BANK FOR INTERNATIONAL SETTLEMENTS

# BIS Working Papers

## No 599

# Banking Industry Dynamics and Size-Dependent Capital Regulation

by Tirupam Goel

Monetary and Economic Department

December 2016

JEL classification: G21, G28, E50, C60

Keywords: Size distribution, Entry & Exit, Heterogeneous agent models, Size dependent policy

This publication is available on the BIS website (www.bis.org).



ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

# Banking Industry Dynamics and Size-Dependent Capital Regulation

Tirupam Goel*
Bank for International Settlements

December 2016

**Abstract**

This paper presents a general equilibrium model with a dynamic banking sector to characterize optimal size-dependent bank capital regulation (CR). Bank leverage choices are subject to the risk-return trade-off: high leverage increases expected return on capital, but also increases return variance and bank failure risk. Financial frictions imply that bank leverage choices are socially inefficient, providing scope for a welfare-enhancing CR that imposes a cap on bank leverage. The optimal CR is tighter relative to the pre-crisis benchmark. Optimal CR is also bank specific, and tighter for large banks than for small banks. This is for three reasons. First, allowing small banks to take more leverage enables them to potentially grow faster, leading to a *growth effect*. Second, although more leverage by small banks results in a higher exit rate, these exits are by the less efficient banks, leading to a *cleansing effect*. Third, failures are more costly among large banks, because these are more efficient in equilibrium and intermediate more capital. Therefore, tighter regulation for large banks renders them less prone to failure, leading to a *stabilization effect*. In terms of industry dynamics, tighter CR results in a smaller bank exit rate and a larger equilibrium mass of better capitalized banks, even though physical capital stock and wages are lower. The calibrated model rationalizes various steady state moments of the US banking industry, and provides general support for the Basel III GSIB framework.

JEL Classification: G21, G28, E50, C60
Keywords: Size distribution, Entry & Exit, Heterogeneous agent models, Size dependent policy

---

*Email: tirupam.goel@bis.org, Address: Bank for International Settlements, Centralbahnplatz 2, 4051, Basel, Switzerland. Acknowledgments: I am extremely grateful to Assaf Razin and Maxim Troshkin for invaluable guidance and encouragement. I am very thankful to Eswar S. Prasad, Karel Mertens, John M. Abowd, Julieta Caunedo and Christopher Huckfeldt for valuable advice and numerous useful discussions. I am also thankful to Frederic Boissay, Markus Brunnermeier, Dean Corbae, Ingo Fender, Benjamin Moll, Guillermo L. Ordonez, Edouard Schaal, Alp Simsek, Julia Thomas, and various seminar participants for useful comments. I thank Cornell Institute for Social and Economic Research for providing state-of-the-art computing resources, and I gratefully acknowledge funding from the Labor Dynamics Institute, and funding from the Department of Economics at Cornell University. Any remaining errors in this paper are mine. Please use this link for the associated Matlab codes: https://sites.google.com/site/tirupam/. Disclaimer: The views expressed here are those of the author and not necessarily those of the Bank for International Settlements.

# 1 Introduction

A bank's leverage decision,[1] where leverage is defined as the ratio of bank assets to bank net worth (i.e. capital), has implications for growth and solvency. On the one hand, leverage increases the potential for a higher return on net worth.[2] On the other hand, leverage increases the variance in returns on net worth and the probability of bank failure (i.e. risk), where failure is defined as the event where a bank's net-worth drops below zero. In this sense, bank leverage decisions have implications for banking industry dynamics that include entry, exit and size distribution (Corbae and D'erasmo [2014]), and for such macroeconomic aggregates as output, investment, and distribution of consumption. In the presence of a friction between bank ownership and management, bank leverage decisions can be socially inefficient, and the heterogeneity of bank size and bank productivity matters for the size of this inefficiency. For example, high leverage by a large bank can lead to greater variance in the return on net-worth compared with the same high leverage by two or more smaller banks obtained by splitting up the large bank – this is because, in the case of small banks, there is some diversification of shock across banks, whereas the case of the large bank would correspond to that of perfectly correlated shocks. Therefore, in the presence of frictions, there are welfare gains from a capital regulation (CR) that imposes a cap on bank leverage, especially if such regulation depends on bank size.[3] To this end, the objective of this paper is to characterize optimal *size-dependent* capital regulation.[4]

Against this backdrop, the question of bank capital regulation has gained considerable importance since the 2008 financial crisis, as high leverage played a critical role in rendering the scale of the crisis. In this regard, among other measures, Basel III (BCBS [2011]) envisions not just a more stringent regulation framework relative to Basel II, but also that global systemically important banks (GSIBs) should face tighter regulation relative to smaller banks. In implementing these two aspects of the Basel III framework, the regulators are facing two key trade-offs. The first relates to the overall tightening of capital regulation, as discussed by, for example Van den Heuvel [2008] and Corbae and D'erasmo [2014], and is also studied in this paper. The second relates to size differentiation. If regulators impose *relatively* tighter capital regulation on smaller banks, this might constrain their growth and end up benefiting larger banks, which might be undesirable from a competition perspective. Also, if larger banks are allowed relatively greater leverage, and if they end up assuming more leverage, larger banks become more prone to failure, which can be socially costly. Figure 1 documents that, when failing, larger banks incur bigger costs, which are often

[1]In this paper, "bank" is a financial intermediary, and represents both commercial banks, and savings and loan associations.

[2]High leverage, i.e., more debt financing relative to equity financing, also implies the tax advantage of debt over equity, but this feature is not pursued in this paper.

[3]In this paper, capital regulation (CR) is defined as a cap on the ratio of bank assets and bank net-worth (i.e. equity). While not directly comparable to the requirements established under Basel III, this formulation is closest in spirit to (the inverse of) the leverage ratio requirement, which is defined as regulatory capital over a risk-insensitive measure of total exposure. Risk-weighted capital requirements would generate comparable results only to the extent that an average risk weight of 100% is being assumed.

[4]Both net worth and assets are appropriate measures of bank size, and there is an equivalence mapping between approaches using either measure. I adopt the former measure for mathematical convenience.

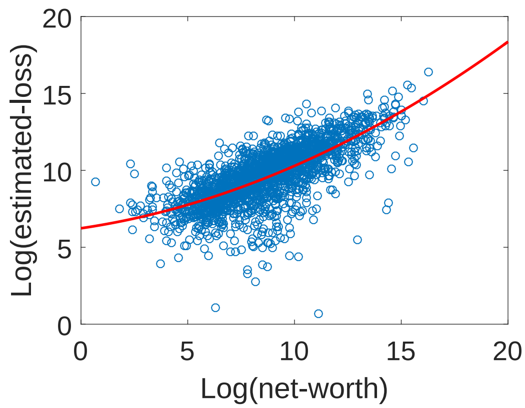


Figure 1: US savings and commercial banks: Log(net-worth) and log(estimated-loss) during 1985 - 2014; *(source: FDIC Historical Statistics on Banking, and Call and Thrift Financial Reports)*

borne by the public (i.e. via deposit insurance).[5] On the other hand, if regulators impose *relatively* tighter capital regulation for large banks, this might reduce the aggregate volume of financial intermediation by restricting bank lending, where large banks play a key role. This might also create disincentives for banks to (learn / adopt technology and) grow large. This can be because banks that better adopt technology, or learn over time, face a better chance of growing large over time, as discussed in Hughes and Mester [2013] and several other empirical studies. Therefore, to determine the most effective size-dependent regulation, it is essential to have a framework in which one can study some of these trade-offs quantitatively.

**Model overview** To this end, I develop a real business cycle (RBC) model designed to examine the relationship between bank leverage, banking industry dynamics, and macroeconomic aggregates. The model features a dynamic banking sector that consists of a continuum of heterogeneous banks. There is no aggregate, but only idiosyncratic uncertainty. The banking industry is modeled in the spirit of Hopenhayn [1992], where banks are intermediaries that accept deposits from households and purchase shares in firms that operate a standard production technology. Banks have idiosyncratic efficiency in managing their equity in firms, and this idiosyncratic efficiency is modeled as a first-order Markov process. To provide a rationale for banks, a limited commitment setup *à la* Gertler and Kiyotaki [2010] is assumed between banks and their depositors. That is, bankers are assumed to be able to divert assets for personal consumption. To ensure that bankers do not divert assets, they are subject to a borrowing constraint that limits the amount of deposits they can raise. The borrowing constraint ensures that the value of diverting assets is not larger than the value of continuing operations. Incumbent banks fail when their net worth

[5]The net worth reported in the figure is as of the last regulatory filing by the bank before its failure. The estimated loss is the amount of FDIC funds disbursed to meet the shortfall in liabilities. In this sense, the estimated loss captures only the idiosyncratic effect of the failure of a bank.

drops below zero, which can happen because of a large shock, or because of high leverage ex-ante. Post exit, new banks enter as per a free entry process.

To quantify the distributional welfare consequences of a size-dependent policy, I consider consumers to consist of workers and bankers.[6] The workers supply labor and also manage banks. The bankers are heterogeneous and are modeled in the spirit of Aiyagari [1994], but with an income endowment process that is endogenously determined. Bankers are bank owners, and receive dividend income from the bank whose shares they own. Since banks are subject to idiosyncratic shocks that govern dividend payments, bankers are subject to idiosyncratic dividend income shocks. The friction that rationalizes capital regulation in this model is the distinction between bank ownership and management – managers do not care about the risk-aversion of bank owners, and choose to maximize bank value by taking socially inefficient levels of risk.[7]

**Key Results** Qualitatively, as regards optimal regulation, the main result of the paper is that there is scope for a tightening of bank regulation relative to the pre-crisis benchmark.[8] Moreover, additional welfare gains are possible if capital regulation is designed in a bank size-specific fashion, providing general support for the Basel III GSIB framework, in particular the idea of a GSIB surcharge.

Quantitatively, in the calibrated model, a tightening of the capital regulation to a level of 3.55 (i.e. asset to net-worth ratio capped at 3.55) results in utilitarian welfare gains of 2.2% in consumption-equivalent (CE) terms relative to the pre-crisis benchmark. Moreover, additional welfare gains of 8.1% in CE terms can be achieved, if capital regulation is bank size specific.[9] In particular, the optimal size dependent regulation is characterized by a relatively tighter regulation of 3.17 for large banks and a relatively liberal regulation of 4.12 for small banks.

Intuitively, an overall tightening of regulation produces welfare gains primarily from a lower industry turnover rate and hence lower entry cost, and higher mean and lower variance of dividend payments by banks. The effect for dividend means is driven primarily by the fact that, in the optimal regulatory regime, there exists a larger mass of better capitalized banks that pay higher dividends in the aggregate. The variance effect is driven by the lower variance of bank dynamics due to a lower leverage. In the optimal regime, these gains dominate the welfare loss due to a lower capital stock, output, and lower

[6] A heterogeneous set of consumers setup is not necessary to analyze the banking industry, but is necessary to understand the welfare implication of change in capital regulation.

[7] There is a large literature that discusses the agency problem between ownership and management of firms. The agency problem can be modeled as a situation where the owners cannot observe the effort put in by managers Holmstrom and Tirole [1997], or the risk taken by managers. In fact, Gabaix and Landier [2006] show that increasing CEO pay is a natural outcome of attempts to address this agency problem.

[8] The rationale for considering the pre-crisis period (1995-2005) as benchmark is elaborated later in the introduction, and also in section 4.

[9] These are large welfare gains relative to those found in other studies in the literature. I argue that these gains are not implausible. For example, the cost of the 2008 financial crisis is of a similar order in permanent consumption terms, and when regulation can avoid the occurrence of such crises, the gains from regulation will be of this order. At the same time, the model's predicted welfare gain will probably be smaller once transition dynamics are taken into consideration.

wages. The welfare gain from a size-dependent regulation follows from three channels. First, allowing small banks to take more leverage enables them to potentially grow faster conditional on survival, leading to a *growth effect*. Second, although more leverage by small banks results in a higher exit rate among these institutions, these exits are by the less efficient banks, leading to a *cleansing effect*. As a corollary, entrant banks have higher mean efficiency relative to exited banks. Third, failures by large banks are more costly because these are more efficient in equilibrium and intermediate more capital. In fact, to replace failed large banks, entrants incur operating costs during the time they need to grow large. Therefore, tighter regulation for large banks renders them less prone to failure, leading to a *stabilization effect* of size-dependent policy.

Finally, with regards to banking industry dynamics, the model generates a steady state equilibrium with entry, exit and a non-degenerate size distribution. In the stationary equilibrium, bigger banks are more highly levered, and there is positive correlation between bank size and bank efficiency. This follows intuitively from the fact that more efficient banks have better future prospects, enabling them to adopt more leverage and potentially grow faster. The paper also documents some results pertaining to how banking industry dynamics respond to changes in regulation. For example, a tightening of regulation results in the following effects (a weakening of regulation results in qualitatively opposite effects). First, the bank failure rate is lower, which increases the mean age of banks. This shifts the size distribution of banks to the right, and also results in a larger equilibrium mass of banks. The latter effect is instrumental in driving up aggregate dividend payments, even though not necessarily for all banks. Second, since the mean leverage of banks is smaller, banks are better capitalized. In particular, the aggregate net worth of banks is larger. Third, tightening regulation results in a less concentrated industry as indicated by a lower Herfindahl Index.

**Data and Stylized facts** In order to characterize capital regulation quantitatively, I use data on US commercial and savings banks from Call and Thrift Financial Reports (CTR) and Historical Statistics on Banking (HSOB), both available with the Federal Deposit Insurance Corporation (FDIC). I draw household asset ownership and inequality data from Survey of Consumer Finances (SCF) and Organization for Economic Co-operation and Development (OECD) statistics, respectively. Aggregate macroeconomic data is obtained from the Federal Reserve Economic Data (FRED) database. I focus on the stationary equilibrium of the model, and estimate/calibrate model parameters to match select data moments.

I begin with some history of *capital regulation* of banks in the US, which has evolved considerably over time along with international minimum standards. In broad terms, regulation has become generally tighter and broader in scope over time.[10] In particular, the cap on leverage has decreased, and the minimum risk-weighted capital ratio requirement has increased (See the left panel of figure 2). Much of this tightening of regulation is motivated by the higher realized or estimated cost of a financial crisis.[11]

[10] For example, see Acharya [2011].

[11] For more detail, refer to Fender and Lewrick [2016], who review the costs of financial crises.

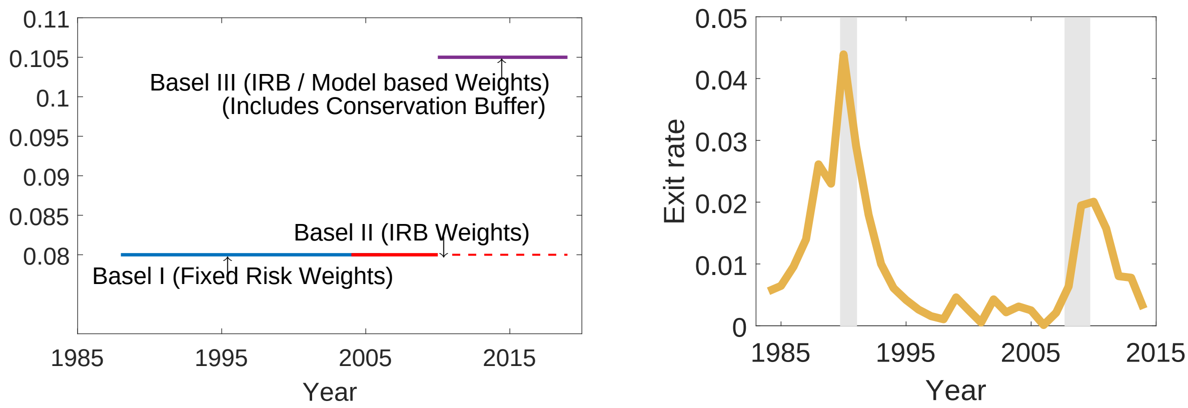


Figure 2: *(Left)* Historical Basel Total (tier1 + tier2) capital requirements *(Right)* US commercial and savings banks: Yearly failure rate (Source: FDIC Historical Statistics on Banking)

I now present some stylized facts observed in the data that I use to discipline my model. In the right panel of figure 2, I graph the yearly exit (failure) rate of US savings and commercial banks, which is one of the targets of my estimation procedure. The shaded regions approximately represent the two business cycle contractions that also featured financial sector crises: (i) the savings and loan crisis of the 1990s and (ii) the financial crisis of 2008. I take note of the following observations. First, the exit rate peaks during the savings and loans crisis and during the financial crisis, and is stable otherwise, i.e., during 1995-2005. Even structurally, the 1995-2005 period is less affected by regulatory changes in the banking industry, because most of the bank branching deregulation happened prior to 1995. In particular, although the Riegle-Neal Interstate Banking and Branching Efficiency Act was implemented in June 1997, US states were most active in removing geographic limits to banking prior to 1995 (Jayaratne and Strahan [1997]). Therefore, since 1995-2005 is a relatively stable period for banking industry dynamics, I use this period for calibration/estimation.

In Figure 3, I document the distribution of inflation-adjusted log net worth at three points in time during the last three decades. For any point in time during this period, the size distribution of bank net-worth in the United States can be fairly well approximated, based on the Kolmogorov-Smirnov Goodness of fit test, by a log-normal distribution. The fit for bank size distribution in 2000:Q4, for example, is shown in the right panel. Two remarks on the tail of the bank size distribution are as follows: (a) log-normal is not adequate to capture the tail, which is better described by a power law distribution, and (b) the model in this paper does not attempt to explain the existence of a long / heavy tail of the bank distribution - it rather focuses on specific percentiles of the size distribution of banks. Over time, the distribution is shifting to the right. However, the inflation-adjusted distributions at any two points in time during 1995-2005 are not statistically significantly different based on the two-sample Kolmogorov-Smirnov (KS) test. In this sense, the size distribution has remained fairly stable during the calibration period.

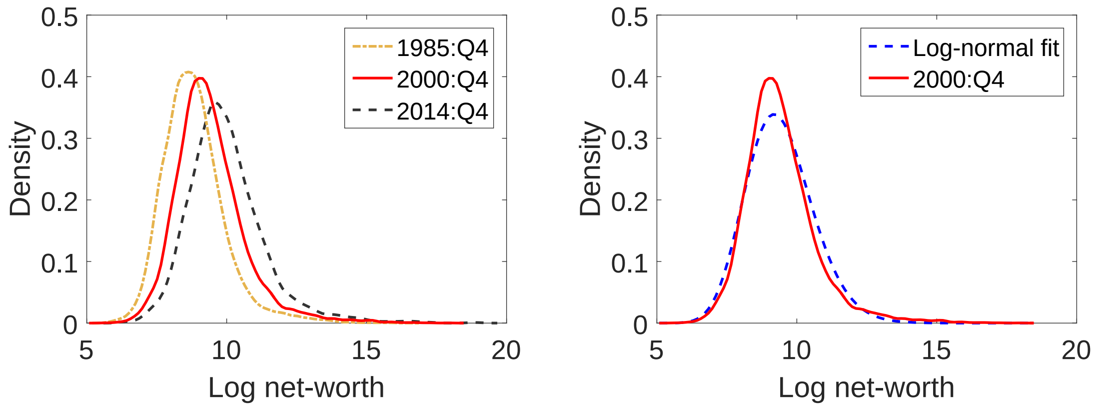


Figure 3: *(Left)* Kernel density estimate of log net-worth distribution; *(Right)* Log-normal fit to 2000Q4 data (source: FDIC Call and Thrift Financial Reports)

**Related Literature** The first strand of literature to which this paper is related, aims to understand how inefficiencies in the banking sector affect the real economy, and to identify which tools regulators could use to address these inefficiencies. The key contribution of this paper helps us understand how, in the presence of frictions, leverage choices affect output, investment, and the distribution of consumption through the banking industry dynamics channel. Few other papers have this feature. For example, Corbae and D'erasmo [2014] develop a quantitative model of an imperfectly competitive banking industry. Bank dynamics (entry and exit) are driven by shocks to borrowers' production technology, and the authors show how competition between banks affects interest rates and borrowers' decisions regarding loan amounts and levels of risk. Covas and Driscoll [2014] study capital and liquidity regulation in a dynamic banking sector, and its effect on loan supply, output, and interest rates. Nguyen [2014] looks into the welfare implications of capital requirements in the presence of moral hazard due to government bailouts. Christiano and Ikeda [2013] study the welfare gains possible due to leverage restrictions in a macroeconomic model with unobservable banker effort. Begenau [2014] shows that higher capital requirements might actually increase bank lending when households value safe and liquid assets, providing another rationale for tighter regulation. Other papers that discuss time-invariant or counter-cyclical capital regulation that is not dependent on bank sizes include Benes and Kumhof [2011], Zhu [2007], De Nicolo et al. [2014] and Christensen et al. [2011].

In this paper, I consider a competitive banking industry that is rationalized by a limited commitment issue between banks and its depositors. By assuming a stylized competitive banking industry instead of a more general imperfectly competitive industry, I am able to achieve tractability in studying endogenously determined size distribution. The fact that banking industry dynamics in general, and size distribution of banks in particular, respond to regulation in this model is the feature which allows me to study size-dependent regulation.

This paper is also related to the literature on macroeconomic models with financial sectors. For example, Gertler and Kiyotaki [2010] and Gertler and Karadi [2011] study unconventional monetary policy and other policy experiments in a macro model in which financial frictions arise due to a limited commitment issue. Adrian and Boyarchenko [2012] use a value-at-risk constraint to generate the endogenously determined borrowing constraints which play the central role in amplifying shocks. In this paper, the credit constraint arises because of a limited commitment issue between banks and their depositors. The main contribution here is to model a heterogeneous financial sector with endogenous entry and exit. In this sense, the paper offers an alternative channel, namely industry dynamics, through which to study the effect of the financial sector on macroeconomic aggregates.

This pursuit is related to another strand in the literature that studies the dynamics of non-financial firms. For example, Cooley and Quadrini [2001], Clementi and Hopenhayn [2006], Albuquerque and Hopenhayn [2004] study the dynamics of firms that borrow subject to borrowing constraints, but do not study the relationship between leverage and the risk of return on net worth or the risk of failure. The implication of borrowing on industry dynamics has been studied theoretically in Miao [2005], and empirically in, for example, Lang et al. [1996], but only in the context of non-financial firms. These studies are not directly applicable to financial firms (equivalently referred to as *banks* in this paper) as the capital structure choice of banks is characterized by higher leverage compared with non-financial firms (Kalemli-Ozcan et al. [2012]). In fact, as DeAngelo and Stulz [2013] show and Hanson et al. [2010] argue, high leverage is optimal for banks. This is because of the nature of the business of financial firms, wherein obtaining cheap funding is key to surviving the competition. Because of the fact that financial firms are more highly levered in general, the interaction between leverage and financial industry dynamics is also more pronounced, more so in the presence of financial frictions. In Section 2, I provide some intuitions for these mechanisms using simple examples. I incorporate these intuitions in my model to understand how financial leverage shapes financial industry dynamics in the presence of financial frictions. Some papers that follow a closely related route are Boyd and De Nicolo [2005], Corbae and D'erasmo [2013], and Corbae and D'erasmo [2014]. The former discusses endogenous bank failures in a static setting, where competition among banks reduces loan rates, incentivising borrowers to take on less risky projects and fail less often - which means fewer losses and failures for banks. The latter use models of imperfect competition to study a range of bank regulations, including capital requirements.

In studying the effect of industry dynamics on the real economy, in a non-financial context, Clementi and Palazzo [2015] investigate how firm entry and exit amplify aggregate fluctuations. The mechanism in their paper is, that after a positive aggregate productivity shock, the number of entrants increases. Of these entrants, those who survive, generate a wider and longer expansion. This insight is closely related to the intuition in my paper, according to which allowing smaller banks to take on more leverage leads to a selection effect such that surviving entrants are able to grow faster on an average.

The remainder of the paper is organized as follows. I present the model in Section 2, the equilibrium in Section 3, the data, estimation, and computations in Section 4, and counterfactual experiments in Section 5. Section 6 concludes.

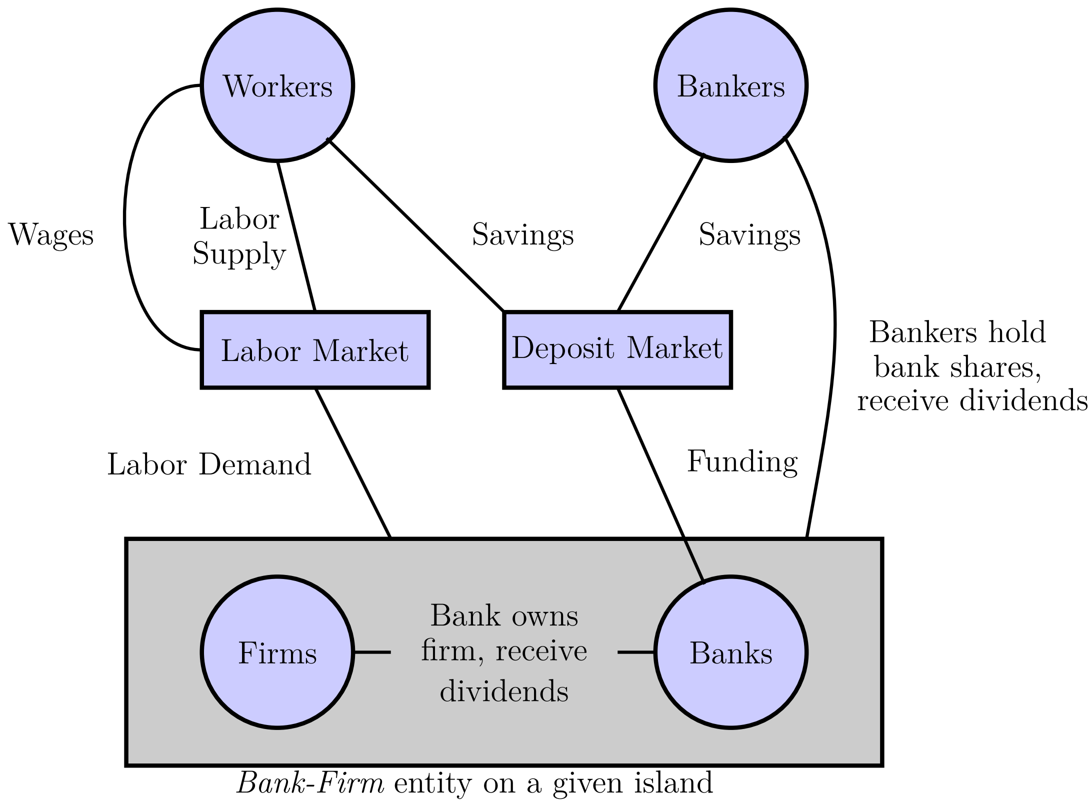


Figure 4: Physical setup of the economy

# 2 Model

**Physical setup** Time is discrete and the horizon is infinite. There is no aggregate uncertainty. The setting is that of an island economy, with the following elements: a unit mass of atomistic consumers, a representative capital goods producing firm (henceforth *capital firm*), a continuum of islands, each with a non-financial firm (henceforth *firms*) and a financial firm (henceforth *banks*), and a government. The mass of consumers is divided into a fraction $\lambda$ of bankers and a fraction $(1 - \lambda)$ of workers.[12] The bankers are best understood as *capitalists* who own the banks and the capital firm, while the workers provide labor to firms and also manage the banks. The capital firm converts consumption goods into capital. The firm on any given island is owned by the bank on that island. Firms produce consumption goods using labor and physical capital, where the latter input is owned by the firm itself, and is not mobile across islands. In order to purchase additional physical capital from capital firms, a firm must issue more shares to the *owner* bank on its island. Physical capital is subject to island-specific *capital quality shocks*. Since firms

[12]The division of consumers into workers and bankers is necessary to quantify distributional aspect of the welfare effect of change in capital regulation. In this sense, the division of consumers into workers and bankers is not necessary for the analysis of bank leverage and its relationship with banking industry dynamics.

are owned by banks, these shocks essentially affect the balance sheet of banks, and I will refer to these shocks as bank *efficiency* shocks. The bank is a financial intermediary that accepts deposits from consumers and uses these funds to purchase shares in the firm on its island. In what follows, since the bank owns the firm on its island, and since there are no contracting frictions assumed between the bank and the firm, I will sometimes refer to the *bank-firm* on an island as a joint entity that is owned by the bankers, and which *jointly* hires workers to produce output. The government runs a deposit insurance scheme by imposing a tax on bankers. The overview of the economy is presented in Figure 4. I will now describe each element of the model in detail.

## 2.1 Consumers

The consumers consist of a fraction $\lambda$ of bankers and a fraction $1-\lambda$ of workers.[13] For both types of consumers, I assume a constant relative risk aversion (CRRA) period utility function. $d$ denotes deposits placed in the banking system and $R$ is the gross interest rate on deposits. The deposit insurance system ensures that deposits are risk-free. $T$ denotes lump sum transfers from the government to the workers, while $\tau$ denotes deposit insurance tax on bankers. I assume that the only way consumers can save or invest in firms is via the deposit market. This can be motivated in numerous ways, one being that households do not have the expertise to monitor borrowers, which banks have. The problems of workers and bankers are given below, where the subscript $b$ indicates banker and the subscript $w$ indicates worker.

**Workers** All the workers are identical, and solve the same decision problem. Each atomistic worker enters into a contract with the *bank-firm*, and supplies one inelastic unit of labor in return for wage $W_t$. I assume that as part of the contract, the worker supplies labor to the firm, and also manages the bank that owns the firm. Although I do not explicitly model how managerial effort enters the *production function* of the bank, managerial effort involves setting bank objective and choosing bank decision variables. This assumption is one parsimonious way of creating a distinction between bank management and ownership by still having only two types of consumers.[14] The assumption implies that even though the worker's wage rate is a function of the marginal product of labor provided, compensation for managerial effort is implicitly contained in the contractual wage payment. In addition to wage income, the worker receives gross interest income on deposits, and receives transfers $T_t$ (or is required to pay taxes). The worker chooses levels of consumption and deposits in order to maximize the present discounted value of lifetime utility, subject to the budget constraint:

[13] The total mass of consumers is normalized to one for simplicity, and is without loss of generality.

[14] In other words, this is a simplifying assumption that spares the need to explicitly model a third type of consumer called bank manager who is risk-neutral and receives a share in bank profits, and also spares the need to explicitly model how managerial effort enters the production process of banks. This assumption is valid precisely because there are no contracting frictions between the bank and the firm, and therefore one could think about these two entities as one entity that performs the functions of the firm and the bank together.

$$\max_{c_{wt},d_{wt}} \quad E_t \sum_{\tau=0}^{\infty} \beta^\tau \left( \frac{c_{wt+\tau}^{1-\gamma}}{1-\gamma} \right)$$
$$c_{wt} + d_{wt} = R_{t-1} d_{wt-1} + W_t + T_t$$

**Bankers** Each atomistic banker holds shares in the banking sector, and receives stochastic dividends $e_t$ from banks. They save through the deposit market, and are subject to an *ad hoc* borrowing constraint $\underline{d}$ *a'la* Aiyagari [1994]. Bankers provide startup funding $n^e$ to new banks , and also incur the entry cost $c^e$ associated with new banks entering the banking industry. Therefore $M_t(n^e + c^e)$ is the total startup expense incurred by bankers, where $M_t$ is the mass of entrants. $\Pi_t$ is the dividends paid by capital firms. When a bank fails, it is replaced by a new bank, and the shareholders of the failed bank become the shareholders of the new bank. Bankers solve the following problem subject to the budget constraint:

$$\max_{c_{bt},d_{bt}\geq \underline{d}} \quad E_t \sum_{\tau=0}^{\infty} \beta^\tau \left( \frac{c_{bt+\tau}^{1-\gamma}}{1-\gamma} \right)$$
$$c_{bt} + d_{bt} = d_{bt-1} + (1-\tau)\Big((R_{t-1}-1)d_{bt-1} + e_{bt}\Big) - M_t(n^e + c^e) + \Pi_t$$

Bankers are subject to idiosyncratic income shocks because dividends are stochastic. In the economy without aggregate uncertainty, the problem of a banker can be cast recursively with state space $(e, d)$. The first dimension of the state space is current dividend income that follows a Markov process induced by the dynamics of the bank paying the dividend. The second dimension of the state space is the amount of savings (deposits) of the banker. Then the lifetime utility $U(e, d)$ of a banker is given as:

$$U(e,d) = \max_{c,d'\geq \underline{d}} \quad u(c) + \beta \sum_{e'} U(e', d')\pi(e, e')$$
$$c + d' = d + (1-\tau)\Big((R-1)d + e\Big) - M(n^e + c^e) + \Pi$$

The consumption and savings policy functions obtained by solving this problem induce a distribution of bankers over the state space, which will be denoted by $\mu_b(e, d)$.

## 2.2 Firms

Each island has a firm, and the firms are identical across islands. They operate a Cobb-Douglas production technology. The firm on island $j$ produces final goods $y_t^j$ by hiring labor $l_t^j$ in the national market at wage rate $W_t$, and by using the capital stock $k_t^j$ that it owns. As assumed above, the labor a firm uses, also provides managerial service to the bank that owns the firm. The idea behind this assumption is that firm production function requires managing the bank in order for production to be successful. The firms solves the following problem, where $A_t$ is aggregate productivity:

$$\max_{l_t^j} \quad A_t (k_t^j)^\alpha (l_t^j)^{1-\alpha} - W_t l_t^j$$

Firm shares are held by the bank on the same island, and each share represents claim to one unit of capital. I assume that there are no financial frictions in the contracting between firms and banks, and that firms pledge all returns to their shareholding banks. Consequently, the price of capital goods is equal to the price of shares in the firm. The firm can issue new shares to the bank on its own island to purchase additional capital. Dividends per unit of firm capital, which is also equal to dividends per share, is given as:

$$Z_t = \frac{y_t^j - W_t l_t^j}{k_t^j}$$

Note that dividend rate $Z_t$ is equalized across the island because labor is hired in the national market at a national wage rate $W_t$. Next, capital stock on a given island is subject to idiosyncratic capital quality shocks, and evolves as follows:

$$k_{t+1}^j = \left(i_t^j + (1-\delta)k_t^j\right)\psi_{t+1}^j$$

where $i_t^j$ is investment in period $t$ and $\psi_{t+1}^j$ is the *capital quality* shock. The shock effectively translates into a shock to return on financial assets (shares) a bank holds. In this sense, capital quality shock can be re-interpreted as the efficiency of a bank in ascertaining the quality of the capital whose claim it holds. In other words, $\psi_t^j$ is a bank's efficiency in managing its equity in the firm. The two interpretations, that $\psi$ is a capital quality shock and that $\psi$ is bank efficiency, are equivalent because banks are owners of the firm, and shares of the firm are one-to-one backed by the physical capital stock of the firm. In particular, if $s_t^j$ is the bank shareholding in the firms on island $j$ at the *end* of period $t$, then it is related to the beginning of a period physical capital stock $k_t^j$ as follows:

$$s_t^j = \left(i_t^j + (1-\delta)k_t^j\right)$$

where $i_t^j$ is investment and $\delta$ is the physical depreciation rate. The above equation implies that shareholding at the *end* of a period is equal to the depreciated capital stock (depreciated because current-period production takes place in the middle of the period) plus new investment undertaken in the current period. The above equation also implies the following simple relationship between capital stock and bank shareholding in firms (bank assets):

$$k_{t+1}^j = \psi_{t+1}^j s_t^j$$

In this sense, $s_t^j$ is a determinant of the expected capital stock at the beginning of period $t+1$, which is $E_t[\psi_{t+1}^j|\psi_t^j]s_t^j$.

So far I have referred to firms by their island indices. However, since (a) firms are identical in their production technology, (b) they operate a CRS Cobb-Douglas production technology, and (c) hire labor in a national market, I can refer to the economy-wide *representative* firm. In other words, I can represent aggregate output $Y_t$ as a function of aggregate capital $K_t$, and not necessarily as a function of the distribution of capital stocks across islands. This also means that I can refer to the representative firm as one that issues equity in the national market and pays a national dividend rate $Z_t$. This interpretation makes computations easier and makes the analysis more tractable, and is therefore employed in the rest of this paper.

## 2.3 Capital Firm

The capital firm produces capital goods from consumption goods, but is subject to convex investment adjustment costs. It maximizes the present discounted value of profits accruing to the consumers (bankers) that own the capital firm:

$$\max_{I_t} \quad E_t \sum_{\tau=t}^{\infty} \Lambda_{t,\tau} \Big( Q_t I_\tau - [1 + f(I_\tau / I_{\tau-1})] I_\tau \Big)$$

where $f(x)$ is a convex cost function and $\Lambda_{t,\tau}$ is the stochastic discount factor between time $t$ and $\tau$. $Q_t$ is the price of capital goods.

## 2.4 Banks

There is a continuum of banks. The banks are owned by the bankers and managed by managers (workers). The objective of an atomistic bank $j$ at time $t$ is to maximize the present discounted value of the dividends $e^j_\tau$ it pays to its shareholders:

$$E_t \sum_{\tau=t}^{\infty} \beta^\tau e^j_{t+\tau}$$

As indicated before, the objective of the bank is set by the manager, and we assume the following friction between bank ownership and bank management. The discount factor employed by the bank (manager) in its objective function is assumed to be $\beta$ – which is independent of the bank owner's (banker's) stochastic discount factor. The motivation behind this assumption is that bank managers care about maximizing profits and dividends, but do not care about the inter-temporal smoothing of dividends, which matters for bankers since they are risk-averse. As a result of this assumption, the bank manager does not need to keep track of the distribution of the stochastic discount factor of bank shareholders, and takes more risk than what is socially efficient.[15]

The time-line for bank decisions is as follows. At the beginning of any period $t$, the bank has equity $s^j_{t-1}$ in the firm (bank assets), and deposits $d^j_{t-1}$ from consumers (bank liabilities). Since the bank owns the firms, and firm capital is subject to idiosyncratic shocks, at the beginning of every period, bank assets are subject to idiosyncratic shocks. As discussed earlier, this shock $\psi^j_t$ is interpreted as a bank's idiosyncratic efficiency in managing its ownership in the firm. This shock captures the notion that a bank that is more *efficient* today makes better asset investment decisions, and hence has a better expected future return on its assets. I assume that $\psi^j_t$ follows a first-order Markov process with transition function $F(\psi^j_t, \psi^j_{t+1})$, which is the conditional distribution of bank efficiency

[15] Another interpretation of this assumption is that the bank manager uses a *representative* discount factor $\beta$ instead of using the distribution of discount factors of its shareholders, which again points to a friction between bank ownership and bank management.

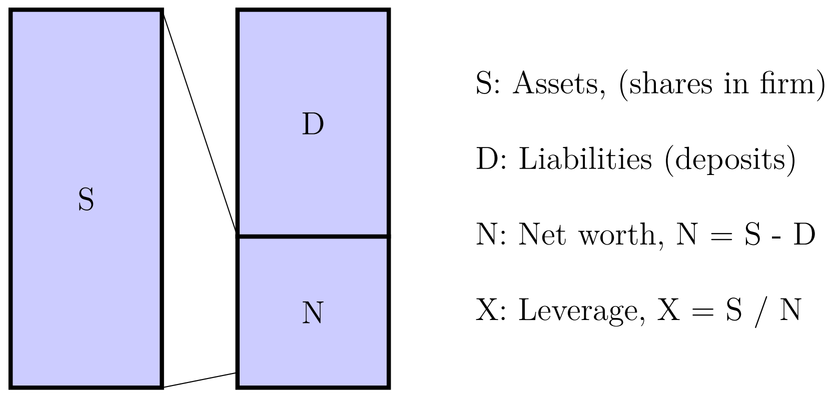


Figure 5: Stylized Bank Balance Sheet

in the next period given bank efficiency in the current period. Later, I will impose more structure on $F$ such that it captures the notion of first-order stochastic dominance (FOSD) in current efficiency: banks with high efficiency today enjoy *better* distribution of efficiencies tomorrow.

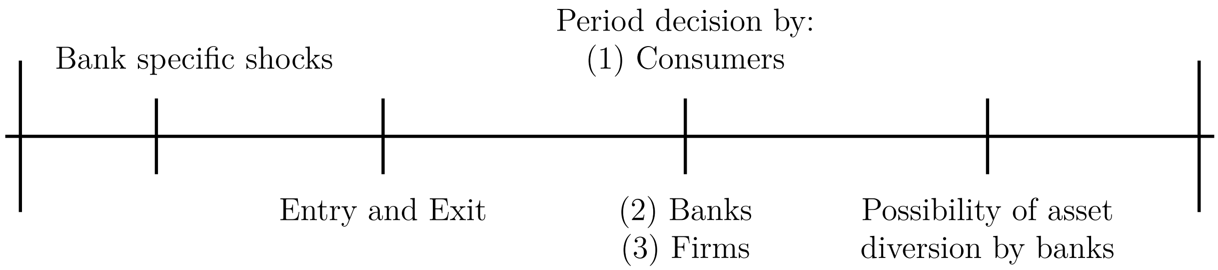


Figure 6: Timeline

**Bank failures** After banks acknowledge efficiency shocks at the beginning of the period, a bank is considered to have failed if its *beginning-of-period liabilities* are at least as large as its *beginning-of-period assets:* [16]

$$\psi_{t+1}^{j} Q_t s_t^j - R_t d_t^j \leq 0$$

The assets of the failed bank are liquidated and used to service as much of the total

[16]Note that beginning-of-period assets are valued at the asset prices of the previous period. The interpretation assumes that such a capital quality shock is realized during the transition from one period to another, and that depositors will make a run on the bank if the bank is insolvent as per the valuation of the bank at the very beginning of the period. Valuation at the beginning of the period is based on previous-period asset prices because current-period asset prices are determined later in the period, only after depositors have ensured that insolvent banks have exited. This intra-period liquidity issue can be motivated by bank creditors calling on the bank (or not willing to roll over the debt) before assets payoff, an issue that was key during the 2008 financial crisis.

amount of liabilities (deposits) as possible. The remaining liabilities are paid off by the deposit insurance system which is funded by a tax on bankers. Once the exiting banks exit, a mass of new banks pay a fixed entry cost $c^e$ and enter the market with start-up capital $n^e$. Upon entry, the new banks draw random idiosyncratic efficiencies from a stationary distribution $G(.)$, and begin operations similar to an incumbent.

**Bank balance sheet** Following bank exit and entry, firms produce output while demanding labor from workers. Firms pay dividends $Z_t$ per unit of equity to banks. The banks decide assets $s_t^j$ and liabilities $d_t^j$ for the current period, given dividends $Z_t$ paid by firms, and taking market prices $Q_t$ and $R_{t-1}$ as given. Banks also pay an operating cost $c(N_t^j)$ as a function of their *end-of-period net worth* $N_t^j$, which for the rest of this paper is the measure of bank size. Banks are subject to the following balance sheet and flow of fund constraints:

$$N_t^j + d_t^j = Q_t s_t^j + e_t^j + c(N_t^j)$$

$$N_t^j = (Z_t + (1-\delta)Q_t)\psi_t^j s_{t-1}^j - R_{t-1}d_{t-1}^j$$

The first equation is the balance sheet identity, which states that the sum of a bank's net worth and liabilities are used in three ways: (a) to purchase assets (b) to pay dividends and (c) to pay operating costs. The second equation is the law of motion for $N_t^j$. The first term consists of dividend gain component $Z_t$ and capital gain component $(1-\delta)Q_t$, adjusted by efficiency shock $\psi_t^j$, per unit of asset holding $s_{t-1}^j$ in the previous period. The second term consists of the cost of liabilities.

**Operating Costs** I assume that banks are subject to a size-dependent convex operating cost. I motivate operating costs as an increasingly higher cost that a bank has to incur in order to operate at larger scales. Expenses that form part of this cost include lobbying, advertising and brand-building efforts, in addition to regular infrastructure and salary expenses. An alternative motivation for convex operating costs comes from the firm dynamics literature, where the 'span of control' argument is used to represent the increasingly limited ability of a manager to manage a firm at larger scales. For banks, this is intuitive because larger banks operate in more diversified and complicated financial markets. Operating costs are critical in this model for showing that (a) a solution exists for the bank's problem, and that (b) an endogenously determined size distribution exists.

However we interpret convex operating costs, they imply a decreasing returns-to-scale (DRS) technology for banks, which is key to the existence of a bank size distribution.[17] Otherwise, with CRS technology the size distribution of banks is indeterminate. A DRS technology is also critical for proving the existence of a unique solution to the bank's problem.[18] I assume the following operating cost function $c(n) = cn^\zeta$, where $c$ is the level

[17] Operating costs are one way to introduce DRS to a bank's technology. I considered dividend adjustment costs, exogenous random destruction of banks, or operating costs that are convex in a bank's assets as alternatives to introduce the necessary curvature to the bank's problem. However, because of analytic convenience, and to keep the trade-off with respect to bank leverage decisions clear, I chose the size-dependent operating cost approach.

[18] A DRS technology is not necessary but sufficient to show that the endogenous state variable in the

and $\zeta$ is the exponent parameter of the cost function, which are to be estimated from the data.

**Limited commitment** In order to introduce a rationale for the banking sector in this model, I introduce the following limited commitment issue between banks and their creditors.[19] In particular, I assume that the banker (bank owner) can divert a fraction of bank assets *a'la* Gertler and Kiyotaki [2010] for personal consumption.[20] The possibility of asset diversion induces a credit constraint on banks:

$$V_t^j \geq \theta Q_t s_t^j$$

which means that the value to a banker of diverting assets must not be greater than the end of period $t$ charter (continuation) value of the bank. This constraint is a manifestation of the fact that creditors, depositors in this case, are aware of the asset diversion possibility, and therefore lend only up to the point at which bankers do not have the incentive to divert assets for personal consumption.

**Leverage** I now introduce the notion of leverage, since leverage decisions are a key focus in this paper.[21] First, I define *effective net worth* $n_t^j$ as net worth minus operating costs: $n_t^j = N_t^j - c(N_t^j)$. Next, I define dividend ratio $h_t^j$ as the fraction of effective net worth paid out as dividends: $h_t^j = e_t^j / n_t^j$. Finally, I define leverage $x_t^j$ as the ratio of assets to residual net worth, i.e., net worth after paying dividends from effective net worth:

$$Q_t s_t^j := (1 - h_t^j) n_t^j x_t^j$$

$$\text{Balance sheet constraint} \implies d_t^j = (1 - h_t^j)(x_t^j - 1) n_t^j$$

The flow-of-funds constraint, written in terms of effective net worth, is as follows, where $g(\omega) = \omega - c(\omega)$, is the function that gives effective net worth as a function of raw net worth:

$$n_{t+1}^j = g\Big( \big(Z_{t+1} + (1-\delta) Q_{t+1}\big) \psi_{t+1}^j s_t^j - R_t d_t^j \Big)$$

$$\implies n_{t+1}^j = g\Bigg( \underbrace{\Big(\frac{Z_{t+1} + (1-\delta) Q_{t+1}}{Q_t}\Big)}_{\substack{R_{kt+1}\text{: Return on} \\ \text{assets}}} \psi_{t+1}^j (1 - h_t^j) n_t^j x_t^j - R_t (1 - h_t^j)(x_t^j - 1) n_t^j \Bigg)$$

dynamic programming setup of the bank problem is bounded, paving way for existence of a unique solution to the bank's problem.

[19] In absence of frictions, the banking sector would be inconsequential, and the model would be equivalent to one without banks.

[20] Other approaches to introducing friction in a macroeconomic model with banks are unobservable effort *a'la* Christiano and Ikeda [2013], or the moral hazard issue due to government bailouts *a'la* Nguyen [2014].

[21] In subsection 8.1 in the appendix, I present simple examples to provide intuition for how leverage is related to bank dynamics.

$$\implies n_{t+1}^j = g\bigg(\Big(R_{kt+1}\psi_{t+1}^j - R_t\Big)x_t^j + R_t)(1-h_t^j)n_t^j\bigg)$$

Then, the cutoff of $\psi_t$ that determines whether a bank fails in the next period, is a function of current-period leverage, and is given as follows:

$$\frac{R_t d_t^j}{Q_t s_t^j} = \frac{R_t(x_t^j - 1)}{x_t^j} =: \psi^*(x_t^j, R_t)$$

Also, the borrowing constraint on banks is written in terms of leverage as:

$$V_t^j \geq \theta(1-h_t^j)n_t^j x_t^j$$

**Capital Regulation** Finally, I introduce capital regulation to address the inefficiency in the model that arises due to the wedge between bank management and ownership. Capital regulation is subject to several trade-offs that are discussed later in the context of optimal capital regulation. Capital regulation is defined as a bank size-dependent constraint on the maximum leverage a bank might assume, where size-dependence of regulation is a key feature of this paper:

$$x_t^j \leq \chi(n_t^j)$$

**Recursive formulation** I cast the bank's problem recursively. The state vector for any bank $j$ at the end of period $t$ is $(\psi_t^j, s_{t-1}^j, d_{t-1}^j; O_t)$, where $O_t$ is the aggregate state vector. In the general model, $O_t$ comprises all market prices, all aggregate shocks, and the distribution of banks. In the baseline model with no aggregate uncertainty, $O_t$ comprises of market prices only. In either version, since the bank takes market prices $Q_t, R_{t-1}$ and dividends $Z_t$ as given, it follows from the flow of funds constraint that $n_t^j$ is a sufficient statistic for $s_{t-1}^j$ and $d_{t-1}^j$. Consequently, a smaller and equivalent state vector for any bank $j$ is $(\psi_t^j, n_t^j; O_t)$. The choice variables for any bank $j$ are $x_t^j$ and $h_t^j$. In writing the recursive problem of an arbitrary bank $j$, I omit the time subscript and bank specific superscript except when necessary. I use a -1 subscript to denote the previous period and a prime superscript to denote the next period. The recursive problem of an incumbent bank is given as follows:

$$V(\psi, n; O) = \max_{x\in[1,\chi(n)], h\in[0,1]} \left(hn + E\Lambda' \int_{\frac{R(x-1)}{x}}^{\overline{\psi}} V(\psi', n'; O')F(\psi, d\psi')\right)$$

$$n' = g\bigg(\Big((R_k'\psi' - R)x + R\Big)(1-h)n\bigg)$$

$$V(\psi, n; O) \geq \theta(1-h)nx$$

The problem indicates that current period value of a bank is the sum of its current period dividend payment and its discounted continuation value. The integral captures the fact that if next period realization of shock $\psi'$ is less than $R(x-1)/x$, then the bank's *beginning* of period net worth is non-positive and the bank fails, which means zero continuation value. Net worth evolves as per a technology $g(.)$ banks posses, which is concave because $g(\omega) = \omega - c(\omega)$ and c(.) is convex. The incentive constraint is based on the current period value function, and imposes an endogenously determined constraint on the choice set. The choice set is bounded in the following sense. $[0, 1]$ is the bounded choice set for $h$ because limited liability[22] $\implies h \geq 0$ and non-negative net worth $\implies h \leq 1$. Leverage, $x$, is bounded below by 1 since deposits are non-negative, and it is bounded above by $\chi(n)$, which is the capital regulation that restricts banks from taking on too much leverage. For the benchmark model that follows and its calibration, I consider a degenerate regulation $\chi(n) = \hat{\chi}$, as is the case in the data for the chosen calibration period. I will now discuss some properties of the bank's problem.

**Proposition 1.** *Given current bank efficiency $\psi$, bank failure probability is strictly increasing in leverage $x$, where the bank solvency probability for a bank characterized by $(\psi, n)$ is given as*

$$p(\psi, n) = \int_{R(x-1)/x}^{\overline{\psi}} F(\psi, \partial\psi')$$

*Proof.* The interval of integration is the bank's endogenously determined region of solvency. The lower limit of this integral is the failure cutoff $\psi^*(x, R) = R(x-1)/x$ i.e. the bank fails if it realizes $\psi' < \psi^*(x)$. Then the proposition follows directly from the fact that $\partial\psi^*(x)/\partial(x) = R/x^2 > 0$. Note, however, that higher leverage does not necessarily correspond to a higher probability of failure, which leads me to the next proposition. ∎

**Proposition 2.** *Assume that $F(\psi, \psi')$ is first-order stochastically increasing in $\psi$. Then, given a bank with current leverage $x$, the probability of failure is weakly decreasing in current efficiency $\psi$.*

*Proof.* This proposition follows directly from the definition of FOSD. If $\psi_1 > \psi_2$, then $F(\psi_1, \psi') \leq F(\psi_2, \psi') \; \forall \; \psi'$ and hence

$$p(\psi_1, n) = \int_{R(x-1)/x}^{\overline{\psi}} F(\psi_1, \partial\psi') \leq \int_{R(x-1)/x}^{\overline{\psi}} F(\psi_2, \partial\psi') = p(\psi_2, n)$$

Intuitively, when a bank's efficiency increases *ceteris paribus*, the distribution of next period efficiency is shifted to the right, and hence reducing the probability of a bad *(where bad depends on the leverage)* shock. In light of these propositions, the trade-off a bank faces with respect to leverage and dividend decisions can be summarized below. ∎

1. The trade-off with respect to leverage is as follows. Increasing leverage increases the variance of *total* premium $(R'_k\psi' - R)x$ a bank earns on its effective post dividend net worth. Depending on the distribution of $\psi'$, and as long as expected premium is positive, increasing leverage raises continuation value. However, increasing leverage also reduces the probability of solvency (increases the probability of failure).

[22] No borrowing from equity holders

2. The trade-off with respect to dividends is as follows. Increasing dividend ratio $h$ increases the current value for bank equity holders, but reduces $n'$. As I prove in the appendix, $V$ is increasing in net worth, which means that increasing $h$ decreases continuation value by reducing retained earnings.

**Entrant bank problem** There are infinitely many potential entrants and there is free entry. Entrants pay a cost $c^e$ before they can enter the market. After entry, they draw their idiosyncratic efficiency from an invariant distribution G. All entrants start operations with a fixed start-up capital $n^e$ provided to them by bankers.[23] Since, after entry, an entrant behaves exactly like an incumbent, the expected value of a potential entrant can be computed using the value function of incumbents as follows:

$$EV^e = \int_{\underline{\psi}}^{\overline{\psi}} V(\psi, n^e) dG(\psi)$$

Therefore, a potential entrant will enter the market if and only if:

$$EV^e - c^e \geq 0$$

But then free entry implies that the net value of entry should not be positive, or else an infinite mass of entrants will enter the market, i.e.:

$$EV^e - c^e \leq 0$$

The two conditions above imply that if the mass $M$ of entrants is strictly positive, it must be true that $EV^e = c^e$. And, if the mass of entrants is zero, it must be the case that $EV^e < c^e$. Now I establish some analytical properties of the bank's problem.

## 2.5 Analytical properties

**Proposition 3.** *If $R_k > R$, then the Bellman Equation specifying the bank's problem is a Contraction Map, and therefore there exists a unique value function corresponding to the bank's problem.*

I prove that the Blackwell Conditions are verified, and ensure they are sufficient for a contraction map in the set up of my problem. *(Proof in appendix).*

**Proposition 4.** *Suppose there are no credit frictions. Suppose $(\psi, n) \ni h(\psi, n) > 0$. Then:*

*1.* $x_1(\psi, n) > 0$

[23] The assumption that banks enter with fixed start-up capital provided by bankers is made to keep the model simple. This assumption essentially abstracts the model away from new equity issuance decisions on part of the banks and equity purchase decisions on part of the bankers. The motivation behind inducing bankers to provide additional capital (start-up capital) to entrants and not to incumbents follows from the concavity of the value function in net worth. It is due to this concavity that marginal value of bank capital is greater for smaller banks than for larger banks. Although in this paper I do not solve for $n^e$ as an equilibrium outcome of the model where the banker's return on providing start-up capital is equalized to banker's return on deposits, I estimate $n^e$ from the data.

2. $h_1(\psi, n) < 0$

3. $x_2(\psi, n) = 0$

4. $h_2(\psi, n) > 0$

The first proposition states that if a bank is more efficient, it takes on more leverage. This follows from the assumption that $F(\psi, \psi')$ is FOSD in $\psi$, which means that a more efficient bank enjoys better distribution of $\psi'$, the next period efficiency. In other words, the possibility of being more efficient in the next period allows the bank to bump up expected net worth by taking on more leverage, with relatively less reduction in the probability of solvency.

The second proposition states that more efficient banks find it optimal to retain earnings by paying lower dividends, precisely because they face better prospects in the next period. The third proposition follows directly from the feature of the model in virtue of which the trade-off with respect to leverage does not depend on size. The fourth proposition states that larger banks pay more dividends. This follows from the concavity of the value function in size. For larger banks, the marginal value of retaining their earnings is relatively lower, rendering it optimal for larger banks to pay more dividends in the current period. (*Outline of proof in appendix*).

## 2.6 Bank distribution

Let $X = [0, 1] \times [0, \bar{n}]$ be the state space of the model. Let $\mathcal{A} = [0, 1] \times [1, \hat{\chi}]$ be the action space of the model. Then, $(x, h) : X \to \mathcal{Y}$ are the policy functions. With this notation, we define the distribution of banks over the state space. The distribution of banks in the *middle* of period $t$ (i.e., after bank entry and exit has occurred but before period decisions have been made), is given by a measure $\mu : \mathcal{B}(X) \to \Re^+$, where $\mathcal{B}$ indicates the set of Borel subsets of $X$.[24] Abusing notation for $\mu$ in the obvious sense, let $\mu_t(\Psi, N)$ denote the time $t$ distribution of banks with idiosyncratic shock $\underline{\psi} \leq \psi \leq \Psi$ and net worth $0 \leq n \leq N$. Then the evolution of bank distribution is given by the following expression:

$$\mu_{t+1}(\Psi, N) = M_{t+1}\mathbb{1}(n^e \leq N) \int_{\underline{\psi}}^{\Psi} dG(\psi)$$
$$+ \int_0^{\bar{n}} \int_{\underline{\psi}}^{\overline{\psi}} \left( \int_{\underline{\psi}}^{\Psi} \mathbb{1}\left( g\Big(\Big((R'_k\psi' - R)x + R\Big)(1-h)n\Big) \leq N \right) \right.$$
$$\left. \mathbb{1}\left( \frac{R(x(\psi, n) - 1)}{x(\psi, n)} \leq \psi' \right) F(\psi, d\psi') \right) d\mu_t(\psi, n)$$

The first term indicates mass $M_{t+1}$ of entrant banks that enter in period $(t+1)$, that have start-up capital less then $N$ and a draw of efficiency less than $\Psi$. Basically, the first term represents the mass of entrants who end up in the following subset of the state space:

[24] Not necessarily a probability measure

$[\underline{\psi}, \psi] \times [0, N]$. The second term represents the flow of incumbent banks into this subset of the state space, net of those which fail. Basically, we "count" all the banks from period $t$ that, while following their respective policy functions, did not fail and landed in the following subset of the state space in period $(t+1)$: $[\underline{\psi}, \psi] \times [0, N]$.

The evolution of bank distribution can be written concisely using the notion of the transition function $W$ on state space $X$. The transition function $W$ is a map from $X \times \mathcal{B}(X) \rightarrow \mathcal{R}^+$, where $W\Big((\psi, n), (\Psi, N)\Big)$ denotes the probability of transition from point $(\psi, n)$ in the state space to the following subset of the state space in the next period: $(\Psi, N) = \{(\psi, n) : \psi \leq \Psi, n \leq N\}$. The transition function and associated distribution evolution equation are as follows:

$$W\Big((\psi, n), (\Psi, N)\Big) = \int_{\underline{\psi}}^{\Psi} \mathbb{1}\bigg(g\Big(\Big((R_k'\psi' - R)x + R\Big)(1-h)n\Big) \leq N\bigg)$$
$$\mathbb{1}\left(\frac{R(x(\psi, n) - 1)}{x(\psi, n)} \leq \psi'\right) F(\psi, \partial\psi')$$
$$\mu_{t+1}(\Psi, N) = M_{t+1}\mathbb{1}(n^e \leq N) \int_{\underline{\psi}}^{\Psi} \partial G(\psi) + \int_X W\Big((\psi, n), (\Psi, N)\Big)\partial\mu_t(\psi, n)$$

The latter can be concisely represented using the *distribution evolution operator* $T$ as $\mu_{t+1}(\Psi, N) = T(\mu_t, M_{t+1}, O_t)$, where $O_t$ is the aggregate state vector comprising equilibrium prices and quantities.

**Induced Dividend payments** Each banker holds shares in the banking industry. Shareholding is modeled as a continuous bijection from the *exogenously specified* mass of bankers to the *endogenously determined* mass of banks. The dividend income stream of a given banker is determined by the dividend payment stream of the banks whose shares the banker holds. In this sense, the bank dividend policy $h(\psi, n)$ and the bank transition function $W\Big((\psi, n), (\Psi, N)\Big)$ induce a dividend distribution transition function $\Pi(e, e') = \Pr(\text{current dividend} = \text{e, future dividend} \leq e')$, where $e(\psi, n) = nh(\psi, n)$. This dividend income transition process is what governs the stream of idiosyncratic dividend income realization for bankers, and also governs a banker's expectation regarding future dividend income. The dividend distribution transition function $\Pi(e, e')$ is computed as follows, from which the dividend transition function $\pi(e, e') = \Pr(\text{current dividend} = \text{e, future dividend} = e')$ can be computed as the marginal of the former:[25]

$$X' = \{(\psi, n) \in X \mid nh(\psi, n) = e\}$$
$$X'' = \{(\psi, n) \in X \mid nh(\psi, n) \leq e'\}$$
$$\Pi(e, e') = \int_{X'} \int_{X''} W\Big((\psi, n), \partial(\Psi, N)\Big)\partial(\psi, n)$$

[25]For the computations, I discretize the state space for dividend payments, and compute the Markov Transition matrix following the approach laid out.

# 3 Stationary General Equilibrium

I now focus on long-run stationary general equilibrium (SGE) in the model. Essentially, I consider the case of no aggregate uncertainty, and establish that equilibrium exists where the aggregate state vector and the bank distribution are time invariant. Although the long run equilibrium entails invariant aggregates, there are interesting dynamics at the individual bank and banker levels to be studied. I begin with the definition of SGE, followed by a discussion of its existence.

**Definition** A Stationary general equilibrium consists of (i) bank value function $V(\psi, n)$ and (ii) bank policy functions $x(\psi, n), h(\psi, n)$ (iii) invariant bank distribution $\mu(\psi, n)$ of net worth and efficiency shocks (iv) entrant mass $M$, (v) banker value function $U(e, d)$ and banker policy functions $d^b(e, d)$ and $c^b(e, d)$ (vi) banker distribution $\mu^b(h, d)$ (vii) worker consumption $c^w$ and savings $d^w$(viii) aggregate consumption $C$, aggregate capital $K$, aggregate deposits $D$, aggregate labor demand $L$, Taxes $T$, Dividends $H$ and (ix) wage rate $W$, interest rate $R$ and dividends $Z$ (these prices form the aggregate state vector $O$) such that given prices:

1. $V(\psi, n)$, $x(\psi, n)$ and $h(\psi, n)$ solve the bank's problem given prices
2. Free entry condition is satisfied: $EV^e \leq c^e$
3. Labor market clears at wage rate $W$: $L = 1 - \lambda$
4. Deposit market clears at interest rate $R$:
$$\int_0^{\bar{n}} \int_{\underline{\psi}}^{\overline{\psi}} d(\psi, n) \partial\mu(\psi, n) = (1 - \lambda) d^w + \lambda \int_{\underline{d}}^{\bar{d}} \int_0^{\bar{e}} d^b(e, d) \partial\mu^b(e, d);$$
5. Asset market clears at price $Q$:
$$K = \int_0^{\bar{n}} \int_{\underline{\psi}}^{\overline{\psi}} \left( \int_{\underline{\psi}}^{\overline{\psi}} \psi' F(\psi, d\psi') \right) s(\psi, n) \partial\mu(\psi, n) = S$$
6. Consumption goods market clears, where effective output is net of entry and operating costs:
$$Y = K^\alpha L^{1-\alpha} = C + I + M_t c^e + c^o;$$
$$c^o = \int_0^{\bar{n}} \int_{\underline{\psi}}^{\overline{\psi}} c(n) \partial\mu(\psi, n);$$
$$C = (1 - \lambda) c^w + \lambda \int_{\underline{d}}^{\bar{d}} \int_0^{\bar{e}} c^b(e, d) \partial\mu^b(h, d);$$
where the interpretation of aggregate investment $I$ in SGE is that it includes 're-covery' of lost capital due to fixed capital depreciation at rate $\delta$ and idiosyncratic stochastic capital quality shocks, such that aggregate capital stock remains constant

7. The bank distribution is a fixed point of the distribution evolution operator: $\mu = T(\mu, M)$

8. $c^b(h,d), d^b(h,d)$ solve the banker's problem; $c^w, d^w$ solve the worker's problem

9. Government budget constraint is satisfied: Taxes, net of transfers, equal the shortfall in liabilities of failing banks:

$$\tau\left(\lambda \int_{\underline{d}}^{\overline{d}} \int_0^{\bar{e}} (e + d^b(e,d)(R-1))\partial\mu^b(e,d)\right) - T =$$

$$-\int_0^{\bar{n}} \int_{\underline{\psi}}^{\overline{\psi}} \left(\int_{\underline{\psi}}^{\overline{\psi}} min(0, \psi' s(\psi,n) - Rd(\psi,n))F(\psi, d\psi')\right)\partial\mu(\psi,n)$$

## 3.1 Existence and solution

I am interested in the SGE with positive entry ($M^* > 0$), the existence of which depends on the fixed parameters of the model, especially the entry cost. To this end, I first solve for equilibrium quantities that can be analytically pinned down. The first-order conditions from the representative banker's problem imply that $R_t = R^* = 1/\beta$ since consumption is a constant $C^*$ in SGE. Similarly, since investment is constant from one period to another, there are no investment adjustment costs, and the capital firm's problem implies that $Q_t = Q^* = 1$ and $\Pi_t = \Pi^* = 0$. Next, deriving the equation from the firm's problem, which takes stock of capital and prices as given, aggregate labor demand $L^*$ is given by:

$$L^* = \left(A\frac{(1-\alpha)}{W^*}\right)^{1/\alpha} K^* \tag{3.1.1}$$

Dividends are then given by the following relationship:

$$Z^* = \frac{Y^* - W^*L^*}{K^*} = \frac{\alpha Y^*}{K^*} = \alpha A(K^*)^{\alpha-1}(L^*)^{1-\alpha} \implies K^* = (\alpha A/Z)^{1/1-\alpha}L^* \tag{3.1.2}$$

Since labor supply equals $1-\lambda$, I can use aggregate labor demand to pin down wages as a function of dividends $Z^*$:

$$W^* = (1-\alpha)A^{1/1-\alpha}\left(\frac{\alpha}{Z^*}\right)^{\alpha/1-\alpha}$$

I now turn to the bank's problem. First I note that if $Z, Q, R$ are known, then the bank's problem can be solved uniquely. Since $R, Q$ are already pinned down, $Z$ is the only equilibrium quantity to be solved for. Using the free entry condition, I can pin down $Z$ as shown in the following proposition.

**Proposition 5.** *Given an entry cost $c^e > n^e$, equilibrium firm dividend rate $Z^*$ is uniquely determined.*

*Proof.* Suppose $Z$ is known, then $R_k$ is also known, and the bank's problem can be solved. From a routine application of the envelope theorem to the Bellman equation for the bank's problem, and since the value function is increasing in net worth, I learn that the value function is strictly increasing in $Z$. I can then define:

$$g(Z) = \int_{\underline{\psi}}^{\overline{\psi}} V(\psi, n^e; Z) dG(\psi)$$

which is the expected value of entry as a function of dividend return $Z$ on bank assets, where $g'(z) > 0$. Now if $Z \to 0 \implies V(\psi, n^e; Z) \to n^e \implies g(Z) \to n^e$, and if $Z \to \infty \implies g(Z) \to \infty$. Then, by the intermediate value theorem (IVT), I know that:

$$\exists\ Z^* \ni g(Z^*) = c^e$$

■

Once $Z^*$ is solved for a given set of parameters of the model, I can solve the bank's problem uniquely, and also pin down $w^*$ using Equation 3.1.2. Then, given an M, and given market prices, I can solve for the fixed point $\mu^*$ of $T$, eventually pinning down all other quantities. Essentially, an equilibrium can be uniquely described by $(Z^*, M^*, \mu^*)$. Based on these insights, I propose the following solution algorithm:

1. Find $Z^*$ using free entry:

$$\int_{\underline{\psi}}^{\overline{\psi}} V(\psi, n^e; Z^*) dG(\psi) = c^e;$$

2. Obtain the bank's value and policy functions, and $W^*$

3. Find that $L^* = 1 - \lambda$ from the labor market clearing condition, and obtain $K^*$ from 3.1.1; also obtain output $Y^*$ from production function

4. Assume $M$, and obtain the steady state bank distribution $\mu$ corresponding to $M$ by iterating on $T$

5. Pin down $M^*$ using the asset market clearing condition; Obtain $\mu^*$

6. Using the bank distribution, obtain aggregates $D, T$

7. Back out the transition matrix for dividends

8. Solve the banker's income fluctuation problem, and the worker's problem

The idea for steps 5 is as follows. Given a mass of entrants M, the bank distribution implies aggregate assets. If there are too many entrants in the equilibrium, then the aggregate assets are higher relative to $K^*$. If there are too few entrants, the opposite is true. Based on this argument, I conjecture that the asset market can clear for some $M > 0$.

Let $K^* - S'(M) = K^e(M)$ be the excess asset supply curve as a function of the mass of entrants M. Then a positive entry SGE exists if:

$$\exists\ M^* > 0 \ni K^e(M^*) = 0$$

The usual strategy for showing this is to use the intermediate value theorem (IVT) on the excess assets supply function. I will show this computationally.[26] Before I proceed to the data and model estimation section, the following proposition is useful for efficient computation in steps 4 and 5.

**Proposition 6.** *The distribution evolution operator $T$ is linearly homogeneous in $(\mu, M)$ i.e.:*

$$\mu(M) = T(\mu(M), M, \theta) \implies M'\mu(M)/M = T(M'\mu(M)/M, M', \theta)$$

*Proof.* Let $\mu(M)$ be the stationary distribution corresponding to $M$. Then,

$$\mu(M)(\Psi, N) = M\mathbb{1}(n^e \leq N)\int_{\underline{\psi}}^{\Psi} dG(\psi) + \int_{\chi} W\Big((\psi, n), (\Psi, N)\Big)d\mu(M)(\psi, n)$$

Multiplying both sides by $M'/M$ gives the following, where $\Big(M'\mu(M)/M\Big)$ is the transformed measure.

$$\Big(M'\mu(M)/M\Big)(\Psi, N) = M'\mathbb{1}(n^e \leq N)\int_{\underline{\psi}}^{\Psi} dG(\psi) +$$

$$\int_{\chi} W\Big((\psi, n), (\Psi, N)\Big)d\Big(M'\mu(M)/M\Big)(\psi, n)$$

■

The key to this linear homogeneity is that since $\mu(M)$ is not a probability measure, multiplication by a scalar simply scales the measure. Intuitively, when more entrants enter the market in equilibrium, each subset of the state space experiences a proportional increase in its equilibrium measure. Another way of looking at this is that when entry mass is higher in stationary equilibrium, the mass of incumbents must rationalize it - the mass of incumbents must be large enough so that it results in a higher failure mass that matches the higher entry mass. Note that *mass* is key here, since failure rate does not change merely due to a greater mass of entrants.

# 4 Data and Estimation

To estimate the model, I use data on US commercial and savings banks from (a) FDIC Call and Thrift Financial (CTR) Reports (b) the FDIC failed banks list and (c) FDIC reports

[26] In order for a stationary general equilibrium to exist where markets clear for some prices, the bank distribution must be continuous in the parameters of the model. This is because if the distribution does not move continuously, there might be 'holes' in the set of aggregate deposits and aggregate consumption obtained by integrating over the bank distribution, causing failure of market-clearing conditions. I do not prove this analytically, but show computationally that equilibrium exists.

of changes in the institution count. I source balance sheet, cash flow and income statement variables from CTR. I use the failed bank list to identify failed banks, and changes in the bank count list to compute failure and entry rates. Consistent time series estimates for individual banks are ensured by following the notes developed by Kashyap and Stein [2000].[27] Macroeconomic aggregates for the US are sourced from FRED and the World Bank Indicators database. As motivated in the introduction, since 1995-2005 is a relatively stable period for banking industry dynamics, is less affected by structural changes, and is pre-2008 financial crisis, I use this period for calibration/estimation. I now discuss the functional forms and parameter estimation.

The functional form for operating costs is $cn^{\zeta}$. To pin down $c$, I suppose $\bar{n} > 0$ is a scaling factor that represents maximum effective net worth on the grid used to solve the model. This factor is simply used to map the scale of the model to scale of the data. Then the parameters of the cost function are chosen such that effective net worth $n - c(n)$ is not greater than $\bar{n}$. This implies that $c = \bar{n}^{(1-\zeta)/\zeta}$. The investment adjustment cost is modeled as a quadratic cost $f(x) = (x-1)^2$. The process for achieving idiosyncratic efficiency is modeled as an AR(1) process with mean $\mu = 1$, persistence $\rho$, and variance $\sigma$:

$$\psi_{it} = \mu(1-\rho) + \rho\psi_{it-1} + \sigma\epsilon_{it}$$

Finally, the distribution of efficiency for entrants, $g(.)$ is the ergodic distribution of $\psi$. One period in the model corresponds to an year in the data. The values of some standard parameters are chosen based on the standard in literature. The discount factor $\beta$ is set such that the interest rate $R = 1/\beta$ in the model matches an annual risk free interest rate close to 4%. The Cobb-Douglas production function parameter $\alpha$ is set such that capital income share in the model implies a capital income share that is close to 0.33. For depreciation rate, a value of 0.025 is used, which is again common in the macroeconomic literature. Total factor productivity A is normalized to 1.

The fraction of bankers $\lambda$ is set to match the pooled mean percentage of households in the US during 1995 - 2005 who held stocks, directly or indirectly, through investment funds or retirement accounts. This estimate is based on data from Survey of Consumer Finances 1995-2005. A break up of asset ownership by households, and stocks in particular, is given in Table 1. The above calibration is rationalized by the fact that bankers in my model are essentially the owners of capital stock, as bankers own banks, banks own firms, and firms own capital.

[27] Consistency of data is an issue with this data because of changes in variable definitions as well as bank mergers or disintegrations.

| Type of financial asset | 1995 | 1998 | 2001 | 2004 | 2007 |
|---|---|---|---|---|---|
| Transaction accounts | 87 | 90.5 | 91.4 | 91.3 | 92.1 |
| Certificates of deposit | 14.3 | 15.3 | 15.7 | 12.7 | 16.1 |
| Savings bonds | 22.8 | 19.3 | 16.7 | 17.6 | 14.9 |
| Bonds | 3.1 | 3 | 3 | 1.8 | 1.6 |
| Stocks | 15.2 | 19.2 | 21.3 | 20.7 | 17.9 |
| Pooled Investment funds | 12.3 | 16.5 | 17.7 | 15 | 11.4 |
| Retirement accounts | 45.2 | 48.8 | 52.2 | 49.7 | 52.6 |
| Cash value life insurance | 32 | 29.6 | 28 | 24.2 | 23 |
| Other managed assets | 3.9 | 5.9 | 6.6 | 7.3 | 5.8 |
| Other | 11.1 | 9.4 | 9.4 | 10 | 9.3 |
| Stocks (Direct plus Indirect) | 40.4 | 48.9 | 52.2 | 50.2 | 51.1 |
| Any Financial Asset | 91 | 92.9 | 93.4 | 93.8 | 93.9 |

Table 1: Financial asset ownership by families in the US, 1995 - 2007 *(Source: Survey of Consumer Finances)*

| Description | Symbol | Value |
|---|---|---|
| Standard parameters | | |
| Coefficient of risk aversion | $\gamma$ | 1.5 |
| Discount factor | $\beta$ | 0.96 |
| Capital exponent in production | $\alpha$ | 0.33 |
| Capital depreciation rate | $\delta$ | 0.025 |
| Benchmark Capital regulation | $\hat{\chi}$ | 29.58 |
| Fraction of bankers | $\lambda$ | 0.475 |
| Joint estimation using SMM | | |
| Efficiency Persistence | $\rho$ | 0.8373 |
| Efficiency Variance | $\sigma$ | 0.0072 |
| Cost exponent | $\zeta$ | 1.6853 |
| Fixed cost of entry | $c^e$ | 3.4060 |
| Asset diversion fraction | $\theta$ | 0.1796 |
| Entrant bank start up capital | $n^e$ | 0.2791 |
| Income tax rate | $\tau$ | 0.2416 |

Table 2: Summary of parameter values.

The pre-crisis benchmark capital regulation $\hat{\chi}$ is calibrated using the statutory risk-weighted capital requirement under Basel I – the international standard effective during most of 1995-2005.[28]. During this period, as per Basel I, banks were required to have an amount of capital that is at least 8% of risk-weighted assets. I compute an approximate implied *non-risk weighted* capital regulation using the fixed statutory risk weights and

[28]Although Basel II was effective in 2004, and a version of leverage ratio requirement was effective in the US prior to 1988, Basel I is closest to the effective regulation in US during 1995-2005, the calibration period chosen for this model (see, for example, Burhouse et al. [2003]).

actual capital ratios of banks. Specifically, since I do not have data on bank securities holdings, I first obtain a range of possible values for the implied capital regulation $\hat{\chi}$. To this end, I estimate the minimum and maximum capital charge for assets of each bank in the data during 1995-2005. Then the mean ratio of total assets to these capital charges gives the maximum and minimum implied capital regulation that is equivalent to the statutory capital regulation during 1995-2005.

$$\hat{\chi}_{max} = 29.58; \quad \hat{\chi}_{min} = 27.37$$

These estimates are consistent with the observation that more than 99 percent of banks have an asset to net-worth ratio of less than 25.30. I choose the more liberal estimate of $\hat{\chi}$ for the benchmark calibration.

The remaining parameters $\mathcal{P} = (\rho, \sigma, \zeta, c^e, \theta, n^e, \tau)$ are estimated jointly using the method of moments. As per this method, I begin with a guess of these parameters, and compute the stationary general equilibrium in the model. From SGE, I compute the target moments listed in Table 3. The vector of these model moments, say $\mathcal{M}^{model}(\mathcal{P})$, is a function of the parameters. The idea is to find the set of parameters $\mathcal{P}^*$ such that:

$$\mathcal{P}^* = arg\min_{\mathcal{P}}(\mathcal{M}^{model}(\mathcal{P}) - \mathcal{M}^{data})'W(\mathcal{M}^{model}(\mathcal{P}) - \mathcal{M}^{data})$$

where $\mathcal{M}^{data}$ is the vector of target data moments, and $W$ is the weighting matrix. I use a pre-specified weighting matrix whose diagonal elements are the inverse of squared data moments. This is essentially equivalent to minimizing the sum of the squared relative difference between model and data moments. This weighting matrix gives me consistent parameter estimates,[29] which are summarized in Table 2. Target moments and estimation results are summarized in Table 3. The choice of these moments is such that key features of banking industry dynamics are captured, namely moments of the size distribution, leverage and dividend behavior, exit rate, size of entrants and relative failure rate of banks. The income tax parameter $\tau$, targets the mean income GINI of US households during 1995-2005.

## 4.1 Computations

In this sub-section, I discuss the computational procedure I use to compute the stationary equilibrium, including the estimation of parameters using the Method of Moments. I use value function iteration with Howard policy improvement to solve for the value and policy functions of the bank.[30] The state space for banks is discretized as follows: (a) $\psi$ has 21 equally spaced grid points in $[\underline{\psi}, \overline{\psi}]$; (b) $n$ has 31 log-spaced grid points in $[0, \bar{n}]$. Spline interpolation is used to evaluate value function at off-grid points. Linear interpolation is used to evaluate policy function at off-grid points. A two-times-finer discretization of the

[29]One could do another step to back out the efficient weighting matrix using the co-variance matrix of initial parameter estimates. Using the efficient weighting matrix would give me the lowest possible standard error on my parameter estimates.

[30]I follow Heer and Maussner [2009] as the key reference for computational methods.

| Moments | Data | Model |
|---|---|---|
| Failure rate | .0026 | .0022 |
| Ratio of size of entrant and incumbent | .2354 | .2374 |
| Correlation of Size and Failure rate | -.4238 | -.3669 |
| Leverage | 10.38 | 12.28 |
| Fraction of banks in quintile 1 | .2000 | .1962 |
| Fraction of banks in quintile 2 | .2000 | .2395 |
| Fraction of banks in quintile 3 | .2000 | .1754 |
| Fraction of banks in quintile 4 | .2000 | .2146 |
| Dividend Ratio | .0939 | .0963 |
| Income Gini | .3641 | .3366 |

Table 3: 1995 - 2005 Data moments vs Model moments. For data moments (except for the quintiles and Gini index), standard errors are $[0.0004, 0.0414, 0.0915, 0.0247, 0.0009]$ respectively. Quintiles are based on size distribution in bottom 95% banks as of 2005:Q4.

state space is used to compute the stationary distribution for banks. On this finer grid, I construct the state space transition matrix using bank policy functions, and compute its ergodic distribution, which is the stationary distribution of banks. To handle entry and exit in this setting, I introduce a *dump* state which represents failures. Since in the stationary equilibrium, entry equals exit, transitions to the dump state represents failures, whereas transitions from the dump state represent entries. The ergodic measure of dump state is therefore interpreted as the failure rate. To solve the banker's income fluctuation problem, policy function iteration is employed. Stationary distribution is computed using a state transition matrix as before.

The Nelder-Mead simplex optimization algorithm is used to compute the Method of Moments parameter estimates.[31] Table 3 documents the result of this procedure. I target nine data moments based on pooled data from 1995 - 2005 to estimate six model parameters. The model does reasonably well at matching moments, except for mean leverage. One reason for not doing well in matching this target is that the actual implied capital regulation is unknown. I obtain an approximation using the statutory capital regulation, which generates the possibility of a tighter capital regulation than the one I use here.

Figure 7 compares the size distribution in the data with the one generated by the model. As also reflected in Table 3, the model captures the observed bank distribution reasonably well, except for the extreme right tail. The figure also shows that larger banks are more efficient in equilibrium. Intuitively, this follows from the dynamics of the industry: banks aim to grow larger by taking on more leverage, but only those that receive a better stream of efficiency shocks actually grow. In this sense, there is a selection effect which is operational. Bank value and policy functions are presented in Figure 18 in the appendix.

**Bank productivity** I define bank productivity as the rate at which a bank converts current period net-worth into next period net-worth, which is mathematically given as

[31] Routine available in MATLAB in the function 'fminsearch'.

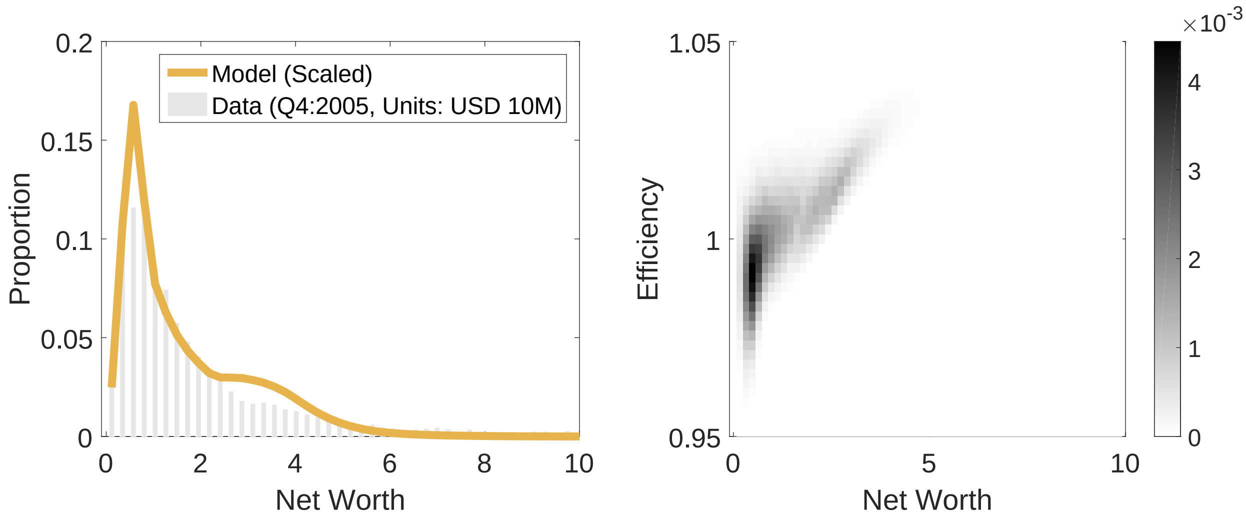


Figure 7: (Left) Distribution of banks over net worth in data and in model (Right) Distribution of banks over net worth and efficiency *(Distribution over state-space)*

follows:

$$q(\psi, n) = E[n']/n = E\bigg[g\bigg(\Big((R'_k\psi' - R)x(\psi, n) + R\Big)(1 - h(\psi, n))n\bigg)\bigg]/n$$

The rationale for this measure is two-fold. First, I would like to combine the notions of technical and economic efficiency of a bank, i.e. asset management efficiency $\psi$ and resource allocation efficiency, into a single measure, which I refer to as bank productivity. Second, I would like to verify if my model is robust to an *economies of scale in banking* result that several empirical papers have found, for example, Hughes and Mester [2013]. To this end, I compute the equilibrium profile of bank productivity in my model, which is given in Figure 8. The shaded band represents one standard deviation. The model prediction that bank productivity is increasing in size of the bank is therefore consistent with empirical findings.

# 5 Counterfactual Experiment

In this section, I discuss the counterfactual experiments I conduct using the model. This exercise has two goals. The first goal is to understand the response of banking industry and macroeconomic aggregates to changes in bank capital regulation. The second goal is to understand the trade-offs a social planner would face when choosing capital regulation to maximize welfare. This allows me to pin down the channels of welfare gains and losses that occur as the regulatory regime changes.

The benchmark economy equilibrium features a size-independent capital regulation regime where $\chi(n) = \hat{\chi} = 29.58$ i.e. all banks face the same limit $\hat{\chi}$ on the level of leverage they can assume. For the counterfactual experiment, I first define a regulation

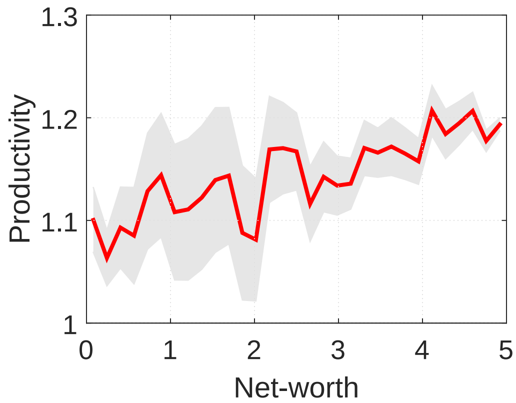


Figure 8: Equilibrium profile of mean bank productivity. The shaded band represents one standard deviation.

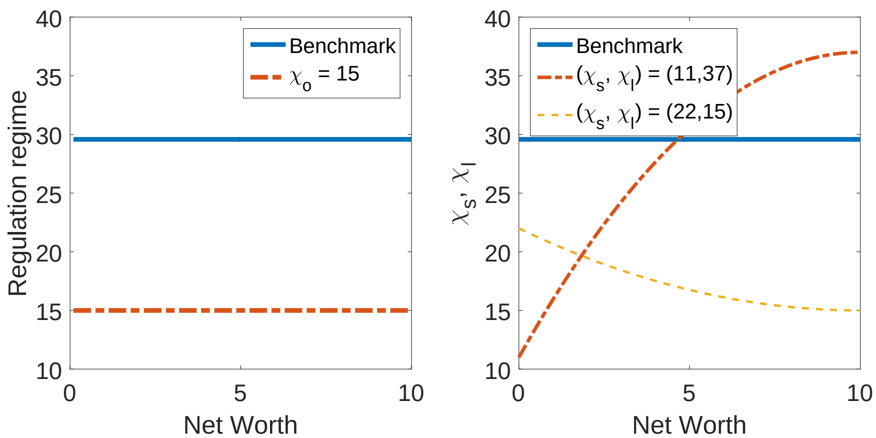


Figure 9: (i) *(Left)* Size independent regulation (ii) *(Right)* Size dependent regulation

function as any continuous, differentiable and monotonic function of bank size $\chi(n)$ that satisfies the following conditions $\chi(n) \geq 1, \lim_{n\to\infty} \chi'(n) = 0$. I consider only monotonic regulation functions to rule out complicated policies. Moreover, optimization over the class of all continuous and differentiable functions is a more difficult problem to handle computationally. The other two conditions imposed on a valid regulation function are intuitive. First, a capital regulation below 1 is absurd as it follows directly from the balance sheet identity that leverage cannot be less than 1. Second, a condition that regulation becomes size independent in the limit is required to ensure that regulation does not diverge to $+\infty$ or $-\infty$. This condition is not restrictive in the sense that size independence of regulation kicks in for banks of very large size only.

As such, the class of regulation functions proposed above is large. For tractability, I consider the class of quadratic regulation functions, which are completely described by three parameters. The choice of quadratic regulation functions is not restrictive. It follows from the *Weierstrass Approximation Theorem*, according to which the class of quadratic regulation functions is representative of the class of all regulation functions up to an error that is bounded from above.

I perform the counterfactual exercise in two steps. The size-independent exercise, where only degenerate regulation functions are considered, allows me to understand the effect of changes in the overall level of regulation, and characterize optimal *size-independent* regulation. The size-dependent exercise allows me to characterize optimal *size-dependent* regulation. I consider the following regulation functions, examples of which are given in Figure 9.

- Size-independent regulation: $\chi(n) \equiv \chi_0 \ni \chi_0 \geq 1$
- Size-dependent regulation:

$$\chi(n) = \chi_2 n^2 + \chi_1 n + \chi_0; \quad \chi(n) \geq 1, \lim_{n \to \infty} \chi'(n) = 0$$

$$\implies \chi(n) = (\chi_s - \chi_l)\left(n/\overline{n}\right)^2 - 2(\chi_s - \chi_l)\left(n/\overline{n}\right) + \chi_s$$

Note that the quadratic regulation function is described by three parameters. The limiting condition imposes restrictions on these parameters such that two parameters are sufficient to describe a regulation function. In fact, the quadratic regulation function can be written in terms of regulation $\chi_s$ for banks of size zero, limiting regulation $\chi_l$, and a choice parameter $\overline{n}$ that controls the upper limit of the grid for bank net worth.

A note about the incidence of regulation is due here. The statutory regulation function is specified in terms of bank size. However, the incidence of regulation on banks depends on the equilibrium distribution of the banks. For example, consider the second alternate regulation function in the left panel in Figure 9. The smallest banks face regulation close to 22, while the largest banks face regulation close to 15. In this regime, if the equilibrium distribution of banks has a support of, say, $[4, 6]$, the incidence of regulation would only vary in an interval around 17.

For each alternative specification, I compute the stationary general equilibrium, and compare banking industry and macroeconomic aggregates. The response of the banking industry to changes in the level of regulation can be summarized as follows.

- Distribution: The distribution of bank net worth moves to the right as regulation becomes tighter (weaker capital regulation), which leads to larger mean bank size, as shown in Figure 10. This result is in line with the observation in the data that as regulation has become tighter over time, the distribution has moved to the right. The variance in size distribution first increases, but eventually decreases as regulation tightens.
- Leverage and exit rate: As shown in Figure 11, mean leverage of banks is lower, and so is the overall failure rate. Wages paid to workers are lower as capital stock is lower (see Figure 20, in appendix), but dividend payment received by bankers are higher.
- Concentration: The concentration of the banking industry, as measured by a decrease in the Herfindahl index (computed assuming 5000 banks) decreases as regulation be-

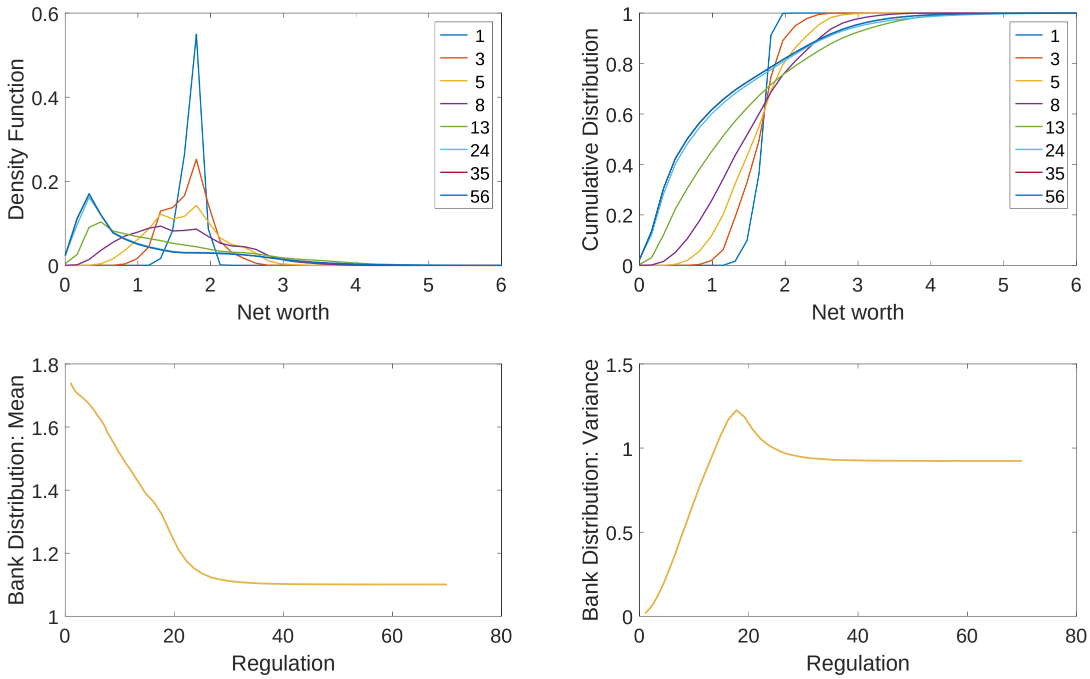


Figure 10: Response of bank distribution to changes in regulation *(In the figures in top row, the legend lists alternate regulation regimes)*

comes tighter. Also the mass of banks in equilibrium (relative to the mass of households, which is normalized to unity) increases, pointing towards a less concentrated industry.

## 5.1 Optimal regulation

In this section, I characterize optimal capital regulation relative to the chosen pre-crisis benchmark calibration of $\hat{\chi} = 29.58$. To this end, optimal regulation $\chi(n)$ is the one that maximizes utilitarian welfare, which in this case is the mass-weighted lifetime utilities of workers and bankers:

$$\max_{\chi \in \mathcal{C}^1(\mathbb{R}^+)} \mathcal{W}(\chi) \quad \equiv \max_{\chi \in \mathcal{C}^1(\mathbb{R}^+)} (1-\lambda)\frac{u\big(c_w(\chi)\big)}{(1-\beta)} + \lambda \int_{\underline{d}}^{\bar{d}} \int_0^{\bar{h}} U\big(h, d; \chi\big) \partial \mu^b\big(h, d; \chi\big)$$

There are both qualitative and quantitative findings. Qualitatively, the optimal *size-independent* regulation is characterized by an overall more stringent regulation relative to the pre-crisis benchmark. The optimal *size-dependent regulation*, in turn, is characterized as being more stringent for large banks relative to small banks.

Quantitatively, the size-independent optimum of 3.55 calls for much lower levels of leverage than under the pre-crisis benchmark (where CR equals 29.58). While not di-

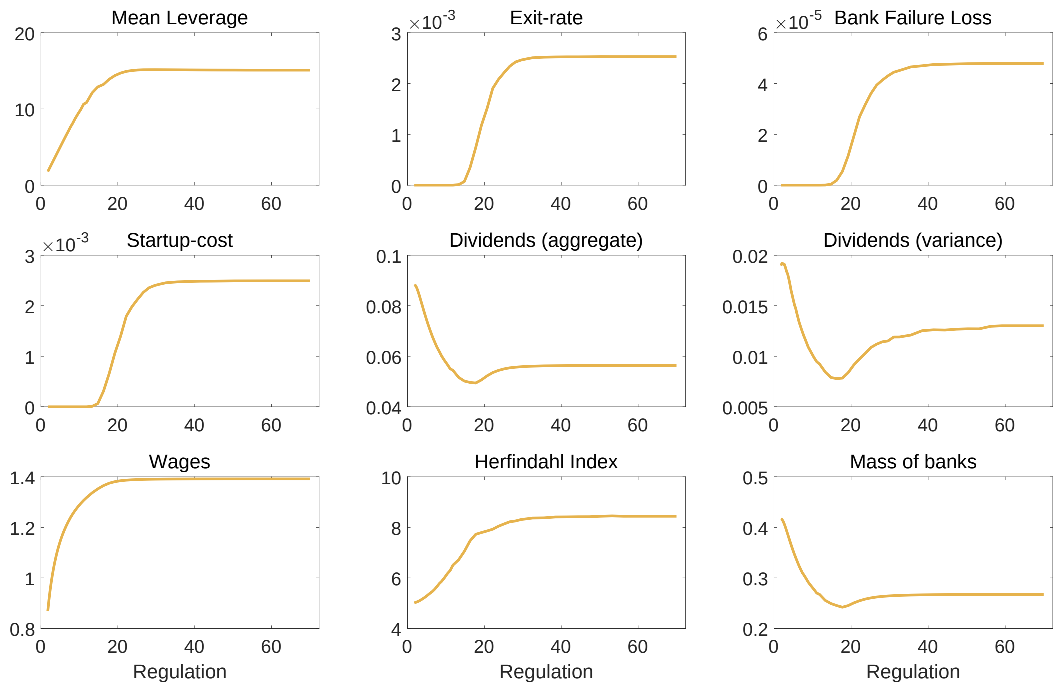


Figure 11: Response of industry to changes in regulation

rectly comparable, this result points in the same direction (which is that of tightening) as that proposed in other studies, including Admati and Hellwig [2014], Begenau [2014], and Nguyen [2014]. The *size-dependent* optimum, in turn, varies quadratically between a level of $\mathcal{X}_s^* = 4.12$ for the smallest banks and $\mathcal{X}_l^* = 3.17$ for the largest banks *(see Table 4)*. In terms of policy, this result reinforces the GSIB surcharges introduced under Basel III, while suggesting that overall capital levels could well be higher (e.g. via pillar 2 requirements) than required under the new international standards.[32]

| Statutory | | Incidence | | | | | |
|---|---|---|---|---|---|---|---|
| Smallest | Largest | Smallest | 10 percentile | Median | 90 percentile | Largest | Mean |
| 4.44 | 1.10 | 4.12 | 3.87 | 3.67 | 3.50 | 3.17 | 3.63 |

Table 4: Optimal Capital Regulation: Statutory value, its incidence on banks of different sizes, and its mean incidence.

**Channels of welfare gain** The rationale for an optimum that implies an overall tighter regulation is multi-dimensional. First, loss due to industry turnover is lower as less levered banks fail less often. This also lowers the start-up costs borne by bankers as a smaller

[32]While not directly comparable, a CR of 4.12 in this model would imply a leverage ratio requirement of about 24%.

mass of entrants enters the market. Second, the mean dividend income is higher, primarily due to a larger equilibrium mass of banks, which dominates the welfare loss from lower wages for workers. Third, the variance of dividend income is lower as banks are less levered, which increases welfare of risk averse bankers. This channel is closely linked with the wedge assumed between bank management and ownership. Fourth, tighter regulation results in higher aggregate bank net worth as banks are better capitalized, even though mean bank size is smaller.

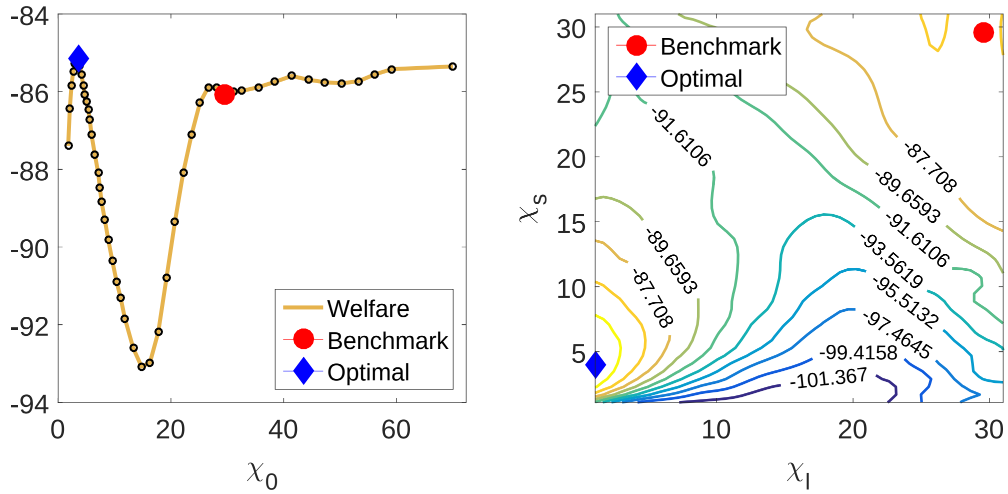


Figure 12: Welfare measures corresponding to *(Left)* Size-independent regulation regimes and *(Right)* Size-dependent regulation regimes (contour plot)

With regards to the profile of size dependent regulation, the following remarks summarize why tighter regulation of large banks relative to small banks improves welfare. I use the median bank size to classify banks as small or large.

**Proposition 7.** *Cleansing effect: The mean efficiency of failed banks is lower than the mean efficiency of entrant banks.*

*Proof.* The mean efficiency of entrants is 1 by calibration of the efficiency process. The mean efficiency of failed banks is given as:

$$m_\psi = \int\int\left(\int_{\underline{\psi}}^{R(x(\psi,n)-1)/x(\psi,n)} \psi' f(\psi,\psi')\partial\psi\right)\partial\mu(\psi,n)$$

Since the upper limit on leverage for all banks is equal to $\chi_s^* = 4.12$, this implies that $x(\psi,n) \le \chi_s^* \;\; \forall\ \psi,n \implies R(x(\psi,n)-1)/x(\psi,n) < 1$. It then follows trivially from the integral that $m_\psi < 1$. ∎

Intuitively, for a given level of current period efficiency, bank failure probability increases with leverage. Therefore, by allowing small banks to be more highly levered, these banks experience a higher failure rate. However the ones that fail are less efficient relative to incumbents. Moreover, the entrants that replace these failed banks draw their efficiency from the ergodic distribution, and have higher mean efficiency. The higher turnover rate among small banks entails costs, but optimally, this improves welfare since less efficient banks are eliminated, and failed small banks are, on average, immediately replaced by entrants which are also smaller.

**Proposition 8.** *Growth effect: For banks with positive expected premium $E[R_k\psi' - R|\psi]$ on intermediation, the expected growth rate for small banks is higher relative to the expected growth rate of large banks.*

*Proof.* The expected growth rate $l(\psi, n)$ of banks of size $n$ can be defined as:

$$l(\psi, n) = \Big[(R_k E[\psi'|\psi] - R)x(\psi, n) + R\Big](1 - h(\psi, n))$$

$$\implies \partial l(\psi, n)/\partial n = \Big[(R_k E[\psi'|\psi] - R)\partial x(\psi, n)/\partial n\Big](1 - h(\psi, n)) + \Big[(R_k E[\psi'|\psi] - R)x(\psi, n) + R\Big](-\partial h(\psi, n)/\partial n)$$

From the profile of optimal regulation, I know that $\partial x(\psi, n)/\partial n < 0$, and from comparative statics, I know that larger banks pay higher dividends, i.e., $\partial h(\psi, n)/\partial n > 0$. Then, since $(R_k E[\psi'|\psi] - R) > 0$, I have $\partial l(\psi, n)/\partial n < 0$. ∎

Intuitively, by allowing the smaller banks to take on more leverage, the regulator allows the small banks to take on more risk, and grow faster *conditional* on survival. Naturally, more small banks end up failing this way, but the failure of small banks is less costly.

**Proposition 9.** *Stabilization effect: The probability of failure for large banks is lower than the probability of failure for small banks.*

*Proof.* Bank failure probability is given as:

$$\hat{p}(\psi, n) = \int_{\underline{\psi}}^{R(x(\psi,n)-1)/x(\psi,n)} f(\psi, \psi')\partial\psi'$$

From the profile of optimal regulation, I know that $\partial x(\psi, n)/\partial n < 0$. Therefore, since the upper limit of the integral is increasing in $x(\psi, n)$, the proposition follows trivially.[33] ∎

By tightening the regulation for large banks, the regulator prohibits them from taking on more risk, therefore rendering them more stable. The downside is that large banks are more efficient in financial intermediation in equilibrium, and limiting the size of their intermediation can be costly. However, their failures are also costly because they are difficult to replace immediately. In particular, an entrant bank needs time to grow large to actually replace a failed large bank, and in each period during this time, additional operating costs are borne.

[33] In fact, the proposition is reinforced due to the property that distribution of $\psi'$ is first order stochastically increasing in $\psi$, and the fact that in equilibrium, larger banks have a first order dominant distribution over $\psi$ relative to smaller banks.

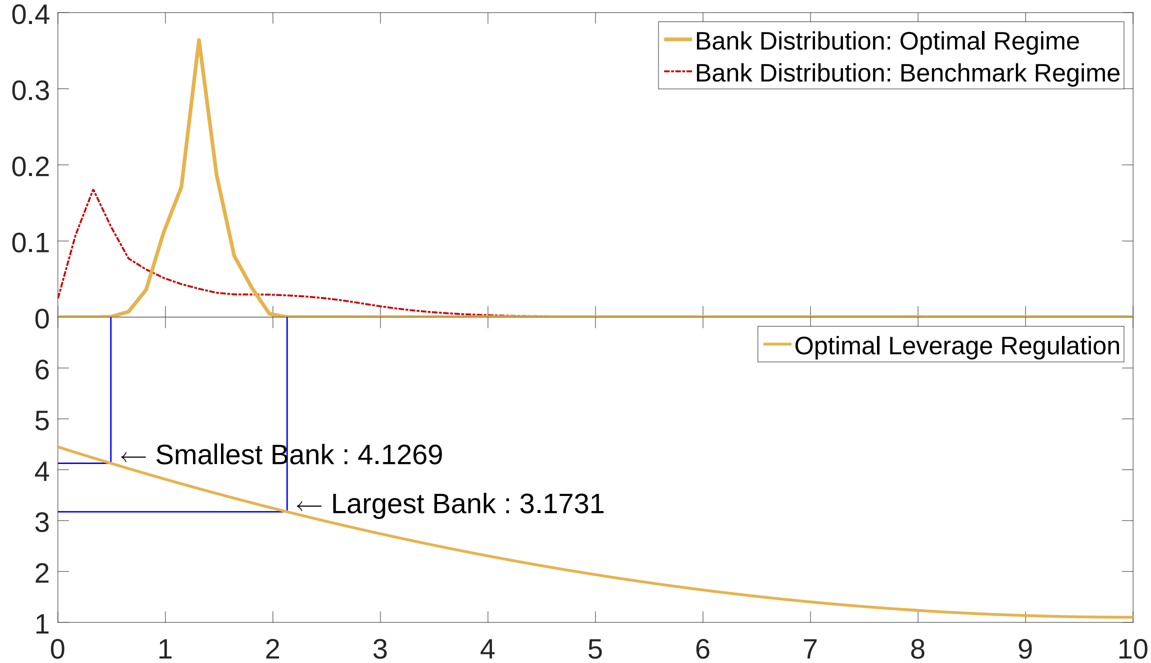


Figure 13: *(Top)* Bank size distribution in the benchmark and the optimal regimes *(Bottom)* Optimal capital regulation, and its incidence on smallest and largest banks

The gain in welfare from imposing regulation can be expressed in terms of consumption equivalence. Let A denote the benchmark regime and let B denote the optimal regime. Then welfare gain in consumption equivalent (CE) units is *defined* as the percentage increase in consumption all agents would receive should they live in the optimal regime forever. Formally, the gain $\omega_U$ is defined as one that satisfies the following equation.

$$(1-\lambda)\frac{u\Big((1+\omega_U)c_w^A\Big)}{(1-\beta)} + \lambda\int_{\underline{d}}^{\bar{d}}\int_0^{\bar{h}} E_0 \sum \beta^t u\Big((1+\omega_U)c_{bt}^A(h,d)\Big)\partial\mu_b^A\Big(h,d\Big) =$$

$$(1-\lambda)\frac{u\Big(c_w^B\Big)}{(1-\beta)} + \lambda\int_{\underline{d}}^{\bar{d}}\int_0^{\bar{h}} E_0 \sum \beta^t u\Big(c_{bt}^B(h,d)\Big)\partial\mu_b^B\Big(h,d\Big)$$

The steady state gain in welfare in CE terms for a size-independent regulation is 2.2%, while that for a size-dependent regulation is 10.5%. These are large gains relative to, for example, the costs of business cycles, but comparable to the contemporaneous GDP cost of the 2008 financial crisis, which is 6.2%, and the conservative lifetime GDP cost estimate provided by Luttrell et al. [2013], which is in the range of $40-90\%$. In other words, these welfare gains are not implausible from a steady state comparison perspective. If the transition of the economy from the benchmark regime to the optimal regime is taken into account, the gain in welfare might be lower. Intuitively, in the initial periods after a

| Parameter Perturbed (Benchmark Value) | Perturbed Value | Statutory | | Incidence | | |
|---|---|---|---|---|---|---|
| | | Smallest | Largest | Smallest | Largest | Mean |
| Benchmark Regime | | 4.44 | 1.10 | 4.12 | 3.17 | 3.63 |
| $\gamma$ (1.5) | 1 | 6.66 | 1.10 | 6.47 | 4.40 | 5.36 |
| | 2 | 4.44 | 1.39 | 4.15 | 3.28 | 3.69 |
| $\lambda$ (0.475) | 0.4 | 4.74 | 1.10 | 4.51 | 3.35 | 3.86 |
| | 0.5 | 4.14 | 1.10 | 3.85 | 2.98 | 3.40 |
| $\tau$ (0.2416) | 0.2 | 4.92 | 1.10 | 4.67 | 3.46 | 4.00 |
| | 0.3 | 3.97 | 1.10 | 3.69 | 2.87 | 3.26 |

Table 5: Sensitivity analysis: A regime is characterized by the value of the parameter that is perturbed relative to benchmark. The shaded row represents benchmark. In the first set of alternative regimes, *ceteris paribus*, $\gamma$ is increased and decreased relative to its benchmark value of 1.5. The tightest allowed capital regulation is 1.10, ensuring a non-trivial problem for the banks.

sudden imposition of the optimal regime, the regulatory constraint binds most, as banks need time to re-optimize. Given that re-optimization requires several periods to achieve, the private decisions of economic agents are sub-optimal during transition, resulting in aggregate inefficiency and welfare loss. This phenomenon provides one of the explanations why a myopic regulator might choose not to adopt such a policy change, even though welfare is eventually higher.

## 5.2 Sensitivity Analysis

In this section, I examine the sensitivity of optimal capital regulation to alternative calibrations. Optimal regulation is especially sensitive to the relative risk aversion of consumers $\gamma$, the fraction of bankers and workers $\lambda$, and the tax rate on bankers $\tau$. I vary these parameters on a discrete set containing the baseline calibration.[34] The result of this exercise is documented in Table 5.

A decrease in relative risk aversion implies a relatively liberal capital regulation. Intuitively, as consumers become less risk averse, the welfare loss from dividend income volatility is smaller. Since dividend payment volatility increases with bank leverage, the optimal regulation is relatively liberal in an economy with lower risk aversion. An increase in risk aversion relative to benchmark does not result in a significant change in either the statutory regulation or its incidence. If the fraction of bankers, who are the consumers directly affected by banking industry dynamics, is smaller, the weight-age of bankers in the welfare function is smaller. Since tighter regulation improves welfare for bankers and liberal regulation improves welfare for workers, a smaller fraction of bankers shifts optimal

[34]Since the computation of optimal regulation for each calibration requires close to 40 hours on a server with 28 parallel processors, only a limited number of such experiments were possible.

regulation in the liberal direction. Similarly, a larger fraction of bankers shifts optimal regulation in the tighter direction.

Finally, as taxes on bankers shift consumption from bankers to workers, lowering tax makes bankers better off. This means that a more liberal regulation would make workers better off, while not making bankers much worse off. Consequently, optimal regulation is relatively liberal. Similarly, higher taxes on bankers make them worse off, shifting the optimal regulation in their favor.

Overall, this exercise shows that the quantitative value of optimal capital regulation is sensitive to the exact calibration in an intuitive manner. However, for small changes in the calibration in the neighborhood of the benchmark calibration, the change in optimal regulation is small, and in this sense, the optimal regulation estimated in this paper is robust. The same applies to the estimated qualitative relationships.

# 6 Conclusion

In this paper, I characterize optimal size-dependent capital regulation for banks. To this end, I develop a stochastic general equilibrium macroeconomic model with a heterogeneous financial sector and financial frictions. A key feature of the model is that bank leverage decisions play a critical role in determining banking sector dynamics and macroeconomic aggregates. Due to the assumed frictions, bank leverage choices are socially inefficient, which rationalizes capital regulation in the model. I show that an overall tighter regulation generates welfare gains relative to the pre-crisis benchmark. In particular, I show that the optimal regulation is characterized by tighter requirements for large banks relative to small banks. In this sense, the paper qualitatively reinforces the Basel III logic that global systemically important banks, which are generally also the larger banks, should face tighter regulation relative to smaller banks. Quantitatively, the results suggest that, from a welfare perspective, there may be room for much higher capital requirements than envisioned under the Basel III minimum standard (e.g. via pillar 2 requirements). This is in line with recent research by Fender and Lewrick [2016], who find that ample room is available for national authorities to further raise regulatory capital.

The paper contributes to the existing literature in two ways. Firstly, this is one of the few papers that take up the question of size-dependent bank regulation in a macroeconomic setting. Secondly, the model developed here captures leverage as the mechanism behind banking industry dynamics, something existing models of firm dynamics do not typically focus on. One reason that this element is missing from the literature is that, for non-financial firms, leverage is not the most critical element in the nature of their business, whereas is it central for banks.

The analysis in this paper can be extended in at least two ways. First, it would be interesting to study the welfare implications of the size dependent regulation when there are other policies present in the model, for example, liquidity regulation. Since the model features an endogenously determined size distribution of banks, independent analysis of size-dependent policies (for example, explicit constraints on bank size), and anti-trust policies is also feasible in this framework. Second, time-invariant capital regulation can have pro-cyclical effects, and can exacerbate recessions. It could thus be insightful to introduce aggregate uncertainty into the model and study countercyclical capital regulation, something that Agarwal and Goel [2016] pursue. Introducing aggregate uncertainty in the current setup would necessitate solution techniques such as those developed in Krusell and Smith [1998] or Reiter [2009], as the distribution of banks would become part of the state space.

# References


Acharya, V. V. (2011). The Dodd-Frank act and basel 3 intentions, unintended consequences, transition risks, and lessons for india.

Admati, A. and M. Hellwig (2014). *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It*. Princeton University Press.

Adrian, T. and N. Boyarchenko (2012). Intermediary leverage cycles and financial stability. *Becker Friedman Institute for Research in Economics Working Paper* (2012-010).

Agarwal, I. and T. Goel (2016). Countercyclical capital regulation: Rules versus discretion.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 659–684.

Albuquerque, R. and H. A. Hopenhayn (2004). Optimal lending contracts and firm dynamics. *The Review of Economic Studies 71*(2), 285–315.

BCBS (2011). Basel 3: A global regulatory framework for more resilient banks and banking systems. *Basel Committee for Banking Supervision*.

Begenau, J. (2014). Capital requirement, risk choice, and liquidity provision in a business cycle model. *Risk Choice, and Liquidity Provision in a Business Cycle Model (October 24, 2014)*.

Benes, J. and M. Kumhof (2011). Risky bank lending and optimal capital adequacy regulation. Technical report.

Boyd, J. H. and G. De Nicolo (2005). The theory of bank risk taking and competition revisited. *The Journal of finance 60*(3), 1329–1343.

Burhouse, S., G. F. John Feid, and K. Ligon (2003). Basel and the evolution of capital regulation: Moving forward, looking back. *FDIC Update on Emerging Issues in Banking*.

Christensen, I., C. Meh, and K. Moran (2011). Bank leverage regulation and macroeconomic dynamic. *CIRANO-Scientific Publications 2011s-76*.

Christiano, L. and D. Ikeda (2013). Leverage restrictions in a business cycle model. Technical report, National Bureau of Economic Research.

Clementi, G. L. and H. A. Hopenhayn (2006). A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics*, 229–265.

Clementi, G. L. and B. Palazzo (2015). Entry, exit, firm dynamics, and aggregate fluctuations. Technical report, National Bureau of Economic Research.

Cooley, T. F. and V. Quadrini (2001). Financial markets and firm dynamics. *American Economic Review*, 1286–1310.

Corbae, D. and P. D'erasmo (2013). A quantitative model of banking industry dynamics.

Corbae, D. and P. D'erasmo (2014). Capital requirements in a quantitative model of banking industry dynamics.

Covas, F. and J. C. Driscoll (2014). Bank liquidity and capital regulation in general equilibrium.

De Nicolo, G., A. Gamba, and M. Lucchetta (2014). Microprudential regulation in a dynamic model of banking. *Review of Financial Studies 27*(7), 2097–2138.

DeAngelo, H. and R. M. Stulz (2013). Why high leverage is optimal for banks. Technical report, National Bureau of Economic Research.

Fender, I. and U. Lewrick (2016). Adding it all up: the macroeconomic impact of basel iii and outstanding reform issues. *BIS Working Papers, no 591*.

Gabaix, X. and A. Landier (2006). Why has ceo pay increased so much? Technical report, National Bureau of Economic Research.

Gertler, M. and P. Karadi (2011). A model of unconventional monetary policy. *Journal of Monetary Economics 58*(1), 17 – 34.

Gertler, M. and N. Kiyotaki (2010). Financial intermediation and credit policy in business cycle analysis. *Handbook of Monetary Economics 3*, 547.

Hanson, S., A. Kashyap, and J. Stein (2010). A macroprudential approach to financial regulation. *Chicago Booth Research Paper* (10-29).

Heer, B. and A. Maussner (2009). *Dynamic general equilibrium modeling: computational methods and applications*. Springer Science & Business Media.

Holmstrom, B. and J. Tirole (1997). Financial intermediation, loanable funds, and the real sector. Technical report.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, 1127–1150.

Hughes, J. P. and L. J. Mester (2013). Who said large banks do not experience scale economies? evidence from a risk-return-driven cost function. *Journal of Financial Intermediation 22*(4), 559–585.

Jayaratne, J. and P. E. Strahan (1997). The benefits of branching deregulation. *Economic Policy Review 3*(4).

Kalemli-Ozcan, S., B. Sorensen, and S. Yesiltas (2012). Leverage across firms, banks, and countries. *Journal of international Economics 88*(2), 284–298.

Kashyap, A. K. and J. C. Stein (2000). What do a million observations on banks say about the transmission of monetary policy? *American Economic Review*, 407–428.

Krusell, P. and A. A. Smith, Jr (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy 106*(5), 867–896.

Lang, L., E. Ofek, and R. Stulz (1996). Leverage, investment, and firm growth. *Journal of financial Economics 40*(1), 3–29.

Luttrell, D., T. Atkinson, and H. Rosenblum (2013). Assessing the costs and consequences of the 2007–09 financial crisis and its aftermath. *Economic Letter 8*.

Miao, J. (2005). Optimal capital structure and industry dynamics. *The Journal of finance 60*(6), 2621–2659.

Nguyen, T. T. (2014). Bank capital requirements: A quantitative analysis. *Available at SSRN 2356043*.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control 33*(3), 649–665.

Stokey, N. L. (1989). *Recursive methods in economic dynamics*. Harvard University Press.

Van den Heuvel, S. J. (2008). The welfare cost of bank capital requirements. *Journal of Monetary Economics 55*(2), 298–320.

Zhu, H. (2007). Capital regulation and banks' financial decisions. *BIS Working Paper*.

# 7 Appendix

## 7.1 Leverage and Bank Dynamics

In this section, in order to develop intuition for how leverage affects bank dynamics, I provide a few examples. Failure is defined as the event where bank net worth drops below zero for the first time. In the examples below, a bank's stylized balance sheet is characterized by net worth and leverage tuple $(n, l)$. As per this notation, assets equal $ln$ and liabilities equal $(l-1)n$. I will assume the initial asset price to be $Q = 1$, so that the number of units of an asset is same as the value of the asset. I focus on bank balance sheet response to asset price shocks.

**Leverage and bank failures** To illustrate the direct effect of leverage on failure, consider two banks of the same size as measured by net worth, holding the same asset. Bank A has balance sheet position $(n, l_A)$ and bank B has balance sheet position $(n, l_B)$. Assume, without loss of generality, $l_B > l_A$. Also assume that the liabilities of a bank do not change due to asset price shocks. Then the percentage drops $\epsilon_A, \epsilon_B$ in asset prices required for bank A and bank B to fail independently are given as:

$$\epsilon_A = 1/l_A, \quad \epsilon_B = 1/l_B \implies \epsilon_B < \epsilon_A$$

Remark 1: *Given any probability distribution of asset price shocks, the probability of failure for a relatively highly levered bank is higher.*

**Leverage and de-leveraging** Consider the same two banks A and B such that $n_A = n_B$ and $l_B > l_A$. Assume that due to some exogenous heterogeneity, $l_A$ and $l_B$ are the optimal levels of leverage for banks A and B, respectively. Now consider a negative shock $\epsilon$ to asset prices, one that is not large enough for any of the banks to fail. Then, the post-shock leverage of the banks is:

$$l'_A = \frac{(1-\epsilon)l_A}{1-\epsilon l_A}, \quad l'_B = \frac{(1-\epsilon)l_B}{1-\epsilon l_B}$$
$$\implies l'_B > l'_A$$
$$\implies l'_B/l_B > l'_A/l_A$$

Remark 2: *A bank that is more highly levered compared with another bank will remain so after a negative asset price shock. The percentage increase in leverage is greater for a bank that is more highly levered in the first place.*

The first implication is intuitive. The second implication is more interesting for the following reason. Assume that, after a negative price shock, both banks re-optimize their balance sheets in order to get back to their respective pre-shock leverage levels. This re-optimization entails de-leveraging, i.e., reducing leverage. De-leveraging can be achieved either by raising fresh equity, or by selling assets to pay off liabilities. Since raising fresh equity is more difficult than selling assets in a business cycle downturn, I will assume that de-leveraging takes place through the sale of assets. Suppose bank $i$ sells $x_i$ units of assets at price $(1-\epsilon)$ to bring its leverage back to the (optimal) pre-shock level. For bank $i$, post-shock assets are $(1-\epsilon)nl_i$, and post-shock net worth is $(1-\epsilon l_i)n$. Then we have the following implication:

$$\frac{(1-\epsilon)nl_i - (1-\epsilon)x_i}{(1-\epsilon l_i)n} = l_i, \quad i \in \{A, B\}$$

$$\implies x_i = \frac{\epsilon n l_i(l_i - 1)}{1-\epsilon}$$
$$\implies x_B > x_A$$

This means that a more highly levered bank would sell more assets for a given shock. However, the more insightful implication follows from the comparison of percentage asset sale, defined as the ratio of units of assets sold to a bank's pre-shock assets. This implication follows trivially from the fact that $l_B > l_A$.

$$\tilde{x}_i = \frac{x_i}{nl_i} = \frac{\epsilon(l_i - 1)}{1-\epsilon}$$

$$\implies \tilde{x}_B > \tilde{x}_A$$

Remark 3: *Given a negative asset price shock, a relatively highly levered bank would sell a larger fraction of its assets while re-optimizing its balance sheet.*

**Bank size matters** In the above examples, I considered banks of the same size. In this example, to illustrate that bank size matters, I consider two banks that are equally levered but are of different sizes, $(n_A, l)$ and $(n_B, l)$ with $n_B > n_A$ and $l_A = l_B$. As before, I will assume that $l$ is the optimal operating leverage ratio for these banks. Suppose there is an asset price shock that depresses asset prices by a fraction $\epsilon$. The updated balance sheets are as follows:

$$A : \left((1-\epsilon l)n_A, \frac{(1-\epsilon)l}{1-\epsilon l}\right); \qquad B : \left((1-\epsilon l)n_B, \frac{(1-\epsilon)l}{1-\epsilon l}\right)$$

The leverage of both banks goes up by the same amount. As above, in order to re-optimize their respective balance sheets, bank $i$ would need to sell assets, say $x_i$ units:

$$x_i = \frac{\epsilon n_i l(l-1)}{1-\epsilon} \quad \forall i \in \{A, B\} \implies x_B > x_A$$

Remark 4: *Given a negative asset price shock, among equally levered banks, the larger banks will sell more assets while re-optimizing their balance sheets.*

## 7.2 Proofs

**Proposition 3: Existence of a unique value function** I reproduce the problem of a bank in an economy without aggregate uncertainty. The aggregate state is degenerate, so we will not write it specifically.

$$V(\psi, n) = \max_{x\in[1,\overline{x}], h\in[0,1]} \left(hn + E\Lambda' \int_{\frac{R(x-1)}{x}}^{\overline{\psi}} V(\psi', n')F(\psi, d\psi')\right)$$

$$n' = g\left(\Big((R_k'\psi' - R)x + R\Big)(1-h)n\right)$$

$$V(\psi, n) \geq \theta(1-h)nx$$

I will follow Stokey [1989] to establish the proof. I will prove that Blackwell conditions are satisfied by the bank's problem, following which, by the Contraction Mapping Theorem (CMT), a unique value function will exist. Some notation is as follows. Let $Y = [1, \underline{\phi}] \times [0, 1]$ be the action space. Let the exogenous state space of efficiency shocks $\psi$ be $[\underline{\psi}, \overline{\psi}] = \Psi$, and let the endogenous state space of net worth be $\mathcal{N} = [0, \overline{n}]$. Then the state space for the dynamic problem can be written as $\chi = \Psi \times \mathcal{N} \subset \Re^2$. By imposing some structure on the operating cost function $c(.)$, we will ensure that the state space is bounded so that the reward function $hn$ is bounded. This is for mathematical convenience in ensuring that Blackwell Conditions are sufficient for the Bellman mapping to be a contraction. Some more notation is useful. Let $\Gamma : \chi \to Y$ be the policy correspondence (i.e. at each point of the state space, what actions a bank can take). Let $g : \chi \times Y \times \Psi \to \mathcal{N}$ be the law of motion for $n$ (given a point in the state space, and given some action, and given a realization of shock tomorrow, where the bank would transit to). Now I verify the assumptions required for Blackwell conditions to be sufficient for CMT.

Firstly, to ensure a bounded state space, no bank must grow larger than $\overline{n}$ for any starting net worth, leverage, or shock. Since $g(.)$ is increasing and concave, the condition for a bounded state space can be written as follows:

$$\lim_{\omega \to \infty} g(\omega) \leq \overline{n} \implies \lim_{\omega \to \infty} \omega - c(\omega) = \omega - c\omega^{\zeta} \leq \overline{n}$$

This condition imposes a constraint on the set of parameters $c, \zeta$ of the cost function and $\overline{n}$. The idea of this condition is that there is an upper limit to how large a bank can grow as for very large sized banks, operating costs grow one to one with net worth. This can be imposed on the parameters as follows:

$$c\zeta\omega^{(\zeta-1)} = 1; \quad \omega - c\omega^{\zeta} = \overline{n}$$

The lower bound on net worth is zero because a bank exits the market as soon as its net worth is less than or equal to zero. This implies a bounded state space for the bank's problem. Secondly, the action space is also bounded as already discussed in the exposition of bank's problem.

The above ensures the following two conditions. (SLP A9.4) $\mathcal{N}$ is a closed, bounded, convex subset of $R$. (SLP A9.5) $\Psi$ is closed, bounded (hence compact) subset of $R$. Moreover, $F$, the transition function for exogenous state $\psi$ should satisfy the Feller property (i.e. maps $C(\Psi) \to C(\Psi)$), which I ensure by assuming a simple AR(1) process for $\psi$ with Gaussian noise. (SLP A9.6) The policy correspondence $\Gamma$ is non-empty, compact valued and continuous, which is satisfied in this case because $Y$ is compact. (SLP A9.7) The reward function $H : \chi \times Y \to \mathcal{R}; H(\psi, n, x, h) = hn$ is bounded because state space is bounded, and is continuous. Also, the discount factor $\beta < 1$. (SLP A9.16) $Y$ is convex and compact subset of $R^2$. (SLP A9.17) $g$ is continuous in all arguments.

Hence I have verified all assumptions required for Blackwell conditions to be sufficient. I begin with the Bellman equation corresponding to the bank's problem[35] . Let $C(\chi)$ be the space of continuous bounded functions on $\chi$ with sup norm. Then, for any $q \in C(\chi)$, the Bellman operator from bank's problem is given as:

$$T(q)(\psi, n) = \max_{\substack{x \in [1,\phi], h \in [0,1] \\ \theta(1-h)nx \leq q(\psi,n)}} \left( hn + \beta \int_{\underline{\psi}}^{\overline{\psi}} q(\psi', n') \mathbb{1}\Big(\psi' \geq \frac{R(x-1)}{x}\Big) F(\psi, d\psi') \right)$$

$$where : n' = g\bigg(\Big(R'_k\psi' - R\Big)x + R\Big)(1-h)n\bigg)$$

Since $q$ is continuous, the maximand is continuous. The action set is closed and bounded, hence compact. For every continuous function on a compact set, a maximum exists, and is attained. Also, since $q$ is bounded, since $F$ satisfies Feller property, and since the reward function $H$ is bounded, $T(q)$ is bounded. Finally, continuity of $H$ and $g$ imply continuity of $T(q)$ by Theorem of the Maximum (SLP Theorem 3.6). Thus I showed that

[35]Which is different from the kind of Bellman equation employed in SLP, and which makes this proof necessary.

$T$ is a map from space of continuous bounded functions to itself. Now I verify the following Blackwell sufficient conditions (BSC) to conclude that $T$ is a contraction mapping.

- **Monotonicity:** $\forall f, q \in C(\chi)$ such that $f(x) \leq q(x)\ \forall x \in \chi \implies$
$$T(f)(x) \leq T(q)(x)\ \forall x \in \chi$$
- **Discounting:** $\exists \beta \in (0,1)$ such that $\forall f \in C(\chi),\ \forall a \geq 0,\ \forall x \in \chi \implies$
$$T(f+a)(x) \leq T(f)(x) + \beta a$$

**Monotonicity** Suppose $f(x) \leq q(x)\ \forall x \in \chi$. Let $(x^f, h^f)$ and $(x^q, h^q)$ be the maximands of $T(f)$ and $T(q)$ respectively at the point $y = (\psi, n)$. This means the following conditions are satisfied:

$$\theta(1-h^q)nx^q \leq q(\psi, n); \quad \theta(1-h^f)nx^f \leq f(\psi, n) \leq q(\psi.n)$$

This means that $(x^f, h^f)$ is a feasible point for the Bellman operator at $q$. Let $y'(x,h) = (\psi', n')$. Note here that $n'$ is a function of $y, x, h, \psi'$, and this is implicit here. Then,

$$T(q)(y) = \left(h^q n + \beta \int_{\underline{\psi}}^{\overline{\psi}} q(y'(x^q, h^q))\mathbb{1}\left(\psi' \geq \frac{R(x^q-1)}{x^q}\right)F(\psi, d\psi')\right)$$

Using definition of the maximum, and the fact that $(x^f, h^f)$ is a feasible point for the Bellman operator at $q$, we get:

$$\implies T(q)(y) \geq \left(h^f n + \beta \int_{\underline{\psi}}^{\overline{\psi}} q(y'(x^f, h^f))\mathbb{1}\left(\psi' \geq \frac{R(x^f-1)}{x^f}\right)F(\psi, d\psi')\right)$$

$$\implies T(q)(y) \geq \left(h^f n + \beta \int_{\underline{\psi}}^{\overline{\psi}} f(y'(x^f, h^f))\mathbb{1}\left(\psi' \geq \frac{R(x^f-1)}{x^f}\right)F(\psi, d\psi')\right)$$

$$\implies T(q)(y) \geq T(f)(y)$$

**Discounting**

$$T(f+a)(y) = \max_{\substack{x\in[1,\phi], h\in[0,1] \\ \theta(1-h)nx \leq (f+a)(y)}} \left(hn + \beta \int_{\underline{\psi}}^{\overline{\psi}} [f(y'(x,h)) + a]\mathbb{1}\left(\psi' \geq \frac{R(x-1)}{x}\right)F(\psi, d\psi')\right)$$

$$\implies T(f+a)(y) \leq T(f)(y) +$$
$$\max_{\substack{x\in[1,\phi], h\in[0,1] \\ \theta(1-h)nx \leq (f+a)(y)}} \int_{\underline{\psi}}^{\overline{\psi}} \beta a \mathbb{1}\left(\psi' \geq \frac{R(x-1)}{x}\right)F(\psi, d\psi')\right)$$

which follows because maximization of a sum of functions is at least as small as maximization of individual functions. Using $x = 1, h = 1$, which is a feasible point, in the second maximization, gives us

$$\implies T(f+a)(y) \leq T(f)(y) + \beta a$$

Hence the Bellman operator is a contraction map, implying that there exists a unique fixed point of the Bellman operator. This fixed point is the value function for bank's problem. Consequently, I can use value function iteration to obtain arbitrarily close estimate of the value function.

**Proposition 4: Comparative statics.** The steady state formulation of bank's problem without the borrowing constraint is given below. The proposition holds for the case with borrowing constraint, but the proof is omitted here as it is more involved.

$$V(\psi, n) = \max_{x \in [1,\overline{x}], h \in [0,1]} \left( hn + \beta \int_{\frac{R(x-1)}{x}}^{\overline{\psi}} V(\psi', n') F(\psi, d\psi') \right)$$

$$n' = g\Bigg( \Big( (R_k \psi' - R)x + R \Big)(1-h)n \Bigg)$$

The optimization problem has four boundary conditions on the choice variables, namely $x \geq 1, x \leq \overline{x}, h \geq 0, h \leq 1$. The comparative statics of the policy functions is trivial if any of the boundary conditions bind[36]. I will therefore derive comparative statics for the general case where none of these conditions bind. For convenience in writing the FOCs, I define the following terms. A numerical subscript denotes derivative with respect to corresponding argument.

$$A(\psi') = R_k \psi' - R$$

$$B(x) = R(x-1)/x; \quad B_1(x) = R/x^2;$$

$$G(x,h,n) = g\Bigg( \Big( A(\psi')x + R \Big)(1-h)n \Bigg); \quad G_1(x,h,n) = g'(.)A(\psi')(1-h)n$$

$$G_2(x,h,n) = -g'(.)(A(\psi')x + R)n; \quad G_3(x,h,n) = g'(.)(A(\psi')x + R)(1-h);$$

$$N(x,h,n) = g\Bigg( \Big( A(B(x))x + R \Big)(1-h)n \Bigg) = g\Bigg( R(R_k - 1)(x-1)(1-h)n \Bigg);$$

[36]In fact, from the structure of the problem, I know that only the second or third constraints might bind in an equilibrium.

The first order conditions and the envelope conditions are as follows:

$$[x]: \quad \beta \int_{B(x)}^{\overline{\psi}} V_2(\psi', n') G_1(x,h,n) f(\psi,\psi') d\psi' - \beta V\Big(B(x), N(x,h,n)\Big) f(\psi, B(x)) B_1(x) = 0$$

$$[h]: \quad n + \beta \int_{B(x)}^{\overline{\psi}} V_2(\psi', n') G_2(x,h,n) f(\psi,\psi') d\psi' = 0$$

$$[\psi]: \quad V_1(\psi, n) = \beta \int_{B(x)}^{\overline{\psi}} V(\psi', n') f_1(\psi,\psi') d\psi'$$

$$[n]: \quad V_2(\psi, n) = h + \beta \int_{B(x)}^{\overline{\psi}} V_2(\psi', n') G_3(x,h,n) f(\psi,\psi') d\psi'$$

Since $-(1-h)G_2(x,h,n) = nG_3(x,h,n)$, $[Env2]$ and $[h]$ imply:

$$V_2(\psi, n) = 1 \implies V(\psi, n) = n + C(\psi) \quad (*)$$

where $C(\psi)$ is some unknown function of $\psi$. I can use this expression to simplify the FOCs $[x]$ and $[h]$:

$$[x]: \quad \beta \int_{B(x)}^{\overline{\psi}} G_1(x,h,n) f(\psi,\psi') d\psi' - \beta \Bigg( N(x,h,n) + C(B(x)) \Bigg) f(\psi, B(x)) B_1(x) = 0$$

$$[h]: \quad n + \beta \int_{B(x)}^{\overline{\psi}} G_2(x,h,n) f(\psi,\psi') d\psi' = 0$$

This gives me a system of two equations $[x]$ and $[h]$ in policies $x, h$ and state space variables $\psi, n$, with respect to which I would like to obtain comparative statics results. Note that the only unknown is $C(.)$, which I will address later. I rename equations as follows: $[x] \to q^1(x,h;\psi,n)$ and $[h] \to q^2(x,h;\psi,n)$. Then, assuming that the policy functions $x(\psi,n), h(\psi,n)$ are differentiable, I can apply the Implicit Function Theorem to obtain $x_\psi(\psi,n), h_\psi(\psi,n), x_n(\psi,n), h_n(\psi,n)$. I begin with the following system of equations:

$$-\begin{bmatrix} g^1_x(x,h;\psi,n) & g^1_h(x,h;\psi,n) \\ g^2_x(x,h;\psi,n) & g^2_h(x,h;\psi,n) \end{bmatrix} \begin{bmatrix} x_i(\psi,n) \\ h_i(\psi,n) \end{bmatrix} = \begin{bmatrix} g^1_i(x,h;\psi,n) \\ g^2_i(x,h;\psi,n) \end{bmatrix} \quad i \in \{\psi, n\}$$

Let $\Delta$ be the determinant of the Jacobian matrix. Then the system can be solved as follows for $i \in \{\psi, n\}$:

$$\begin{bmatrix} x_i(\psi,n) \\ h_i(\psi,n) \end{bmatrix} = -\begin{bmatrix} g^1_x(x,h;\psi,n) & g^1_h(x,h;\psi,n) \\ g^2_x(x,h;\psi,n) & g^2_h(x,h;\psi,n) \end{bmatrix}^{-1} \begin{bmatrix} g^1_i(x,h;\psi,n) \\ g^2_i(x,h;\psi,n) \end{bmatrix}$$

$$= -\frac{1}{\Delta} \begin{bmatrix} g^2_h(x,h;\psi,n) & -g^1_h(x,h;\psi,n) \\ -g^2_x(x,h;\psi,n) & g^1_x(x,h;\psi,n) \end{bmatrix} \begin{bmatrix} g^1_i(x,h;\psi,n) \\ g^2_i(x,h;\psi,n) \end{bmatrix}$$

Now, since $x(\psi,n)$ and $h(\psi,n)$ are the policy functions, they represent optimal decisions at the given point of the state space, and therefore the Jacobian matrix associated with the first order conditions must have a non-positive determinant, $\Delta \leq 0$. The individual

| Call report variable name | Variable label | Call report variable name | Variable label |
|---|---|---|---|
| RCFD2170 | Total Assets | RCFD0010 | Cash |
| RCFD2948 | Total Liabilities | RIAD4460 | Dividend component A |
| RCFD3210 | Total Equity Capital | RIAD4470 | Dividend component B |
| RIAD4000 | Total operating Income | RCFD3230 | Common Stock |
| RIAD4107 | Interest Income | RIAD4073 | Total Interest expense |
| RIAD4079 | Non Interest Income | RIAD4074 | Net Interest Income |
| RIAD4340 | Net income | RIAD4093 | Total Non Interest Income |
| RIAD4130 | Total operating expense | RSSD9950 | Date of Incorporation |

Table 6: Call report data

terms of the RHS can be simplified using the FOCs. I also assume that $C'(.) > 0$, which essentially implies that the value function is some increasing function of $\psi$. Then, using the conditional distribution of $\psi'$ given $\psi$, and by plugging in the calibrated value of the known parameters $R, \beta, R_k$ I can sign the partial derivatives of the policy functions.

## 7.3 Data analysis

I source aggregate time series data from the Federal Reserve Bank of St. Louis Economic Research Database (FRED). I obtain Individual bank data from the Federal Reserve Bank of Chicago Commercial Bank datasets. Financial asset ownership data is obtained from Survey of Consumer Finances. Finally, bank entry and exit data is sourced from the Historical Statistics on Banking database at the Federal Deposit Insurance Corporation. In constructing target moments of banking industry dynamics, I focus on the variables from Call Reports given in Table 6.

Leverage is computed as the ratio of Total Assets and Total Equity Capital (net worth). Mean dividend payout ratio is computed as the ratio of sum of two dividend components and net worth. To compute size ratio of entrants and incumbents, the date of incorporation of a bank is used to identify entrants, and net worth is used as the measure of size, as is the case through out this paper. For exit rate, total number of failures in a given year is divided by the total number of banks at the start of the year. Assisted / unassisted mergers or bailouts are not counted as exits to be consistent with the interpretation of failures in the model. The distribution of bank net worth is based on inflation adjusted total equity capital. I use the total operating expense profile in data to motivate the assumption that banks are subject to convex operating costs in the model.

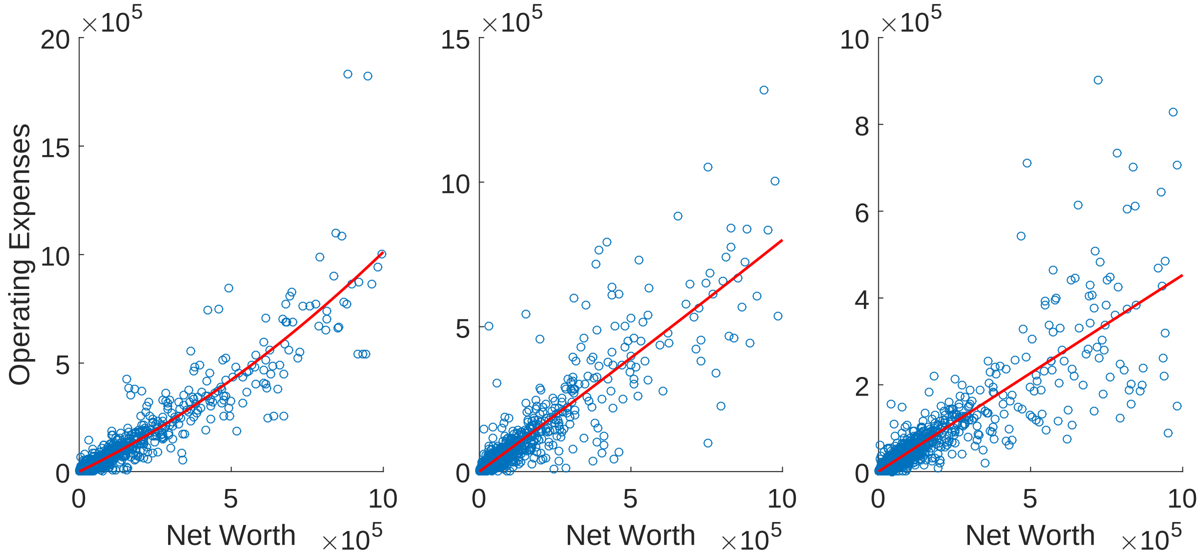


Figure 14: Operating cost and net worth scatter for 1995Q4, 2000Q4, 2005Q4

The result from Kolmogorov-Smirnov (KS) two sample test is given in Table 7. The null hypothesis is that the empirical distribution functions are drawn from the same distribution. I consider various time intervals in the calibration period, and run the test on the distribution of inflation adjusted log net worth of banks as of the beginning (Year 1) and end (Year 2) of the interval. I cannot reject the null, even if the distributions are 11 years apart.:

| Year 1 | Year 2 | Distance | P-value |
|---|---|---|---|
| 1995 | 2005 | 10 | 0.3863 |
| 1997 | 2005 | 8 | 0.1550 |
| 1995 | 2003 | 8 | 0.8727 |
| 2000 | 2005 | 5 | 0.0623 |
| 1997 | 2002 | 5 | 0.4812 |
| 1995 | 2000 | 5 | 0.4526 |
| 2003 | 2005 | 2 | 0.6052 |
| 1995 | 1997 | 2 | 0.9568 |

Table 7: KS two sample test

## 7.4 Aggregate shocks

Although there is no aggregate uncertainty in my model, in this subsection, I consider a simple exercise wherein banks receive an unexpected aggregate shock to their efficiency $\psi$ at date 1. Assuming that the economy is in the steady state at date 0, I compute the mass of bank failures that occur at the beginning of date 1 as a result of this shock. I then compare the incremental bank failures that occur due to the aggregate shock across

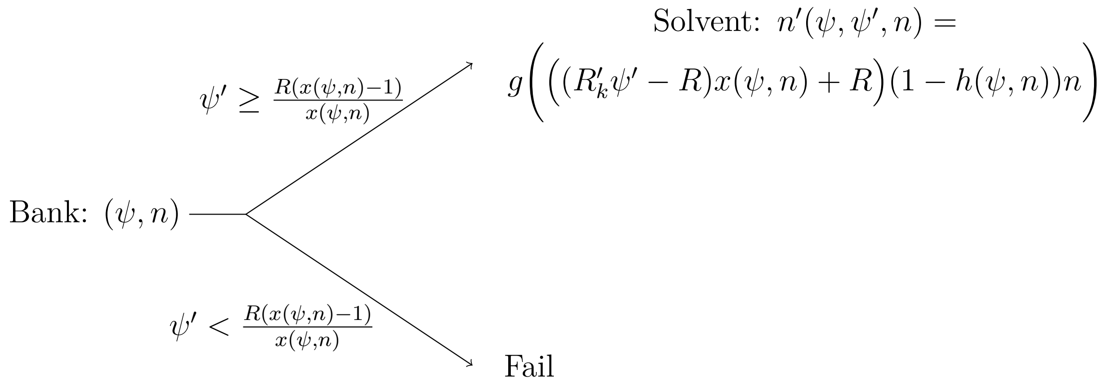


Figure 15: Dynamics of an individual bank

the benchmark and optimal regimes. The objective of this exercise is to show that the optimal regulation reduces the incidence of an aggregate shock by reducing the failure rate of banks, and also by reducing the *distress rate* of (surviving) banks as measured by the expected percentage decrease in net worth as a result of the aggregate shock. In other words, this exercise shows that a less levered banking industry responds less aggressively to an aggregate shock.

In the benchmark economy, the steady state failure rate of banks is 0.0022. An unexpected aggregate shock of $\epsilon = 5\%$ at the beginning of date 1 would result in a higher failure rate as more number of banks will find their leverage too high to sustain the shock and remain solvent i.e. maintain a positive net-worth. I must clarify here that using this approach I can only compute the instantaneous mass of failures that occur at the beginning of date 1. To compute the response of the economy thereafter, I would need to compute the full transition dynamics of the economy back to the steady state, taking into consideration the persistence of the shock. However, such computation is outside the scope of this paper. In order to compute the effective failure rate at date 1, I begin with the steady state equilibrium, and compute the next period solvency and net-worth positions of individual banks as follows: Without an aggregate shock, $\psi$ follows a standard AR(1) process with mean $\mu = 1$. With an aggregate shock that affects the mean of the conditional distribution of bank efficiency, the resulting distribution looks as follows: I now compute the failure and distress rates in the economies with the benchmark regulation (left panel) and the optimal regulation (right panel).

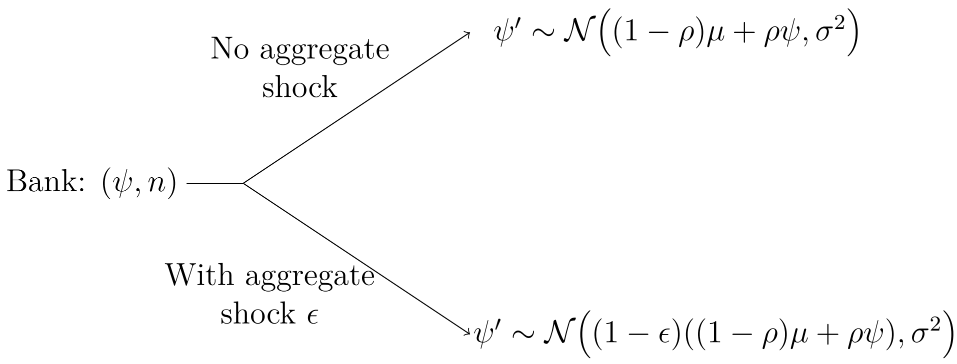


Figure 16: Distribution of realized shocks

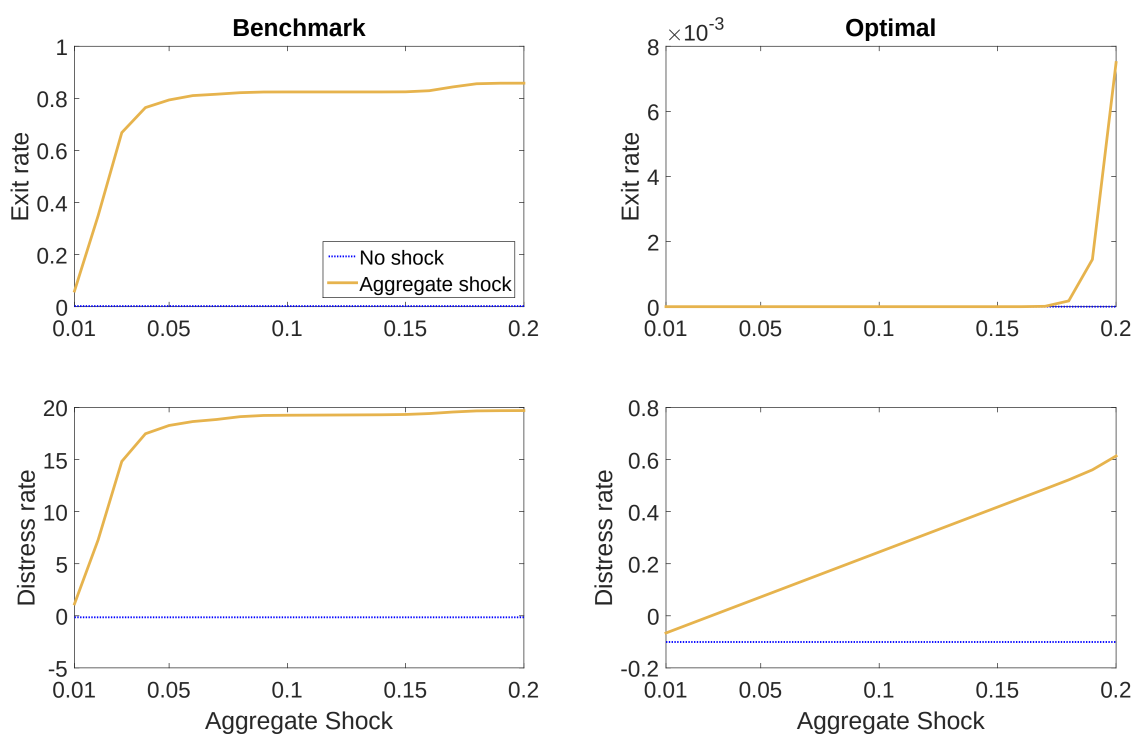


Figure 17: *(Top)* Exit rate; *(Bottom)* Distress rate *(percentage decrease in net worth)*

## 7.5 Additional Figures

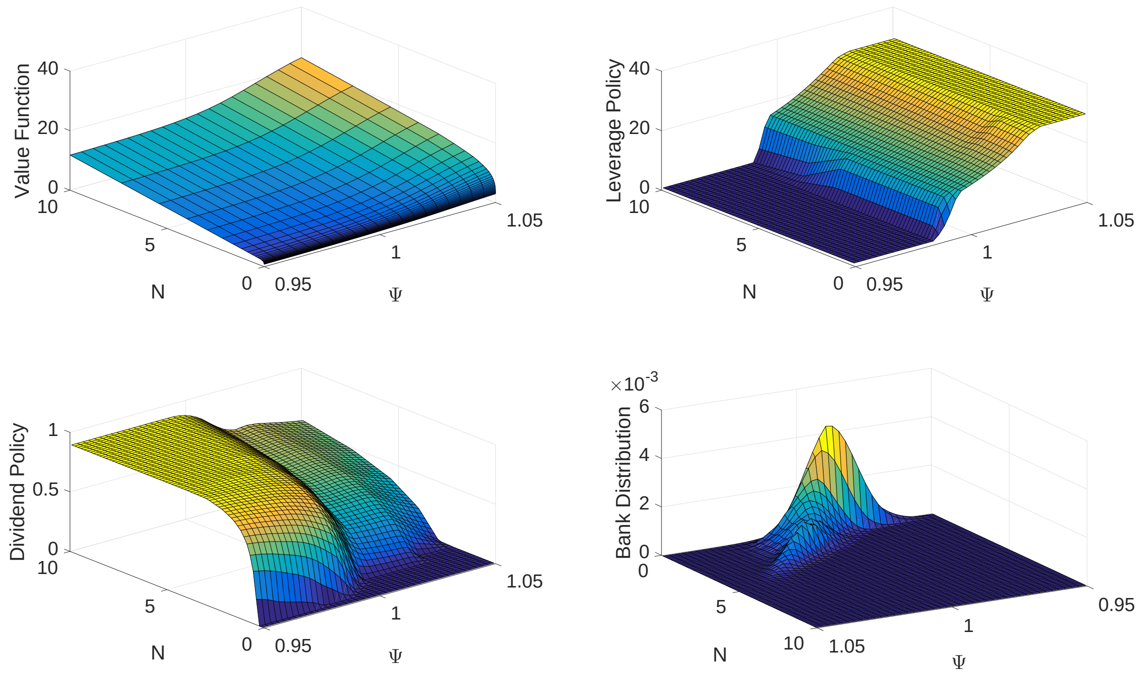


Figure 18: Solution to the problem of the banks

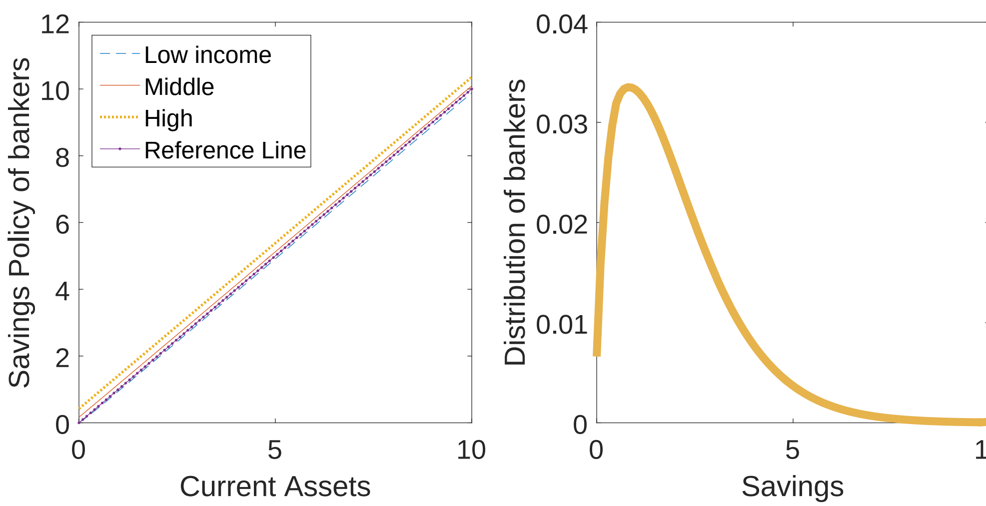


Figure 19: Solution to the problem of the bankers

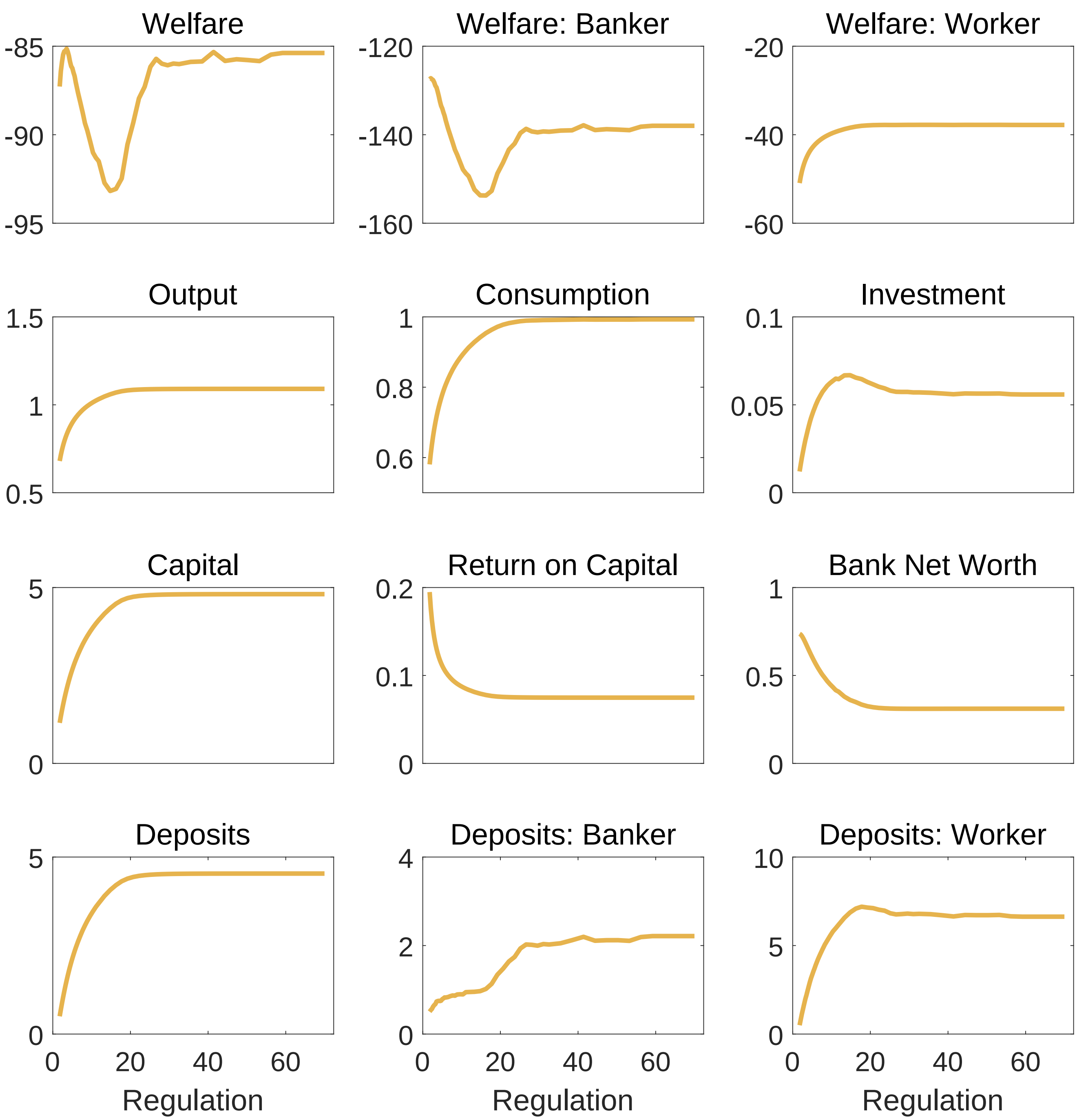


Figure 20: Response of macroeconomic aggregates to change in level of regulation

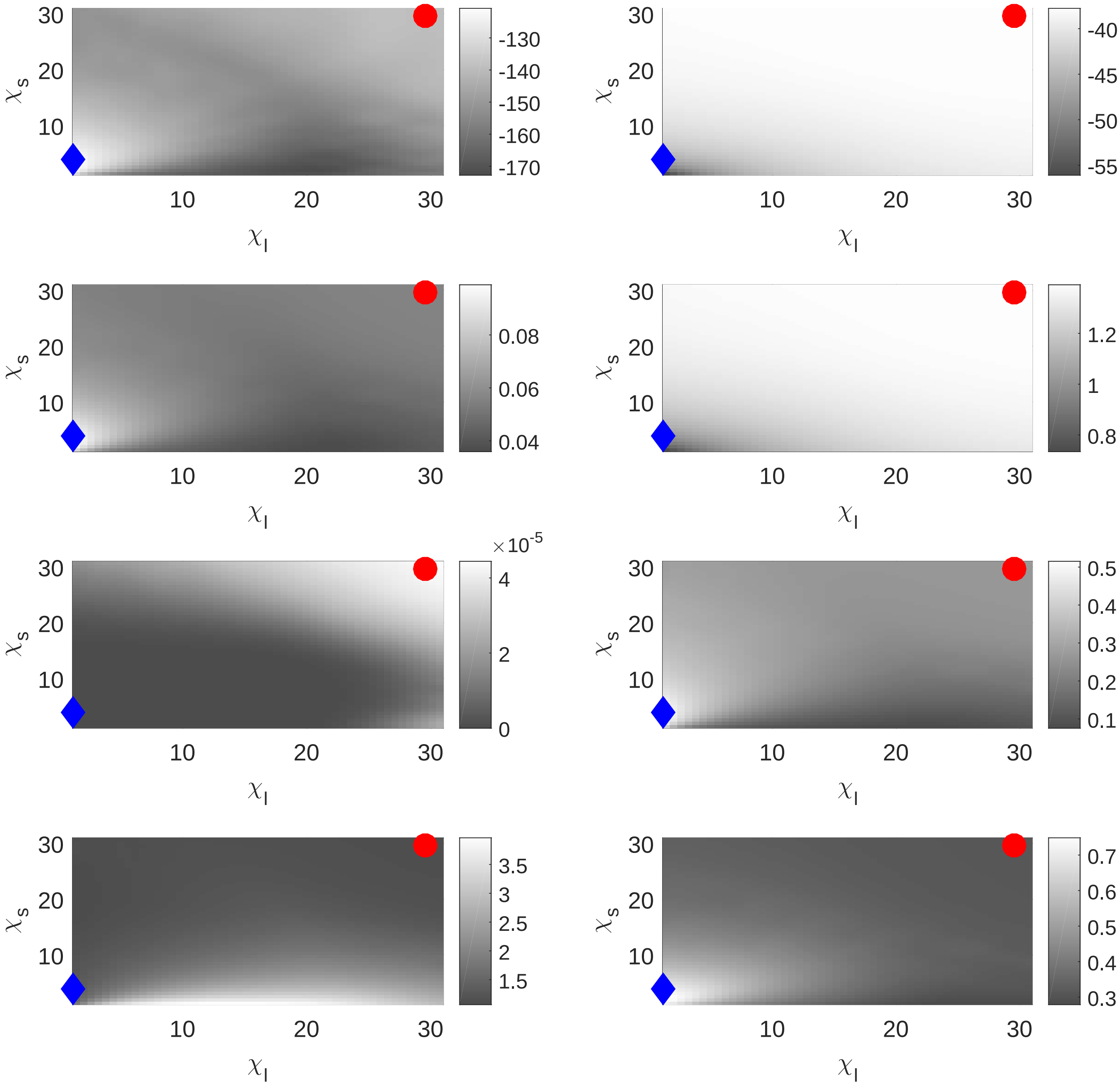


Figure 21: Response of macroeconomic aggregates to change in regulation regime. $\chi_l$ is the limiting bank regulation, while $\chi_s$ is the smallest bank regulation.

## Previous volumes in this series

| No | Title | Author |
|---|---|---|
| 598<br>December 2016 | Did the founding of the Federal Reserve affect the vulnerability of the interbank system to contagion risk? | Mark Carlson and David C Wheelock |
| 597<br>December 2016 | Bank Networks: Contagion, Systemic Risk and Prudential Policy | Iñaki Aldasoro, Domenico Delli Gatti, Ester Faia |
| 596<br>December 2016 | Macroeconomics of bank capital and liquidity regulations | Frédéric Boissay and Fabrice Collard |
| 595<br>December 2016 | Bank lending and loan quality: the case of India | Pallavi Chavan and Leonardo Gambacorta |
| 594<br>December 2016 | A quantitative case for leaning against the wind | Andrew Filardo and Phurichai Rungcharoenkitkul |
| 593<br>December 2016 | The countercyclical capital buffer and the composition of bank lending | Raphael Auer and Steven Ongena |
| 592<br>November 2016 | The dollar, bank leverage and the deviation from covered interest parity | Stefan Avdjiev, Wenxin Du, Catherine Koch and Hyun Song Shin |
| 591<br>November 2016 | Adding it all up: the macroeconomic impact of Basel III and outstanding reform issues | Ingo Fender and Ulf Lewrick |
| 590<br>October 2016 | The failure of covered interest parity: FX hedging demand and costly balance sheets | Vladyslav Sushko, Claudio Borio, Robert McCauley and Patrick McGuire |
| 589<br>October 2016 | International prudential policy spillovers: a global perspective | Stefan Avdjiev, Cathérine Koch, Patrick McGuire and Goetz von Peter |
| 588<br>October 2016 | Macroprudential policies, the long-term interest rate and the exchange rate | Philip Turner |
| 587<br>October 2016 | Globalisation and financial stability risks: is the residency-based approach of the national accounts old-fashioned? | Bruno Tissot |
| 586<br>September 2016 | Leverage and risk weighted capital requirements | Leonardo Gambacorta and Sudipto Karmakar |
| 585<br>September 2016 | The effects of a central bank's inflation forecasts on private sector forecasts: Recent evidence from Japan | Masazumi Hattori, Steven Kong, Frank Packer and Toshitaka Sekine |
| 584<br>September 2016 | Intuitive and Reliable Estimates of the Output Gap from a Beveridge-Nelson Filter | Güneş Kamber, James Morley and Benjamin Wong |
| 583<br>September 2016 | Exchange rate pass-through: What has changed since the crisis? | Martina Jašová, Richhild Moessner and Elől Takáts |
| 582<br>September 2016 | Global inflation forecasts | Jonathan Kearns |
| 581<br>September 2016 | Near-Money Premiums, Monetary Policy, and the Integration of Money Markets: Lessons from Deregulation | Mark Carlson and David C Wheelock |

All volumes are available on our website www.bis.org.