“Too-Small-To-Survive” versus “Too-Big-To-Fail” banks: The two sides of the same coin

Theoharry Grammatikos[a], Nikolaos I. Papanikolaou[b*]

[a] *Luxembourg School of Finance, University of Luxembourg, 4 rue Albert Borschette, L-1246, Luxembourg*

[b] *University of Sussex, School of Business, Management and Economics, Brighton, BN1 9SL, UK*

Abstract

In the recent crisis, the U.S. authorities bailed out numerous banks through TARP, whilst let many others to fail as going concern entities. Even though both interventions fully protect depositors, a bail out represents an implied subsidy to shareholders, which is not yet the case with closures where creditors are not subsidised. We investigate this non-uniform policy, demonstrating that size and not performance is the decision variable that endogenously determines one threshold below which banks are treated as TSTS by regulators and another one above which are considered to be TBTF. Our results suggest that regulators do not bailout the shareholders or the other uninsured creditors of a distressed bank if the bank is considered to be TSTS. Further, that the more complex a bank is the more likely is to be bailed out and, hence, to have all of its creditors protected. Banks which are perceived as being TBTF are also found to be too-complex-to-fail.

Keywords: Too-Big-To-Fail; Too-Small-To-Survive; threshold modelling; bank size; complexity

JEL classification: D02; G01; G21; G28

---

* Corresponding author. Nikolaos I. Papanikolaou: n.papanikolaou@sussex.ac.uk; Tel: +44 (0) 1273678738
Theoharry Grammatikos: theoharry.grammatikos@uni.lu; Tel: +352 4666446939

1. Introduction

In October 2008, the U.S. Congress passed the Emergency Economic Stabilization Act (EESA) and authorised the Department of the Treasury to launch the Troubled Asset Relief Program (TARP) with the purpose to offer emergency financial aid to corporate firms but, most importantly, to bolster the resiliency of banking institutions. Not surprisingly, many banks that had been largely affected by the turmoil in credit markets triggered by the U.S. subprime mortgage crisis received TARP funds via the Capital Purchase Program (CPP), which was the key component of TARP. Through CPP, the U.S. Treasury invested up to $250 billion in the preferred equity of banks to enhance their capital ratios. The primary aim of this rescue package was the prevention of the sudden and simultaneous collapse of a large number of distressed banks, which would have had destructive effects on the entire financial system.

Nonetheless, every coin has two sides: on 28 September 2007, NetBank was the first banking firm to fail as a going concern entity in the U.S. in the recent financial crisis. The Federal Deposit Insurance Corporation (FDIC) took receivership of NetBank and all the insured deposit accounts were transferred to an assuming institution. Some days later, on 4 October 2007, Miami Valley Bank was also shut down by the authorities. The collapse of Miami Valley Bank was followed by those of Douglas National Bank and Hume Bank in early 2008. Importantly, the number of failures increased rapidly from 2008 onwards. In total, more than 500 collapses were recorded during the recent crisis. The FDIC was appointed receiver of the bankrupt institutions and this inflicted a total loss of $76 billion on the system.

Accordingly, the U.S. federal authorities -like European and other national authorities worldwide- provided substantial financial aid to a number of troubled banking organisations during the crisis while, at the same time, allowed many others to go bankrupt as going concern entities. The FDIC-backed resolution mechanism is designed to cope with the insolvency of distressed banks under either normal or unstable economic conditions, whereas TARP is viewed as an emergency mechanism designed to cope with bank fragility. Even though under both interventions small bank depositors remain fully protected, TARP government bailouts represent an implicit subsidy to the bank's shareholders, which is not yet the case with FDIC-backed failures where shareholders are not subsidised. As documented in Gandhi and Lustig (2015), government guarantees in the form of bailouts protect the shareholders of large banks, but not those of small banks in a financial disaster. In a similar vein, Veronesi and Zingales (2010) calculate the costs and benefits of TARP from the perspective of big banks' shareholders and conclude that these firms receive large subsidies. Hence, the question should not be bailout vs. failure, but rather subsidy for stockholders (via TARP) vs. only protecting depositors (via FDIC-backed failures) with severe consequences for the uninsured creditors and the stockholders of the failed banks.

Such a differential treatment of distressed banks raises some important questions: do regulators consider some institutions as being very important for the system -or alternatively 'Too-Big-To-Fail' (TBTF)- in the sense that a collapse of any of them has to be deterred for not to trigger contagious defaults in the entire banking network, whereas some others are perceived as being 'Too-Small-To-Survive' (TSTS) in that their failure as going concern entities has no material impact on their counterparts, let alone on the system as a whole? Is it the size of financial institutions the key determinant that makes the authorities to treat distressed banks differently, or it is that the failed banks lag behind in terms of performance compared to those that the authorities decide to financially support via TARP? To put it differently, is it that regulators are reluctant to support the uninsured creditors and the shareholders of some distressed banks because they consider these banks as being TSTS? Crucially, is there any specific threshold size below which a banking firm is viewed as TSTS by regulatory authorities? And, if so, is there a relevant threshold size for TBTF banks? And, finally, what is the role of bank complexity in regulators' interventions in the case of financial distress and how complexity interacts with size?

Admittedly, size lies in the very centre of banking research. Banks of different sizes follow diverse business models which are related to various levels of risk, increased or reduced earnings, higher or lower failure probabilities and carry a different weight for the financial system. Notwithstanding the fact that these variations have been well-documented in the extant literature, the still growing crisis literature, within which our study falls, has not paid the necessary attention to the role that size and complexity play in the decision of regulatory authorities of how to treat a distressed bank and its creditors. Hence, in this paper, we focus on the recent financial crisis aiming to provide concrete answers to the aforementioned questions.

We collect data for the entire population of U.S. commercial and savings banks and distinguish our sample banks into two key groups: the non-distressed banks, and the distressed banks composed by the TARP and the FDIC-backed (failed) banks. We first conduct a univariate analysis to compare the average size and the performance of each banking group in the years prior to the onset of the crisis, accounting also for the organisational and operational complexity of banks given the systemic implications of complexity in resolving distressed banks and the relative importance of complexity vs. size. The measurement of bank performance relies on the rating system, which has been utilised by U.S. regulators for more than two decades now to monitor the safety and soundness of banks. We, then, employ a multivariate technique which can endogenously define one or more threshold levels of an observed variable to examine whether we can specify one threshold size below which banks are considered by the authorities as being TSTS and a second one above which banks are considered to be TBTF. To put it differently, what we are modelling is whether FDIC funds or TARP funds are used to support the distressed banks in the recent crisis, and whether the bailout extends beyond bank debts.

The univariate analysis suggests that the decision of regulatory authorities to choose between a TARP-assisted bailout and a FDIC-backed failure is more influenced by bank's size and less influenced by bank's performance, complexity, and risk-taking profile. The regression results of the multivariate threshold analysis reveal that the failure and TARP probabilities are, in essence, determined by bank size. In this context, our threshold technique endogenously specifies two cut-off points for size: one for the TSTS banks and another one for the TBTF banks, which are considered to be the two sides of the same coin. Regulators choose not to bailout the shareholders and the uninsured creditors of a distressed bank if the bank is considered to be TSTS. From a market equilibrium viewpoint, the existence of these size regimes means that the free market outcome cannot be reached because it is the size and not the performance of banks which is the key decision variable for the failure and TARP probabilities. The complexity of banks are found to be negatively related to the probability of failure across the different size regimes, implying that the more complex a bank is the more likely is to receive TARP assistance and, hence, to have all of its creditors protected. The impact of complexity on failure is considerably stronger in the TBTF regime and the estimates are highly statistically significant in this regime, showing that banks which are perceived as TBTF are also too-complex-to-fail.

The paper proceeds as follows. Section 2 reviews the importance of size and complexity in the banking literature. The aim of this Section is not to provide an extensive review of the past and the current literature; rather, its key purpose is to present how size is intertwined with bank performance, and discuss the relevance of bank size and complexity through the lenses of the most influential studies. Section 3 presents our data set and outlines how it is constructed; the relevant key variables and the univariate analysis we conduct are also presented in this Section. The multivariate threshold econometric technique we employ is fully described in Section 4. The regression results are presented and discussed in Section 5; a set of interesting policy implications drawn from the results are also offered. Section 6 is devoted to sensitivity analysis, whereas Section 7 provides a brief summary of our main findings and offers some concluding remarks.

2. Bank size and complexity

In September 1984, the Office of the Comptroller of the Currency (OCC) in U.S. made, for the first time, a public distinction between TBTF and non-TBTF banking institutions. It, specifically, announced that the biggest 11 from a total of approximately 14,000 banks that were in operation at that time were considered as being TBTF and as such they would be offered a full deposit insurance, whereas all the other banks would remain only partially covered. After that announcement, the spotlight of the relevant literature turned to shine large banking organisations and to examine the importance of size for the smooth functioning of the whole financial system. Of the most prominent studies in the early TBTF banking literature were those of O'Hara and Shaw (1990), Boyd and Runkle (1993), Demsetz and Strahan (1997), and Galloway et al. (1997).

A significant part of the current banking literature, which has been sparked by the emergence of the global financial crisis, focuses on the relevance of TBTF banks in the propagation of the crisis and its subsequent dissemination throughout the global economy. For instance, Huang et al. (2009) construct a framework for measuring and stress testing the systemic risk of 12 U.S. major commercial and investment banks. Adrian and Shin (2010) examine the procyclicality in leverage of the 5 biggest U.S. investment banks before the outbreak of the crisis. Patro et al. (2013) uses the 22 largest commercial and investment banks in U.S. to analyse the relevance of stock return correlations in assessing the level of systemic risk. In a similar context, Papanikolaou and Wolff (2014) focus on 20 U.S. systemically important banks to study how the aforementioned modern activities affect the overall risk profile of banks as well as the level of systemic risk before and after the onset of the global financial crisis. More recently, Kanno (2015) focuses on banks with over $50 billion in total assets to assess systemic risk based on interbank exposures in the global financial system.

From a somewhat different perspective, De Haan and Poghosyan (2012) investigate whether the volatility of bank earnings during the recent crisis depends on the size of banks and the level of concentration in the banking industry. They document that larger banking institutions, which are located in highly concentrated markets are those that experience higher volatility. In a similar vein, Bertay et al. (2013) use an international sample of banks to examine the extent to which a bank's risk profile, profitability, activity mix, funding strategy, and the level of market discipline depend on both its absolute and systemic size. They conclude that bank returns increase with absolute size and decrease with systemic size; also, that large banks are subjected to greater market discipline compared to smaller banks. Gandhi and Lustig (2015) investigate the asset pricing implications of financial shocks based on historical data of bank stock returns in the U.S. They document the existence of a factor in the component of bank returns, which measures the size-dependent exposure to bank-specific tail risk.

Size is also employed in the literature to investigate the likely differences in the business models and the performance between the large banking firms and their smaller counterparts. For instance, size has been found to be amongst the key factors in the decision of a bank to follow a specific business model. Focusing on the U.S. banking market and classifying banks into different size classes, DeYoung et al. (2004) show that the deregulation process and the technological changes of the ’80s and the ’90s gave birth to two main bank size groups: the first group consists of big banks, whose operation is characterised by the use of ‘hard’ information, impersonal relationships with their customers, low unit costs, and standardised loans; the second group contains small banks that collect and make use of ‘soft’ information, develop more personal relations with their customers, face higher unit costs, offer non-standardised loans, and provide the bulk of financing to small business firms. By the same token, Berger et al. (2005) find that small banks have a comparative advantage in making loans based on ‘soft’ information and this is due to the different sets of incentives in the organisational structures of small and large banks. Further, Carter and McNulty (2005) document an inverse relationship between the size of banking firms and the net return on small business lending, suggesting that smaller banks perform better than larger banks in the relevant loan market. On the other hand, larger banks are found to have a comparative advantage in credit card lending, a market characterised by impersonal relationships and standardised loans.

Over the past two decades or so, banks have turned to adopt sophisticated organisational structures and to follow complex business models. In the U.S., the spectrum of banking activities has been expanded dramatically into near- and non-bank business lines since the repeal of the Gramm-Leach-Bliley Act in 1999. Indeed, banks have been diversified away from the traditional intermediation services of deposit-taking and loan-granting into structured, market-based products like securitised assets and financial derivatives. Rime and Stiroh (2003) show that large banks are very prone to the so-called ‘universal activities’ in contrast to small and mid-sized institutions, which are less diversified and resemble single-line businesses. Indeed, banking activities are nowadays characterised by a higher degree of complexity and duress, and the business models of banks have become more opaque. Moreover, a number of banking institutions have multiple branches across regions and prefectures and are largely involved in cross-border activity areas and, hence, are subject to different legal and regulatory jurisdictions. An important policy implication is that these complex operational and organisational structures do not facilitate the orderly resolutions of troubled institutions.

The literature on bank complexity is rather scarce. Interest on complexity has been sparked only after the outbreak of the global financial crisis. Caballero and Simsek (2013) conceptualise complexity as the uncertainty of banks about their cross exposures: banks know their own exposures but they are uncertain about the exposures and the health of their counterparties in their business network. DeYoung et al. (2013) examine the relationship between the complexity of failed banks and the relevant resolution process. They document a too-complex-to-fail resolution strategy, which lies in the inability of regulators to credibly commit to closing insolvent complex banks thus encouraging banks to increase their level of complexity. Cetorelli et al. (2014) introduce two broad measures of organisational complexity and business diversification and empirically assess the complexity of a large set of global banks. Their results reveal a steady growth in the average complexity over time. As regards the studies of Berger and Bouwman (2013) and Berger and Roman (2015), they both account for the degree of complexity of their sample banks in their empirical analyses using measures that rely on the number of bank branches and the number of U.S. states that a bank is active. Recently, Eisenbach et al. (2016) examined the trade-offs between the benefits and costs of supervision to interpret the relationship that holds between supervisory efforts and bank characteristics. Their findings show that more supervisory resources are spent on larger, more complex, and riskier banks.

3. Data, key variables, and univariate analysis

3.1. Data

We focus on U.S. commercial and savings banking institutions that file a Report on Condition and Income (also known as Call Report). Thrifts -*i.e.*, savings and loans associations- are excluded from our empirical analysis because they file a different report (the Thrift Financial Report).[1] Data are of quarterly frequency and extend from the beginning of 2002 (2002q1) to the end of 2012 (2012q4) when the banking crisis in the U.S. is generally thought to have come to a halt. We consider the fourth quarter of 2007 (2007q4) to be the starting point of the crisis. Indeed, that was the time when the TED spread (the difference between the yield on the three-month London Interbank Offered Rate -i.e., LIBOR- and the yield on three-month U.S. Treasury bills) which is one of the most widely-used indicators of credit risk, widened to almost 200 basis points relative to a historically stable range of 10-50 basis points. We do not examine the years prior to 2002 because the two international financial crises which erupted in East Asia and in Russia towards the end of the '90s combined with the Long Term Capital Management (LTCM) crisis in late 1998 and the dot-com bubble crisis of the early 2000s all had a considerable destabilising impact on the operation of international financial markets and on the U.S. banking system.

We begin with 8,905 active commercial and savings banking institutions that filed a Call Report in 2002q1. We make a distinction between non-distressed and distressed (failed and TARP) banks in the crisis period. After checking the data for reporting errors and other inconsistencies, we end up with a total of 7,704 banks of which 6,431 are non-distressed and the remaining 1,273 are distressed banks. Of the distressed banks, in turn, 449 were allowed to fail as going concern entities and 824 survived as going concern entities through TARP government support.

[1] With the implementation of the Dodd-Frank Act and the establishment of the Office of Thrift Supervision in July 2011, all thrifts were required to file and submit a Call Report from March 2012.

3.2. Non-distressed banks

The group of non-distressed banks consists of all the banking firms that stayed afloat as going concern entities during the recent financial meltdown. More specifically, these banks neither failed, nor received a TARP assistance, nor merged with or acquired by some other institution throughout the entire sample period. It is important to mention here that the 46 banks which either failed or merged with or acquired by another bank at some later point in time that is not covered in our data period -that is, from 2013q1 to 2015q4- are excluded from our set of non-distressed banks.

3.3. Distressed banks

Distressed banks are those which either failed as going concern entities during the crisis or received TARP assistance as we discuss below.

3.3.1. Failed banks

Failed banks are defined as the insured commercial and savings banks that were closed requiring disbursements by the FDIC from the onset of the crisis in mid-to-late 2007 through the end of our data period. Generally, a bank is closed when regulatory authorities determine that it is critically undercapitalised and deem it unable to meet its obligations to depositors and others. The FDIC acts as a receiver and is in charge of the failure resolution process. There are two main failure resolution mechanisms: the 'purchase-and-assumption' transaction and the 'deposit payoff'. Under the former one, which is the preferred and the most common resolution mechanism, the failed banking institution's insured deposits are transferred to a successor institution, and its charter is closed. The acquiring bank may also assume additional liabilities (mainly part or all of the uninsured deposits) and purchase the assets (primarily, loans) of the failed bank. Insured depositors become depositors of the assuming bank and obtain immediate access to their insured funds. The FDIC usually provides assistance to the acquirer most often in the form of loan loss sharing agreements. In the case of remaining assets and liabilities, these are liquidated and the liquidation costs are internalised. In several purchase-and-assumption transactions, the acquiring bank compensates the FDIC for the franchise value from the failed bank's established customer relationships, which helps reduce the insurer's resolution cost. In a deposit payoff, on the other hand, the FDIC pays depositors the full amount of their insured deposits directly and the failed bank's charter is closed. Deposit payoffs occur when there is no bank acquirer. Accordingly, insured depositors are fully protected under both failure resolution mechanisms even though the failed bank's charter is terminated. Typically, after insured depositors are paid, uninsured depositors are paid next, followed by creditors and then stockholders. In most failure cases, however, general creditors and stockholders are not protected, thus realising little or no recovery.

In total, for the period starting from 2007q4 and extending to 2012q4, there have been recorded 396 failures of commercial banks and 53 failures of savings banks based on the relevant data collected from the FDIC web site.[2] Out of these 449 failures, 427 were purchase-and-assumption transactions and 22 were deposit payoff transactions. Hence, in the vast majority of failures, the distressed bank was acquired by another bank via FDIC assistance.

[2] The names of the banks, their distribution across the U.S. states and cities, the date that every failed institution ceased to exist as a privately-held going concern entity, the estimated assets and deposits of each institution at the time of failure, and the cost of every individual failure for FDIC are all available upon request.

3.3.2. TARP banks

TARP has been the largest U.S. government bailout programme in history. It authorised the Treasury to inject loads of capital into distressed banks by purchasing senior preferred shares. Those injections were intended to restore the health and increase the soundness of the participated banks by helping them to address liquidity shortages and strengthen their capital base. Banks were scheduled to repay or redeem the preferred stock at an undetermined time, but the programme required them to pay an established dividend rate and interest rate to the Treasury as long as the securities were outstanding. In the context of TARP, the bailout extended beyond the debts of the bank in the sense that it not only protected depositors as a whole, but also sheltered creditors and shareholders who enjoyed implicit government guarantees.

The literature identifies two key phases of TARP.[3] In the first phase, nine of the largest financial institutions were arm twisted by the authorities to participate in the programme. Indeed, on October 14, 2008 that the Treasury announced CPP, the nine banks, which together accounted for the 55 percent of US banks' assets, announced that they would subscribe to the facility in a total amount of $125 billion. The nine institutions were Bank of America, Citigroup, JP Morgan Chase, Wells Fargo, Morgan Stanley, Goldman Sachs, Bank of New York Mellon, State Street, and Merrill Lynch. During the second phase of TARP which ended in November 14, 2008, all other publicly held financial institutions were eligible to apply for financial assistance. Accordingly, in the first phase, participation in CPP was mandatory, whereas, in the second phase, banks were not forced but chose to issue preferred stock after having voluntarily applied and being approved for issuance.

To construct the sample of TARP banks, we refer to the complete list of TARP recipients (i.e., both voluntary and involuntary recipients) as obtained from the U.S. Department of the Treasury.

[3] See Calomiris and Kahn (2015) and Kim and Stock (2012) for a thorough analysis of the different phases of TARP.

This list discloses all the financial institutions that received TARP funds via CPP together with the respective transaction dates and investment amounts. We trace all commercial and savings banks which participated in the programme either directly, or through their parent (bank holding) companies. In total, we identify 736 TARP investment transactions excluding any multiple transactions, *i.e.*, transactions in which a bank is involved in more than once. Out of these 736 institutions that received capital injections, 47 were thrifts which, as earlier mentioned, are excluded from our analysis. This leaves 689 institutions in our sample, out of which 596 are Bank Holding Companies (BHCs) and 93 are commercial and savings banks. We assume that if a BHC was approved to participate in TARP, its subsidiary banks would have received some fraction of TARP funds. Out of 596 BHCs that participated in TARP, 56 were multi-BHCs, while the remaining 540 were mono-BHCs. We match all BHCs to their subsidiary commercial and savings banks by hand-matching the relevant information found in the Consolidated Financial Statements for Bank Holding Company Report (FR Y9-C Report) to the 'higher-holder' codes of the examined banks found in Call Reports. By doing so, we obtain a total of 731 FDIC-insured banks that received TARP funds via their parent holding companies. We add to this figure the 93 commercial and savings banks which are not linked to some BHC to construct the final sample of 824 banks that received TARP support.[4]

3.4. Key variables

We can now turn to describe the key variables that we employ in our empirical analysis. All the balance sheet variables are of quarterly frequency and are collected from Call Reports as found in the website of the Federal Reserve Bank of Chicago and that of the Federal Financial Institutions

[4] The detailed list of these banks is available upon request.

Examination Council (FFIEC) Central Data Repository's Public Data Distribution. Interest rates and yields are collected from the Federal Reserve Board and the U.S. Department of the Treasury and are also of quarterly frequency. All variables and the relevant data sources are summarised in Appendix A.

Bank performance relies on the individual components of the CAMELS rating system, which has been utilised by U.S. regulators for more than two decades now to monitor the safety and soundness of individual banks. CAMELS system consists of the following six components: Capital adequacy, Asset quality, Management expertise, Earnings strength, Liquidity, and Sensitivity to market risk. We follow the relevant literature (see, e.g., Stojanovic et al., 2008; Duchin and Sosyura, 2012; Klomp and de Haan, 2012) to construct a vector of bank performance and risk-taking measures, which is designed to resemble the original CAMELS components. We use the standard equity-to-assets ratio as an indicator of bank capital strength (*CAP1*); asset quality is measured by the ratio of non-performing loans to total loans and leases (*ASSETQLT1*); the quality of bank management is measured by managerial efficiency as calculated by the input-oriented Data Envelopment Analysis (*MNGEXP1*);[5] the return on assets expressed as the ratio of total net income (given by the difference between total interest plus non-interest income and total interest plus non-interest expense) to total assets is applied as a measure of earnings strength (*EARN1*); the ratio of cash and balances due from depository institutions to total deposits reflects the degree of bank liquidity (*LQDT1*); lastly, sensitivity to market risk (*SENSRISK1*) is proxied by the change in the slope of the yield curve (given by the change in the quarterly difference between the 10-year U.S. T-bill rate and the 3-month U.S. T-bill rate) divided by total earning assets.

[5] The calculation of *MNGEXP1* is described in Appendix B.

Further, we adopt two metrics of bank complexity. We measure organisational complexity (*ORGCOMPL*) by the log of the product of the number of branches that each sample bank has and the number of U.S. states in which the bank has branches because banks which are more decentralised with a greater number of branches are characterised by more complex organisational structures (see Berger and Bouwman, 2013; Berger and Roman, 2015). Our second measure of complexity (*BUSINCOMPL*) captures the scope and diversity of bank business lines and relies upon the Bank for International Settlement methodology for the designation of globally systemically important banks that measures complexity using the notional value of Over-The-Counter (OTC) derivatives, the balance sheet presence of "Level 3" assets (i.e., assets for which prices cannot be inferred by either markets or models), and the size of the trading and available-for-sale securities (BCBS, 2014). We measure *BUSINCOMPL* as the sum of the notional amount of outstanding derivative contracts and the amount of credit exposure arising from recourse or other credit enhancements provided to the purchasers of the securitised loans, leases, and other assets divided by total assets (*BUSINCOMPL*).

3.5. Univariate analysis

We present and discuss the summary statistics on CAMELS components (*i.e.*, *CAP1*, *ASSETQLT1*, *MNGEXP1*, *EARN1*, *LQDT1*, and *SENSRISK1*), bank complexity (*ORGCOMPL*, *BUSINCOMPL*) as well as on bank size (*SIZE*) measured by the book value of total assets. We proceed to make pairwise comparisons between the three groups of banks prior to the onset of the crisis. Towards this, we conduct a univariate analysis of the mean differences of the aforementioned variables among the three bank groups based on average quarterly data over the pre-crisis period (2002q1-2007q3).

Table 1
Summary statistics and univariate analysis

	Non-distressed (obs=6,431)	Failed (obs=449)	TARP (obs=824)	Non-distressed vs Failed	Non-distressed vs TARP	Failed vs TARP
Variable	**Mean (Stdev)**	**Mean (Stdev)**	**Mean (Stdev)**	**Mean diff. (*t*-statistics)**	**Mean diff. (*t*-statistics)**	**Mean diff. (*t*-statistics)**
CAP1 (%)	12.82 (6.50)	10.20 (4.31)	9.23 (6.79)	2.62*** (5.32)	3.59*** (6.34)	0.97*** (7.40)
ASSETQLT1 (%)	0.51 (8.90)	1.38 (11.34)	1.92 (16.73)	-0.87*** (-4.98)	-1.41*** (-6.71)	-0.54*** (-7.30)
MNGEXP1	0.76 (2.38)	0.75 (1.75)	0.64 (9.41)	0.01 (1.34)	0.12** (1.99)	0.11** (1.86)
EARN1 (%)	0.84 (3.04)	0.29 (1.43)	0.14 (7.87)	0.55*** (3.48)	0.70*** (4.51)	0.15*** (3.65)
LQDT1 (%)	4.92 (4.52)	3.04 (5.64)	2.01 (10.43)	1.88*** (3.28)	2.91*** (3.87)	1.03*** (4.53)
SENSRISK1 (%)	10.68 (7.43)	10.63 (7.99)	17.18 (10.63)	0.05 (1.54)	-6.50** (-2.54)	-6.55** (-2.51)
SIZE (in $bn)	0.74 (241.84)	0.82 (368.89)	9.39 (529.23)	-0.08 (-1.43)	-8.65*** (-8.10)	-8.57*** (-10.53)
ORGCOMPL	1.25 (144.29)	1.42 (113.24)	1.64 (91.59)	-0.17 (-1.98)**	-0.39 (-3.61)***	-0.22 (-1.72)*
BUSINCOMPL (%)	17.98 (96.30)	18.92 (87.50)	35.80 (17.38)	-0.94 (-1.67)*	-17.82 (-6.42)***	-16.88 (-2.38)**

This table presents the summary statistics, reporting the means of the six components of CAMELS ratings, *i.e.*, capital strength (*CAP1*), asset quality (*ASSETQLT1*), quality of management (*MNGEXP1*), earnings strength (*EARN1*), degree of liquidity (*LQDT1*), sensitivity to market risk (*SENSRISK1*), as well as those of size (*SIZE*), organisational complexity (*ORGCOMP*), and business complexity (*BUSINCOMP*). The standard deviations of means for the aforementioned variables are reported in parentheses. The table also presents the results of a univariate analysis of the mean differences for the six CAMELS components, size, and complexity among the groups of non-distressed, failed, and TARP banks; the values of a *t*-test which captures the statistical differences in the means are reported in parentheses. All observations are on bank level and constitute average bank-quarter observations over the pre-crisis period (2002q1-2007q3). The description of variables and the relevant data sources are provided in Appendix A.
***, **, and * correspond to 1%, 5%, and 10% significance levels for a two-tailed test, respectively.

As reported in Table 1, the non-distressed banks were on average well-capitalised in the years preceding the crisis with a mean equity capital ratio (*CAP1*) of 12.82%. The mean value for the capital ratio of failed banks is equal to 10.20%, while that of TARP banks is 9.23%, showing that the latter group experienced a relatively lower capital cushion compared to their peers prior to the crisis. The mean differences are all statistically significant at the 1% level. Turning to examine the asset quality indicator (*ASSETQLT1*), figures reveal that the asset portfolio of non-distressed banks was the least risky compared to the relevant portfolios of the other two groups. In specific, the mean of *ASSETQLT1* was equal to 0.51% for non-distressed banks, 1.38% for failed banks, and 1.92% for TARP banks. Therefore, failed banks experienced a better asset quality if compared to that of TARP banks as they had 0.54% less non-performing loans compared to TARP institutions. The pairwise differences in means for *ASSETQLT1* are all significant at the 1% level. Moreover, non-distressed banks shared very similar managerial efficiency scores (*MNGEXP1*) with failed banks (0.76 and 0.75, respectively); the reported difference of 0.01 points is found not to be statistically significant. On the other hand, the management of TARP banks is found to be less efficient by 0.12 points and 0.11 points than that of non-distressed and failed banks respectively, where the mean differences are statistically significant at the 5% level. Focusing on *EARN1*, we observe that TARP banks were the least profitable institutions amongst the examined ones prior to the outbreak of the financial crisis: they earned 0.70% less than non-distressed banks and 0.15% less than failed banks. Both mean differences are significant at the 1%. Further, the profitability of failed banks was significantly lower by 0.55% if compared to that of non-distressed banks. In specific, failed banks earned 0.55% less than non-distressed banks. As regards the mean liquidity ratio (*LQDT1*), this was equal to 4.92% for non-distressed banks, 3.04% for failed banks, and 2.01% for TARP banks. That is, failed banks held fewer liquid assets than non-distressed banks, while TARP banks held the most illiquid portfolio of assets amongst its peers. The corresponding mean differences are all significant at the 1% level. To continue, non-distressed banks were, on average, almost equally sensitive to market risk with failed institutions the years prior to the crisis with an average *SENSRISK1* of 10.68% and 10.63%, respectively. The reported mean difference of 0.05% is not found to be statistically significant. On the other hand, the average sensitivity of TARP banks to market risk was equal to 17.18%, revealing that this group of banks was highly exposed to market-based activities. The mean differences (-6.50% and -6.55%) are significant at the 5%.

Importantly, non-distressed banks had almost the same average size (*SIZE*) with the failed institutions: \$0.74 billion and \$0.82 billion, respectively. In fact, the mean difference of \$0.08 is not statistically significant. As regards TARP banks, they had total assets of \$9.39 billion, being, on average, more than 11 times larger compared to the non-distressed or the failed banks. The relevant differences in means are found to be highly significant. It is no surprise that size was a crucial factor in providing a troubled institution with TARP money. We know that TARP -and, mostly, its first phase- specifically targeted only the largest banking firms to avoid any problems of continuity in their operations, or departures of their staff. This reflects the fact that it would have been very hard for the FDIC to arrange timely acquisitions of these large and complex banks (although it seems that they were able to do so for some very large banks during the crisis, so long as they were not very complex in their activities).[6]

As regards the organisational complexity (*ORGCOMPL*) of the three groups of banks, TARP banks are found to be the most complex ones, whereas non-distressed banks are the least complex institutions. Notably, the level of organisational complexity of failed banks, even though it is lower compared to that of TARP banks in numerical terms, it is not substantially different from a statistical viewpoint. This shows that the two groups of banks shared some similarities in their organisational structure in the years preceding the crisis. Turning to the business model complexity (*BUSINCOMPL*), failed banks are found to have been engaged in derivatives and securitisation activities to an almost equal degree with non-distressed banks. More concretely, the mean proportion of *BUSINCOMPL* is equal to 17.98% for non-distressed banks. This percentage is only

[6] We are thankful to Charles Calomiris for making this comment.

0.94% lower compared to that of failed banks (18.92%) and the reported difference is marginally statistically significant. On the other hand, the mean value of *BUSINCOMPL* for TARP institutions equals to 35.80%, which is 17.82% and 16.88% higher than the relevant means for the groups of non-distressed and failed banks, respectively. The reported mean differences are highly statistically significant.

In sum, we can postulate that the size and performance of TARP banks were significantly different from those of their peers during the pre-crisis period. TARP banks were much larger institutions, which experienced lower capital ratios, riskier portfolios of assets, weaker managerial efficiency, lower profitability, increased illiquid assets, and higher degree of sensitivity to market risk. In terms of complexity, TARP banks demonstrated more complex organisational and business structures compared to the non-distressed banks, but not so highly different structures from a statistical perspective compared to those of the failed banking firms.

Accordingly, TARP banks were less financially sound compared to the non-distressed banks in the years prior to the outburst of the crisis, but, most importantly, less sound than failed banks. Therefore, regulators appear to have put a less heavy weight on the overall performance and soundness of banks in their decision to save a distressed bank or to let it go down as a going concern entity. As regards the level of complexity, this seems to have played an important role in the decision of authorities to choose between a TARP-assisted bailout and a FDIC-backed failure, but it does not appear to be the key determinant for this decision. Rather, authorities have put a heavier weight on size in taking the relevant decision, as size is considered to be the most crucial determinant of systemic importance. Our findings are in line with Buiter (2009) who argues that "…the real issue is size; a complex but small business is no threat to systemic stability; neither is a highly international but small business; size is the core of the problem."

The evidence provided by the univariate analysis we conduct paves the way for our research to employ a multivariate technique which can endogenously define one or more size threshold levels and can help us to examine the following two questions which are viewed as being the two sides of the same coin: how small a bank should be so as to be classified as TSTS by regulators, and how big should be in order to be considered as TBTF.

4. The threshold regression model

Threshold models, whose origins can be traced in the threshold autoregressive model of Tong (1983), have become increasingly popular in econometric practice both in time series and cross-section as well as in panel data applications. Much of the relevance of threshold modelling in empirical research is explained by the preference policy makers have for threshold-related policies. To give some examples, Lensink and Hermes (2004) demonstrate how the entry of foreign banks in a domestic banking market depends on a threshold development level of the domestic economy. Fatum and Yamamoto (2014) conduct a threshold analysis on the impact of policy interventions of various intensities on the JPY/USD exchange rate over a twenty-year period and find that only interventions above some specific threshold are effective. In a very recent study, Hossfeld and MacDonald (2015) show that a currency can qualify as 'safe haven currency' on the basis of different country-specific threshold values for financial stress. A thorough review of papers which either contribute to the theory of threshold estimation and inference, or provide significant applications in economics and finance is provided by Hansen (2011).

Threshold regression models specify that individual observations can be divided into distinct regimes based on the value of some observed variable. In our research, we ask how small a bank should be to be classified as TSTS by regulators, and how big should be in order to be classified as TBTF. That is, bank size is our threshold variable. The panel threshold technique we employ in our empirical analysis relies on that of Hansen (1999), which is appropriate for panel data and considers multiple thresholds. This technique allows us to divide our sample of failed and TARP banks into zero, one, two or more different regimes based on the certain threshold values of size. We first consider the single threshold model and then extend our analysis to its multiple threshold counterpart.

Our single threshold structural model has the following form:

$$y_{it} = a_i + \beta'_1 x_{it} I(q_{it} \leq \gamma) + \beta'_2 x_{it} I(q_{it} > \gamma) + \delta' w_{it} + \varepsilon_{it} \tag{1}$$

An alternative way of writing Eq. (1) is:

$$y_{it} = \begin{cases} a_i + \beta'_1 x_{it} + \delta' w_{it} + \varepsilon_{it}, & q_{it} \leq \gamma \\ a_i + \beta'_2 x_{it} + \delta' w_{it} + \varepsilon_{it}, & q_{it} > \gamma \end{cases} \tag{2}$$

In Eqs. (1) and (2), it holds that: *i*=1, 2,…, *N* individuals and *t*=1, 2,…, *T* time periods; $y_{it}$is a scalar; the regressor $x_{it}$is a *k*-dimensional vector; $w_{it}$is a *m*-dimensional vector; the threshold variable q_{it} is a scalar; γ stands for the threshold; a_i is the individual fixed-effects; ε_{it} is the unobserved error term with mean zero and finite variance σ^2; finally, in Eq. (1), $I(\cdot)$ is an indicator function that takes the value of 1 or 0 depending on whether q_{it} falls short of or exceeds γ.

In Eqs. (1) and (2), the observations are divided into two distinct regimes depending on whether the value of the threshold variable, q_{it}, is smaller or larger than the threshold γ. The two regimes are characterised by different regression slopes, $\beta_1$and β_2. To identify $\beta_1$and β_2, it is required that the elements of x_{it} are not time invariant. In a similar vein, q_{it} and w_{it} are also assumed not to be time invariant. If $q_{it}$is below or above a certain value of γ, then x_{it} has a different impact on the dependent variable of the model, y_{it}, with $\beta_1 \neq \beta_2$.

If we set $x_{it}(\gamma) = \binom{x_{it}I(q_{it} \leq \gamma)}{x_{it}I(q_{it} > \gamma)}$ and $\beta = (\beta_1', \beta_2')'$, Eq. (1) can be rewritten as follows:

$$y_{it} = a_i + \beta' x_{it}(\gamma) + \varepsilon_{it} \tag{3}$$

After eliminating the individual bank fixed effects a_i by removing the individual-specific means, the slope coefficient β can be estimated for any given γ by Ordinary Least Squares (OLS):

$$\hat{\beta}(\gamma) = (X^*(\gamma)' X^*(\gamma))^{-1} X^*(\gamma)' Y^* \tag{4}$$

where $X^*(\gamma)$ and Y^*denote the data stacked over all individual banks. The vector of regression residuals is given by:

$$\hat{\varepsilon}^*(\gamma) = Y^* - X^*(\gamma)' \hat{\beta}(\gamma) \tag{5}$$

Hence, the sum of squared errors can be written as follows:

$$S_1(\gamma) = \hat{\varepsilon}^*(\gamma)' \hat{\varepsilon}^*(\gamma) = Y^{*\prime}(I - X^*(\gamma)'(X^*(\gamma)'^{X^*}(\gamma))^{-1} X^*(\gamma)')Y^* \tag{6}$$

The estimation of γ by least squares can be achieved by the minimisation of the concentrated sum of squared errors:

$$\hat{\gamma} = argmin_{\gamma} S_1(\gamma) \quad (7)$$

Once we obtain $\widehat{\gamma}$, we can estimate the slope coefficient by $\hat{\beta} = \hat{\beta}(\gamma)$. The residual vector is $\hat{\varepsilon}^* = \hat{\varepsilon}^*(\gamma)$ and the residual variance is defined as follows:

$$\hat{\sigma}^2 = \frac{1}{N(T-1)} \hat{\varepsilon}^{*\prime} \hat{\varepsilon}^* = \frac{1}{N(T-1)} S_1(\hat{\gamma}) \quad (8)$$

To determine whether the threshold effect is statistically significant, we test the hypothesis of no threshold effect H_0: $\beta_1 = \beta_2$. Under the null hypothesis, the model is:

$$\mathrm{y}_{it} = a_i + \beta'_1 x_{it} + \varepsilon_{it} \quad (9)$$

Based on the fixed effects transformation, Eq. (9) can be written as:

$$y_{it}^* = \beta'_1 x_{it}^* + \varepsilon_{it}^* \quad (10)$$

The OLS estimator of $\beta_1$is $\tilde{\beta}_1$, the residuals are$\tilde{\varepsilon}_{it}^*$, and the sum of squared errors is $S_0 = \tilde{\varepsilon}_{it}^{*\prime} \tilde{\varepsilon}_{it}^*$. Then, the likelihood ratio test of H_0 is based on:

$$F_1 = \frac{S_0 - S_1(\hat{\gamma})}{\hat{\sigma}^2} \quad (11)$$

Hansen (1996) suggests a bootstrap methodology to simulate the asymptotic distribution of the likelihood ratio test. He shows how to attain the first-order asymptotic distribution, so *p*-values constructed from the bootstrap are asymptotically valid. For the bootstrap procedure, the regressor x_{it} and the threshold variable q_{it} are given, in that their values are fixed in repeated bootstrap samples. We obtain a sample of size *nT* with replacement from the empirical distribution and create a bootstrap sample under the null of no threshold. This bootstrap sample is used to estimate Eq. (1) under H_0 and H_1 and to calculate the bootstrap value of the likelihood ratio statistic F_1. This procedure is frequently repeated and the bootstrap estimate of the asymptotic *p*-value for F_1under H_0 is the percentage of draws for which the simulated likelihood ratio statistic exceeds the actual statistic. If the *p*-value is smaller than the desired critical value, then H_0 is rejected implying that a threshold exists.

In case of a threshold effect, $(\beta_1 \neq \beta_2)$, the estimate $\widehat{\gamma}$ is consistent for the true value of γ, say γ_0. Since the asymptotic distribution of the threshold estimate $\widehat{\gamma}$ is highly non-standard, Hansen (2000) uses the likelihood ratio statistic for tests on γ to form the relevant confidence intervals for γ. The null hypothesis is $H_0: \gamma = \gamma_0$ and the likelihood ratio statistic is:

$$LR_1(\gamma) = \frac{S_1(\gamma) - S_1(\hat{\gamma})}{\hat{\sigma}^2} \tag{12}$$

The null is rejected for large values of $LR_1(\gamma_0)$. In specific, the test rejects $H_0: \gamma = \gamma_0$ at the asymptotic level α if $LR_1(\gamma_0)$>0. The asymptotic (1 - α) confidence interval for γ is the set of values of γ with $LR_1(\gamma) \leq c(\alpha)$.

In several applications of the Hansen (1999) threshold technique, more than one thresholds are identified. For instance, Wang and Huang (2009) estimate the cost efficiency scores of the Taiwanese commercial banking sector and find that the technologies that the banks employ in their production function can be splitted into three different regimes based on a double threshold they specify. Similarly, Ben Cheikh and Louhichi (2016) examine the role of inflation in explaining the degree of exchange-rate pass through into import prices and the related monetary policy issues based on 63 economies and provide evidence of two inflation threshold levels. In what follows, we extend the single threshold model (Eq. 1) to its double threshold counterpart in order to test the existence of a TSTS and a TBTF double size threshold. Our double threshold model takes the following form:

$$\mathrm{y}_{it} = a_i + \beta'_1 x_{it} I(q_{it} \leq \gamma_1) + \beta'_2 x_{it} I(\gamma_1 < q_{it} \leq \gamma_2) + \beta'_3 x_{it} I(\gamma_2 < q_{it}) + \delta' w_{it} + \varepsilon_{it} \quad (13)$$

The two thresholds, $\gamma_1$and γ_2, are ordered so that $\gamma_1 < \gamma_2$. Eq. (13) can be estimated by OLS, since it is linear in slopes $(\beta_1, \beta_2, \beta_3)$ for given (γ_1, γ_2).The sum of squared errors $S(\gamma_1, \gamma_2)$ can be calculated based on Eq. (6) in the single threshold model and the joint least squares estimates of (γ_1, γ_2) are by definition the values that jointly minimise $S(\gamma_1, \gamma_2)$. More concretely, the following sequential estimation procedure is proposed by Hansen (1999).

Let $S_1(\gamma)$be the single threshold sum of squared errors as defined in Eq. (6) and let $\hat{\gamma}_1$ be the threshold estimate that minimises $S_1(\gamma)$. Fixing the first-stage estimate $\hat{\gamma}_1$, the criterion for the second stage is:

$$S_2^r(\gamma_2) = \begin{cases} S(\hat{\gamma}_1, \gamma_2) & if\ \hat{\gamma}_1 < \gamma_2 \\ S(\gamma_2, \hat{\gamma}_1) & if\ \gamma_2 < \hat{\gamma}_1 \end{cases} \quad (14)$$

Hence, the second-stage threshold estimate is:

$$\hat{\gamma}_2^r = argmin_{\gamma_2} S_2^r(\gamma_2) \quad (15)$$

If we hold the second-stage estimate $\hat{\gamma}_2^r$ fixed, we obtain the following criterion:

$$S_1^r(\gamma_1) = \begin{cases} S(\gamma_1, \hat{\gamma}_2^r) & if\ \gamma_1 < \hat{\gamma}_2^r \\ S(\hat{\gamma}_2^r, \hat{\gamma}_1) & if\ \hat{\gamma}_2^r < \hat{\gamma}_1 \end{cases} \qquad (16)$$

and the relevant estimate is:

$$\hat{\gamma}_1^r = argmin_{\gamma_1} S_1^r(\gamma_1) \qquad (17)$$

We can now turn to determine the number of thresholds in Eq. (13): there will be either no thresholds, one threshold, or two thresholds. The same process applies if more than two thresholds are to be determined. Like we did in the single threshold case, we resort to F_1as given by Eq. (11) to test the null of no threshold. If the null hypothesis is rejected, we need an additional test to distinguish between one or two thresholds. The minimised sum of squared errors form the second stage threshold estimate is $S_2^r(\hat{\gamma}_2^r)$ with variance estimate:

$$\hat{\sigma}^2 = \frac{S_2^r(\hat{\gamma}_2^r)}{N(T-1)} \qquad (18)$$

Therefore, the likelihood ratio statistic for a test of one versus two thresholds is:

$$F_2 = \frac{S_1(\hat{\gamma}_1) - S_2^r(\hat{\gamma}_2^r)}{\hat{\sigma}^2} \qquad (19)$$

A bootstrap procedure is followed that produces the relevant sample. From the bootstrap sample, F_2 is calculated, and this procedure is repeated multiple times to calculate the bootstrap *p*-value. The hypothesis of one threshold is rejected in favour of two thresholds if $F_2$is large.

We finally turn to construct the confidence intervals for the two threshold parameters (γ_1, γ_2). Let:

$$LR_2^r(\gamma) = \frac{S_2^r(\gamma) - S_2^r(\hat{\gamma}_2^r)}{\hat{\sigma}^2} \tag{20}$$

and

$$LR_1^r(\gamma) = \frac{S_1^r(\gamma) - S_1^r(\hat{\gamma}_1^r)}{\hat{\sigma}^2} \tag{21}$$

where $S_2^r(\gamma)$and $S_1^r(\gamma)$are defined in Eq. (14) and Eq. (16), respectively. The asymptotic (1 - α) confidence intervals for the threshold estimates are the set of values of γ with $LR_2^r(\gamma) \leq c(\alpha)$ and $LR_1^r(\gamma) \leq c(\alpha)$.

5. Estimation and results

Our estimates are produced using the sample of distressed banks (*i.e.*, failed and TARP banks). To avoid violating the exogeneity assumption of the regressors in our estimation as imposed by Hansen (1999), our threshold variable as well as the rest of our regressors are introduced in Eq. (13) with their past realisations under the thought that the latter are given before the current values are realised. In fact, Hansen (1999) also uses lagged regressors to ensure that the exogeneity condition is met in his empirical application.[7] The lag structure in our analysis is determined by

[7] A recent application of Hansen's threshold model technique that also resorts to past realisations of regressors is that of Hossfeld and MacDonald (2015).

two of the most popular selection criteria, namely the Akaike Information Criterion and the Schwarz-Bayesian Information Criterion. Both criteria specify a 4-quarter lag (*i.e.*, $t-4$) structure to be followed. Hence, Eq. (13) is written as follows:

$$y_{it} = a_i + \beta'_1 x_{it-4} I(q_{it-4} \leq \gamma_1) + \beta'_2 x_{it-4} I(\gamma_1 < q_{it-4} \leq \gamma_2) + \\ + \beta'_3 x_{it-4} I(\gamma_2 < q_{it-4}) + \delta' w_{it-4} + \varepsilon_{it} \tag{22}$$

where i=1, 2,…, N (N=1273) distressed banks, *i.e.*, 449 failed and 824 TARP banks, and t=1, 2,…, T (T=44) quarters; y_{it} is a binary scalar, which is equal to 1 if the sample bank i failed as a going concern entity at t and 0 if it received TARP money at t and survived the crisis; the vector x_{it} contains the six components of CAMELS ratings (*CAP1*, *ASSETQLT1*, *MNGEXP1*, *EARN1*, *LQDT1*, and *SENSRISK1*) as well as the two complexity measures (*ORGCOMPL* and *BUSINCOMPL*); w_{it} contains a set of bank-specific control variables that we present below; the size (*SIZE*) threshold variable is given by q_{it}; γ_1 and γ_2 stand for the two thresholds of *SIZE*; a_i is the individual bank fixed-effects; ε_{it} is the unobserved error term with mean zero and finite variance σ^2; and, $I(\cdot)$ is an indicator function that takes either the value of 1 or 0 depending on whether the size threshold variable q_{it} is higher or lower than γ_1, γ_2.[8]

In what follows, we present the bank-specific control variables which are included in w_{it}. Appendix A provides a description of these variables and the relevant data sources. To begin with, the relevant literature has demonstrated that connections with regulators and policy-makers have a considerable impact on the decision of authorities to save a bank through the extension of a TARP facility. We use a set of variables to capture these connections. First, we follow Blau et al.

[8] The notation followed herein fully complies with those in Eqs. (1) and (2).

(2013) and resort to the Center for Responsive Politics (CRP)'s Revolving Door database to construct an indicator variable (*POLCON*) to proxy the connections that our sample banks may have with policy-makers. *POLCON* is equal to unity if a sample bank has employed, or is currently employing an individual who is also employed or has been employed in the federal government or appointed to a government advisory board, a congressional or presidential cabinet entity, or an independent commission. Second, we identify any connections that banks may have with regulatory and supervisory authorities. We follow Bayazitova and Shivdasani (2012), Duchin and Sosyura (2012), Li (2013), and Berger and Roman (2015) to construct an indicator variable (*FEDCON*) that is equal to unity if an executive at a sample bank was on the board of directors of one of the 12 Federal Reserve Banks or one of their branches either in 2008 or 2009. We first obtain the relevant data on the top executives of our sample BHCs from BoardEx and then match them to the list of directors from the Fed's website. Third, we use House of Representatives committee data and follow Berger and Roman (2015) and Duchin and Sosyura (2014) to construct a dummy variable (*COMMIT*) that equals one if a sample bank is headquartered in a district of a House member who served on the key finance committees involved in drafting and amending TARP, i.e. the Subcommittee on Financial Institutions, or the Subcommittee on Capital Markets of the House Financial Services Committee, either in 2008 or 2009. We resort to data from the U.S. Census Bureau and the U.S. Library of Congress to match the sample banks with the relevant congressional districts using the zip codes of their headquarters. And, forth, as an additional measure of the ties that may exist between the financial services industry and politicians, we refer to the contributions of banks to federal political campaigns (*CAMP*). We collect data from the Federal Election Commission that cover contributions from Political Action Committees (PACs) to candidates' election campaigns. Following Duchin and Sosyura (2012) and Bayazitova and

Shivdasani (2012), *CAMP* is an indicator that takes the value of one if a sample bank has made PAC contributions in the election cycle for the 2008 congressional election to the members of the Subcommittee on Financial Institutions and the Subcommittee on Capital Markets.

A number of TARP banks played the role of acquirers in the M&A deals that took place during the examined period but, mainly, after the outbreak of the crisis. We, therefore, resort to the relevant files of the Federal Reserve Bank of Chicago to investigate whether a bank has been involved in a M&A transaction as an acquirer to control for the effect on our dependent variable.[9] Towards this, we introduce a dummy variable in our model (*MA*), which is equal to unity when the acquirer bank *i* is involved in a M&A transaction and remains equal to one until the end of our data period. For example, if an acquisition occurred on April 15 2008 then this transaction is recorded in the second quarter of 2008, meaning that the binary variable *MA* takes the value of one in 2008q2 and remains as such for all the subsequent quarters.

Based on the geographical characteristics of our sample banks, we note that TARP banks are headquartered and located in terms of branching activity near salt water, that is, near the West and, mostly, the East U.S. Coasts. As regards the distribution of failures, the states of Arizona, California, Georgia, and Nevada are amongst those with the highest number of FDIC-supported bankruptcies. Most of the Northeastern and Southeastern states (excluding California) had either no or a few bank failures, whereas the Western U.S. states, which experienced a relatively larger decline in economic performance as measured by the GDP growth rate and the unemployment rate, had the highest failure rates. Further, a number of failed banks are located in rather distant, sparsely populated geographical districts. We, therefore, follow Jordan et al. (2011) and Berger and Roman (2015) and introduce a dummy indicator (*MSA*) which is equal to one if a bank is

[9] The relevant data can be found in the following web page: https://www.chicagofed.org/banking/financial-institution-reports/merger-data

located in a Metropolitan Statistical Area -an integrated economic and social unit with a recognised large population nucleus- and zero otherwise. The geographical location of each sample bank is identified through Call Reports; detailed data for the Metropolitan Statistical Areas are taken from the U.S. Office of Management and Budget.

It is well-documented in the banking literature that the behaviour and performance of the newly chartered banks substantially differ from those of banks in operation over a relatively long period of time. More specifically, once a bank first enters the market, its financial performance tends to lag by a considerable margin compared to that of the existing banking firms.[10] That said, we account for the so-called *de novo* banks, defined as banks less than five years old by including a dummy (*DENOVO*) in our model.

We follow Berger and Roman (2015) and construct an indicator variable (*PUBLIC*) that captures if a bank is listed on the stock exchange. Since the decision-making units we examine are not holding companies, the subsidiaries of publicly traded BHCs are considered to be public. Banks with private placements of shares with a Committee on Uniform Securities Identification Procedures (CUSIP) number, banks without a stock exchange listing, and banks whose bank holding company is not listed at the stock exchange are treated as non-public. The data on trading and listing are derived from the Center for Research in Security Prices (CRSP) database. Lastly, a dummy variable (*BHC*) showing whether a sample bank is a subsidiary of a BHC is also considered in our empirical analysis as in Jordan et al. (2011) and Berger and Roman (2015).

The summary statistics which are shown in Table 2 reveal substantial and statistically significant differences between TARP and failed banks. We find that *POLCON* is significantly larger at the 1% level for banks that received TARP money than those that were backed by the

[10] See, e.g., DeYoung and Hasan (1998), and DeYoung (2003) for a thorough analysis on the operational behaviour of *de novo* banks.

FDIC. Specifically, 7.38% of TARP institutions have employed, or are currently employing at least one individual, who is affiliated or has been affiliated with the federal government or some other cabinet entity, whereas the relevant percentage for the FDIC-backed banks is only 1.84%. Similarly, if we turn to examine *FEDCON*, we observe that TARP banks are more closely linked to Fed regulators and supervisors compared to their failed peers (6.31% and 1.93%, respectively). The difference in the means is found to be statistically significant at the 1% level. Further, 9.36% of the TARP banks and 3.12% of the failed banks are headquartered in a district of a House member who served on the key finance committees (*COMMIT*); the reported difference is significant at the 5% level. Regarding the contributions of the two groups of banks to federal political campaigns (*CAMP*), 5.42% of the TARP banks and 1.13% of the failed banks made such contributions and the relevant difference is highly significant.

An average of 27.24% of TARP banks has been involved in at least one M&A transaction as acquirer during the sample period, whereas the relevant percentage of failed banks is only 3.20%. The difference in the means of *MA* for the two groups of banks is significant at the 5% level. To continue, roughly half of the failed banks (52.27%) are located in a Metropolitan Statistical Area (*MSA*). The relevant percentage for banks that received TARP money is significantly higher at the 5% level and is equal to 71.41%. An additional considerable difference of failed compared to TARP banks is that more than twice of the former group of banks are newly-chartered banks (*DENOVO*) compared to the latter group (7.61% vs. 3.20%, respectively), and that the reported difference in means is significant at the 1% level. Moreover, the summary statistics for *PUBLIC* show that the percentage of listed failed banks is equal to 3.29%, whereas that of listed TARP banks is equal to 7.56%; the mean difference is found to be statistically significant at the 5% level. Lastly, 12.80% of the failed banks are, on average, affiliated with a Bank Holding Company

(*BHC*). The corresponding percentage for the assisted institutions is much higher and equals to 62.68%; the reported difference in the relevant means is significant at the 5% level.

Table 2
Summary statistics for the control variables.

	Failed banks (obs=449)			TARP banks (obs=824)		
Variable	**Mean**	**Median**	**Stdev**	**Mean**	**Median**	**Stdev**
POLCON	0.0184***	0.0000	6.73	0.0738	0.0000	12.54
FEDCON	0.0193***	0.0000	14.60	0.0631	0.0000	10.95
COMMIT	0.0312**	0.0000	19.02	0.0936	0.0000	7.68
CAMP	0.0113***	0.0089	4.20	0.0542	0.0493	3.15
MA	0.0032**	0.0000	2.29	0.2724	1.0000	3.02
MSA	0.5227**	0.0000	7.45	0.7141	1.0000	10.69
DE NOVO	0.0761***	0.0000	27.90	0.0320	0.0000	31.84
PUBLIC	0.0329**	0.0000	9.43	0.0756	0.0000	12.70
BHC	0.1280**	0.0000	7.85	0.6268	1.0000	5.62

This table presents the summary statistics, reporting the means, medians, and standard deviations for the control variables, which are contained in the vector w_{it}: a dummy capturing the political connections of banks (*POLCON*); a dummy for the connections of banks with the regulatory and supervisory authorities (*FEDCON*); a dummy that shows if a sample bank is headquartered in a district of a House member who served on the key finance committees (*COMMIT*); a dummy for banks which made PAC contributions in the 2008 elections (*CAMP*); a dummy for the acquirer banks in M&A transactions (*MA*); a dummy showing whether a bank is located in a MSA or in a rural county (*MSA*); a dummy for newly-chartered banks (*DENOVO*); a dummy for banks which are listed on the stock exchange (*PUBLIC*); and a dummy indicating whether a bank is a subsidiary of a BHC (*BHC*). The description of the control variables and the relevant data sources are provided in Appendix A.
** The mean of failed banks is significantly different from that of TARP banks at the 5% level.
*** The mean of failed banks is significantly different from that of TARP banks at the 1% level.

We estimate Eq. (22) by linear probability OLS regression, which is robust to model misspecifications and, importantly, allows us to accommodate the fixed effects nature of Hansen (1999)'s panel threshold model. A limitation is that the error term of the linear probability model

ε_{it} is heteroskedastic. We deal with this problem by obtaining estimates of the standard errors that are robust to heteroskedasticity by employing the White (1980)'s heteroskedasticity-consistent covariance matrix.

The number of thresholds as well as the levels of our threshold variable (*SIZE*) are endogenously determined. For the number of thresholds to be determined, our estimation allows for (sequentially) zero, one, two, and three thresholds. The test statistics F_1, F_2, and F_3 for testing zero against one, one against two, and two against three thresholds along with their bootstrap *p*-values, are reported in Table 3 below. We follow Hansen (1999) and apply 300 bootstrap replications for each of the three bootstrap tests.

Table 3
Tests for determining the number of thresholds.

H_0: no threshold vs one threshold	
F_1	29.830
p-value	0.001
(10%, 5%, 1% critical values)	(11.97, 15.18, 30.17)
H_0: one threshold vs two thresholds	
F_2	25.620
p-value	0.008
(10%, 5%, 1% critical values)	(12.81, 16.04, 31.26)
H_0: two vs three thresholds	
F_3	9.005
p-value	0.461
(10%, 5%, 1% critical values)	(10.05, 11.70, 20.18)

The test statistics F_1, F_2, F_3, their asymptotic bootstrap *p*-values, and the relevant critical values at 10%, 5%, 1% levels are presented in this Table. 300 bootstrap replications are applied for each of the three bootstrap tests.

The test for a single threshold F_1 is strongly statistically significant with a bootstrap *p*-value of 0.001. The test for a double threshold F_2 is also highly significant with a bootstrap *p*-value of 0.008. However, the test for a third threshold F_3 is far from being significant since the relevant bootstrap *p*-value is equal to 0.461. In sum, the sequential test procedure provides robust evidence for two thresholds for bank size in our model. We will, therefore, work with this double threshold model in the remainder of our analysis.

We now turn to obtain the estimates of the two threshold levels for our threshold variable. These estimates are obtained by searching through values of γ that equal the distinct values of *SIZE* in our sample. Following Hansen (1999, 2000), we ensure that a minimum number of observations fall into one or the other regime. We thus restrict the search to values of *SIZE* such that not less than 5% of the observations lie in each regime. The $\widehat{\gamma}$ which minimises the sum of squared residuals is selected.

As presented in Table 4 that follows, the point estimates of the two thresholds for *SIZE* are \$0.402bn and \$2.850bn, respectively. Hence, two thresholds are endogenously specified: one for the TSTS banks which is equal to \$0.402bn and a second one for the TBTF banks that equals to \$2.850bn. Accordingly, our sample banks are allocated to the following three size regimes: a TSTS regime that contains all banks with total assets up to \$0.402bn; an intermediate regime that consists of all banks with total assets from \$0.402bn to \$2.850bn; and a TBTF regime, which includes all institutions with more than \$2.850bn. In the TSTS regime, it is only the insured depositors and, in some cases, a part of uninsured depositors, debtholders, and other stakeholders who are fully bailed out. In the TBTF regime, on the other hand, all the stakeholders together with the shareholders are fully bailed out. The asymptotic 95% confidence intervals for each threshold shown in Table 4 are tight, reflecting little uncertainty about the nature of this clustering.

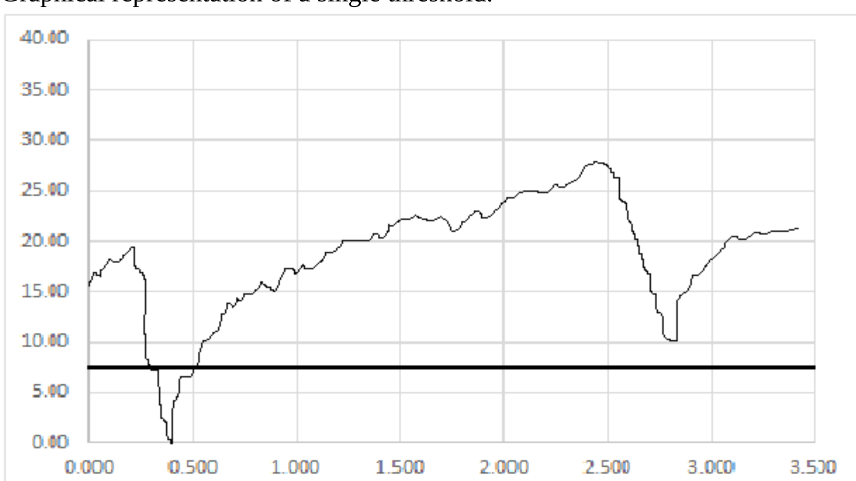
Table 4
Threshold estimates.

	Estimate	**95% confidence interval**
$\hat{\gamma}_1$	\$0.402bn	[0.329, 0.456]
$\hat{\gamma}_2$	\$2.850bn	[2.644, 3.039]

This Table reports the point estimates of the two size thresholds and their asymptotic 95% confidence interval. Estimates are expressed in US\$ bn.

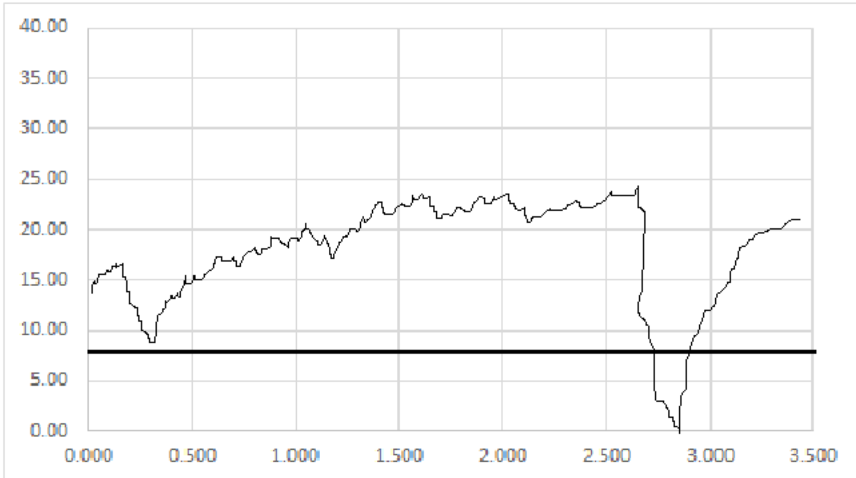
To obtain a graphical representation of the threshold estimates, we can draw the plots of the concentrated likelihood ratio functions $LR_1(\gamma)$, $LR_2^r(\gamma)$, and $LR_1^r(\gamma)$ against $\widehat{\gamma}_1, \widehat{\gamma}_2^r$, and $\hat{\gamma}_1^r$, respectively. The point estimates are the values of γ as shown in the horizontal axis in Figures 1-3; the vertical axis displays the values of the likelihood ratio. The 95% confidence intervals for γ_2 and γ_1 can be obtained from $LR_2^r(\gamma)$ and $LR_1^r(\gamma)$ by the values of γ for which the likelihood ratio lies beneath the bold line. Figure 1 presents the first-step likelihood ratio function $LR_1(\gamma)$ that is computed when we test for a single threshold. As earlier mentioned, the first-step threshold estimate $\widehat{\gamma}_1$ is equal to \$0.402bn. At this particular size level $LR_1(\gamma)$ is zero, which confirms the existence of a single threshold at this point. A second major dip in the likelihood ratio occurs around the second-step threshold estimate which equals to \$2.850bn. Hence, the single threshold likelihood estimation, as reflected in Figure 1, suggests that there is a second threshold size in the regression. The existence of the second threshold is confirmed in Figure 2: the relevant likelihood ratio is equal to zero for $\widehat{\gamma}_2$= \$2.850bn. Figure 3, in turn, suggests that a third threshold is not likely to exist.

Figure 1
Graphical representation of a single threshold.



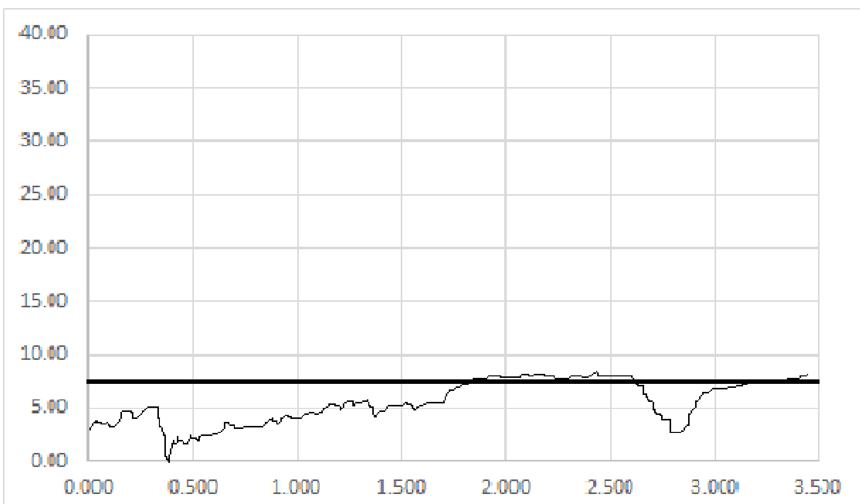
In this Figure, the concentrated likelihood ratio function $LR_1(\gamma)$ is plotted against $\widehat{\gamma}_1$. The point estimates are the values of γ as shown in the horizontal axis; the vertical axis displays the values of the likelihood ratio. The 95% confidence intervals for γ_2 and γ_1 can be found from $LR_2^r(\gamma)$ and $LR_1^r(\gamma)$ by the values of γ for which the likelihood ratio lies beneath the bold line.

Figure 2
Graphical representation of a double threshold.



In this Figure, the concentrated likelihood ratio function LR_2^r is plotted against $\hat{\gamma}_2^r$. The point estimates are the values of γ as shown in the horizontal axis; the vertical axis displays the values of the likelihood ratio. The 95% confidence intervals for γ_2 and γ_1 can be found from $LR_2^r(\gamma)$ and $LR_1^r(\gamma)$ by the values of γ for which the likelihood ratio lies beneath the bold line.

Figure 3
Graphical representation of a triple threshold.



In this Figure, the concentrated likelihood ratio function $LR_1^r(\gamma)$ is plotted against $\hat{\gamma}_1^r$. The point estimates are the values of γ as shown in the horizontal axis; the vertical axis displays the values of the likelihood ratio. The 95% confidence intervals for γ_2 and γ_1 can be found from $LR_2^r(\gamma)$ and $LR_1^r(\gamma)$ by the values of γ for which the likelihood ratio lies beneath the bold line.

As shown in Table 5, a total of 302 failed banks (67.26%) and a total of 98 TARP banks (11.90%) are classified in the TSTS regime; 122 failed banks (27.17%) and 231 TARP banks

(5.57%) fall into the intermediate regime, whereas 25 failed banks (28.03%) and 495 TARP banks (60.07%) are classified in the TBTF regime.

Table 5
Threshold size clustering of failed and TARP banks.

Size regime	Failed banks	TARP banks
TSTS regime: *SIZE* ≤ \$0.402bn	302 (67.26%)	98 (11.90%)
Intermediate regime: \$0.402bn < *SIZE* ≤ \$2.850bn	122 (27.17%)	231 (28.03%)
TBTF regime: \$2.850bn < *SIZE*	25 (5.57%)	495 (60.07%)

This Table reports the number of failed and TARP banks which are classified into each of the three size regimes. The relevant percentages are reported in parentheses.

In Table 6, we present the OLS estimation of our linear probability regression model (Eq. 22). The most commonly recognised flaw in the linear probability model (LPM) is that the fitted probabilities may not be bounded on the unit interval in that they can take values below zero or above unity. Wooldridge (2002) suggests that a straightforward check on the LPM is to test how many of the fitted values do not lie between zero and one. In our model, there are not more than 1.4% of the total fitted values that lie outside the unit interval.

Table 6
Threshold regression results.

	TSTS regime	**Intermediate regime**	**TBTF regime**
	SIZE ≤ \$0.402bn	\$0.402bn < *SIZE* ≤ \$2.850bn	\$2.850bn < *SIZE*
Variable			
CAP1	-0.049** (0.024)	-0.050*** (0.010)	-0.071** (0.032)
ASSETQLT1	0.133** (0.062)	0.138*** (0.037)	0.086** (0.040)
MNGEXP1	-0.041** (0.020)	-0.038** (0.017)	-0.059** (0.029)
EARN1	-0.103** (0.047)	-0.095*** (0.024)	-0.147** (0.073)
LQDT1	-0.033* (0.018)	-0.035*** (0.010)	-0.069** (0.033)

SENSRISK1	0.102** (0.047)	0.098*** (0.022)	0.052* (0.028)	
ORGCOMPL	-0.054* (0.029)	-0.032 (0.038)	-0.115*** (0.017)	
BUSINCOMPL	-0.071 (0.052)	-0.040 (0.037)	-0.144*** (0.025)	
POLCON				-0.108*** (0.025)
FEDCON				-0.129*** (0.037)
COMMIT				-0.052** (0.023)
CAMP				-0.042** (0.017)
MA				-0.026** (0.010)
MSA				-0.074*** (0.020)
DENOVO				0.040** (0.017)
PUBLIC				-0.054*** (0.008)
BHC				-0.010 (0.013)
R^2		0.19		

This table presents the estimation results of the multiple threshold regression model (Eq. 22) for the three bank size (*SIZE*) regimes which are endogenously determined by our model: *SIZE* ≤ \$0.402bn, \$0.402bn < *SIZE* ≤ \$2.850bn, and \$2.850bn < *SIZE*. The dependent variable is equal to 1 if a sample bank failed as a going concern entity and 0 if it received TARP money and survived the crisis. The main explanatory variables are: capital strength (*CAP1*), asset quality (*ASSETQLT1*), quality of management (*MNGEXP1*), earnings strength (*EARN1*), degree of liquidity (*LQDT1*), and sensitivity to market risk (*SENSRISK1*). The set of control variables includes: organisational complexity of banks (*ORGCOMPL*); bank business model complexity (*BUSINCOMPL*); a dummy capturing the political connections of banks (*POLCON*); a dummy for the connections of banks with the federal regulatory and supervisory authorities (*FEDCON*); a dummy that shows if a sample bank is headquartered in a district of a House member who served on the key finance committees (*COMMIT*); a dummy for the banks which made PAC contributions in the 2008 election cycle (*CAMP*); a dummy for acquirer banks in M&A transactions (*MA*); a dummy showing whether a bank is located in a MSA or in a rural county (*MSA*); a dummy for newly-chartered banks (*DENOVO*); a dummy variable for banks which are listed on the stock exchange (*PUBLIC*); and a dummy indicating whether a bank is a subsidiary of a BHC (*BHC*). All observations are on bank level, constitute bank-quarter observations, and cover the entire data period, which extends from 2002q1 to 2012q4. All the explanatory variables are lagged by four quarters to address possible endogeneity concerns. The description of each variable and the relevant data sources are included in Appendix A. White heteroskedasticity-robust standard errors are reported in parentheses.
***, **, * correspond to 1%, 5%, and 10% level of significance respectively for a two-tailed distribution

As Table 6 displays, the estimated coefficients on CAMELS components are all statistically significant and their signs remain unchanged across the three size regimes. Considering that a positive (negative) sign indicates an increase (decrease) in failure probability, the regression results demonstrate that more capitalised banks are less likely to fail, or, to put it differently, banks which are highly levered are more likely to go bankrupt. Further, low asset quality is found to increase the failure probability, while managerial efficiency has a negative impact on the examined probability. The latter relationship is also true for banks which are more profitable as well as for those that hold a larger proportion of liquid assets in their portfolios. Lastly, increased sensitivity to market risk is found to be positive linked with the likelihood of failure.

The linear probability model has the advantage of allowing a straightforward interpretation of the regression coefficients. The estimated parameters measure the percentage change in the probability of failure resulting from a unit change in the variable under scrutiny, holding all other factors fixed. This can be interpreted as the partial effect on the failure probability. If we focus our discussion of the regression results on the two size regimes of primary interest, that is, the TSTS and the TBTF regimes, we note that, although the estimates of the CAMELS components share the same signs in both regimes, the magnitude of the estimates of each of the six components substantially differs between the two regimes. In specific, TSTS banks experience higher estimated coefficients on the CAMELS components which are positively related with the failure probability (*i.e.*, *ASSETQLT1* and *SENSRISK1*), and lower coefficients on the components which are negatively related with the relevant probability (*i.e.*, *CAP1*, *MNGEXP1*, *EARN1*, and *LQDT1*). For instance, if *ASSETQLT1* increases by 1 unit, the failure probability is enhanced by 13.3% for TSTS banks but only by 8.6% for TBTF banks. On the other hand, if *CAP1* increases by 1 unit, the probability of failure decreases by 4.9% for TSTS banks and by 7.1% for TBTF banks. This is to say, a TSTS bank which has exactly the same overall performance with a TBTF bank based on the examined regulatory ratings system is more likely to fail due to the different weights put on each CAMELS component and which are in favour of TBTF banks. Our threshold variable, bank size (*SIZE*), which specifies the three size regimes is the one that provides us with the different weights that regulators implicitly assign to CAMELS components for each regime. Hence, regulators are viewed as using these inferred weights to prune out the TSTS banks as going concern entities by assigning a higher weight to all the 'bad' CAMELS components (*i.e.*, *ASSETQLT1* and *SENSRISK1*) and a lower weight to all the 'good' CAMELS components (*i.e.*, *CAP1*, *MNGEXP1*, *EARN1*, and *LQDT1*) of these banks. In the case of TBTF banks, on the other hand, they assign the opposite weights to the relevant CAMELS components in order to keep TBTF banks alive as going concerns, thus bailing out their shareholders together with other creditors. Hence, we can argue that the regulatory framework unduly affects the TSTS banks and their shareholders, even though these banks have performed relatively better compared to TBTF banks as demonstrated earlier in our univariate analysis.

We further observe that the estimates of the CAMELS components in the TSTS and TBTF regimes are less statistically significant if compared with the relevant estimates in the intermediate regime. This shows that the overall performance of banks plays a slightly less important role for the failure and TARP probabilities in the two extreme regimes compared to the role that performance indicators play in the intermediate regime.

Importantly, the existence of the two extreme size regimes implies that market clearance is not likely to occur in the case of a financial turmoil under the existing bank resolution and intervention mechanisms. The market outcome is indeed expected to be distorted as it is not the performance of banks that mainly determines the probability of a bank to fail or to stay afloat. Poorly performed banks remain alive due to an exogenous intervention, banishing others that perform relatively better. In other words, performance is not the key indicator for a bank to survive or to fail; to the contrary, it is bank size, or, more accurately, the size regime in which a bank is clustered, which essentially determines the failure probability.

In sum, we document that size is the key determinant that classifies banks into the two different regimes and makes the authorities to treat distressed banks differently. Indeed, regulators appear to follow different standards in supporting distressed banks based on their size. They appear to be reluctant to help some distressed banks to survive as going concern entities since they consider them as being TSTS regardless of their relative performance. On the other hand, they provide financial support to some other banks, which are perceived as being TBTF thereby helping them to survive as going concern entities even though their performance is relatively worse compared to that of TSTS banks. This key finding is in line with the main argument of Goodhart and Huang (2005) according to which it is optimal for supervisory and regulatory authorities to rescue those banks whose size is above some threshold level. We thus suggest that regulators should revise their implied weighting scheme on the ratings system they utilise to evaluate banks' overall performance so as to "push" banks above the TSTS threshold and below the TBTF threshold.

As regards the complexity variables (*ORGCOMPL*, *BUSINCOMPL*), both are negatively linked to the probability of failure across the three size regimes. This implies that the more complex a bank is, the more likely is to receive TARP assistance thus having all of its creditors protected. However, the magnitude of complexity on failure is considerably higher in the TBTF regime. Indeed, the organisational (operational) complexity of a TBTF bank reduces the failure likelihood by 11.5% (14.4%), whereas the relevant percentages for TSTS banks are 5.4% and 7.1%, respectively. Notably, estimates are highly statistically significant only for the TBTF regime, showing that banks which are perceived as TBTF are also too-complex-to-fail.

Banks have lately turned to be too complex for insiders, but mostly for outsiders like regulators to understand. Bank officers, having to deal with these complexities, may struggle to manage every aspect of their business effectively. Generating a report that an outsider could comprehend and use as the basis for regulation may not be an easy task, so that additional agency problems are introduced into complex financial companies. Further, considering that shareholders are those who affect the risk appetite of their banks as evidence in the literature shows that boards' decisions create value for shareholders and largely represent their risk preferences, the latter agents may not only tolerate but also support the increase in size and complexity, as they feel that in case their bank gets into trouble, regulators will be more likely to keep their bank afloat. It is indeed unclear to regulators what the consequences of complex bank failures would be, so that, when push comes to shove, complex TBTF banks are more likely to be bailed out in full as we demonstrate in our model, which is not what happens for complex TSTS financial institutions. In short, the too-complex-to-fail problem may exacerbate the problems caused by the TBTF problem. This issue can be addressed partly through improved resolution procedures for TBTF distressed banks based on the relevant size threshold specified in our model; yet, it is rather unlikely that very large and extremely complex fragile banks will ever be treated precisely as small institutions are.

A serious problem stemming from the difficulty of regulators to implement formal resolution practices on TBTF financial institutions as a result of the size and the complexity of these institutions. The implicit subsidy which is encrypted in TARP-type bailouts creates strong incentives to the shareholders in TBTF banks to ask for even higher risks and for leverage maximisation. Our results show that this moral hazard phenomenon has been further amplified in recent years as banks have been able to push into modern and more complex organisational structures and activities thus broadening their scope. The nature of these complexities has made it very difficult for regulatory authorities to keep pace with the changes and analyse the implications of the failure of a TBTF bank thus choosing to keep these banks alive at the expense of shareholders of TSTS banks for which a FDIC-backed intervention is deemed to be preferable.

In a similar vein, bank managers also have strong incentives to respond to the existence of TSTS and TBTF size regimes by adopting a moral hazard behaviour. More specifically, it is in the interest of managers to shape and follow strategies that focus on the aggressive size growth of their banks knowing that the bigger and the more complex a bank becomes the more likely is to receive a TARP-style and not a FDIC-backed assistance and, therefore, not losing its charter. In line with this, Hakenes and Schnabel (2010) establish a link between size and performance by showing that banks which are not considered by authorities to be systemically important may turn to take higher risk, especially when the bailout probability of those banks which are protected by the system is increased.

We now turn to examine the effect of control variables on the likelihood of failure. A bank's political connections as captured by *POLCON* exert a significantly negative impact on failure as they lower the relevant probability by 10.8%. In line with this finding, Dunchin and Sosyura (2012) suggest that the political connections of distressed banks was a major determinant in the distribution of TARP funds. By the same token, Bayazitova and Shivdasani (2012) find that the TARP infusions were provided to those banks that posed systemic risk, faced high expected financial distress costs, and were politically well-connected. Hence, the shareholders of banks which are involved in the political process to a greater extent are more likely to receive a favourable treatment and being bailed out by the authorities together with uninsured depositors and other junior creditors. Along the same lines, we find that when a bank is more connected to regulators (*FEDCON*) then the failure probability of this banks is significantly reduced by 12.9%. Moreover, a connection to a House member serving on key finance committees involved in drafting and amending TARP (*COMMIT*) is associated with a statistically significant decrease of 5.2 percentage points in the likelihood of failure. In addition, our results reveal that contributions to political parties campaigns (*CAMP*) significantly lower the chance of a bank being let to fail as a going concern by 4.2%.

When a bank is involved as an acquirer in a M&A transaction (*MA*), this significantly reduces its failure likelihood by 2.6%. If a sample bank is located in some MSA, then it is less likely to fail. Indeed, the failure likelihood is reduced by 7.4%. The latter finding is confirmed by the geographical characteristics of our data set: a large number of failed banks is located in rural counties and not near the East and West Coasts of the U.S. These banks concentrate their activity in the mainland and, more specifically, in states like Iowa, Nebraska, and Utah. Most of the Northeastern and Southeastern states (excluding California) had either no or a few bank failures, whereas the Western U.S. states, which experienced a relatively larger decline in economic performance as measured by the GDP growth rate and the unemployment rate, had the highest bank failure rates.

As expected, newly-chartered banks are more likely to fail (4.0%), whereas banks which are publically traded are found to be less likely to fail (-5.4%). The latter result is in line with the results obtained thus far: TBTF banks are those which are typically publically traded in contrast with their TSTS counterparts which are not listed on the stock exchange market. Further, there is no statistically significant relation between a bank which is a subsidiary of a BHC and the probabilities under scrutiny.

6. Sensitivity analysis

We now move to examine the sensitivity of our baseline regression results. To this end, we use a set of alternative variables to construct CAMELS ratings. The main reason of doing so is because the components of CAMELS are kept confidential from regulators and, hence, it is crucial to test the sensitivity of our baseline regression results to a set of alternative CAMELS variables. Capital adequacy is measured by the ratio of Tier 1 regulatory capital to total risk-weighted assets (*CAP2*); asset quality is captured by the restructured and outstanding balances of loans and lease financing receivables that the bank has placed in nonaccrual status divided by total loans and leases (*ASSETQLT2*); management expertise is proxied by the total operating income calculated by the sum of interest income and non-interest income as a fraction of the total earning assets (*MNGEXP2*) which is a typical measure of operating efficiency in the banking literature (see, e.g., Lane et al., 1986); the return on equity given by the ratio of total net income to total equity capital is utilised to measure banks' earnings (*EARN2*); the ratio of federal funds purchased and securities sold under agreements to repurchase to total assets (*LQDT2*) is employed to measure the degree of liquidity of the sample banking firms; and the sensitivity to market risk (*SENSRISK2*) is proxied by the market interest rate risk defined as the quarterly standard deviation of the day-to-day 3-month U.S. T-bill rate divided by total earning assets. We also replace *ORGCOMPL* and *BUSINCOMPL* with a measure of complexity that accounts for the cross-border activities of the sample banks. This is given by the ratio of balances due from banks in foreign countries and foreign central banks to cash and balances from depository institutions (*CROSSCOMPL*).

Clearly, the first phase of TARP was driven by the very large size of banks and perhaps the interaction of size and complexity. We account for the impact of the involuntary participation in TARP by excluding the nine banks of the first phase from our analysis to alleviate any concerns that the decision of U.S. Treasury to force those banks to receive financial assistance was based on different motivations. Hence, our sample now consists of 815 instead of 824 TARP banks.

We again estimate Eq. (22) by linear probability OLS regression based on heteroskedasticity-robust standard errors ensuring that not less than 5% of the total observations would fall into any of the estimated regimes.[11] Instead of 300 bootstrap applications, we now apply 1,000 replications for each of the three bootstrap tests. As Table 7 reveals, the test for a single threshold F_1 and that of a double threshold F_2 are both significant with bootstrap *p*-values of 0.001, and 0.009, respectively. However, as in our baseline regression analysis, the test for a third threshold F_3 is far from being significant since the relevant bootstrap *p*-value equals to 0.527. We, therefore, confirm the existence of two thresholds for bank size in our model.

Table 7
Tests for determining the number of thresholds.

H_0: no threshold vs one threshold	
F_1	27.104
p-value	0.001
(10%, 5%, 1% critical values)	(11.86, 13.79, 28.96)
H_0: one threshold vs two thresholds	
F_2	21.792
p-value	0.009
(10%, 5%, 1% critical values)	(11.05, 14.74, 25.28)
H_0: two vs three thresholds	
F_3	7.017
p-value	0.527
(10%, 5%, 1% critical values)	(8.33, 10.08, 17.53)

The test statistics F_1, F_2, F_3, their asymptotic bootstrap *p*-values, and the relevant critical values at 10%, 5%, 1% levels are presented in this Table. 1,000 bootstrap replications are applied for each of the three bootstrap tests.

Further, as shown in Table 8, the point estimates of the two thresholds for *SIZE* we obtain are equal to \$0.384bn and \$2.919bn, respectively. Therefore, our sample banks are endogenously allocated to the following three size regimes: a TSTS regime that contains all banks with total

[11] As an additional robustness test, we allow the minimum number of observations that lie in each regime to be equal to the 2.5% of total observations. The results we obtain are very similar to those obtained in the baseline analysis. The estimated TSTS size threshold is found to only slightly decrease, whist the TBTF threshold is found to slightly increase.

assets up to \$0.384bn; an intermediate regime that consists of all banks with total assets from \$0.384bn to \$2.919bn; and a TBTF regime which includes all institutions with more than \$2.919bn.

Table 8
Threshold estimates.

	Estimate	**95% confidence interval**
$\hat{\gamma}_1$	\$0.384bn	[0.305, 0.418]
$\hat{\gamma}_2$	\$2.919bn	[2.659, 3.024]

This Table reports the point estimates of the two size thresholds and their asymptotic 95% confidence interval. Estimates are expressed in US\$ bn.

As presented in Table 9, 296 failed banks (65.92%) and 93 TARP banks (11.41%) are classified in the TSTS size regime; 128 failed banks (28.51%) and 242 TARP banks (29.69%) fall into the intermediate regime; and, 25 failed banks (5.57%) and 480 TARP banks (58.90%) are classified in the TBTF regime.

Table 9
Threshold size clustering of failed and TARP banks.

Size regime	**Failed banks**	**TARP banks**
TSTS regime: *SIZE* ≤ \$0.384bn	296	93
	(65.92%)	(11.41%)
Intermediate regime: \$0.384bn < *SIZE* ≤ \$2.919bn	128	242
	(28.51%)	(29.69%)
TBTF regime: \$2.919bn < *SIZE*	25	480
	(5.57%)	(59.90%)

This Table reports the number of failed and TARP banks which are classified into each of the three size regimes. The relevant percentages are reported in parentheses.

The estimation results of our sensitivity analysis are consistent with those of our baseline analysis. As Table 10 reports, the estimates of the CAMELS components are all statistically significant and their signs remain homogeneous across the three size regimes showing that banking firms with illiquid and risky assets, inadequate equity capital, poor management, low levels of earnings, and high sensitivity to market risk are more likely to fail as going concerns. Importantly, the TSTS banks are found to experience higher estimated coefficients on the CAMELS components which are positively linked to the failure probability and lower coefficients on the components which are negatively related with the relevant probability. This implies that regulators treat TSTS banks differently from TBTF banks: the failure probability is higher for a TSTS bank compared to that of a TBTF bank even if the performance of the two banks is the same. This is due to the different weights that regulators implicitly put on each CAMELS component and which are in favour of the TBTF banks. Moreover, in line with the results of our baseline analysis, the estimates of the CAMELS components in the TSTS and TBTF size regimes are found to be less statistically significant if compared with the relevant estimates in the intermediate regime. This implies that size -and not performance- is the key factor that makes the authorities to treat distressed banks differently. Overall, size appears to be the most important determinant of failure and TARP probabilities in the two extreme regimes.

The complexity variable (*CROSSCOMP*) is negatively linked to failure probability across the three size regimes, showing that banks which are more involved in cross-country transactions are more likely to receive TARP money and to remain alive as going concerns. The impact of complexity on failure is substantially higher in the TBTF regime compared to the other two regimes, and estimates are statistically significant only for banks that belong to this regime. Citi is maybe the most tangible example that provides strong support to our results. It had nearly 2,500 subsidiaries prior to the crisis and operated in 84 countries (Herring and Carmassi, 2010). It was so troubled that it was allowed by regulators to participate in both phases of TARP and also to receive additional funding through more junior forms of investment by the government.[12]

Our estimation results remain robust in respect to all the control variables we employ in our model. Importantly, we corroborate that better-connected banks are significantly more likely to receive TARP money. A bank's connections with politicians, political parties, or regulators exert

[12] Calomiris and Khan (2015) describe the Treasury's investments in Citi in detail.

a significantly negative impact on failure as they lower the relevant probability. This is to say, regulators are more likely to provide a TARP support to a distressed banking firm which is well-connected, thus saving all of its creditors. This result holds for all the four relevant variables either at the 1% level of significance (*POLCON* and *FEDCON*), or at the 5% level (*COMMIT* and *CAMP*).

Table 10
Threshold regression results-sensitivity analysis.

Variable	**TSTS regime** *SIZE* ≤ \$0.384bn	**Intermediate regime** \$0.384bn < *SIZE* ≤ \$2.919bn	**TBTF regime** \$2.919bn < *SIZE*	
CAP2	-0.052** (0.025)	-0.050*** (0.012)	-0.080** (0.038)	
ASSETQLT2	0.130** (0.053)	0.129*** (0.025)	0.076** (0.033)	
MNGEXP2	-0.044** (0.021)	-0.035** (0.017)	-0.068** (0.032)	
EARN2	-0.111** (0.052)	-0.110*** (0.027)	-0.155* (0.083)	
LQDT2	-0.029* (0.016)	-0.031*** (0.008)	-0.070** (0.029)	
SENSRISK2	0.109** (0.039)	0.104*** (0.016)	0.058** (0.028)	
CROSSCOMPL	-0.078 (0.049)	-0.053 (0.044)	-0.130*** (0.036)	
POLCON				-0.126*** (0.034)
FEDCON				-0.132*** (0.041)
COMMIT				-0.049** (0.024)
CAMP				-0.051** (0.022)
MA				-0.020** (0.009)
MSA				-0.073*** (0.017)

DENOVO		0.042** (0.019)
INVOL		
PUBLIC		-0.050*** (0.013)
BHC		-0.011 (0.015)
R^2	0.22	

This table presents the estimation results of the multiple threshold regression model (Eq. 22) for the three bank size (*SIZE*) regimes which are endogenously determined by our model: *SIZE* ≤ \$0.384bn, \$0.384bn < *SIZE* ≤ \$2.919bn, and \$2.919bn < *SIZE*. The dependent variable is equal to 1 if a sample bank received no TARP funds and went bankrupt and 0 if a bank received TARP money and survived the crisis. The alternative explanatory variables we use in the sensitivity analysis we conduct are: capital strength (*CAP2*), asset quality (*ASSETQLT2*), quality of management (*MNGEXP2*), earnings strength (*EARN2*), degree of liquidity (*LQDT2*), and sensitivity to market risk (*SENSRISK2*). The set of control variables includes: cross-border bank complexity (*CROSSCOMPL*); a dummy capturing the political connections of banks (*POLCON*); a dummy for the connections of banks with the federal regulatory and supervisory authorities (*FEDCON*); a dummy that shows if a sample bank is headquartered in a district of a House member who served on the key finance committees (*COMMIT*); a dummy for the banks which made PAC contributions in the 2008 election cycle (*CAMP*); a dummy for acquirer banks in M&A transactions (*MA*); a dummy showing whether a bank is located in a MSA or in a rural county (*MSA*); a dummy for newly-chartered banks (*DENOVO*); a dummy variable for banks which are listed on the stock exchange (*PUBLIC*); and a dummy indicating whether a bank is a subsidiary of a BHC (*BHC*). All observations are on bank level, constitute bank-quarter observations, and cover the entire data period, which extends from 2002q1 to 2012q4. All the explanatory variables are lagged by four quarters to address possible endogeneity concerns. The description of each variable and the relevant data sources are included in Appendix A. White heteroskedasticity-robust standard errors are reported in parentheses.
***, **, * correspond to 1%, 5%, and 10% level of significance respectively for a two-tailed distribution

7. Concluding remarks

In the recent financial crisis, the U.S. regulatory authorities provided substantial financial support to some distressed banks through TARP, whilst at the same time let several others to go bankrupt as going concern entities via FDIC-backed failures. Even though under both government interventions small retail depositors do not lose a penny, TARP government bailouts represent an implicit subsidy to the bank’s shareholders, which is not yet the case with FDIC-backed failures where shareholders as well as other creditors are not subsidised. In this paper, we aim to shed ample light on this non-uniform policy by focusing on size as well as the complexity, performance and risk profile of distressed banks with the utmost purpose to establish a double size threshold that can explain why these banks are treated differently by regulators.

The results of the univariate analysis we conduct confirm that authorities do not follow a homogeneous treatment of distressed banks. We demonstrate that banks that the FDIC took down have performed better than those that received TARP assistance and remained afloat. Further, we show that the latter banks were larger and more complex compared to the former ones, even though their degree of complexity was not highly significantly different from that of failed banks. Taken together, our results suggest that the decision of regulatory authorities to choose between a TARP-assisted bailout and a FDIC-backed failure is more influenced by the bank's size and whether this falls below some threshold level and less influenced by the bank's complexity, performance and risk appetite.

Our multivariate threshold regression analysis lends strong support to the aforementioned postulate by revealing that the failure and TARP probabilities are essentially determined by bank size, where two size thresholds are endogenously specified: one for the TSTS banks and a second one for the TBTF banks, which are considered to be the two sides of the same coin. Our threshold variable, bank size, provides us with the different weights that regulators implicitly assign to CAMELS components. A TSTS bank that has the same overall performance with a TBTF bank is more likely to fail as a going concern entity due to the different weights put on each CAMELS component and which are in favour of TBTF banks. That is, regulators appear to be reluctant to bailout the shareholders and the uninsured creditors of a distressed bank if the bank is considered to be TSTS. From a market equilibrium point of view, the existence of the two extreme size regimes means that the free market outcome cannot be reached. This occurs because it is the size of banks and not their performance which is the key decision variable for the failure and TARP probabilities, implying that banks of lower performance remain alive due to an exogenous intervention, banishing others that perform relatively better.

The organisational and operational complexity of banks are found to be negatively linked to the probability of failure across the different size regimes, implying that the more complex a bank is the more likely is to receive TARP assistance and, hence, to have all of its creditors fully protected. However, the magnitude of complexity on failure is considerably higher in the TBTF regime. In addition, estimates are highly statistically significant only in the TBTF regime, showing that banks which are perceived as TBTF are also too-complex-to-fail. The latter findings combined with those on bank size provide strong incentives to bank managers to shape strategies towards the size expansion and the increase in complexity of their banks thus adopting a moral hazard pattern of management and administration. In other words, banks are incentivised to make themselves big and complex enough not to be left to fail by regulators as going concern entities. In fact, this moral hazard implication is twofold: on the one hand, managers will deliberately try to escape placing their banks into the TSTS regime, and, on the other, the shareholders and the institutional creditors of big and complex banks will reward them with lower borrowing costs for being placed in the TBTF regime.

When bank size takes values smaller than the critical TSTS threshold size or exceeds the TBTF threshold, the banking sector enters a zone of vulnerability. In the case of a considerable financial turmoil, like that of 2007-8, regulatory authorities may confront the dilemma to let a number of TSTS banks to collapse which incurs a considerable cost for the economy or to provide substantial financial support to TBTF banks. In view of this dilemma, authorities can resort to the two size thresholds we determine in our research to formulate the necessary regulatory and supervisory policies to reduce the bankruptcy and bailout burdens and mitigate the relevant risks that exert a considerable destabilising impact on the entire system by taking pre-emptive measures to avoid bank size from crossing the specified critical threshold values.

Acknowledgements: We are indebted to Bruce Hansen for his guidance, valuable comments and suggestions. Also, we would like to express our sincere thanks to Charles Calomiris, for his incisive comments. Further, we wish to thank the participants in the 19th Global Finance Conference in Chicago (U.S.A), the FMA European Meeting 2013 in Luxembourg, the IFABS 2013 in Nottingham (UK), the 26th Australasian Finance and Banking Conference in Sydney, and the 10th BMRC-DEMS Conference in Brunel University. The paper has been benefited from discussions with Manapol Ekkayokkaya, Franco Fiordelisi, Iftekhar Hasan, Martien Lubberink, Gordon Roberts, Anthony Saunders, David Storey, Sunti Tirapat, and with colleagues at the Chulalongkorn University in Thailand. Also, we would like to thank the Fonds National de la Research (FNR) in Luxembourg for its financial support and Mounir Sahl for his valuable help in data collection.

References

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

Basel Committee on Banking Supervision (BCBS), 2014. The G-SIB assessment methodology-score calculation. Bank for International Settlements.

Bayazitova, D., Shivdasani, A., 2012. Assessing TARP. The Review of Financial Studies 25, 377-407.

Ben Cheikh, N., Louhichi, W., 2016. Revisiting the role of inflation environment in exchange rate pass-through: A panel threshold approach. Economic Modelling 52, 233-238.

Berger, A.N., Bouwman, C.H.S., 2013. How does capital affect bank performance during financial crises? Journal of Financial Economics 109, 146-176.

Berger, A.N., Humphrey, D.B., 1991. The dominance of inefficiencies over scale and product mix economies in banking. Journal of Monetary Economics 28, 117-148.

Berger, A.N., Humphrey, D.B., 1997. Efficiency of financial institutions: International survey and directions for future research. European Journal of Operational Research 98, 175-212.

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

Berger, A.N., Roman, R.A., 2015. Did TARP banks get competitive advantage? Journal of Financial and Quantitative Analysis 50, 1199-1236.

Bertay, A.C., Demirguc-Kunt, A., Huizinga, H., 2013. Do we need big banks? Evidence on performance, strategy, and market discipline. Journal of Financial Intermediation 22, 532-558.

Blau, B.M., Brough, T.J., Thomas, D.W., 2013. Corporate lobbying, political connections, and the bailout of banks. Journal of Banking and Finance 37, 3007-3017.

Boyd, J.H., Runkle, D.E., 1993. Size performance of banking firms: Testing the predictions of theory. Journal of Monetary Economics 31, 47-67.

Buiter, W., 2009. Lessons from the global financial crisis for regulators and supervisors, paper presented at the 25th anniversary Workshop 'The global financial crisis: Lessons and outlook' of the advanced studies program of the IFW. Kiel, 8-9 May, 2009.

Caballero, R.J., Simsek, A., 2013. Fire sales in a model of complexity. The Journal of Finance 68, 2549-2587.

Calomiris, W.C., Kahn, U., 2015. An assessment of TARP assistance to financial institutions. Journal of Economic Perspectives 29, 53-80.

Carter, D.A., McNulty, J.E., 2005. Deregulation, technological change, and the business-lending performance of large and small banks. Journal of Banking and Finance 29, 1113-1130.

Cetorelli, N., McAndrews, J., Traina J., 2014. Evolution in bank complexity. Economic Policy Review, Federal Reserve Bank of New York, 85-106.

Coelli, T., Rao, D.S.P., O'Donnell, C.C., Battese, G.E., 2005. An introduction to efficiency and productivity analysis. Springer: New York.

De Haan, J., Poghosyan, T., 2012. Bank size, market concentration, and bank earnings volatility in US. Journal of International Financial Markets, Institutions, and Money 22, 35-54.

DeYoung, R., 2003. De novo bank exit. Journal of Money, Credit, and Banking 35, 711-728.

DeYoung, R., Hasan, I., 1998. The performance of de novo commercial banks. Journal of Banking and Finance 22, 565-587.

DeYoung, R., Hunter, W.C., Udell, G.F., 2004. The past, present, and probable future for community banks. Journal of Financial Services Research 25, 85-133.

DeYoung, R., Kowalik, M., Reidhill, J., 2013. A theory of failed bank resolution: Technological change and political economics. Journal of Financial Stability 9, 612-627.

Demsetz, R.S., Strahan, P.E., 1997. Diversification, size, and risk at Bank Holding Companies. Journal of Money, Credit and Banking 29, 300-313.

Duchin, R., Sosyura, D., 2012. The politics of government investment. Journal of Financial Economics 106, 24-48.

Duchin, R., Sosyura, D., 2014. Safer ratios, riskier portfolios: Banks' response to government aid. Journal of Financial Economics 113, 1-28.

Eisenbach, T.M., Lucca, D.O, Townsend, R.M., 2016. The economics of bank supervision. Federal Reserve Bank of New York Staff Reports, no. 769.

Fatum, R., Yamamoto, Y., 2014. Large versus small foreign exchange interventions. Journal of Banking and Finance 43, 114-123.

Galloway, T.M., Winson, B.L., Roden, D.M., 1997. Banks' changing incentives and opportunities for risk taking. Journal of Banking and Finance 21, 509-527.

Gandhi, P., Lustig, H., 2015. Size anomalies in U.S. bank stock returns. The Journal of Finance 70, 733-768.

Goodhart, C.A.E., Huang, H., 2005. The lender of last resort. Journal of Banking and Finance 29, 1059-1082.

Hakenes, H., Schnabel, I., 2010. Banks without parachutes: Competitive effects of government bail-out policies. Journal of Financial Stability 6, 156-168.

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413-430.

Hansen, B.E., 1999. Threshold effects in non-dynamic panels: Estimation, testing, and inference. Journal of Econometrics 93, 345-368.

Hansen, B.E., 2000. Sample splitting and threshold estimation. Econometrica 68, 575-603.

Hansen, B.E., 2011. Threshold autoregression in Economics. Statistics and Its Interface 4, 123-127.

Herring, R., Carmassi, J., 2010. "The corporate structure of international financial conglomerates: Complexity and its implications for safety and soundness." In *The Oxford Handbook of Banking*, ed. Berger, A.N., Molyneux, P., Wilson, J.O.S., 195-229. Oxford: Oxford University Press.

Hossfeld, O., MacDonald, R., 2015. Carry funding and safe haven currencies: A threshold regression approach. Journal of International Money and Finance 59, 185-202.

Huang, X., Zhou, H., Zhu, H., 2009. A framework for assessing the systemic risk of major financial institutions. Journal of Banking and Finance 33, 2036-2049.

Jordan, D.J., Rice, D., Sanchez, J., Wort, D.H., 2011. Explaining bank market-to-book ratios: Evidence from 2006 to 2009. Journal of Banking and Finance 35, 2047-2055.

Kim, D.H., Stock, D., 2012. Impact of the TARP financing choice on existing preferred stock. Journal of Corporate Finance 18, 1121-1142.

Kanno, M., 2015. Assessing systemic risk using interbank exposures in the global banking system. Journal of Financial Stability 20, 105-130.

Klomp, J., de Haan, J., 2012. Banking risk and regulation: Does one size fit all? Journal of Banking and Finance 36, 3197-3212.

Lane, W., Looney, S., Wansley, J., 1986. An application of the Cox proportional hazards model to bank failure. Journal of Banking and Finance 10, 511-531.

Lensink, R., Hermes, N., 2004. The short-term effects of foreign bank entry on domestic bank behaviour: Does economic development matter? Journal of Banking and Finance 28, 553-568.

Li, L. 2013. TARP funds distribution and bank loan supply. Journal of Banking & Finance 37, 4777-4792.

O'Hara, M., Shaw, W., 1990. Deposit insurance and wealth effects: the value of being 'too big to fail'. Journal of Finance 45, 1587-1600.

Papanikolaou, N.I., Wolff, C.P., 2014. The role of on- and off-balance-sheet leverage of banks in the late 2000s crisis. Journal of Financial Stability 14, 3-22.

Patro, D.K., Qi, M., Sun, X., 2013. A simple indicator of systemic risk. Journal of Financial Stability 9, 105-116.

Rime, B., Stiroh, K.J., 2003. The performance of universal banks: evidence from Switzerland. Journal of Banking and Finance 27, 2121-2150.

Stojanovic, D., Vaughan, M.D., Yeager, T.J., 2008. Do federal home loan bank membership and advances increase bank risk-taking? Journal of Banking and Finance 32, 680-698.

Tong, H., 1983. *Threshold models in non-linear time series analysis: Lecture notes in Statistics*, 21. Berlin: Springer.

Veronesi, P., Zingales, L., 2010. Paulson's gift. Journal of Financial Economics 97, 339-368.

Wang, M.H., Huang, T.H., 2009. Threshold effects of financial status on the cost frontiers of financial institutions in non-dynamic panels. Applied Economics 41, 3389-3401.

Wooldridge, J. M., 2002. *Econometric analysis of cross section and panel data*, MIT Press, Cambridge, MA.

Appendix A
Variables and data sources.

Variable	Abbreviation	Definition	Data source
CAMELS components			
Capital adequacy	*CAP1*	The ratio of book equity capital to total assets	Call Reports
	CAP2	The ratio of regulatory (Tier 1) capital to total risk-weighted assets	
Asset quality	*ASSETQLT1*	The ratio of non-performing loans to total loans and leases	
	ASSETQLT2	The ratio of restructured and outstanding balances of loans and lease financing receivables that the bank has placed in nonaccrual status to total loans and leases	
Management expertise	*MNGEXP1*	Managerial efficiency calculated using the input-oriented DEA model	
	MNGEXP2	The ratio of total operating income calculated as the sum of interest income and non-interest income to total earning assets	
Earnings strength	*EARN1*	The ratio of total net income given by the difference between total interest plus non-interest income and total interest plus non-interest expense to total assets	
	EARN2	The ratio of total net income given by the difference between total interest plus non-interest income and total interest plus non-interest expense to total equity capital	
Liquidity	*LQDT1*	The ratio of cash and balances due from depository institutions to total deposits	
	LQDT2	The ratio of federal funds purchased and securities sold under agreements to repurchase to total assets	
Sensitivity to market risk	*SENSRISK1*	The change in the slope of the yield curve (given by the change in the quarterly difference between the 10-year U.S. T-bill rate and the 3-month U.S. T-bill rate) divided by total earning assets.	Federal Reserve Board & U.S. Department of the Treasury
	SENSRISK2	Market interest rate risk (defined as the quarterly standard deviation of the day-to-day 3-month U.S. T-bill rate) divided by total earning assets.	

Threshold variable			
Bank size	*SIZE*	The book value of total assets	Call Reports
Bank complexity			
Organisational complexity	*ORGCOMPL*	The log of the product of the number of branches that each sample bank has and the number of U.S. states in which the bank	
Business complexity	*BUSINCOMPL*	The sum of the notional amount of outstanding derivative contracts and the amount of credit exposure arising from recourse or other credit enhancements provided to the purchasers of the securitised loans, leases, and other assets divided by total assets	Call Reports
Cross-border complexity	*CROSSCOMPL*	The ratio of balances due from banks in foreign countries and foreign central banks to cash and balances from depository institutions	
Managerial efficiency			
Total loans	*u1*	The sum of commercial, construction, industrial, individual and real estate loans	Call Reports
Total deposits	*u2*	The sum of total transaction deposit accounts, non-transaction savings deposits, and total time deposits	

Other earning assets	*u3*	The sum of income-earned assets other than loans and the net deferred income taxes	
Total non-interest income	*u4*	The sum of income from fiduciary activities, service charges on deposit accounts, trading fees and income from foreign exchange transactions and from assets held in trading accounts, and other non-interest income	
Securitisation activity	*u5*	The value of the outstanding principal balance of loans, leases, and all relevant assets securitised and sold to other financial institutions with recourse or other credit enhancements divided by total assets	
Price of borrowed funds	*v1*	The ratio of total interest expense to total deposits and other borrowed money	
Price of labour	*v2*	The ratio of total salaries and benefits to the number of full-time employees	
Price of physical capital	*v3*	The ratio of expenses for premises and fixed assets to the dollar amount of premises and fixed assets	
Control variables			
Political connections	*POLCON*	A dummy that equals one if a bank has employed, or is currently employing an individual who is also employed or has been employed in the federal government or appointed to a government advisory board, a congressional or presidential cabinet entity, or an independent commission	Centre for Responsive Politics (CRP)'s Revolving Door
Federal connections	*FEDCON*	A dummy that is equal to unity if an executive at a sample bank was on the board of directors of one of the 12 Federal Reserve Banks or one of their branches either in 2008 or 2009	Federal Reserve & BoardEx

Political commitments	*COMMIT*	A dummy that equals one if a sample bank is headquartered in a district of a House member who served on the key finance committees involved in drafting and amending TARP either in 2008 or 2009	House of Representative, U.S. Census Bureau & U.S. Library of Congress
Campaign contributions	*CAMP*	A dummy that takes the value of one if a sample bank has made PAC contributions in the election cycle for the 2008 congressional election to the members of the Subcommittee on Financial Institutions and the Subcommittee on Capital Markets	Federal Election Commission Political Action Committees (PACs)
M&A transactions	*MA*	A dummy which is equal to unity if a bank is involved in a M&A transaction as an acquirer	M&As database/Federal Reserve Bank of Chicago
Bank location	*MSA*	A dummy showing whether a bank is located in a Metropolitan Statistical Area or not	Call Reports & U.S. Office of Management and Budget
Newly-chartered bank	*DENOVO*	A dummy capturing the banks which are less than five years old	Call Reports
Listed bank	*PUBLIC*	A dummy which is equal to unity if bank *i* is listed on the exchange market	Call Reports & Center for Research in Security Prices (CRSP)
BHC affiliation	*BHC*	A dummy variable indicating whether a sample bank is a subsidiary of some BHC	Call Reports

This Appendix presents all the variables that we use in the baseline econometric analysis as well as in the sensitivity analysis. The abbreviation of each variable and the sources we use to collect the data are reported.

Appendix B

To calculate managerial efficiency (*MNGEXP1*), we employ the Data Envelopment Analysis (DEA) model. DEA model can be computed either as input- or output-oriented. The input-oriented DEA model shows by how much input quantities can be reduced without varying the output quantities produced. Similarly, the output-oriented DEA model assesses by how much output quantities can be proportionally increased without changing the input quantities used. Both output- and input-oriented models identify the same set of efficient/inefficient bank management. Nevertheless, even though the two approaches provide the same results under constant returns to scale, they give different values under variable returns to scale.[13]

We assume that for the N sample banks, there exist P inputs producing M outputs. Hence, each bank i uses a nonnegative vector of inputs denoted by $v^i = (v_1^i,\ v_2^i, \dots, v_p^i) \in R_+^P$ to produce a nonnegative vector of outputs, denoted by $u^i = (u_1^i,\ u_2^i, \dots, u_m^i) \in R_+^M$, where: $i = 1, 2,\dots, N$; $p = 1, 2,\dots, P$; and, $m = 1, 2,\dots, M$. The production technology, $F = \{(u, v): v\ can\ produce\ u\}$, describes the set of feasible input-output vectors. The input sets of production technology, $L(y) = \{v: (u, v) \in F\}$, describe the sets of input vectors which are feasible for each output vector.

To measure the variable returns to scale managerial cost efficiency (*MNGEXP1*), we resort to the following input-oriented DEA model, where inputs are minimised and outputs are held at constant levels. Below, we sketch out the optimisation (minimisation) problem of bank$_1$'s (i=1) cost inefficiency. Note that each bank i faces the same optimisation problem.

$$MNGEXP1_1^* = \min(-MNGEXP1_1),\ \ s.t.\ \ \sum_{i=1}^{N} \lambda_i v_{ip} \leq (MNGEXP1_1)\big(v_{1p}\big) \qquad \text{(B1)}$$

[13] For a detailed discussion on the differences between input- and output-oriented DEA models, the interested reader can refer to Coelli et al. (2005).

$$\sum_{i=1}^{N} \lambda_i u_{im} \geq u_{1m} \tag{B2}$$

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{B3}$$

$$\lambda_i \geq 0 \tag{B4}$$

In Eq. (B1- B4), v_{1p} and u_{1m} are the *p*th input and *m*th output for bank_1, respectively; the convexity constraint, $\sum_{i=1}^{N} \lambda_i = 1$, accounts for variable returns to scale, where λ_i stands for the activity vector and denotes the intensity levels at which the total observations are conducted. This approach, through the convexity constraint, forms a convex hull of intersecting planes, since the frontier production plane is defined by combining a set of actual production planes. If $MNGEXP1_1^*$ is equal to unity, then the optimal efficiency score is achieved for bank_1. This shows that the levels of inputs used cannot be proportionally improved given the output levels, indicating that bank_1 lies upon the cost efficiency frontier. If, on the other hand, $MNGEXP1_1$ is less than unity the management of bank_1 is considered to be inefficient. The more $MNGEXP1_1$ deviates from the unity, the less efficient the management of bank_1 becomes.

An important concern in the estimation of *MNGEXP1* is the definition of inputs and outputs. This essentially depends on the specific role that deposits play in the overall business model of banks. The relevant literature addresses this issue by traditionally referring to two approaches: the intermediation (or asset) approach, and the production (or value-added) approach.[14] Under the former approach, financial firms are viewed as intermediaries which transform deposits and purchased funds into loans and other earning assets. That is, liabilities and physical factors are treated as inputs, while assets are treated as outputs. The production approach, on the other hand,

[14] See Berger and Humphrey (1997) for a detailed analysis of the advantages and disadvantages of each of the two approaches.

regards financial institutions as producers of services for account holders, measuring output with the number of transactions or documents processed over a given period of time. Therefore, deposits are encompassed in the output and not in the input vector, which exclusively consists of physical entities.

Berger and Humphrey (1991) proposed a third approach, the modified production approach, which, contrary to the aforementioned traditional approaches, captures the dual role of bank deposits. This third approach is regarded as a combination of the intermediation and production approaches, as it enables the consideration of both the input and output characteristics of deposits in the cost function. More specifically, the price of deposits is considered to be an input, whereas the volume of deposits is accounted as an output. Under this specification, banks are assumed to provide intermediation and loan services as well as payment, liquidity, and safekeeping services at the same time. Hence, it can be argued that the latter approach describes the key bank activity of deposit-taking in a more complete manner thereby providing a closer representation of reality.

We adopt the modified production approach to define inputs and outputs in the estimation of *MNGEXP1*. We specify five variable outputs in total of which traditional banking activities are captured by three outputs, namely, total loans (u_1) calculated as the sum of commercial, construction, industrial, individual and real estate loans; total deposits (u_2) which is the sum of total transaction deposit accounts, non-transaction savings deposits, and total time deposits; and, other earning assets (u_3), expressed as the sum of income-earned assets other than loans and the net deferred income taxes. Non-traditional banking activities are proxied by two outputs: total non-interest income (u_4), which is the sum of income from fiduciary activities, service charges on deposit accounts, trading fees and income from foreign exchange transactions and from assets held in trading accounts augmented by any other non-interest income; and, securitisation activity (u_5)

measured as the value of the outstanding principal balance of loans, leases, and all relevant assets securitised and sold to other financial institutions with recourse or other credit enhancements divided by total assets.

Regarding the inputs we employ in the estimation of *MNGEXP1*, we consider borrowed funds, labour, and physical capital. The price of borrowed funds (v_1) is defined as the ratio of total interest expense scaled by total deposits and other borrowed money; the price of labour (v_2) is calculated by dividing total salaries and benefits by the number of full-time employees; and, lastly, the price of physical capital (v_3), which is equal to the expenses for premises and fixed assets divided by the dollar amount of premises and fixed assets.