddard et al. (2002a), the form of η_{it} does not present any problems to test H$_0$ because $\eta_{it} = \varepsilon_{it}$ under H$_0$. To do so, we use the panel unit root test devised by Breitung and Meyer (1994) and an Ordinary Least Squares (OLS) test. Using Monte Carlo simulations, Bond et al. (2005) show that these two tests are the most powerful for testing unit roots in micro panels across a set of alternative tests.

The hypothesis of no persistence in growth rates (P2) can be tested as $H_0 : \rho = 0$ under the alternative that $H_1 : \rho \neq 0$. Finally, following Goddard et al. (2002b), we investigate the existence of homoscedastic growth rates (P3) by applying a standard heteroscedasticity test to the residuals of each estimated equation. A Lagrange multiplier (LM) test based on an auxiliary regression of the squared residuals on the squared lagged size measure produces a heteroscedasticity test statistic distributed χ_1^2 under the null hypothesis of homoscedasticity. When the null is rejected, the sign of the estimated coefficient on the squared initial size measure in the auxiliary regression indicates the direction of the relationship between bank size and the variability of growth.

If we assume homogeneity ($\alpha_i = \alpha$), equation 3.5 can be estimated by pooling the data across banks and over time, and using OLS with time dummies to allow for fixed time effects. As long as this underlying assumption is correct, the resulting estimator $\hat{\beta}_{OLS}$ is unbiased. The OLS test is based on the corresponding t-statistic on $\hat{\beta}_{OLS} - 1$ and is asymptotically normal. However, if the individual effects are heterogeneous, the resulting estimator is again biased and inconsistent.

As a response, Breitung and Meyer (1994) modified the OLS test using the following transformation:

$$s_{it} - s_{it-1} = \alpha_i(1 - \rho) + \delta_t + (\beta - 1)(s_{it-1} - s_{i0}) + \rho(s_{it-1} - s_{it-2}) + \xi_{it} \qquad (3.6)$$

where $\xi_{it} = \eta_{it} + \alpha_i(1-\rho) + (\beta-1)s_{i0}$. The corresponding estimator is denoted by $\hat{\beta}_{BM}$ and is unbiased under the null that β=1. Under the alternative that β≠1, $\hat{\beta}_{BM}$ is upward biased, with a bias equal to $\beta + (1-\beta)/2$. However, its properties remain unaffected by heterogeneity in α_i.

3.5 Empirical results

To investigate the evolution of the size distribution, we report empirical results for different estimation periods. First, for commercial and savings banks, estimates for the full sample period 1970–2006, the sub-periods 1980-2006, 1990-2006, the decades 1970-1979, 1980-1989, 1990-1999 and the period 2000-2006 are provided. Then, we report estimates for the whole sample of commercial, savings and cooperative banks for the periods 1980-2006, 1990-2006, the two decades 1980-1989, 1990-1999 and the period 2000-2006. In the conducted estimations, the logarithmic size and growth variables based on the assets, deposits and loans size measures are calculated from nominal data. The effects of inflation as well as any factor which is common to all banks are captured entirely by the intercept and the time dummies. The estimation of the parameters is therefore unaffected by the choice between nominal or real size definitions.

3.5.1. Commercial and savings banks (1970-2006)

Table 3.4 reports the results of the tests for commercial and savings banks over the full sample period, 1970-2006, and the 1980-2006 and 1990-2006 periods for the three size measures considered. The intercept and the coefficients of the dummy variables are not reported. The estimate of β provides the test for P1, that growth rates and size are independent. For the 1970-2006 and 1980-2006 periods, both the OLS and Breitung-Meyer (B-M) tests reject the LPE for each of the measures considered, with the exception of the B-M test for the loans measure, which provides a non-significant estimate over the 1980-2006 period. A pattern signalling that small banks tend to grow faster than their larger counterparts emerges. If we consider the period 1990-2006 instead, the B-M tests are not significant for none of the size measures. Overall, the results tend to indicate an increasing rate of growth for larger banks relative to smaller banks through time.

Table 3.4. Estimation results for commercial and savings banks. Full sample, 1980-2006 and 1990-2006 periods.

	Assets			Loans			Deposits	
	OLS	B-M		OLS	B-M		OLS	B-M
1970-2006								
$\hat{\beta}$	0.9769**	0.97551**		0.9764**	0.9888*		0.9779**	0.9773**
	(-7.85)	*(-6.97)*		*(-6.12)*	*(-2.43)*		*(-5.81)*	*(-4.75)*
$\hat{\rho}$	0.0619**	0.0675**		0.0406**	0.0385**		-0.0380**	-0.0316**
	(5.53)	*(5.97)*		*(3.22)*	*(2.98)*		*(-3.28)*	*(-2.63)*
Heterosc.	-0.0024**	-0.0025**		-0.0065**	-0.0067**		-0.0169**	-0.0170**
	330.18	*27.85*		*61.84*	*63.81*		*192.60*	*140.23*
1980-2006								
$\hat{\beta}$	0.9756**	0.9646**		0.9783**	0.9989		0.9798**	0.9837**
	(-5.90)	*(-6.66)*		*(-4.45)*	*(-0.17)*		*(-4.64)*	*(-3.00)*
$\hat{\rho}$	0.0579**	0.0744**		0.0417**	0.0325*		-0.0310*	-0.0275*
	(4.30)	*(5.36)*		*(2.76)*	*(2.05)*		*(-2.26)*	*(-1.99)*
Heterosc.	-0.0046**	-0.0046**		-0.0111**	-0.0113**		-0.0204**	-0.0206**
	45.90	*46.88*		*91.82*	*92.11*		*132.68*	*134.15*
1990-2006								
$\hat{\beta}$	0.9780**	0.9927		0.9808**	1.0070		0.9826**	0.9872
	(-5.62)	*(-0.98)*		*(-3.28)*	*(0.80)*		*(-3.36)*	*(-1.52)*
$\hat{\rho}$	0.0314	0.0345		0.0977**	0.0833**		-0.0634**	-0.0608**
	(1.85)	*(1.88)*		*(4.82)*	*(3.78)*		*(-3.66)*	*(-3.27)*
Heterosc.	-0.0050**	-0.0052**		-0.0121**	-0.0124**		-0.0201**	-0.0202**
	66.19	*65.07*		*65.23*	*65.50*		*99.82*	*100.11*

**/* indicate that the coefficient is significantly different from zero at the 1%/5% level. OLS and B-M refer to the ordinary least squares and Breitung-Meyer unit root tests. *t*-statistics are shown beneath in brackets. The reported *t*-statistics for $\hat{\beta}$ refer to the coefficient (β-1). 'Heterosc' is the estimated coefficient on the lagged squared size variable in the auxiliary regression used to obtain the heteroscedasticity LM test statistic, shown beneath in italics. Intercept and time dummy variables are not reported.

The estimate of ρ provides the tests for P2, that growth rates in two consecutive periods are independent of each other. As noted in section 3.2, one would expect to find positive persistence in banking, implying that the advantages acquired in one period, carry over to the next. The assets and loans measures show positive and significant persistence, except for the assets size measure in the 1990-2006 period where no relationship between consecutive growth rates is found. On average, banks that achieve above average growth in the credit market in one period tend to grow faster in the next. However, the persistence in deposits is found to be negative in all the sub-periods considered. This result could be reflective of the intense and growing competition in Spanish deposit markets and would be consistent with the finding of a progressive increase in competition in Spanish (Carbó et al., 2005; Ayuso and Martínez, 2006) and European deposit markets (Bikker, 2003).

The conducted LM tests reject the null hypothesis of homoscedastic growth at the 1% level in all the estimations and periods. Overall, the negative estimated coefficients on the squared lagged size variables indicate an inverse relationship between size and the variability of growth. Large banks may benefit from diversification advantages that make them less vulnerable to abnormal high fluctuations in growth (Wilson and Williams, 2000). Similarly, Jovanovic (1982) argues that larger firms are likely to be older than smaller firms and may benefit from learning economies of scale enabling them to avoid costly mistakes. Tschoegl (1983) and Goddard et al. (2002b) find a similar result in their studies.

In order to further investigate the evolution of the size-growth relationship over time, the whole sample period is divided into four sub-periods: 1970-1979, 1980-1989, 1990-1999 and 2000-2006. As we pointed out in section 3.3, each decade virtually corresponds to different situations of the deregulation, liberalization and integration processes. The corresponding results are reported in table 3.5.

.

Table 3.5. Estimation results for commercial and savings banks. 1970-1979, 1980-1989, 1990-1999 and 2000-2006 periods.

	Assets			Loans			Deposits	
	OLS	B-M		OLS	B-M		OLS	B-M
1970-1979								
$\hat{\beta}$	0.9708**	1.0012		0.9704**	0.9948		0.9748**	0.9834
	(-8.36)	(0.16)		(-7.87)	(-0.51)		(-4.56)	(-0.96)
$\hat{\rho}$	0.0715**	0.0743**		0.1497**	0.1705		0.1575**	0.1965**
	(3.85)	(3.58)		(-6.81)	(6.89)		(4.54)	(5.12)
Heterosc.	-0.0012**	-0.0013**		-0.0014**	-0.0015**		-0.0032*	-0.0033*
	143.38	32.83		28.48	30.30		4.12	4.27
1980-1989								
$\hat{\beta}$	0.9733**	0.9201**		0.9658**	0.9483**		0.9507**	0.9473**
	(-2.87)	(-4.17)		(-3.14)	(-2.86)		(-4.24)	(-3.22)
$\hat{\rho}$	0.0889**	0.1495**		-0.0528*	-0.0284		0.0133	0.0373
	(3.34)	(4.64)		(-1.96)	(-0.93)		(0.53)	(1.29)
Heterosc.	-0.0054*	-0.0053**		-0.0099**	-0.0101**		-0.0237**	-0.0245**
	5.86	5.86		14.99	15.98		30.68	32.70
1990-1999								
$\hat{\beta}$	0.9584**	0.9786		0.9724*	1.0322		0.9784*	1.0239
	(-6.39)	(-1.53)		(-2.15)	(1.91)		(-2.60)	(1.62)
$\hat{\rho}$	0.0595*	0.0776*		-0.0416	-0.0833**		-0.0222	-0.0459
	(2.15)	(2.27)		(-1.52)	(-2.65)		(-0.85)	(-1.56)
Heterosc.	-0.0049**	-0.0055**		-0.0120**	-0.0121**		-0.0181**	-0.0181**
	36.89	39.93		50.02	50.79		80.28	79.48
2000-2006								
$\hat{\beta}$	0.9893	1.0456**		0.9845	1.0426*		0.9958	1.0379
	(-1.52)	(2.69)		(-1.92)	(2.06)		(-0.54)	(1.72)
$\hat{\rho}$	0.0122	-0.0382		0.2456**	0.1782**		-0.0584*	-0.1037**
	(0.46)	(-1.19)		(6.61)	(3.73)		(-2.00)	(-2.85)
Heterosc.	-0.0060*	-0.0060**		-0.0142**	-0.0146**		-0.0209**	-0.0212**
	5.48	22.99		17.79	17.71		28.43	27.91

**/* indicate that the coefficient is significantly different from zero at the 1%/5% level. OLS and B-M refer to the ordinary least squares and Breitung-Meyer unit root tests. *t*-statistics are shown beneath in brackets. The reported *t*-statistics for $\hat{\beta}$ refer to the coefficient (β-1). 'Heterosc' is the estimated coefficient on the lagged squared size variable in the auxiliary regression used to obtain the heteroscedasticity LM test statistic, shown beneath in italics. Intercept and time dummy variables are not reported.

Over 1970-1979, we obtain significance on the OLS test with $\hat{\beta}_{OLS}<1$ for all the size measures considered but the coefficients for the B-M test are not significant. During this period commercial banks had freedom to open branches nationwide from 1974 onwards but savings banks had severe restrictions to expansion, even in their own traditional regions, increasing growth costs and generating a high heterogeneity in growth rates across different institutions.

For the 1980-1989 period, $H_0:\beta=1$ is rejected in favour of $H_1:\beta\neq1$ at the 1% level by both tests. Since all the size-growth estimates are less than unity, small banks tend to grow faster than large banks on average. In this decade, commercial banks still benefit from expansion advantages but to a lesser extent than in the previous decade. In fact, in 1985, the investment coefficients for commercial and savings banks are equalized and savings banks are free to open branches in their local regions. Later, in 1987, interest rates and commissions are fully liberalized. This permitted banks to set deposit and loan rates in response to market conditions and to compete for deposit market share and banking relationships.

During the 1990-1999 period, the coefficients corresponding to the OLS test are less than unity $\hat{\beta}_{OLS}<1$ and the coefficients for the B-M test are not significant. This evidence could suggests that size and growth are still negatively related but to a lower extent than in the 80s. The strong consolidation among savings banks in the early 90s may have allowed large institutions to increase their growth rate. Finally, over 2000 to 2006, both tests show evidence that size and growth are unrelated or positively related, implying, in the latter case, that larger banks grow faster than the smaller ones. The OLS test provides non-significant coefficients for all the size measures. The B-M test, however, provides evidence that larger banks grow faster than the smaller in total assets and total loans, but not in total deposits, where size and growth appear to be independent
The result that larger banks grow faster than the smaller, has been reported in some previous studies resulting, from the greater role played by growth through acquisitions and mergers among the larger banks (Sutton, 1997). In fact, integration processes have been tenser during this decade and have accelerated the growth of larger banks.

Overall, the results seem to indicate that relationship between bank size and growth rates tends to evolve over time, as liberalization and deregulation processes occur. Larger banks tend to increase their rate of growth through time, obtaining efficiency gains and benefiting from scale and scope economies, consistently with a growth orientation in the larger institutions. After 1998, there is a general increase of the average size of the M&A in Spain, a phenomenon paralleled in the EU. This fact shows the increased concern of banks during the last decade with size and the need to become large players due to competitive pressures (Cabral et al., 2002). However, the fact that the vast majority of mergers are domestic may make small institutions disappear in favour of the larger ones. In this respect, the importance of savings banks in several EU countries, including Spain, which create a barrier to foreign banks by reducing the number of institutions available for acquisition, may have been a contributing factor (Belaisch et al., 2001).

The dynamic evolution of the persistence follows different patterns depending on the size measure employed in the estimations. For the asset size measure, growth persistence is positive but decreasing for the periods 1970-1979, 1980-1989 and 1990-1999. In the period 2000-2006 we find no persistence in asset growth rates. Liberalization progressively reduces entry barriers and technology advantages become more available to all banks, independently of size. In addition, new advantages may increasingly come from other sources not reflected in the balance sheet such as loan commitments and other off-balance sheet items as well as customer relationships.

Both loans and deposit measures show positive persistence in the 70s. However, the trends of both measures diverge in subsequent decades. The loans measure shows negative and significant persistence in the 80s but follows a decreasing trend in the following periods. In 2000-2006, the coefficients turn out to be positive and highly significant, indicating that advantages in loan markets are carried over time. Overall, this view is consistent with the increased importance of relationship lending on several scores. Information advantages, long-term contractual relationships and improved monitoring activity, can provide market power and make it difficult for new banks to gain market share[31]. The persistence in the deposits measure follows the opposite pattern again, from

---

[31] See Boot (2000) for an extensive discussion of the costs and benefits of relationship lending.

positive but insignificant in the 1980-1980 and 1990-1999 periods, to negative and significant coefficients in the 2000-2006 period.

Finally, homoscedasticity is again rejected for all the size measures and decades. The negative coefficients indicate higher growth variability in small firms. In addition, heteroscedasticity tends to amplify over time showing that banks are increasing their risk as a response to competitive pressures.

3.5.2. Commercial, savings and cooperative banks (1980-2006)

Table 3.6 presents the estimation results for the whole population of commercial, savings and cooperative banks over the 1980-2006 and 1990-2006 periods. For the 1980-2006 period, the OLS tests supports the proposition that size and growth are negatively related for all the size measures. The larger banks in the sample grow in average more slowly than the smaller ones. This result is confirmed by the B-M test for the assets and deposits measures, but not for the loans measure whose coefficient turns out to be not significant, implying that size and growth are unrelated. In the 1990-2006 period, OLS estimates are highly significant for all the size measures, but B-M estimates are not significant. These results are broadly similar to those attained with the previous sample, which did not include cooperative banks.

The results on persistence are similar to those attained without cooperatives, with positive and significant persistence in total assets and negative and significant persistence in the deposits size measure. However, persistence effects are now more pronounced since we find higher coefficients in assets and loans and lower (more negative) coefficients in deposits for both the 1980-2006 and 1990-2006 periods. In fact, the estimated coefficient in assets is now significant at the 5% level.

Table 3.6. Estimation results for commercial, savings and cooperative banks. 1980-2006 and 1990-2006 periods.

	Assets			Loans			Deposits	
	OLS	B-M		OLS	B-M		OLS	B-M
1980-2006								
$\hat{\beta}$	0.9875**	0.9716**		0.9831**	1.0017		0.9827**	0.9850**
	(-5.83)	*(-6.56)*		*(-5.84)*	*(0.37)*		*(-4.95)*	*(-3.83)*
$\hat{\rho}$	0.0585**	0.0738**		0.0464**	0.0378*		-0.0319**	-0.0281*
	(5.37)	*(6.58)*		*(3.88)*	*(3.00)*		*(-2.83)*	*(-2.39)*
Heterosc.	-0.0015**	-0.0015**		-0.0043**	-0.0044**		-0.0144**	-0.0146**
	13.54	*14.29*		*38.01*	*38.74*		*115.34*	*116.85*
1990-2006								
$\hat{\beta}$	0.9870**	0.9952		0.9848**	1.0071		0.9853**	0.9881
	(-4.54)	*(-0.78)*		*(-3.82)*	*(0.99)*		*(-3.40)*	*(-1.64)*
$\hat{\rho}$	0.0325*	0.0339*		0.0993**	0.0862**		- 0.0657**	-0.0627**
	(2.29)	*(2.22)*		*(5.98)*	*(4.77)*		*(-4.42)*	*(-3.92)*
Heterosc.	-0.0025**	-0.0025**		-0.0065**	-0.0066**		-0.0178**	-0.0180**
	37.13	*37.13*		*42.04*	*42.04*		*112.10*	*112.52*

**/* indicate that the coefficient is significantly different from zero at the 1%/5% level. OLS and B-M refer to the ordinary least squares and Breitung-Meyer unit root tests. *t*-statistics are shown beneath in brackets. The reported *t*-statistics for $\hat{\beta}$ refer to the coefficient (β-1). 'Heterosc' is the estimated coefficient on the lagged squared size variable in the auxiliary regression used to obtain the heteroscedasticity LM test statistic, shown beneath in italics. Intercept and time dummy variables are not reported.

As can be expected, the Lagrange-Multiplier tests reject the null hypothesis of homoscedastic growth with an inverse relationship between size and growth variability. Heteroscedasticity also increases over time. However, the variability of growth rates when cooperative banks are not included in the estimations more than doubles those of the full sample. This result is somewhat surprising given that cooperative banks are fairly small on average and one would expect higher degrees of growth variability in smaller banks. A plausible explanation for this finding is that Spanish cooperative banks tend to specialize more in relational lending than commercial banks. In practice, credit cooperatives lend to closer and smaller borrowers and make more use of collateralized loans than commercial banks (Delgado et al., 2007).

The analyses for the 1980-1989, 1990-1999 and 2000-2006 periods, reveal the same pattern that large banks increase their rate of growth over time relative to small banks, but this effect occurs now at a higher pace. Regardless the chosen size measure almost all the size-growth estimates are strictly higher when cooperative banks are included in the sample. The results are shown in table 3.7. As before, the 1980-1989 estimates reveal that small banks tend to grow slower than larger banks. Over 1990-1999, all the estimates tend to get closer to one, and B-M tests are insignificant. Finally, over the 2000-2006 period, all the B-M estimates turn out to be higher than one and significant at the 5% level, implying that larger banks grow more rapidly than their smaller counterparts.

The estimates of ρ including credit cooperatives turn out again to be quite similar to those attained in the previous section. The persistence in assets follows the same decreasing pattern over time which indicates that scale economies advantages erode quickly due to new competitive pressures. As expected, the ρ estimates for loans are slightly higher when cooperative banks are included in the sample in almost all periods. Cooperatives lend to closer and smaller borrowers and hence should engage in closer lending relationships carrying lending advantages from one period to the next. The fact that the estimates for 2000-2006 are not significant is consistent with the recent view that competition in Spanish loan markets has decreased over time. Finally, the estimates for the deposits size measure evolve from positive but not significant in the 80s to negative and significant over 2000-2006.

Table 3.7. Estimation results for commercial, savings and cooperative banks. 1980-1989, 1990-1999 and 2000-2006 periods.

	Assets			Loans			Deposits	
	OLS	B-M		OLS	B-M		OLS	B-M
1980-1989								
$\hat{\beta}$	0.9853**	0.9301**		0.9767**	0.9610**		0.9666**	0.9481**
	(-3.74)	(-4.90)		(-4.96)	(-2.97)		(-4.68)	(-4.09)
$\hat{\rho}$	0.0843**	0.1472**		-0.0376	-0.0128		0.0081	0.0408
	(4.24)	(5.99)		(-1.89)	(-0.55)		(0.43)	(1.82)
Heterosc.	-0.0001	-0.0001		-0.0009	-0.0009		-0.0088**	-0.0089**
	0.00	0.00		0.54	0.54		11.01	11.55
1990-1999								
$\hat{\beta}$	0.9767**	0.9808		0.9801**	0.9803		0.9807**	1.0251
	(-5.75)	(-1.66)		(-2.83)	(-1.79)		(-2.80)	(1.95)
$\hat{\rho}$	0.0600**	0.0766**		-0.0345	-0.0756**		-0.0226	-0.0470
	(2.60)	(2.94)		(-1.52)	(-2.89)		(-1.00)	(-1.85)
Heterosc.	-0.0021**	-0.0020**		-0.0060**	-0.0060**		-0.0153**	-0.0153**
	15.41	16.36		27.73	28.19		72.97	72.49
2000-2006								
$\hat{\beta}$	0.9970	1.0222*		0.9850*	1.0427*		0.9974	1.0387*
	(-0.61)	(2.36)		(-2.46)	(2.37)		(-0.41)	(2.07)
$\hat{\rho}$	0.0142	-0.0103		0.2446**	0.1784**		-0.0574*	-0.1028**
	(0.64)	(-0.43)		(8.13)	(4.62)		(-2.33)	(-3.33)
Heterosc.	-0.0037**	-0.0037**		-0.0087**	-0.0089**		-0.0195**	-0.0197**
	18.23	16.69		13.71	13.71		36.57	36.19

**/* indicate that the coefficient is significantly different from zero at the 1%/5% level. OLS and B-M refer to the ordinary least squares and Breitung-Meyer unit root tests. *t*-statistics are shown beneath in brackets. The reported *t*-statistics for $\hat{\beta}$ refer to the coefficient (β-1). 'Heterosc' is the estimated coefficient on the lagged squared size variable in the auxiliary regression used to obtain the heteroscedasticity LM test statistic, shown beneath in italics. Intercept and time dummy variables are not reported.

Finally, homoscedasticity is again rejected for the sub periods 1990-1999 and 2000-2006, where the negative and highly significant coefficients indicate higher growth variability in small firms. This result is also found for 1980-1989 for the deposits measure, but not for the assets and loans measures. Again, the fact that heteroscedasticity increases over time confirms that banks progressively increase their risk as financial integration occurs and competition increases.

3.6 Conclusion

We have estimated the size-growth relationship for the whole population of Spanish commercial, savings and cooperative banks over the 1970-2006 period, and different sub-periods corresponding to different scenarios of the liberalization, deregulation and integration processes. Three different size measures have been used: total assets, total loans and total deposits. The panel data estimation technique used to test the hypothesis overcomes several limitations of the classical methodology based on cross-sectional regressions.

Rather than a time invariant size-growth relationship, our findings point to the existence of an evolving relationship as the banking industry goes through a liberalization or consolidation process. When Spanish banking was highly regulated, we find that smaller banks grew faster than their larger counterparts, indicating that there were diseconomies of scale, monopoly rents and political costs to grow (Tschoegl, 1983). In recent years, however, we find that larger banks tend to grow at the same rate or faster than smaller banks, a phenomenon consistent with the importance of economies of scale and scope to gain efficiency and the recent waves of M&A. Thus, our study points to the conditioned nature of the size-growth relationship and the size distribution of banks, as emphasized by Berger et al. (1995), Ennis (2001), Jones and Critchfield (2005) and Janicki and Prescott (2006) for the US banking system.

Evidence suggests that bank-specific growth persistence also follows a dynamic pattern, but this pattern depends on the selected size measure, highlighting the different characteristics of the banking business. Liberalization and consolidation tend to erode size advantages when measured by total assets. The fact that liberalization progressively

reduces entry barriers and technological advantages may increasingly become more available to all banks may explain this result.

The fact that size advantages measured by total assets erode quickly, contrasts with the general concern with size shown by banks in recent years and the general increase of the average size of the M&A deals that occurs after 1998 in European banking. An alternative explanation for merger activity is based on the expected efficiency gains through profitability improvements. Recent studies support this view for the European banking industry[32].

Persistence in loans growth rates tends to increase through liberalization, a finding that is consistent with the increased importance of relationship lending in modern banking and with recent studies that show that competition in Spanish lending markets has decreased over time. Spanish savings and cooperative banks tend to specialize more in relational lending than commercial banks. However, whether relationship banking has potential benefits or costs for borrowers and the overall economy remains an open issue.

Persistence in deposits growth rates tends to decrease over time, and becomes negative in recent years. The fierce competition with respect to funding has been identified in recent studies at the Spanish and European settings. The introduction of internet banking may have been a crucial boost factor.

The decline of variability of growth rates with size confirms that larger banks are more capable of diversifying away part of their total risk. However, variability tends to increase over time, indicating that banks are more prone to taking risks as competitive pressures tend to erode market share.

As has been pointed out earlier, the evolution of the bank size distribution depends, ultimately, on the nature of the size-growth relationship. My analysis shows that this relationship changed as deregulation proceeded. Future research should focus on gaining a deeper understanding of the economic process by which liberalization and deregulation influence the growth rates of different sized banks.

---

[32] See Focarelli et al. (2002), Diaz et al. (2004) and Campa and Hernando (2006).

Chapter 4: Banks versus Markets: The joint effects of market participation and aggregate risk

4.1 Introduction

The US economy experienced a rapid recovery from the credit crunch compared to Europe despite the initial shock being similar in size and nature in both economies and equity prices, house price inflation, bank lending, and residential construction all falling at similar sharp rates over 2007 and 2008. Over the period 2011-13, the US economy grew by about 6 percentage points more. Even taking into account the increasing demographic differential, the US economy grew 4.5 percentage points more on a per capita basis (World Economic Forum, 2014).

Several commentators have pointed to an overreliance on bank intermediaries in Europe as a key reason for such striking difference. As illustrated by Dr. Adam Posen, former member of the Monetary Policy Committee of the Bank of England in a speech given in 2012 at the National Institute of Economic and Social Research: "Why the US economy has had a better recovery than the UK economy from the global financial crisis so far? Corporate investment rebounded much more in the US than the UK because there were more non-bank options available to provide financing for investment" (Bank of England, 2012).

The European financial system is traditionally bank-oriented and, unlike other advanced countries, has accentuated this characteristic in the last decades (ESRB, 2014). Importantly, the size of the European banking system (as measured by total assets) grew significantly in the run-up to the financial crisis, particularly after the introduction of the euro. The total assets of the EU banking sector amounted to 334% of GDP, compared with 192% in Japan and 145% in the US (ESRB, 2014). Iceland's bank assets, for example, rose to over eight times GDP by 2007, six times the 2001 ratio. In 2008 the Icelandic banking sector collapsed, followed by its own currency, the Icelandic Krona. Banks of Greece, Cyprus, Spain, Ireland or the UK also grew exponentially leading to several bank failures and government bail-outs. Since many European firms have limited or no access to public debt markets, this led to further financing constraints contributing

to a slow recovery (CEPR, 2014). These events have provided for further debate considering the merits and robustness of market versus bank-based economies.

The role of market-based and financially intermediated systems as liquidity providers and the related underlying mechanisms that make banks vulnerable to a loss of depositors' confidence have been examined extensively in the literature. In pioneering work by Bryant (1980) and Diamond and Dybvig (1983), and extensions by Jacklin (1987) and Bhattacharya and Gale (1987), among others, investors convey asset allocation decisions by trading their liquidity risk against the high return of long-term investments. Compared to markets, banks are able to provide additional insurance to risk-averse investors by aggregating wealth and offering demand deposits that allow contingent withdrawals according to the consumption needs of investors. This risk-sharing mechanism however, leaves banks exposed to bank runs, which may decrease welfare and make equity contracts a better instrument for risk-sharing than bank deposits unless preventing policies are put in place. Diamond (1997), however, demonstrates that the banks' ability to provide liquidity depends on the size of the subset of agents that are able to participate in markets. If there exists limited market participation, markets alone may not be able to provide additional liquidity and a mixed financial system where banks and markets coexist and influence each other provides superior welfare. Qian et al. (2004) obtain similar results in an overlapping-generations framework. These studies do not consider however, the existence of bank runs and their implications on the overall welfare achieved.

This chapter revisits the liquidity provision role of financial markets and intermediaries when both are active, there is limited participation in markets, and there exists aggregate technology risk in the economy. We show that bank runs may occur if a subset of non-participating agents receive information about the future return of the underlying assets. The negative impact on welfare of bank runs may however, not be sufficient to tilt the scale in markets favour. We argue that the level of the underlying aggregate technology risk and of limited market participation jointly determine the superiority of the mixed economy over pure equity contracts.

We use a similar setup to Diamond (1997), with differences in the manner the different technologies and information structures around them are modelled. In our setting, there

are two technologies available, a short-term safe one and a long-term risky one, and we classify agent types according to three different groups whose aggregate size is known at date 0, but not its composition. Therefore, ex-ante people do not know when they need to consume or whether they are able to participate in the market. Those who will not participate in the market are a subset of the patient agents. Agents learn their type at date 1, when a market opens for trade. Information about the future returns of the risky technology is, initially, only available to participating agents.

We first investigate the performance of a market-based economy. Participating individuals will trade short-term for long-term claims at date 1 whereas non-participating individuals will reinvest any holdings of short-term claims, resulting in inefficient re-investment of liquid assets and overall underinvestment in long-term risky assets (and consequently lower prices of long-term claims) compared to a full participation situation. A bank offering a deposit contract however can aggregate wealth to avoid inefficient reinvestment and provide risk-sharing against the uncertainty in depositors' liquidity needs at the same time. In addition the bank can trade claims in the market at date 1, increasing liquidity provision and normalising market prices.

Non-incentive compatible withdrawals by patient non-participating agents may occur if a subset of these receive information about the future return of the risky asset, and the long-term return is small enough. Depending on the number of agents that receive information and decide to withdraw early, how small is the return, and the amount of excess liquidity that the bank holds at date 1, the bank may be forced to undertake fire sales of long-term assets, depressing the market price and reducing overall consumption.

Our analysis features a unique equilibrium where the probability of bank runs can be endogenously determined as a result of information acquisition or observation by non-participating agents. This approach is consistent with the business cycle view of bank runs, where the occurrence of bank runs is related to economic fundamentals, as evidenced in empirical studies of Gorton (1988), Calomiris and Gorton(1991), and theoretical work by Allen and Gale (1998), Goldstein and Pauzner (2005), and Zhu (2001) among others.

The single equilibrium result of the previous models contrasts with the multiple equilibria outcome of models such as Diamond and Dybvig (1983), Chari and Jaganathan (1988), Chang and Velasco (2000, 2001), or Jeitschko and Taylor (2001). On one hand, if no agent expects that a bank run will happen, the risk-sharing mechanism provided by the banking sector functions well and the economic resources are allocated in an efficient way. On the other hand, if all agents anticipate a bank run, then they all have the incentive to withdraw their deposits immediately and a bank run occurs as expected. Which of the two equilibria happens depends on the expectations of agents, which, unfortunately, are not addressed in their models.

Our analysis evidences that reducing market participation increases the cross-subsidy that banks provide causing the banking sector to expand, primarily through increased holdings of long-term assets, whenever the is a relatively low level of aggregate risk. Under this scenario, the welfare effects of bank runs may not be sufficient to overcome the positive effects of cross-subsidisation and efficient channel of investment. Increasing aggregate risk, however, decreases the scale of the banking sector and reduces overall cross-subsidisation under the mixed economy, reducing the short-term return available to investors that banks provide relative to secondary markets. In addition, higher levels of aggregate risk increase the impact of bank runs reducing the market price of the long-term asset and generating lower bank returns. Whenever aggregate risk exceeds a certain threshold, then markets may perform better than banks even for low or null levels of market participation. In fact our results show, that markets may perform better than banks the lower the market participation under some circumstances.

The observation that deposit contracts that are subject to bank runs tend to be better for financing low-risk assets is consistent with previous literature. Jacklin and Bhattacharya (1988) demonstrate that the choice between deposit or equity contracts depends on the riskiness of the underlying assets held by the intermediary as well as the nature and availability of information about those assets. In their model, there is a threshold risk dispersion level beyond which individuals run to the bank making equity contracts over-perform deposit contracts. Alonso (1996) modifies the initial setting of Chari and Jagannathan (1988) and focuses on the possible ex-ante risk-sharing contracts that the bank can develop. Under this setting, she finds that equity arrangement over-perform the equity arrangement when the underlying investment projects are very risky. However,

this is the case only when banks do not adjust the deposit contract so as that runs are prevented. In this sense a number of studies consider that banks may find profit maximising the design of contracts that runs cannot occur. In our model, the bank cannot achieve the necessary scale to prevent bank runs in the low state of nature since doing so could prove too costly, leading to avoidance of trade unless depositor pay-offs are altered in each state of nature.

Our model has important policy implications. Recent regulatory reforms have focused on constraining the risk undertaken by imposing new capital and liquidity requirements. Our findings, however, imply that such efforts should be accompanied by further action aimed at improving market access, particularly in Europe, where recovery from the financial crisis was sluggish.

The rest of this chapter proceeds as follows. Section 4.2 formulates the model. Section 4.3 analyses the performance of markets when there are no banking options available in the economy. Section 4.4 investigates the roles of banks and markets when both are active and evidences that bank runs can occur. Section 4.5 presents the main results. Section 4.6 concludes with a summary of the main results and a discussion of avenues for future research. The appendix, in section 4.7, collects the proofs.

4.2 The Model

There are three dates in the economy $t = 0, 1, 2$ and a continuum of risk- averse agents that are ex-ante identical. Agents are endowed with one unit of capital at date 0 and have no more endowment in the subsequent periods. At $t = 1$ a secondary market for assets is open. Agents are subject to a privately observed uninsurable risk at $t = 1$ which affects their demand for liquidity and participation in the secondary market. Individuals can be of either three types: type 1, type 2A and type 2B. Type 1 agents derive utility for consumption only at date 1 and type 2A and 2B agents derive utility for consumption only at date 2. Type 1 and type 2A agents are active in the secondary market while type 2B agents are not. There is a probability q_1, q_{2A} and q_{2B} of being a type 1, type 2A or type 2B agent respectively and there is no aggregate uncertainty regarding these probabilities. At the beginning of date 1 every agent discovers his type.

The economy is endowed with two indefinitely divisible investment technologies that differ in the returns and liquidity that they provide. The first one is a one-period storage technology with per-unit payoff of 1. The second is a long-term risky technology that generates a random return after two periods. One unit invested at $t = 0$ yields a nil return at $t = 1$, and a random return of $R(s)$ *at t = 2*, where $s \in S$ is the state of nature which is realized at $t = 2$. Hence, this technology is completely illiquid.

There are two possible states of nature at $t = 2$ denoted by l and h. The ex-ante probability for each of the states is $p(s)$. It is assumed that $R(h) > R(l)$ and that $p(l)R(l) + p(h)R(h) > 1$ which implies that the net present value of investing in the long-term technology is positive.

At the beginning of date 1, all the agents that are active in the secondary market (type 1 and type 2A) receive perfect information about the realization of s. A timetable of events is given in Table 4.1.

Denote by c_1^s, c_{2A}^s and c_{2B}^s the consumption of type 1, type 2A and type 2B agents respectively conditional on the realization of s. The ex-ante expected utility function of a representative agent is given by:

$$E(u) = E_{R(s)}[q_1 u(c_1^s) + q_{2A} u(c_{2A}^s) + q_{2B} u(c_{2B}^s)] \qquad (4.1)$$

where $u(\cdot)$ is a twice continuously differentiable, increasing, strictly concave utility function that satisfies Inada conditions $u'(0) = \infty$ and $u'(\infty) = 0$. For analytical derivations and numerical computations we will assume that $u(c) = c^{1-\gamma}/(1-\gamma)$ with $\gamma > 1$, where γ is the coefficient of relative risk aversion.

Table 4.1. Timing of events

Date 0	Date 1	Date 2
Each agent receives an endowment of one to allocate between the short-term asset, the long-term asset, and bank deposits if available.	All agents learn their type and some of them receive perfect information about the future realization of s. Agents choose their withdrawal from bank and some agents may trade in the secondary market.	State s occurs.
	Type 1 agents sell any long-term claims or withdraw from the bank and consume all wealth.	**Type 1** agents do nothing.
	Type 2A agents sell any short-term claims or withdraw from the bank to buy long-term claims or fund new one-period investment.	**Type 2A** agents withdraw from the bank and consume all wealth.
	Type 2B agents take any short-term claims or withdraw from the bank to fund new one-period investment.	**Type 2B** agents withdraw from the bank and consume all wealth.

4.3 The performance of markets alone

In our economy, agents would like to obtain insurance against the idiosyncratic preference risk they face at birth and the aggregate shock that affect the return of the long-term technology. We assume now that investors cannot rely on any financial intermediary. Each agent invests at date 0 a proportion $\propto$ of his endowment in the short-term technology and $1-\propto$ in the long-term risky technology. Type 1 and type 2A agents have the possibility to trade at $t = 1$ after they have received information about the future return of the long-term technology. Trading at date 1 leads to type 1 investors to trade their date 2 claims for date 1 claims of type 2A investors. Denote by $b(l)$ and $b(h)$ the date 1 price of a unit claim on the technology that matures at date 2 if the state occurring in date 2 is low or high. Type 2B agents have no access to markets so they take their maturing short-term claims at date 1 and invest in a new short-term asset, yielding 1 of date 2 consumption, in addition to their original holding of long-term assets. Ex-ante expected utility is given by:

$$E(u) = p(l)\left[q_1 u\left(c_1^l\right) + q_{2A} u\left(c_{2A}^l\right) + q_{2B} u\left(c_{2B}^l\right)\right] + p(h)\left[q_1 u\left(c_1^h\right) + q_{2A} u\left(c_{2A}^h\right) + q_{2B} u\left(c_{2B}^h\right)\right]$$

(4.2)

where consumption of each agent in each state is given by:

$$c_1^l = \propto +(1-\propto)b(l)R(l); \qquad c_1^h = \propto +(1-\propto)b(h)R(h);$$

$$c_{2A}^l = \frac{\propto}{b(l)} + (1-\propto)R(l); \qquad c_{2A}^h = \frac{\propto}{b(h)} + (1-\propto)R(h);$$

$$c_{2B}^l = \propto +(1-\propto)R(l); \qquad c_{2B}^h = \propto +(1-\propto)R(h);$$

Since all the agents that participate in markets receive perfect information about the future state of nature that will occur in the interim period, the clearing prices will reveal all the information available. Clearing prices must be positive and less or equal than $1/R(s)$, otherwise decreasing $\propto$ would increase the consumption of all types, increasing expected utility. This leads to the following result.

Lemma 1 *The date 1 prices of a unit claim on date 2 consumption are given by* $b(l) = \frac{q_{2A}\alpha}{q_1(1-\alpha)R(l)} \leq \frac{1}{R(l)}$ *and* $b(h) = \frac{q_{2A}\alpha}{q_1(1-\alpha)R(h)} \leq \frac{1}{R(h)}$ *depending on whether the future state of nature turns out to be low or high.*

Lemma 1 implies that optimal consumption levels are given by:

$$c_1 = \frac{q_1 + q_{2A}}{q_1}\alpha;$$

$$c_{2A}^l = \frac{R(l)}{R(h)}c_{2A}^h; \qquad\qquad c_{2A}^h = \frac{q_1 + q_{2A}}{q_{2A}}(1-\alpha)R(h);$$

$$c_{2B}^l = \alpha + (1-\alpha)R(l); \qquad c_{2B}^h = \alpha + (1-\alpha)R(h);$$

and $q_1 \leq \alpha \leq \frac{q_1}{q_1 + q_{2A}}$. To see this, substitute $\alpha = q_1 > 0$ into the prices equations, yielding $b(l) = \frac{q_{2A}}{(q_{2A}+q_{2B})R(l)}$ and $b(h) = \frac{q_{2A}}{(q_{2A}+q_{2B})R(h)}$. When all agents participate in markets (i.e. $q_{2B} = 0$), $\alpha = q_1$. Whenever there is limited participation, the proportion invested in the short-term technology increases and therefore $\alpha > q_1$, implying that expected utility is increasing in α. As individuals become more risk-averse, the investment in short-term assets increases even more and type 2A individuals consume less in favour of those that that need liquidity and those that do not participate in markets. Since the market clearing prices are fully revealing, the aggregate level of risk in the economy does not affect the individuals' investment allocation.

Assuming $u(c) = c^{1-\gamma}/(1-\gamma)$, the optimal problem for a representative agent can be written as:

$$\max_{\{c_1, c_{2A}^h, c_{2B}^h\}}(u) = q_1 \frac{(c_1)^{1-\gamma}}{1-\gamma} + A\left[q_{2A}\frac{(c_{2A}^h)^{1-\gamma}}{1-\gamma}\right] + p(l)q_{2B}\frac{(c_{2B}^l)^{1-\gamma}}{1-\gamma} + p(h)q_{2B}\frac{(c_{2B}^h)^{1-\gamma}}{1-\gamma} \quad (4.3)$$

s.t.

$$c_1 = \frac{q_1 + q_{2A}}{q_1}\alpha \qquad\qquad\qquad (4.4)$$

$$c_{2A}^h = \frac{q_1 + q_{2A}}{q_{2A}}(1-\alpha)R(h) \qquad\qquad (4.5)$$

$$c_{2B}^h = \alpha + (1-\alpha)R(h) \qquad\qquad\qquad (4.6)$$

$$q_1 \leq \propto \leq \frac{q_1}{q_1 + q_{2A}} \tag{4.7}$$

$$c_1, c_{2A}^h, c_{2B}^h \geq 0 \tag{4.8}$$

where $A = p(l) \left[\frac{R(l)}{R(h)} \right]^{1-\gamma} + p(h)$.

The allocation obtained in this section will serve as a benchmark to compare with an economy where banks and markets co-exist.

4.4 Banks and markets

In this section, we investigate the role of banks and markets when they co-exist in our economy. We first characterise the role of intermediaries and markets by solving for the optimal set of incentive-compatible consumption opportunities that agents could choose at date 0, the first-best allocation. As in Diamond (1997) we then show how a mixed economy where markets and banks coexist can implement this allocation.

4.4.1 The first-best allocation

In order to characterize the role of banks and markets, we solve for the optimal mechanism in which each investor reveals his or her type and is given contingent consumption on each date depending on its type and the realization of the state of nature, subject to resource and incentive-compatibility constraints.

The optimal risk-sharing problem can be written as follows:

$$\max_{\{c_1, c_{2A}^s, c_{2B}^s; s \in S\}} E_{R(s)}[q_1 U(c_1) + q_{2A} U(c_{2A}^s) + q_{2B} U(c_{2B}^s)] \tag{4.9}$$

s.t.

$$q_1 c_1 + \frac{q_{2A} c_{2A}^s + q_{2B} c_{2B}^s}{R(s)} \leq 1 \tag{4.10}$$

$$c_1 = c_{2A}^s \frac{1}{R(s)} \tag{4.11}$$

$$c_{2B}^s \geq c_1 \tag{4.12}$$

$$c_1, c_{2A}^s, c_{2B}^s \geq 0 \tag{4.13}$$

$$s \in S = \{l, h\} \tag{4.14}$$

Equation 4.10 is the resource constraint for each of the states of nature, limiting cost-adjusted consumption to the initial endowment per capital. Equations 4.11 and 4.12 are incentive compatibility constraints. As shown by Diamond (1997), at date 0 individuals could form alternative mechanisms (i.e. competing banks), when they realise that agents that turn out to be type 1 or type 2A are able to trade in the same market at date 1 anonymously. Since setting up a bank is costless in our model, asset holdings can be replicated by a competing bank. This ability imposes the constraint that the cost-adjusted rate of return of claims intended for agents that trade in the market is equalised, as reflected in equation 4.11. Equation 4.12 ensures type 2B agents do not have incentives to take the proceeds of date 1 claims and reinvest in new short-term investments maturing at date 2. This condition is loosest for low levels of aggregate risk and can be binding.

Proposition 1 *Assuming* $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, *the solution to the optimal mechanism is given by:*

$$c_1 = \frac{(q_1 + q_{2A}B)^{\frac{1}{\gamma}}}{q_{2B}[(q_1 + q_{2A})B]^{\frac{1}{\gamma}} + (q_1 + q_{2A})(q_1 + q_{2A}B)^{\frac{1}{\gamma}}};$$

$$c_{2A}^l = \frac{R(l)}{R(h)} c_{2A}^h; \qquad c_{2A}^h = c_1 R(h);$$

$$c_{2B}^l = \frac{R(l)}{R(h)} c_{2B}^h; \qquad c_{2B}^h = [1 - (q_1 + q_{2A})c_1]\frac{R(h)}{q_{2B}};$$

where $B = p(l)R(l)^{1-\gamma} + p(h)R(h)^{1-\gamma}$.

In the particular case that constraint 4.12 is binding, i.e. $c_{2B}^l = c_1$, *it can be shown that:*

$$c_1 = \frac{1}{q_1 + q_{2A} + \frac{q_{2B}}{R(l)}};$$

$$c_{2A}^l = \frac{R(l)}{R(h)} c_{2A}^h; \qquad c_{2A}^h = c_1 R(h);$$

$$c_{2B}^l = c_1; \qquad c_{2B}^h = [1 - (q_1 + q_{2A})c_1]\frac{R(h)}{q_{2B}}.$$

4.4.2 The banking contract

The first-best allocation shown above can be implemented in a mixed financial system where banks and markets coexist. Banks can offer risk-sharing contracts to individuals that are uncertain about their liquidity needs allowing agents to adjust their pattern of

withdrawals to their consumption needs. [33] These contracts are in the form of a demand deposit, and require an initial investment δ at date 0 in exchange for the right to withdraw either a fixed payment of w_1 at date 1, or a random payment of w_2^s at date 2. The uncertain second period payment reflects the fact that having invested in a risky asset the bank may not be able to make its promised payments at date 2. As shown by Jacklin (1987), the demand deposit contract optimally combines the two types of deposits that banks usually hold, a time deposit and a more typical demand deposit contract.

Individuals can thus invest δ in bank demand deposits, and $(1 - \delta)$ in the available technologies. The bank itself would need to allocate δ between the short term and long-term technologies. Let's represent by δ_s the amount invested by the bank in short-term assets and δ_L the amount invested in long-term assets, and thus $\delta = \delta_s + \delta_L$. In order to avoid inefficient reinvestment of short-term assets at date 1 by type 2B investors, the bank could hold all the short-term assets. A lower bound on the bank's short-term investments would therefore be $\delta_s = q_1 c_1$, because total date 1 consumption is c_1 by a fraction q_1 of investors, with any additional bank holdings $\delta = \delta_L$ being comprised of long-term assets. Each individual investor would hold a proportion $(1 - \delta)$ of date 2 claims and the bank's promised payments would be $w_1 = c_1 - (1 - \delta)$ and $w_2^s = c_{2B}^s - (1 - \delta)R(s)$.

At date 1, type 1 and 2A agents withdraw from the bank. Once deposits are withdrawn, type 1, type 2A depositors and the bank will trade in markets[34]. Type 2A agents will use their all their deposits to buy long-term assets from type 1 agents and the bank, whereas type 1 depositors will sell all their long-term assets to type 2A depositors and the bank. Hence, the consumption levels of each agent are given by:

$$c_1 = w_1 + (1 - \delta)b(s)R(s) \tag{4.15}$$

$$c_{2A}^s = \frac{w_1}{b(s)} + (1 - \delta)R(s) \tag{4.16}$$

$$c_{2B}^s = w_2^s + (1 - \delta)R(s) \tag{4.17}$$

where $s \in S = \{l, h\}$.

---

[33] As in Diamond (1997), the bank designs a contract that allow the implementation of the first-best allocation in section 4.4.1 as opposed to maximising a specific objective function.

[34] Note that Diamond (1997) considers two different cases where the bank either participates or does not participate in markets. Here we assume the bank always trades in markets.

The total demand of deposits experienced at date 1 by the bank from type 1 and type 2A deposits is given by $(q_1 + q_{2A})w_1$. The overall bank's supply of short-term assets is given by δ_s. Let's denote by E_1 the excess supply $(\delta_s - (q_1 + q_{2A})w_1 = E_1 > 0)$, or excess demand $(\delta_s - (q_1 + q_{2A})w_1 = E_1 < 0)$ of such deposits at date 1. Whenever there is excess supply, the bank sells short-term assets in the secondary market to type 1 agents. Whenever there is excess demand, the bank uses its holdings of long-term assets to buy liquidity from deposits withdrawn by type 2A agents. At date 2, the bank needs to meet withdrawals by type 2B agents, given by $q_{2B}w_2^s$. If the bank had excess supply of deposits at date 1, the bank would meet this date 2 demand with the market proceeds from sales of short-term assets as well as the returns from its holdings of long-term assets (if any). If the bank had experienced excess demand of deposits at date 1, then the bank would meet this date 2 demand with the returns from the remaining holdings of long-term assets that were not sold at date 1 to meet deposit demands.

Proposition 2 *When bank and markets coexist, the date 1 prices of a unit claim on date 2 are given by* $b(l) = \frac{q_{2A}w_1 + E_1}{q_1(1-\delta)R(l)} = \frac{1}{R(l)}$ *and* $b(h) = \frac{q_{2A}w_1 + E_1}{q_1(1-\delta)R(h)} = \frac{1}{R(h)}$ *depending on whether the future state of nature turns out to be low or high.*

Proposition 2 allows to restate the consumption levels as follows:

$$c_1 = w_1 + (1 - \delta) \tag{4.18}$$
$$c_{2A}^s = w_1 R(s) + (1 - \delta)R(s) \tag{4.19}$$
$$c_{2B}^s = w_2^s + (1 - \delta)R(s) \tag{4.20}$$

As noted earlier, the bank will promise payments at $t=1$ and $t=2$ given by $w_1 = c_1 - (1 - \delta)$ and $w_2^s = c_{2B}^s - (1 - \delta)R(s)$ which can also be expressed as $w_1 = c_1 - (1 - \delta_s - \delta_L)$ and $w_2^s = c_{2B}^s - (1 - \delta_s - \delta_L)R(s)$, on the basis that $\delta = \delta_s + \delta_L$. The bank's holdings of short-term assets are given by $\delta_s = q_1 c_1$ with any additional holdings $\delta = \delta_L$ comprised of long-term assets. Thus we can state:

$$w_1 = c_1 - (1 - q_1 c_1 - \delta_L) \tag{4.21}$$
$$w_2^s = c_{2B}^s - (1 - q_1 c_1 - \delta_L)R(s) \tag{4.22}$$

Incentive compatible self-selection of bank withdrawals for each state of nature, $s=h$ and $s=l$, requires setting $w_2^h > w_1$ and $w_2^l > w_1$. Otherwise, type 2B investors could choose to withdraw early, at date 1, and invest the proceeds in the short-term technology for one period, resulting in inefficient reinvestment. By equating $w_2^s = w_1$ and solving for δ_L using equations 4.21 and 4.22, it can be seen that the condition $w_2^s > w_1$ would be satisfied as long as $\delta_L > 1 - \frac{c_{2B}^s - c_1}{R(s)-1} - q_1 c_1$. In order to make the allocation incentive compatible under the high state of nature (i.e. setting $w_2^h > w_1$), the bank would need to invest in a minimum amount of long-term assets given by $\delta_L = \max\left[0, 1 - \frac{c_{2B}^h - c_1}{R(h)-1} - q_1 c_1\right]$.

Providing incentive compatible bank withdrawals in both the low and high states of nature would require a higher investment in long-term assets given by $\delta_L = \max\left[0, 1 - \frac{c_{2B}^l - c_1}{R(l)-1} - q_1 c_1\right] \geq \max\left[0, 1 - \frac{c_{2B}^h - c_1}{R(h)-1} - q_1 c_1\right]$. Achieving such level of investment may not be possible since, as shown in section 4.4.1, $c_{2B}^l \to c_1$ for sufficiently low market participation and low levels of aggregate risk, and thus constraint 4.12 can be binding $(c_{2B}^l = c_1)$, resulting in $\delta_L = 1 - q_1 c_1 = 1 - \delta_S$ which implies that $\delta = 1$. The intermediary may need to hold all short and long-term assets available in order to provide incentive compatible withdrawals in both states of nature with individuals holding no assets, resulting in no trade in the interim period. If there are variable costs to scale, the bank will be able to provide incentive compatible withdrawals in the high state of nature $(s=h)$, holding short-terms assets of $\delta_S = q_1 c_1$ and long terms assets of $\delta_L = \max\left[0, 1 - \frac{c_{2B}^h - c_1}{R(h)-1} - q_1 c_1\right]$.

The mixed financial system creates more liquidity than the one with secondary markets alone by two ways. First, banks can centralize the holding of short-term assets and thus eliminate inefficient investment of the excess of liquidity at t=1 by agents that do not participate in the market. Second, banks can cross-subsidize investors who need liquidity at the expense of those who can wait. In the next sub-section, we show however that banks can be subject to costly bank runs.

4.4.3 Bank runs

In the previous section we have assumed that only participating agents receive information about the future state of nature. If a proportion ρ of the total type 2B agents receive information about the realisation of s at date 1, these informed agents may have an incentive to withdraw their deposits early at date 1 if $R(s) = R(l)$. This is because the negative information shock could signal that in addition to their direct holdings of long term assets, reinvestment of bank withdrawals at date 1 may provide a higher pay off than withdrawals at date 2 if the initial self-selection pattern of type specific withdrawals is undermined under the low state of nature. According to the first-come first-served rule, the bank would honour its promise to pay w_1 as long as it has, or is able to obtain in the market, enough liquidity to satisfy demand at date 1, at the expense of date 2 withdrawals.

Following our discussion in section 4.4.2, early withdrawal by a subset of type 2B agents results in total demand of bank deposits at date 1 of $(q_1 + q_{2A} + \rho q_{2B})w_1$. At date 1 the bank will sell short-term assets in the secondary market to type 1 agents whenever there is excess supply, or will buy short-term asset from deposits withdrawn by type 2A agents whenever there is excess demand. We denote by E_1^{RUN} the bank's excess supply ($E_1^{RUN} > 0$), or excess demand ($E_1^{RUN} < 0$) of such deposits under such ban run scenario. This leads to the following result.

Lemma 2 The date 1 price of a unit claim on date 2 under a bank run is given by
$$b(l)^{RUN} = \frac{q_{2A}w_1 + E_1^{RUN}}{q_1(1-\delta)R(l)}.$$

Note that whenever $\rho > 0$, then the secondary market price $b(l)^{RUN}$ is strictly lower than $\frac{q_{2A}w_1 + E_1}{q_1(1-\delta)R(l)}$ and therefore strictly lower than $\frac{1}{R(l)}$. Under a bank run, the bank needs to use additional reserves of short-term assets or liquidate a higher amount of long-term assets at date 1 that it would have should a bank run had not occurred to satisfy the increased demand of liquidity by informed type 2B agents $E_1 - E_1^{RUN} = \rho q_{2B}w_1$ which is strictly positive as long as $\rho > 0$.

If the bank originally had excess supply of deposits prior to the occurrence of the run, increasing ρ would increase the amount of short term assets required to meet the

additional demand of deposits by informed type 2B agents. Any excess short-term assets after meeting such demand would be sold in the secondary markets to type 1 agents in exchange for long-term assets. Increasing ρ would progressively decrease the bank's supply of short-term assets in the secondary market up to the point where the bank exhausts it short-term assets reserves and needs to liquidate its holdings of long-term assets (if any) to buy short-term liquidity.

For a sufficiently high value of ρ, the bank could exhaust all its reserves of both short-term and long-term assets (if any) at date 1. We can denote as $\hat{\rho}$ the minimum value of ρ for which this occurs. Note that whenever $\rho > \hat{\rho}$, then the bank will not longer trade in markets at date 1 or operate at date 2, resulting in a total fraction $(1 - \hat{\rho})$ of type 2B depositors not being able to withdraw and obtain liquidity from the bank, either at date 1 if informed $(\rho - \hat{\rho})$ or at date 2 if uninformed $(1 - \rho)$.

4.5 Banks versus markets: welfare comparisons

In this section we provide some numerical examples in order to compare the ex-ante expected utility obtained in the mixed economy, where banks and markets co-exist, and the economy where agents have access to the secondary market and no intermediaries exist.

Figures 4.1-4.3 plot the expected utility under the mixed economy minus the expected utility under the market economy as a function of q_{2B} for different levels of expected return. The same ex-ante expected utility function is maximised in both economies. We focus our attention at the case with risk-averse individuals (i.e. the coefficient of relative risk aversion exceeds one). As discussed in the introduction, this is a de minimis condition to guarantee that the mixed economy can create more liquidity than secondary markets so that the cross-subsidy that banks provide to investors that need to consume early is valuable.

Figure 4.1. Expected utility of the mixed economy less expected utility of markets economy a function of market participation measured by q_{2B}. Parameter values: $\gamma = 3$; $q_1 = 0.9$; $p(l) = 0.3$; $\rho = 0.5$.
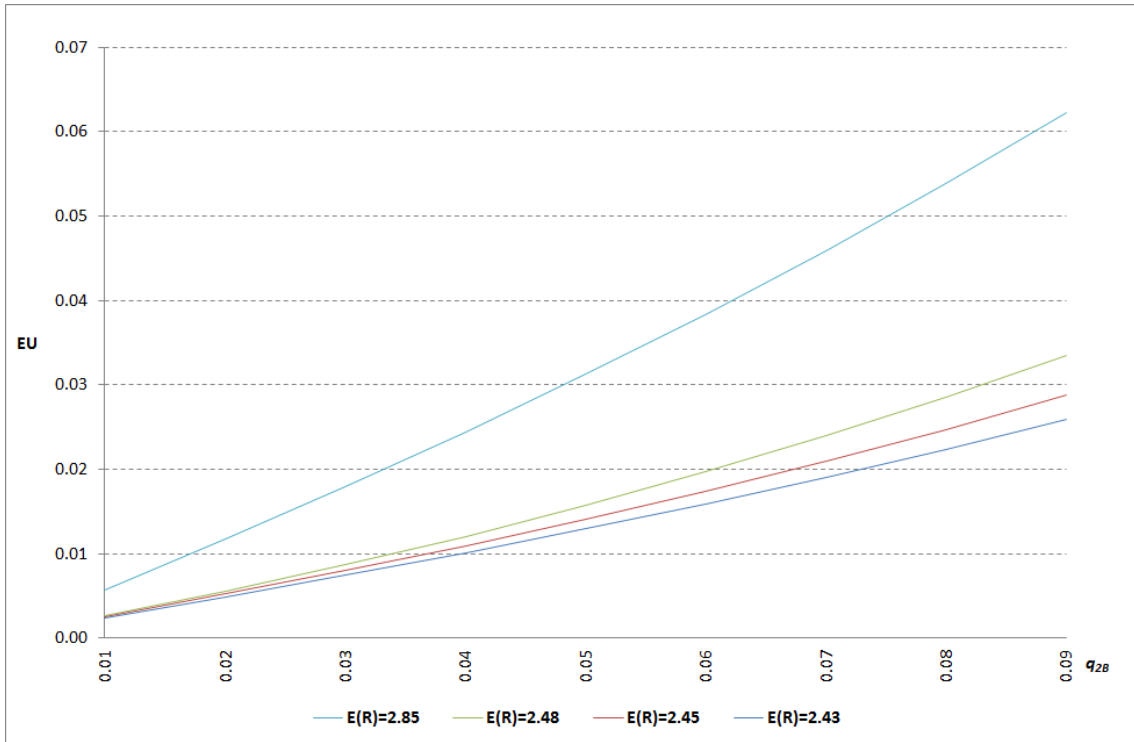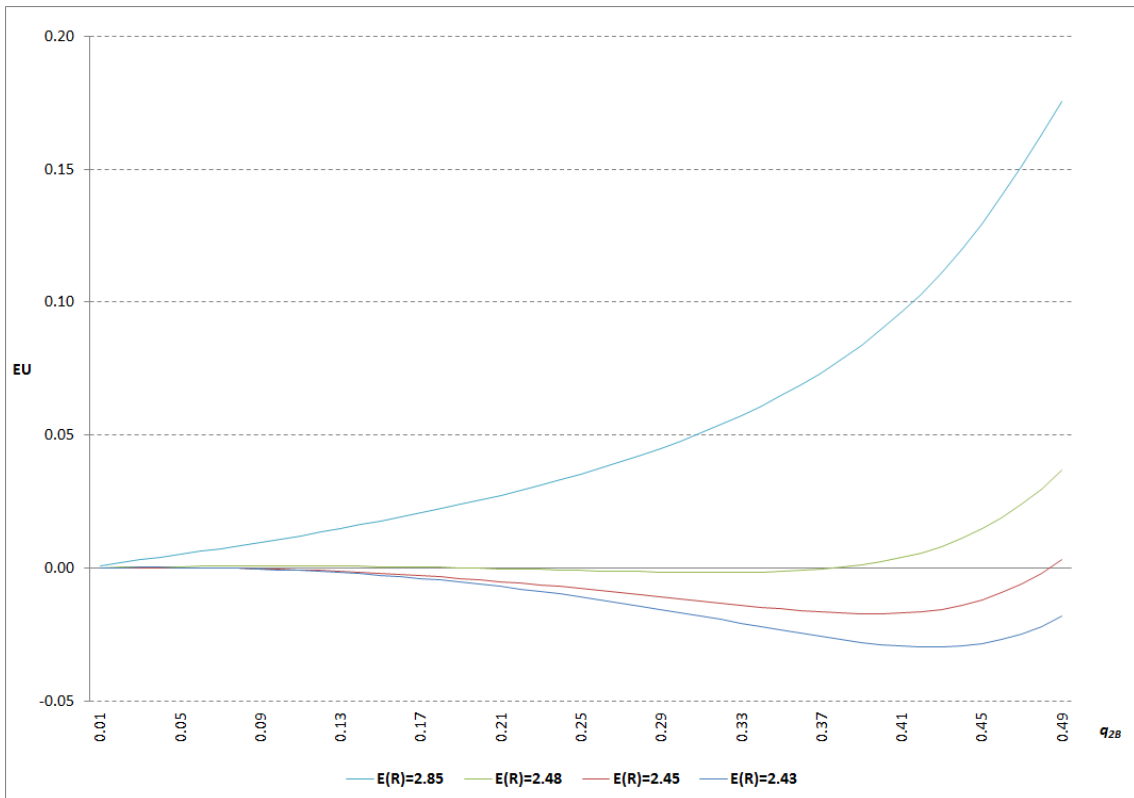

Figure 4.2. Expected utility of the mixed economy less expected utility of markets economy a function of market participation measured by q_{2B}. Parameter values: $\gamma = 3$; $q_1 = 0.5$; $p(l) = 0.3$; $\rho = 0.85$.
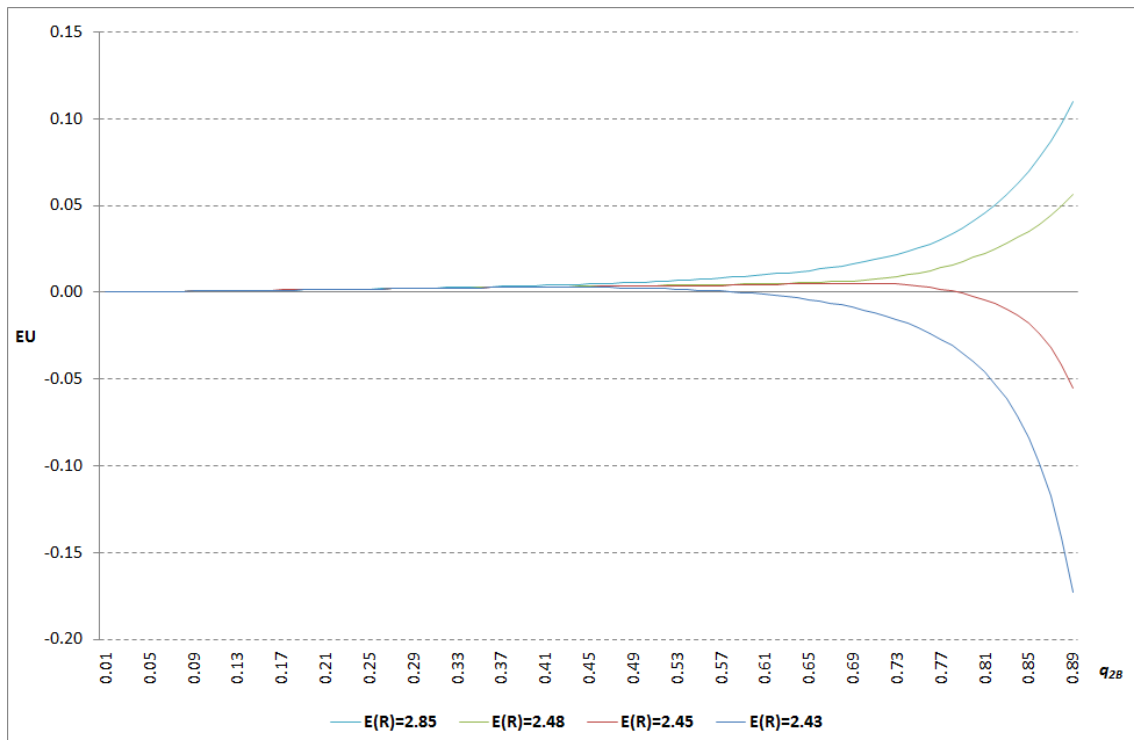
Figure 4.3. Expected utility of the mixed economy less expected utility of markets economy a function of market participation measured by q_{2B}. Parameter values: γ = 3; $q_1 = 0.1$; $p(l) = 0.3$; $\rho = 0.5$.

A common feature to the examples is that when all agents participate in secondary markets, both economies tend provide the same level of risk sharing regardless of the level of aggregate risk in the economy. As individuals' participation decreases the utility that banks and markets provide diverges depending on the proportion of agents participating in markets as well as the overall aggregate risk that is present in the economy.

Reducing market participation increases the cross-subsidy that banks provide to investors that need to consume early (type 1) at the expense of those who cannot trade (type 2B) causing the banking sector to expand, primarily through increased holdings of long-term assets. Banks also centralise the holding of liquid assets reducing the opportunity cost of excess date 1 liquidity held by non-participating investors (type 2B) and provide additional liquidity in markets.

Increasing aggregate risk decreases the scale of the banking sector and reduces overall cross-subsidisation under the mixed economy, reducing the short-term return available to

investors that banks provide relative to secondary markets. In addition, higher aggregate risk increases the negative impact of bank runs by reducing bank reserves and decreasing the market price of the long-term asset. Lower market participation implies that a higher proportion of type 2B agents would intent to withdraw early in a run, forcing the bank to sub-optimally liquidate holding of short-term assets and long-term assets, which in turn results in lower secondary market prices and overall welfare.

Whenever aggregate risk exceeds a certain amount, then markets may perform better than banks even for low levels of market participation. In fact our results show, that markets may perform better than banks the lower the market participation under some circumstances. Our results point to the level of aggregate risk and limited participation jointly determining the superiority of the mixed economy over pure market-driven equity contracts, and their critical values depend in turn on the exogenous parameters of the model.

It may be concluded that *if agents have smooth preferences, represented by an additive utility function with risk aversion above one, there is exists aggregate risk in the economy, and a sufficient fraction of investors do not participate $q_{2B} > \widehat{q_{2B}}$ and withdraw early in the low state of nature (bank run state), then the allocation obtained in the market economy is welfare superior with respect to the one achieved in the mixed economy whenever aggregate risk is above a value $E(R) > \widehat{E(R)}$.*

4.6 Conclusion

In this paper we compare the welfare properties of banks and markets when there is limited market participation and liquidity and technology risks are present in the economy. By extending the set-up in Diamond (1997), we show that bank runs may occur if a subset of non-market participating agents receive information about the future return of the underlying assets. The welfare effects of bank runs may however, not be sufficient to tilt the scale in markets favour. We show that the level of the underlying aggregate technology risk and of limited market participation jointly determine the superiority of the mixed economy over pure equity contracts.

Our analysis evidences that reducing market participation increases the cross-subsidy that banks provide causing the banking sector to expand, primarily through increased holdings of long-term assets, whenever the is a relatively low level of aggregate risk. Under this scenario, the welfare effects of bank runs may not be sufficient to overcome the positive effects of cross-subsidisation and efficient channel of investment. Increasing aggregate risk, however, decreases the scale of the banking sector and reduces overall cross-subsidisation under the mixed economy, reducing the short-term return available to investors that banks provide relative to secondary markets. In addition, increased aggregate risk increases the impact of bank runs by reducing the market price of the long-term asset and generating lower bank returns. Whenever aggregate risk exceeds a certain threshold, then markets may perform better than banks even for low or null levels of market participation.

Our model has important policy implications. Recent regulatory reforms have focused on constraining the risk undertaken by, for example, imposing new capital and liquidity requirements. Our findings, however, imply that such efforts should be accompanied by further action aimed at improving market access and participation, particularly in Europe, where recovery from the financial crisis was sluggish.

Our model can be extended to study the intergenerational aspect of liquidity provision. Qi (1994), Bhattacharya and Padilla (1996) and Fulghieri and Rovelli (1998) have studied the liquidity provision of banks and markets in dynamic overlapping generation (OLG) economies. Qi (1994) examines banks' liquidity service and stability in an overlapping generations version of Diamond and Dybvig model. He concludes that the intergenerational transfers enable an intergenerational bank to achieve interest rate smoothing and provide depositors with liquidity insurance without Diamond and Dybvig's assumption of no side trades. Instead, the need for side trades is removed by imposing an incentive-compatible constraint that the long-term interest rate is no less than the long-term real investment return. However, this incentive constraint implicitly assumes a market consumption level that could be achieved in a finite model. Whether or not this constraint will actually eliminate trades in the overlapping generations framework remains an open question. Given two technologies (storage technology and long-term profitable technology), it is not possible for an agent to obtain consumption level higher than the long-term real return in a finite trading model. It is not certain, however, that

with intergenerational trading possible, investors cannot obtain higher returns than the returns offered by banks.

Qian et al. (2004) demonstrate that the full-participation market with intergenerational trading can provide more liquidity than one without. Insurance is provided through wealth transfer across generations, instead of cross-subsidization across contemporaneous types as is the case in the finite economy. Given a full-participation market that allows trading across generations, only banks with initial capital can provide additional liquidity. In a limited-participation market with uncertainty about trading types, an intergenerational bank (with or without initial capital) provides additional insurance to investors. The need for trade is eliminated. Finally, if there is no uncertainty about trading types, then an intergenerational bank with initial capital eliminates the need for trading and improves welfare for all. An intergenerational bank without initial capital improves welfare for people who do not trade.

Finally, our model can be extended to explore the welfare effects of government policies aimed at preventing or dealing with bank runs. Theoretical papers have analysed the ex-ante effects of regulatory policies inlude Freixas and Gabillon (1999), Samartin (2002) or Zhu (2005).

4.7 Appendix

Proof of proposition 1

The problem's constraints allow to express the expected utility function in terms of c_1. We consider two cases. The F.O.C can be expressed as follows:

$$\frac{\partial EU}{\partial c_1} = q_1 c_1^{-\gamma} + q_1 c_1^{-\gamma} B - p(l) \left[\frac{[1-(q_1+q_{2A})c_1]R(l)}{q_{2B}} \right]^{-\gamma} [R(l)(q_1 + q_{2A})] -$$

$$p(h) \left[\frac{[1-(q_1+q_{2A})c_1]R(h)}{q_{2B}} \right]^{-\gamma} [R(h)(q_1 + q_{2A})] = 0 \qquad (4.23)$$

where we denote $B = p(l)R(l)^{1-\gamma} + p(h)R(h)^{1-\gamma}$. It follows that:

$$\frac{\partial EU}{\partial c_1} = c_1^{-\gamma}(q_1 + q_{2A}B) - \left[\frac{[1-(q_1+q_{2A})c_1]}{q_{2B}} \right]^{-\gamma}(q_1 + q_{2A})B = 0 \qquad (4.24)$$

The optimal solution to this problem yields:

$$c_1 = \frac{(q_1+q_{2A}B)^{\frac{1}{\gamma}}}{q_{2B}[(q_1+q_{2A})B]^{\frac{1}{\gamma}}+(q_1+q_{2A})(q_1+q_{2A}B)^{\frac{1}{\gamma}}};$$

$$c_{2A}^l = \frac{R(l)}{R(h)} c_{2A}^h; \qquad\qquad c_{2A}^h = c_1 R(h);$$

$$c_{2B}^l = \frac{R(l)}{R(h)} c_{2B}^h; \qquad\qquad c_{2B}^h = [1 - (q_1 + q_{2A})c_1]\frac{R(h)}{q_{2B}}.$$

where $c_{2A}^l, c_{2A}^h, c_{2B}^l, c_{2B}^h$ are obtained by substituting c_1 in the problem's constraints.

The solution for the particular case that $c_{2B}^l = c_1$ can be obtained by substituting $c_1 = c_{2B}^l$ in the problem constraints yielding:

$$c_1 = \frac{1}{q_1+q_{2A}+\frac{q_{2B}}{R(l)}};$$

$$c_{2A}^l = \frac{R(l)}{R(h)} c_{2A}^h; \qquad\qquad c_{2A}^h = c_1 R(h);$$

$$c_{2B}^l = c_1; \qquad\qquad c_{2B}^h = [1 - (q_1 + q_{2A})c_1]\frac{R(h)}{q_{2B}}.$$

Q.E.D.

Proof of proposition 2

As per section 4.4.2, the market clearing prices must be positive and equalise the supply of short-term and long-term assets at date 1. In addition, the resulting consumption levels according to the incentive-compatible bank withdrawals are as follows:

$$c_1 = w_1 + (1 - \delta)b(s)R(s) \tag{4.25}$$

$$c_{2A}^S = \frac{w_1}{b(s)} + (1 - \delta)R(s) \tag{4.26}$$

$$c_{2B}^S = w_2 + (1 - \delta)R(s) \tag{4.27}$$

where $s \in S = \{l, h\}$.

Equation 4.11 requires that $c_1 = c_{2A}^S \frac{1}{R(s)}$. It follows that

$$w_1 + (1 - \delta)b(s)R(s) = \frac{w_1}{b(s)R(s)} + (1 - \delta) \tag{4.28}$$

Let's denote $X = b(s)R(s)$ and assume $X \neq 1$. Since $b(s) > 0$ *and* $R(s) > 0$, it follows that $X > 0$.

From equation 4.20 if follows that:

$$(1 - \delta)X^2 + [w_1 - (1 - \delta)]X - w_1 = 0 \tag{4.29}$$

$$X = \frac{1}{2} - \frac{w_1}{2(1-\delta)} \pm \frac{(1-\delta)+w_1}{2(1-\delta)} \tag{4.30}$$

where $X = 1$ or $X = \frac{1}{2} + \frac{(\delta-1)-2w_1}{2(1-\delta)} = \frac{-w_1}{(1-\delta)} < 0$ are possible solutions.

Since $w_1 > 0$ and $0 \leq \delta \leq 1$, it follows that $X = 1$ is the only possible solution and therefore $b(s) = 1/R(s)$. Q.E.D.

This page is intentionally left blank.

References

Adrian, T. and H.S. Shin (2010), "Liquidity and leverage", Journal of Financial Intermediation 19(3), 418-437.

Affinito, M. and E. Tagliaferri (2010), "Why do (or did?) banks securitize their loans? Evidence from Italy", Journal of Financial Stability 6(4), 189–202.

Ahnert, T., K. Anand, P. Gai and J. Chapman (2016), "Asset encumbrance, bank funding and financial fragility", Deutsche Bundesbank Discussion Paper 17/2016.

Alhadeff, D.A.and C.P. Alhadeff (1964), "Growth of Large Banks, 1930–1960". Review of Economics and Statistics 46, 356–363.

Allen, F. and D. Gale (1998), "Optimal financial crises," Journal of Finance 53, 1245-1284.

Alonso, I. (1996), "On avoiding bank runs", Journal of Monetary Economics 37, 73-87.

Ayuso, J. and J. Martínez (2006), "Assessing banking competition: An application to the Spanish market for (quality changing) deposits", Working Paper No. 0626, Banco de España.

Babihuga, R. and M. Spaltro (2014), "Bank funding costs for international banks", IMF Working Paper WP/10/153.

Bank of England (2012), "Medium-term risks to financial stability", Financial Stability Report No. 31, 30–46.

Bank of England (2012), "Why is their recovery better than ours? (Even though neither is good enough)", speech by Adam Posen at the National Institute of Economic and Social Research, London, 27 March 2012.

Bank of England (2014), "Bank funding costs: what are they, what determines them and why do they matter?", Quarterly Bulletin Q4.

Basel Committee on Banking Supervision (BCBS) (2013), "Basel III: The liquidity coverage ratio and liquidity risk monitoring tools".

Belaisch, A., L. Kodres and J. Levy (2001), "Euro-area banking at the crossroads", Working Paper no.01/28, International Monetary Fund.

Berger, A.N., R. Demsetz and P.H Strahan (1999), "The consolidation of the financial services industry: Causes, consequences and implications for the future", Journal of Banking and Finance 24, 135–194.

Berger, A.N., W.L. Hunter and S.G. Timme (1993), "The efficiency of financial institutions: a review and preview of research past, present and future", Journal of Banking and Finance 17, 221-249.

Berger, A.N., A.K. Kashyap and J.M. Scalise (1995), "The transformation of the U.S. banking industry: What a long, strange trip it's been", Brookings Papers on Economic Activity 2, 55–201.

Berger, A.N., N.H. Miller, M.A. Petersen, R.G. Rajan, and J.C. Stein (2005), "Does function follow organizational form? Evidence from the lending practices of large and small banks", Journal of Financial Economics 76, 237-269.

Bhattacharya, S. and D. Gale (1987), "Preference shocks, liquidity, and central bank policy". In: W.A. Barnett and K.J. Singleton, (Eds.), New Approaches to Monetary Economics, Cambridge University Press.

Bikker, J.A. (2003), "Testing for imperfect competition on EU deposit and loan markets with Bresnahan's market power model", Research Series Supervision No. 52, De Nederlandsche Bank.

Bond, S., C., Nauges and F. Windmeijer (2005), "Unit roots: Identification and testing in micro panels", Cemmap Working Papers CWP07/05.

Boot, A. (2000), "Relationship banking: What do we know?", Journal of Financial Intermediation 9, 7–25.

Breitung, J. and W. Meyer (1994), "Testing for unit roots in panel data: Are wages on different bargaining levels cointegrated?", Applied Economics 26, 353–361.

Brunnermeier, M. K. and L.H. Pedersen (2009), "Market Liquidity and Funding Liquidity", Review of Financial Studies 22, 2201-2238.

Brunnermeier, M. K., Gorton, G. and A Krishnamurthy (2014), "Liquidity Mismatch". Risk Topography. University of Chicago Press.

Bryant, J. (1980), "A model of reserves, bank runs and deposit insurance", Journal of Banking and Finance 4, 335-344.

Cabral, I., F. Dierick and J. Vesala (2002), "Banking integration in the Euro area", European Central Bank Ocassional Paper Series No. 6, European Central Bank.

Calomiris, C.W. and G. Gorton (1991), "The origins of banking panics: models, facts, and bank regulation". In R. Glenn Hubbard (Ed.), Financial Markets and Financial Crises, University of Chicago.

Caminal, R., J., Gual and X. Vives (1993), "Competition in Spanish banking". In: J. Dermine (Ed.), European Banking in the 1990s, Basil Blackwell, Oxford.

Campa, J.M. and I. Hernando (2006), "M&A performance in the European financial industry", Journal of Banking and Finance 30, 3367-3392.

Carbó-Valverde, S., J. Fernández de Guevara, D. Humphrey and J. Maudos (2005), "Estimating the intensity of price and non-price competition in banking. An application to the Spanish case", Working Paper No. 2, Fundación BBVA.

Carbó-Valverde, S., D. Marques-Ibanez and F. Rodríguez-Fernández (2012), "Securitization, risk-transferring and financial instability: The case of Spain", Journal of International Money and Finance 31(1), 80–101.

Centre for Economic Policy Research (2014), "Restarting European Long-Term Investment Finance", Green Paper Discussion Document.

Chang R. and A. Velasco (2000), "Financial fragility and the exchange rate regime" Journal of Economic Theory 92, 1-34.

Chang R. and A. Velasco (2001), "A model of financial crises in emerging markets," Quarterly Journal of Economics 116, 489-517.

Chari, V.V. and R. Jagannathan. (1988), "Banking panics, information and rational expectations equilibrium", Journal of Finance 42, 749-61.

Chesher, A. (1979), "Testing the law of proportionate effect", Journal of Industrial Economics 27, 403–411.

Chiaramonte, L. and B. Casu (2013), "The determinants of bank CDS spreads: evidence from the financial crisis", European Journal of Finance 19(9), 861–887.

De Nederlandsche Bank (2016), "Keeping encumbrance to a minimum", April 2016 Newsletter.

Committee on the Global Financial System (CGFS) (2011), "The impact of sovereign credit risk on bank funding conditions", CGFS Papers (Bank for International Settlements) No. 43.

Committee on the Global Financial System (CGFS) (2013), "Asset encumbrance, financial reform and the demand for collateral assets", CGFS Papers (Bank for International Settlements) No. 49.

Cooper, R. and T.W. Ross (1998), "Bank runs: Liquidity costs and investments distortions", Journal of Monetary Economics 41, 27-38.

Demsetz, R. and P. Strahan (1997), "Diversification, size and risk at bank holding companies", Journal of Money Credit and Banking 29, 300–313.

Delgado, J., V. Salas and J. Saurina (2007), "Joint size and ownership specialization in bank lending", Journal of Banking and Finance 31(12), 3563–3583.

Diamond, D.W. (1997), "Liquidity, banks and markets". Journal of Political Economy 105, 928-956.

Diamond, D.W. and P.H. Dybvig, (1983), "Bank runs, deposit insurance and liquidity", Journal of Political Economy 91, 401-419.

Díaz, B., M. García and S. Sanfilippo (2004), "Bank acquisitions and performance: Evidence from a panel of European credit entities", Journal of Economics and Business 56, 377–404.

Dunne, P. and A. Hughes (1994), "Age, size growth and survival: UK Companies in the 1980s", Journal of Industrial Economics 42, 115–140.

Dunne, T., M. Roberts, M. And L. Samuelson (1989), "The growth and failure of US manufacturing plants", Quarterly Journal of Economics 96, 671–698.

Eisenbach, T., T. Keister, J. McAndrews and T. Yorulmazer (2014), "Stability of funding models: an analytical framework". FRBNY Economic Policy Review 20 No. 1.

Ennis, H.M. (2001), "On the size distribution of banks", Federal Reserve Bank of Richmond Economic Quarterly 87, 1–25.

European Banking Authority (2014), "Guidelines on disclosure of encumbered and unencumbered assets", EBA/GL/2014/03.

European Banking Authority (2016), "EBA report on asset encumbrance", June 2016.

European Banking Authority (2017), "EBA risk dashboard", April 2017.

European Central Bank (2002), "Structural analysis of the banking sector", November.

European Central Bank (2006), "Report on EU banking structures", October.

European Central Bank (2012), "Changes in bank funding patters", April.

European Central Bank (2016), "Report on financial structures", October.

European Commission (2015), "Commission implementing regulation (EU) 2015/79 of 18 December 2014".

European Systemic Risk Board (2013), "Recommendations of the European Systemic Risk Board of 20 December 2012 on funding of credit institutions".

European Systemic Risk Board (2014), "Is Europe Overbanked?", Report of the Advisory Scientific Committee No. 4, June.

Evans, D.S. (1987a), "Tests of alternative theories of firm growth", Journal of Political Economy 95, 657–674.

Evans, D.S. (1987b) "The relationship between firm growth, size and age estimates for 100 manufacturing industries", Journal of Industrial Economics 35, 567–581.

Fariñas, J.C. and L. Moreno (2000), "Firms' growth, size and age: A nonparametric approach", Review of Industrial Organization 17, 249–265.

Focarelli, D., F. Panetta. F. and C. Salleo (2002), "Why do banks merge?", Journal of Money, Credit and Banking 34, 1047–1066.

Freixas, X. and E. Gabillon (1999), "Optimal regulation of a fully insured deposit banking system", Journal of Regulatory Economics 16, 111-134.

Gai, P., A.G. Haldane, S. Kapadia and B. Nelson (2013), "Bank funding and financial stability", in A. Heath et al. (eds), Liquidity and Funding Markets, Proceedings of the 2013 Annual Conference of the Reserve Bank of Australia.

Geanakoplos, J. (2010), "The Leverage Cycle", in D. Acemoglu, K. Rogoff, and M. Woodford (eds.), NBER Macroeconomics Annual 2009, Vol. 24, University of Chicago Press, 1-65.

Geroski, P.A. (1999) "The growth of firms in theory and practice", CEPR discussion paper No. 2092.

Gibrat, R. (1931), "Les Inegalities Economiques", Sirey, Paris.

Goddard, J.A., J.O.S. Wilson and P. Blandon (2002a), "Panel tests of Gibrat's law for Japanese manufacturing", International Journal of Industrial Organization 20, 415–433.

Goddard, J.A., D.G. McKillop and J.O.S Wilson (2002b), "The growth of US credit unions", Journal of Banking and Finance 26, 2327-2356.

Goddard, J.A, P. Molineux and J.O.S. Wilson (2004), "Dynamics of growth and profitability in banking", Journal of Money, Credit and Banking 36, 1069-1090.

Goldstein, I. and A. Pauzner, (2005), "Demand-deposit contracts and the probability of bank runs," Journal of Finance 60, 1293-1327.

Gorton, G. (1988), "Banking panics and business cycles," Oxford Economic Papers No. 40, 751-81.

Gorton, G. and A. Metrick (2012), "Securitized banking and the run on repo," Journal of Financial Economics 104(3), 425-451.

Gual, J. (1993), "La competencia en el mercado español de depósitos bancarios", Moneda y Crédito 196, 143-179.

Haldane, A.G. (2012), "Financial arms races", speech delivered at the Institute for New Economic Thinking, Berlin.

Hall, B. (1987), "The relationship between firm size and firm growth in the US manufacturing sector", Journal of Industrial Economics 35, 583–606.

Hardy, D.C. (2014), "Bank resolution costs, depositor preference and asset encumbrance", Journal of Financial Regulation and Compliance 22(2), 96-114.

Hasan, I., L. Liu and G. Zhang (2015), "The determinants of global bank credit-default-swap spreads", Journal of Financial Services Research 50(3), 1–35.

Helberg, S. and S. Lindset (2014), "How do asset encumbrance and debt regulations affect bank capital and bond risk?", Journal of Banking and Finance 44(C), 39-54.

Houben, A. and J.W. Slingenberg (2013), "Collateral scarcity and asset encumbrance: implications for the European financial system", Banque de France Financial Stability Review 17, 197-206.

International Monetary Fund (IMF) (2013), "Changes in bank funding patterns and financial stability risks", Global Financial Stability Report, Chapter 3, October.

International Capital Markets Association (ICMA) (2013), "Does repo 'encumber' a borrower's assets?", FAQ available at www.icmagroup.org.

International Swaps and Derivatives Association (ISDA) (2015), "ISDA margin survey 2015".

Jacklin, C.J. (1987), "Demand deposits, trading restrictions, and risk sharing". In: E.C. Prescott and N. Wallace (Eds.), Contractual Arrangement in Intertemporal Trade. University of Minnesota Press, Minneapolis.

Jacklin, C. J. and S. Bhattacharya (1988), "Distinguishing panics and information-based bank buns: Welfare and policy implications", Journal of Political Economy 96, 568-92.

Janicki, H.P. and E.C. Prescott (2006), "Changes in the size distribution of U.S. banks: 1960–2005", Federal Reserve Bank of Richmond Economic Quarterly 92, 291–316.

Iannotta, G., Nocera, G., Sironi, A. (2013). "The Impact of Government Ownership on Bank Risk", Journal of Financial Intermediation 22(2): 152-176.

Jeitschko, T. D. and C. R. Taylor (2001), "Local discouragement and global collapse: A theory of information avalanches", American Economic Review 91, 208-24.

Jones, K.D. and T. Critchfield (2005), "Consolidation in the U.S. banking industry: Is the long, strange trip about to end?", FDIC Banking Review 17, 31–61.

Jovanovic, B. (1982), "Selection and evolution of industry", Econometrica 50, 649–670.

Juks, R. (2012), "Asset encumbrance and its relevance for financial stability", Sveriges Riksbank Economic Review 3, 1-13.

Kanagaretnam, K., G. Zhangb and S.B. Zhangd (2016), "CDS pricing and accounting disclosures: Evidence from U.S. bank holding corporations around the recent financial crisis", Journal of Financial Stability 22, 33–44.

Krishnamurthy, A., S. Nagel, and D. Orlov (2014), "Sizing up repo", Journal of Finance 69(6), 2381-2417.

Le Leslé, V. (2012), "Bank debt in Europe: 'Are funding models broken'", IMF Working Paper WP/12/299.

Matta, R. and E. Perotti (2015), "Insecure debt", SRC Discussion Paper No 41.

Maudos, J. and X. Vives (2016), "Banking in Spain". In: Thorsten Beck and Barbara Casu. The Palgrave Handbook of European Banking. UK: Palgrave Macmillan, 567 - 601.

Merton, R.C. (1977), "An analytical derivation of the cost of deposit insurance and loan guarantees: An application of modern option pricing theory", Journal of Banking and Finance 1, 3-11.

Ötker-Robe, I. and J. Podpiera (2010), "The fundamental determinants of credit default risk for European large complex financial institutions", IMF Working Paper WP/10/153.

Panetta, F. and A. Pozzolo (2010), "Why do banks securitize their assets? Bank-level evidence from over one hundred countries". In Proceedings of the 2010 Financial Intermediation Research Society Conference (1483).

Park, S. (1997): "Risk-taking behavior of banks under regulation", Journal of Banking and Finance 21, 491-507.

Peek, J. and E.S. Rosengren (1996), "Small business credit availability: How important is size of lender?". In: Saunders, A., Walter, I. (Eds.), Financial System Design: The Case for Universal Banking. Irwin (Richard D.), Burr Ridge, IL.

Perotti, E. (2010), "Systemic liquidity risk and bankruptcy exceptions", CEPR Policy Insight No. 52.

Rhoades, S.A. and A.J. Yeats (1974), "Growth, consolidation and mergers in banking", Journal of Finance 29, 1397–1405.

Qian, Y., K. John and T. John (2004), "Financial system design and liquidity provision by banks and markets in a dynamic economy", Journal of International Money and Finance 23, 385-403.

Salas, V. and J. Saurina (2003), "Deregulation, market power and risk behaviour in Spanish banks", European Economic Review 47, 1061-1075.

Samartín, M. (2002), "Should bank runs be prevented?", Journal of Banking and Finance, 27(5), 977-1000.

Singh, M. (2010), "Collateral, Netting and Systemic Risk in OTC Derivatives Market" IMF Working Paper WP/10/99.

Stein, J.C. (2002), "Information production and capital allocation: Decentralized vs. hierarchical firms", Journal of Finance 57, 1891–1921.

Sutton, J. (1997), "Gibrat's legacy", Journal of Economic Literature 35, 40–59.

Tschoegl, A.E. (1983), "Size, growth and transnationality among the world's largest banks", Journal of Business 56, 187–201.

van Rixtel, A. and G. Gasperini (2013), "Financial crises and bank funding: recent experience in the euro area", BIS Working Paper No. 406.

World Economic Forum (2014), "Why has the US recovered more quickly than Europe?", article by Daniel Gros, Director of the Brussels-based Center for European Policy Studies.

Wilson, J.O.S. and J. Williams (2000), "The size and growth of banks: Evidence from four European countries", Applied Economics 32, 1101–1109.

Yeats, A.J., E.D. Irons and S.A. Rhoades (1975), "An analysis of new bank growth", Journal of Business 48, 199–203.

Zhu, H. (2001), "Bank runs without self-fulfilling prophecies", BIS Working Paper No. 106.