Depositing Trust
Banking Demand and Too-Big-to-Fail

Andrew Perry*

DRAFT: April 20, 2026
DO NOT DISTRIBUTE

ABSTRACT

This paper investigates the competitive distortions introduced by "Too-Big-To-Fail" (TBTF) guarantees within the U.S. banking sector, specifically examining the flight-to-safety dynamic that redistributes funding during periods of systemic distress. To formalize the mechanism through which implicit government backstops alter deposit market competition, I develop a continuous-time structural model that integrates spatial differentiation with endogenous default risk. The model yields a "scissors" dynamic, predicting that while distressed small banks must compress margins to retain uninsured depositors, TBTF incumbents strategically widen spreads to monetize their perceived safety. To empirically test these predictions across the entire size spectrum of the banking system, I construct a novel, continuous measure of default probability for the universe of regional and local banks using a Weighted Ensemble machine learning framework. Utilizing a difference-in-differences identification strategy, I find robust evidence of a post-GFC regime shift: uninsured deposit demand has become highly elastic to bank solvency, but with a stark size hierarchy. A one-percentage-point increase in default probability is associated with an approximately 48 percent increase in uninsured deposit share for Global Systemically Important Banks (GSIBs), while local banks experience an essentially flat or slightly negative relationship. These findings suggest that regulatory reforms have not eliminated the TBTF subsidy but have instead repriced it, allowing systemic institutions to act as liquidity sinks that absorb franchise value from smaller competitors during economic contractions.

JEL classification: G01, G21, G28, H12, D43, C45.

*Andrew Perry is with the Haas School of Business, U.C. Berkeley. Acknowledgments:

I. Introduction

Commercial banks are the engines of credit creation, but they run on a specific fuel... deposits. Deposits comprise over 90% of liabilities and three-quarters of total funding for the average bank. Consequently, the stability and cost of this funding source are the primary determinants of bank profitability and solvency. The United States banking sector is uniquely segmented in how it gathers this fuel, split between thousands of small "community banks" or "local banks" serving single towns, larger regional institutions, and a handful of Globally Systemically Important Banks. This paper investigates a central question of modern financial architecture: Did the government interventions during the Global Financial Crisis (GFC) distort the competitive landscape of this market, effectively subsidizing the largest banks at the expense of their smaller competitors?

The Financial Stability Board annually designates a set of Global Systemically Important Banks (GSIBs) on the basis of five criteria reflecting systemic risk: size, cross-jurisdictional activity, interconnectedness, substitutability, and complexity Financial Stability Board (2025). The rationale is that the disorderly failure of such an institution would disrupt critical financial services and impose cascading losses on the broader economy. In the United States, four domestic commercial banks currently hold this designation: JPMorgan Chase, Bank of America, Citigroup, and Wells Fargo. Collectively, these four institutions hold approximately $10 trillion in total assets Federal Reserve System (2026). Wells Fargo is the smallest GSIB at $1.8 trillion in assets which is three times bigger than the next largest bank, US Bank.

The broader U.S. commercial banking sector is organized in a pronounced size hierarchy. Beneath the four GSIBs sits a tier of large regional banks, institutions with assets typically ranging from roughly $100 billion to $500 billion, followed by thousands of community and local banks whose operations are largely confined to single metropolitan areas or rural counties. As of recent years, over 4,000 federally insured commercial banks operate across the United States (Egan, Hortaçsu, and Matvos, 2017). This extreme size dispersion defines the competitive landscape studied in this paper: institutions of vastly different scale draw from the same national deposit base yet face markedly different regulatory obligations and perceived safety profiles.

Figure 1 gives motivating evidence for this project. The top panel plots the aggregate savings deposit rates for GSIBs versus all other commercial banks. While a spread existed pre-GFC, the relationship fundamentally shifted following the crisis. To control for the dramatic decline in the interest rate environment, the bottom panel plots the "Savings Rate Ratio", the spread between non-GSIB and GSIB rates scaled by the Federal Funds Rate.

The data reveals a stark regime change; prior to 2008, the scaled spread was relatively stable. Post-GFC, however, the ratio spikes dramatically, indicating that non-GSIB banks were forced to pay a significantly higher premium for deposits relative to GSIBs. Crucially, this advantage appears counter-cyclical: the premium paid by small banks spikes during periods of intense systemic distress, most notably during the GFC and the onset of the COVID-19 pandemic—and mean-reverts during periods of calm. While one might argue this dispersion is a mechanical artifact of the Zero Lower Bound (ZLB), the ratio exhibits significant variation between 2012 and 2016 even as the policy rate

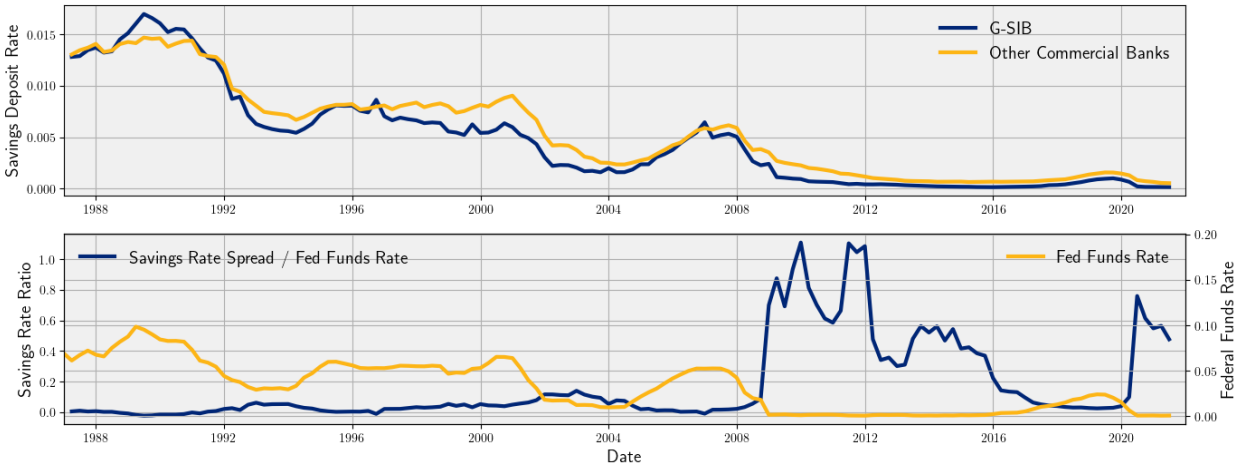


Figure 1. The top panel shows the deposit weighted average of the savings deposit rate for the four domestic GSIBS and all other commercial banks in the sample. The bottom panel shows the spread between the savings deposit rates of all other commercial banks minus the GSIBs scaled by the Federal Funds rate. Additionally, the right axis of the bottom panel shows the path of the Federal Funds rate. The data is collected quarterly between 1987 and 2022. The savings deposit rate is calculated from Call Reports by dividing the total interest expense on savings deposits by the total amount of savings deposits.

remained pinned at zero, suggesting the driver is economic, not mechanical.

This paper argues that this widening spread represents the price of safety. The architecture of the modern banking system is predicated on a fundamental tension, the transformation of liquid, short-term liabilities into illiquid, long-term assets. As formalized by Diamond and Dybvig (1983), this maturity mismatch exposes institutions to self-fulfilling runs where the mere expectation of insolvency creates actual insolvency. To mitigate this, governments have erected safety nets, implicitly guaranteeing that GSIBs are "Too-Big-To-Fail" (TBTF). The stylized facts in Figure 1 suggest that this guarantee distorts market discipline. During times of crisis, uninsured depositors flee to the safety of GSIBs, allowing these giants to lower their funding costs exactly when the rest of the market faces a liquidity squeeze.

The distinction between insured and uninsured depositors is central to this dynamic. The Federal Deposit Insurance Corporation (FDIC) guarantees deposits up to $250,000 per depositor per institution, effectively eliminating credit risk for the majority of retail account holders. Below this threshold, a depositor has no incentive to withdraw preemptively; their principal is protected regardless of the bank's condition. Uninsured depositors face a materially different problem. Their claims above $250,000 are subject to potential losses in the FDIC receivership process, exposing them to genuine default risk. It is this cohort that responds to deteriorating bank fundamentals by reallocating funds toward perceived safe havens Egan et al. (2017), and their capacity for coordinated withdrawal constitutes the primary source of banking system fragility.

The structural fragility of the U.S. banking system is perhaps best illustrated by the counterfactual analysis of Wachovia Bank during the 2008 financial crisis. Egan et al. (2017) estimate a structural model of the U.S. banking sector that confirms the existence of multiple equilibria. In March 2008, the market-implied probability of default for Wachovia, based on its fundamental

asset quality and leverage, was a manageable 3.3%. However, the authors identified a second, "fragile" equilibrium consistent with those exact same fundamentals where the default probability skyrocketed to 52%. This stark divergence highlights the central thesis of this paper: large banks, inoculated by implicit government guarantees, persistently live in the "good" equilibrium. Even if their fundamentals deteriorate, the expectation of a bailout prevents the coordinated withdrawal of uninsured funds. Conversely, smaller banks, lacking this credible guarantee, remain perpetually exposed to the "bad" equilibrium.

This disparity has significant implications for the political economy of bank stability. Following the GFC, regulatory frameworks such as the Dodd-Frank Act explicitly aimed to end the TBTF regime. However, the effectiveness of these reforms remains debated. Berndt, Duffie, and Zhu (2025) estimate that while creditors now price in some risk of failure, the market still assigns an overwhelming 87% probability that the government will save a GSIB in a crisis. The implicit guarantee has not been eliminated; it has merely been repriced.

This persistent expectation of support drives the "Flight to Safety" dynamic investigated here. Testing the magnitude of this distortion, however, presents a significant identification challenge. A naive comparison of deposit rates is confounded by non-regulatory factors, such as the "sophistication paradox" identified by d'Avernas, Eisfeldt, Huang, Stanton, and Wallace (2023), where wealthy consumers accept lower rates at large banks in exchange for superior technology (e.g., mobile apps, expansive ATM network). This paper resolves this ambiguity by exploiting the temporal dynamics of the safety premium. Technological advantages are static over short horizons, whereas the TBTF guarantee is highly counter-cyclical. By using a difference-in-differences methodology that isolates the change in deposit spreads during stress events, I disentangle the safety premium from the technology premium.

To formalize the channel through which TBTF guarantees distort deposit market competition, I develop a structural model that integrates the spatial differentiation framework of Hotelling (1929) with the endogenous default mechanics of Merton (1974) and Leland (1994). In this economy, banks compete for funding not only on the basis of offered rates and spatial convenience but also on their perceived safety. The model posits a fundamental asymmetry: while the large "National" bank effectively enjoys a regulatory backstop that insulates its depositors from loss, the smaller "Local" bank is subject to default risk, forcing it to pay a credit risk premium to retain uninsured depositors. This framework generates a sharp, testable implication; the deposit spread between large and small banks should not be constant, but rather a convex function of the smaller bank's financial health. Specifically, as a small bank's probability of default rises, the model predicts a "flight-to-safety" dynamic where the small bank must aggressively raise rates to compensate for risk, while the TBTF incumbent strategically widens its net interest margin, effectively monetizing the systemic distress. This theoretical prediction provides a structural foundation for my empirical tests, which aim to identify whether the sensitivity of deposit flows to bank solvency is indeed magnified for unprotected institutions during periods of aggregate stress.

Operationalizing these model implications requires a reliable, continuous measure of default

probability for the universe of small banks. Unlike GSIBs, whose risk is priced in liquid Credit Default Swap (CDS) markets, most regional banks lack market-based risk indicators. To overcome this, I construct a novel proxy for bank distress using a supervised machine learning framework. Building on the bankruptcy prediction literature, I employ a weighted ensemble model that aggregates predictions from multiple machine learning architectures. This approach addresses the limitations of traditional linear models by capturing the complex, non-linear interactions inherent in banking data.

The model is trained on a rich dataset of quarterly regulatory filings spanning 2002 to 2023. The target variable is the market-implied probability of default derived from the CDS spreads of the largest U.S. banks. By mapping the accounting fundamentals of these large banks to their market-priced risk, the model learns a transfer function that can be applied to smaller banks to estimate their default probability. This methodological contribution allows for the first comprehensive analysis of deposit elasticity across the entire size spectrum of the U.S. banking system.

The empirical analysis validates the flight-to-safety mechanism predicted by the theoretical model, particularly through the quantity channel. As presented in Table III, I find a stark hierarchy in how uninsured depositors respond to bank distress in the post-GFC regime. With local banks as the reference category, the post-GFC triple interaction terms for both regional banks and GSIBs are positive and statistically significant, indicating that larger institutions attract rather than repel uninsured capital as their default risk rises. The effect is largest for GSIBs: a one-percentage-point increase in default probability is associated with an approximately 48 percent increase in a GSIB's uninsured deposit share relative to the outside option, while local banks experience an essentially flat or slightly negative relationship. Crucially, this sensitivity is effectively absent for insured deposits, isolating the mechanism to those liabilities most exposed to credit risk, and the result survives a within-bank triple-difference that controls for time-varying bank quality.

The results for the pricing channel, presented in Table IV, offer corroborating but statistically weaker evidence. While the model predicts that GSIBs should be able to lower their deposit rates during distress, the empirical identification of this effect is constrained by the structural characteristics of the deposit market. As shown by the high R^2 values exceeding 0.99 in the spread regressions, the vast majority of variation in deposit rates is absorbed by time and bank fixed effects, leaving little residual variation to be explained by idiosyncratic risk. This lack of statistical power is consistent with the well-documented "stickiness" of deposit rates Drechsler, Savov, and Schnabl (2017); because banks are hesitant to adjust rates frequently due to administrative costs and customer inertia, the primary margin of adjustment during stress events is deposit flows rather than interest rates. Consequently, while the signs of the coefficients align with the TBTF distortion, suggesting GSIBs enjoy a pricing advantage, the flight-to-safety is principally observed through the massive reallocation of volume shown in Table III.

The remainder of this paper is organized as follows. Section II reviews the related literature on banking stability, market discipline, and TBTF distortions. Section III develops the continuous-time structural model of deposit competition and derives the equilibrium pricing conditions. Section

IV details the empirical strategy, including the construction of the machine learning default model and the difference-in-differences specification. Section V presents the main results. Section VI concludes with a discussion of policy implications.

II. Related Literature

The theoretical foundation of banking fragility rests on the seminal model of Diamond and Dybvig (1983), who characterize bank runs as equilibrium outcomes driven by self-fulfilling panic rather than solely by fundamental insolvency. While this framework supports multiple equilibria, Goldstein and Pauzner (2005) refine the theory using global games to demonstrate that runs are triggered when fundamentals deteriorate below a specific threshold, re-establishing the link between asset quality and depositor flight. This transmission of distress is not purely financial; Kelly and Gráda (2000) and Iyer and Puri (2012) document that social networks amplify contagion, a dynamic accelerated by modern technology as evidenced by the "digital runs" documented by Cookson, Fox, Gil-Bazo, Imbet, and Schiller (2023).

A structural assessment of this fragility is provided by Egan et al. (2017), who confirm that the U.S. banking sector is characterized by multiple equilibria. Their work reveals a stark bifurcation in market discipline: while insured depositors are insensitive to risk, the demand for uninsured deposits is highly elastic. Their estimates suggest that a one percentage point increase in a bank's risk-neutral probability of default leads to a 12% decline in its market share of uninsured deposits. This elasticity creates a "bad" equilibrium where fear of default becomes self-fulfilling. Their counterfactual analysis of Wachovia Bank in 2008 illustrates this mechanism, while Wachovia's fundamentals implied a default probability of only 3.3%, the authors identified an alternative equilibrium consistent with the same fundamentals where the default probability was 52%. This finding is critical to this paper, as it suggests that smaller banks without implicit guarantees remain structurally exposed to this "fragile" equilibrium.

While government guarantees stabilize the system against such runs, they introduce profound moral hazard. Stern and Feldman (2004) argue that TBTF subsidies distort risk pricing, encouraging banks to leverage up and capture the upside of risky bets while taxpayers bear the tail risk. This subsidy is empirically robust across asset classes. In the bond market, Santos (2014) provides evidence that large banks enjoy a significant funding cost advantage that is counter-cyclical; the spread between TBTF and non-TBTF costs widens dramatically during crises when the guarantee is most valuable. **?** corroborate this, finding that prior to the Dodd-Frank Act, bond spreads for large financial institutions were remarkably insensitive to risk, implying a near-total implicit guarantee.

This distortion extends to equity markets and corporate strategy. Gandhi and Lustig (2015) identify a size anomaly where large bank stocks yield significantly lower risk-adjusted returns than small bank stocks. They argue this reflects an "insurance premium" investors pay for the government's tail-risk protection, which accrues to the bank as a lower cost of equity capital. Consequently,

banks are incentivized to pursue scale not for efficiency, but to secure this subsidy; Penas and Unal (2004) document that bond returns following mergers reflect the attainment of "TBTF status" rather than purely economic synergies.

Recent literature debates whether post-2008 regulatory reforms have successfully eliminated these distortions. Berndt et al. (2025) provide a current assessment, utilizing a reduced-form asset pricing model to decompose credit spreads into insolvency risk and bailout probability. They find that while the risk-neutral probability of not getting bailed out rose from approximately 5% pre-crisis to 13% post-crisis, the market still assigns an overwhelming 87% probability that the government will save a GSIB. They estimate the present value of this remaining subsidy at 1.6% of GDP. This persistent expectation confirms that despite statutory changes, the flight to safety dynamic remains a potent competitive force, driving uninsured deposits toward GSIBs during stress events.

Identifying this safety premium is complicated by non-regulatory factors. d'Avernas et al. (2023) argue that the funding advantage of large banks may stem from a sophistication paradox, where wealthy clients accept lower rates in exchange for superior mobile technology and liquidity services. This paper addresses this identification challenge by exploiting the temporal variation in distress; while technological advantages are static, the safety premium is counter-cyclical. Finally, to operationalize this analysis for smaller banks lacking market data, we build upon the bankruptcy prediction literature, evolving from the static discriminant analysis of **?** and **?** to modern Machine Learning techniques that better capture non-linear distress signals.

III. Model

In this section, we develop a structural model of deposit competition that integrates the spatial differentiation framework of Hotelling (1929) with the endogenous default risk mechanics of Merton (1974) and Leland (1994). The primary objective is to formalize the mechanism through which TBTF guarantees distort deposit pricing. By moving to a continuous-time setting, we can analyze how the equilibrium deposit spread evolves dynamically as a function of the smaller bank's distance-to-default.

A. Model Setup

Let $(\Omega, \mathbb{F})$ be the given probability space and canonical filtration. We consider a continuous-time, infinite-horizon economy populated by a unit mass of risk-neutral agents. The banking sector consists of two competing depository institutions: a National Bank (N) and a Local Bank (L). Following the standard spatial competition literature, these banks are located at the extremes of a unit interval $[0, 1]$, with bank N at $x = 0$ and bank L at $x = 1$.

The fundamental asymmetry in the model arises from the regulatory treatment of these institutions. The National Bank is designated as a Systemically Important Financial Institution (SIFI). Consistent with the "implicit guarantee" literature (e.g., O'Hara and Shaw (1990); Gandhi and

Lustig (2015)), we assume this designation renders the National Bank "Too-Big-To-Fail." Consequently, its deposits are perceived by agents as risk-free, regardless of the bank's underlying asset performance. Conversely, the Local Bank does not enjoy this protection. Its deposits are subject to credit risk, and in the event of default, uninsured depositors suffer a loss of principal.

The Local Bank holds a portfolio of risky assets, V_t, which evolves according to a Geometric Brownian Motion (GBM) under the risk-neutral measure $\mathcal{Q}$. Following Merton (1974), the dynamics of the bank's asset value are given by

$$\frac{dV_t}{V_t} = (r_f - \delta)dt + \sigma dZ_t, \tag{1}$$

where r_f is the constant net risk-free rate, δ represents the total payout rate (dividends and coupons) to claimholders, σ is the constant volatility of the asset portfolio, and Z_t is a standard Brownian motion. δ prevents the bank from growing indefinitely and ensures a stationary leverage ratio in the absence of shocks, consistent with the stationary debt structure in Leland (1994). The initial asset value is V_0, and we assume V_0 is sufficiently high such that the bank is initially solvent.

A critical departure from the simple static formulation is the treatment of default. Rather than a fixed exogenous probability, default is an endogenous stopping time driven by the solvency decisions of the Local Bank's shareholders. Following Black and Cox (1976), default occurs the first time the bank's asset value V_t touches a lower barrier V_B. We define the default time τ as

$$\tau := \inf\{t \geq 0 \,:\, V_t \leq V_B\} \tag{2}$$

The barrier V_B is not imposed by debt covenants as in Merton (1974) but is chosen endogenously by equity holders to maximize the value of their claim, a feature central to the trade-off models of Leland (1994) and Leland and Toft (1996). We derive the optimal V_B in Section A.

B. Valuing Default Risk

Before analyzing the deposit market equilibrium, we must establish how risk-neutral depositors value the contingent loss associated with the Local Bank. Depositors face a potential loss given default (LGD) of $\alpha \in (0, 1]$. If the Local Bank defaults at time τ, depositors recover $(1 - \alpha)$ of their deposit face value.

To price this risk, we utilize the standard pricing kernel for a "first-passage-time" security. Let $P(V_t)$ denote the present value of a claim that pays \$1 at the moment default occurs. Specifically, $P(V_t) = \mathbf{E}_{\mathcal{Q}}[e^{-r_f \tau} \mid V_t]$. For the GBM process specified in Equation (1), the closed-form solution for this security is

$$P(V_t) = \left(\frac{V_t}{V_B}\right)^{-\gamma}, \tag{3}$$

where $-\gamma$ is the positive root of the fundamental quadratic equation derived from the generator of the diffusion process

$$\frac{1}{2}\sigma^2\gamma(\gamma+1)+(\delta-r_f)\gamma-r_f=0. \tag{4}$$

The term $P(V_t)$ can be interpreted as the risk-neutral probability of default discounted at the risk-free rate. It captures the "distance-to-default" in value space. As $V_t \to \infty$, $P(V_t) \to 0$, and as $V_t \to V_B$, $P(V_t) \to 1$. This valuation serves as the basis for the "credit premium" depositors demand from the Local Bank.

C. Deposit Market

We model the deposit market using a continuous-time extension of the Hotelling framework. There is a unit mass of uninsured depositors uniformly distributed along the line $x \in [0,1]$. Depositors seek to maximize their instantaneous flow utility, which consists of the interest income net of transport costs and expected credit losses.

Let r_N and r_L denote the instantaneous deposit rates offered by the National and Local Banks, respectively. Depositors at the local bank have their transport cost normalized to 1 and the national bank has a convenience yield $\phi \in (0,1]$. This parameter ϕ proxies for the degree of differentiation between banks, capturing factors such as branch density, digital app quality, or brand loyalty for the national bank d'Avernas et al. (2023).

For a depositor located at x, the flow utility from depositing at the risk-free National Bank is simply the net yield

$$U_N(x) = r_N - \phi x. \tag{5}$$

For the Local Bank, the depositor faces credit risk. The flow utility must account for the instantaneous intensity of the expected loss. We define the instantaneous credit spread required to compensate for the default claim $P(V_t)$ as $\Psi(V_t, V_B)$. Thus, the utility from the Local Bank is

$$U_L(x) = r_L - \alpha\Psi(V_t, V_B) - (1-x). \tag{6}$$

Here, $\alpha\Psi(V_t, V_B)$ represents the flow cost of credit risk. Intuitively, this is the insurance premium a depositor would need to pay in a perfectly competitive market to insure their deposit against the loss α.

The marginal depositor, located at x^*, is indifferent between the two banks when $U_N(x^*) = U_L(x^*)$. Solving this equality for x^* yields the market share of deposits at the national bank,

$$M_N(V_t, V_B) = \frac{r_N - r_L + \alpha\Psi(V_t, V_B) + 1}{1+\phi}, \tag{7}$$

with the local bank taking the other $M_L = 1 - M_N$ mass of deposits.

Equation 7 reveals the central tension of the model. The Local Bank's market share is determined by three forces: (1) the baseline spatial split $(1/(1+\phi))$, (2) the interest rate differential $(r_N - r_L)$, and (3) the TBTF distortion term $\alpha\Psi(V_t, V_B)$. As the Local Bank's asset value V_t

declines, the risk term $\Psi(V_t, V_B)$ rises, forcing the bank to either accept a lower market share or compress its net interest margin by raising r_L.

D. Bank Competition

We assume the Local Bank maximizes shareholder value. This involves two distinct optimization problems: a static decision regarding the deposit rates r_N and r_L to offer at each instant, and a dynamic decision regarding when to exercise the option to default, for the local bank.

Shareholders hold a residual claim on the bank's assets. Let $E(V_t)$ denote the value of the Local Bank's equity. The bank pays a total flow of coupons to depositors, denoted C. For tractability, we approximate the total interest expense as a constant proportional to the steady-state deposit base, $C = r_L\overline{D}$.

The default problem for the shareholders is an optimal stopping problem. In the continuation region $V_t > V_B$, the value of equity must satisfy the standard asset pricing differential equation

$$\frac{1}{2}\sigma^2 V^2 E''(V) + (r_f - \delta)VE'(V) - rE(V) + \delta V - C = 0. \tag{8}$$

The term $(\delta V - C)$ represents the net cash flow to equity holders. To solve this second-order ordinary differential equation (ODE), we impose three boundary conditions standard in the structural credit risk literature Leland (1994)

- Limited Liability: At the default barrier, equity value is zero

$$E(V_B) = 0.$$

- Solvency: As asset values grow arbitrarily large, the probability of default vanishes, and equity value approaches the risk-free value of assets net of debt

$$\lim_{V\to\infty} E(V) = V - \frac{C}{r_f}.$$

- Smooth Pasting: Shareholders choose the default barrier V_B optimally. This requires the derivative of the equity value to be continuous at the boundary

$$E'(V_B) = 0.$$

Solving the Euler-Cauchy ODE subject to the first two boundary conditions yields the general solution

$$E(V) = V - \frac{C}{r_f} + \left(\frac{C}{r_f} - V_B\right)\left(\frac{V}{V_B}\right)^{-\gamma}. \tag{9}$$

Applying the smooth pasting condition, we derive the closed-form solution for the endogenous default barrier

$$V_B = \frac{C}{r_f}\left(\frac{\gamma}{\gamma+1}\right). \tag{10}$$

Equation 10 provides a micro-founded threshold for default. The term $\frac{\gamma}{\gamma+1} < 1$ indicates that shareholders will keep the bank operating even when the value of assets falls below the face value of liabilities. This "gambling for resurrection" behavior arises from the option value of limited liability.

We assume banks compete in rates continuously to maximize instantaneous expected profits, taking the asset value V_t and the default barrier V_B as given. The profit maximization problem for bank $k \in \{N, L\}$ is

$$\max_{r_k \in \mathbb{R}_{++}} \Pi_k \equiv (R - r_k)M_k \tag{11}$$

where R is the expected gross return on assets.

The FOCs for the bank optimization problem relates the marginal benefit of attracting more depositors and the marginal cost of the additional cost of capital that each depositor earns.

E. Equilibrium

Combining the demand side and supply side, we arrive at the main theoretical result of the paper, the Nash equilibrium deposit spread. In a symmetric Hotelling model without risk, the equilibrium spread only depends on the amenity value that the national bank offers. However, the presence of the TBTF distortion increases the deposit spread. The equilibrium condition requires the Local Bank to fully compensate the marginal depositor for the expected credit loss to maintain its market share.

Proposition 1 (Equilibrium Deposit Rates):
In the Nash equilibrium, the optimal deposit rates set by the National Bank (r_N^*) and the Local Bank (r_L^*) are given by:

$$r_N^*(V_t, V_B) = R - \frac{2+\phi}{3} - \frac{\alpha}{3}\Psi(V_t, V_B) \tag{12}$$

$$r_L^*(V_t, V_B) = R - \frac{1+2\phi}{3} + \frac{\alpha}{3}\Psi(V_t, V_B) \tag{13}$$

where $\Psi(V_t, V_B)$ is the instantaneous credit yield derived from the structural model

$$\Psi(V_t, V_B) = r\left(\frac{(V_t/V_B)^{-\gamma}}{1-(V_t/V_B)^{-\gamma}}\right).$$

As V_t approaches V_B, the term $(V_t/V_B)^{-\gamma}$ approaches 1, causing the denominator in Ψ to vanish and the spread $\Psi(V_t, V_B)$ to diverge. This convexity is a key prediction of the model; the TBTF subsidy is relatively small during good economic times but expands non-linearly during periods of

financial distress.

The equilibrium deposit rates exhibit a distinct "scissors" dynamic that reflects the strategic exploitation of the TBTF subsidy. In the absence of credit risk, both banks set rates based on the marginal product of capital, R, markdown by the spatial differentiation parameter, ϕ, which serves as a measure of the National Bank's convenience yield. However, as the Local Bank's distance-to-default shortens, the pricing strategies diverge. The Local Bank is forced to raise its rate, r_L, to compensate the marginal depositor for the escalating intensity of expected credit losses, $\alpha\Psi(V_t, V_B)$. Simultaneously, the National Bank optimally lowers its rate, r_N. This markdown occurs because the Local Bank's increasing risk profile reduces its substitutability, effectively softening the competitive constraint on the National Bank; the TBTF incumbent exploits this reduced demand elasticity to widen its net interest margin, monetizing the flight to safety by offering lower yields while still retaining market share.

Corollary 1 (Equilibrium Market Structure and Profitability): *Given the Nash equilibrium deposit rates r_N^* and r_L^*, the resulting equilibrium deposit spread S^*, market shares M_k^*, and profit functions Π_k^* are characterized as follows:*

(i) The equilibrium deposit rate spread between the local and national banks is

$$S^*(V_t, V_B) := r_L^* - r_N^* = \frac{1-\phi}{3} + \frac{2}{3}\alpha\Psi(V_t, V_B). \tag{14}$$

(ii) The mass of depositors retained by the National Bank (M_N^*) and the Local Bank (M_L^*) are:

$$M_N^*(V_t, V_B) = \frac{2+\phi+\alpha\Psi(V_t, V_B)}{3(1+\phi)} \qquad M_L^*(V_t, V_B) = \frac{1+2\phi-\alpha\Psi(V_t, V_B)}{3(1+\phi)}.$$

(iii) The equilibrium profits for each bank are given by

$$\Pi_N^*(V_t, V_B) = (1+\phi)\left(\frac{2+\phi+\alpha\Psi(V_t, V_B)}{3(1+\phi)}\right)^2$$

$$\Pi_L^*(V_t, V_B) = (1+\phi)\left(\frac{1+2\phi-\alpha\Psi(V_t, V_B)}{3(1+\phi)}\right)^2$$

It is important to note that the model dynamics permit corner solutions during periods of extreme systemic stress. As the Local Bank's probability of default rises sufficiently high, the flight-to-safety becomes absolute, driving the National Bank's market share toward unity, $M_N^* \to 1$. In this limit, the Local Bank is effectively crowded out of the funding market; its deposit base evaporates, and its expected profits collapse to zero, precipitating an immediate default and market exit. Simultaneously, having captured the entire market, the National Bank transitions to an unconstrained monopoly pricing strategy. Crucially, the model assumes that depositors face a prohibitive cost to accessing outside options, such as money market funds or physical storage (the "mattress"). Without a viable alternative store of value to enforce a participation constraint, the National Bank can optimally set deposit rates to zero or negative levels to extract maximal

consumer surplus from the captive depositor base, monetizing the systemic instability.

The decomposition of the equilibrium deposit spread in (i) reveals that the pricing differential between GSIBs and smaller institutions is driven by two distinct economic forces: a structural convenience yield and a state-dependent risk premium. The first component, $(1-\phi)/3$, represents the baseline "convenience premium" that the National Bank enjoys solely due to its superior network density and technological sophistication; even in the absence of default risk, the Local Bank must pay a premium to overcome the National Bank's non-pecuniary amenity value. The second component, $\frac{2}{3}\alpha\Psi$, captures the TBTF distortion. Notably, the pass-through of credit risk to the spread is incomplete (two-thirds rather than unity) because the Nash equilibrium involves mutual adjustment; the Local Bank absorbs one-third of the shock through margin compression to retain depositors, while the National Bank strategically accommodates the remaining two-thirds by widening the spread from the bottom, lowering its own rate to exploit its rival's distress.

This decomposition provides a structural interpretation for the behavior of insured versus uninsured deposit spreads. For fully insured deposits, the depositor faces no credit loss in the event of default, implying an effective loss-given-default of zero ($\alpha = 0$). In this limiting case, the risk premium term vanishes $\frac{2}{3}\alpha\Psi \equiv 0$, and the deposit spread collapses to the constant convenience yield, $(1-\phi)/3$. This result aligns with the findings of d'Avernas et al. (2023), who argue that the persistent spread between large and small banks is driven by differences in "banking services" and technology rather than credit risk. Our model nests this service quality view within the convenience parameter ϕ but demonstrates that for uninsured depositors, who cannot ignore Ψ, the TBTF subsidy drives a wedge beyond what quality differences alone can explain, particularly during periods of financial distress.

Item (ii) quantifies the flight-to-safety dynamic, demonstrating that the National Bank's market share expands linearly with the Local Bank's credit intensity, Ψ, despite the National Bank offering lower deposit rates. This result highlights a breakdown in the standard price elasticity of demand during periods of stress; as the probability of default rises, the safety attribute of the National Bank becomes the dominant factor in depositor utility, overwhelming the yield differential. Consequently, depositors effectively become price-insensitive to the National Bank's rate cuts, allowing the TBTF incumbent to hoard liquidity precisely when its competitor is experiencing a runoff. This mechanism formalizes the empirical observation that deposit flows during financial crises are driven not by the pursuit of yield, but by the avoidance of tail risk, granting TBTF institutions a distinct funding advantage that accelerates as systemic stress deepens.

Finally, the profit functions in (iii) illustrate the convex nature of the TBTF subsidy. Because equilibrium profits are proportional to the square of market share, the National Bank acts as the beneficiary of its competitor's fragility, experiencing a compounding increase in profitability as the Local Bank approaches the default barrier. This double dipping, gaining both higher deposit volume and wider net interest margins, implies that TBTF protections do not merely shield large banks from their own failures, but actively transfer franchise value from distressed smaller banks to the protected sector. Conversely, the Local Bank faces a convex penalty; as it attempts to defend

its deposit base with higher rates, its net interest margin collapses faster than its market share, potentially accelerating the depletion of equity capital and creating a self-reinforcing death spiral that justifies the endogenous decision to default earlier than book-value insolvency would suggest.

The following comparative statics elucidate the economic mechanisms driving the flight-to-safety dynamics and the competitive distortion introduced by TBTF guarantees.

The most direct determinant of the deposit spread is the financial health of the Local Bank, proxied by its current asset value V_t. Differentiating the equilibrium spread equation with respect to V_t yields

$$\frac{\partial S^*}{\partial V_t}(V_t, V_B) = \frac{2}{3}\alpha \frac{\partial \Psi}{\partial V_t}(V_t, V_B) < 0.$$

Since the credit intensity Ψ is proportional to the first passage time probability $(V_t/V_B)^{-\gamma}$, and the exponent $-\gamma$ is negative, this derivative is strictly negative. Mathematically, this confirms that as the Local Bank's asset value declines towards the endogenous default barrier V_B, the deposit spread widens non-linearly. This result generates the flight-to-safety dynamic observed in the empirical data. In "good times" (when $V_t \gg V_B$), the default probability is negligible, and the spread is driven almost entirely by the structural convenience yield, $(1-\phi)/3$. However, as the bank enters distress, the TBTF distortion activates. The convexity of the function implies that a small deterioration in asset value for a healthy bank has little impact on funding costs, whereas the same deterioration for a distressed bank causes a sharp spike in the spread. Importantly, in our Nash equilibrium, this widening is driven by two simultaneous reactions: the distressed Local Bank raising rates to retain wary depositors, and the TBTF National Bank opportunistically lowering rates to exploit its rival's weakness.

While the analysis of deposit spreads illustrates how prices adjust to distress, the sensitivity of market shares, M_k^*, reveals the flow of funds and the resulting concentration of systemic liquidity. Differentiating the equilibrium market share of the National Bank, M_N^*, with respect to the Local Bank's asset value yields

$$\frac{\partial M_N^*}{\partial V_t}(V_t, V_B) = \frac{\alpha}{3(1+\phi)} \frac{\partial \Psi}{\partial V_t}(V_t, V_B) < 0.$$

Since $\partial_{V_t}\Psi < 0$ the inverse relationship implies that as the Local Bank's asset value falls (distress increases), the National Bank's market share expands. This result demonstrates that the price mechanism is insufficient to clear the market during a crisis. In a standard competitive market, the Local Bank would raise rates until demand stabilizes. However, in our Nash equilibrium, despite the Local Bank raising rates and the National Bank lowering them, the Local Bank still suffers a net outflow of deposits. This indicates that the demand elasticity with respect to safety overwhelms the elasticity with respect to yield. The National Bank acts as a liquidity sink, absorbing deposits purely due to its TBTF status. This dynamic suggests that TBTF distortions lead to pro-cyclical deposit concentration: precisely when the banking system is most fragile, funding flees the periphery and concentrates in the systemic core, potentially exacerbating TBTF risks by making the largest

institutions even larger during downturns.

The parameter ϕ captures the convenience yield or non-pecuniary amenity value of the National Bank, such as a superior branch network or digital infrastructure. We examine the sensitivity of both the equilibrium spread and market share to this structural friction

$$\frac{\partial S^*}{\partial \phi}(V_t, V_B) = -\frac{1}{3} < 0 \qquad \frac{\partial M_N^*}{\partial \phi}(V_t, V_B) = -\frac{1 + \alpha\Psi(V_t, V_B)}{3(1+\phi)^2} < 0.$$

The negative derivative of the spread isolates the structural component of the National Bank's funding advantage; a lower ϕ (indicating higher convenience) increases the baseline spread $(1-\phi)/3$, allowing the National Bank to pay consistently lower rates independent of credit risk. Simultaneously, the sensitivity of market share reveals a crucial interaction effect: the presence of the risk term $\alpha\Psi$ in the numerator implies that the volume gains from superior technology are magnified during periods of distress. When depositors are sensitized to default risk, they become hyper-sensitive to frictions. Consequently, large banks that invest in friction-reducing technologies secure a double dividend, capturing a disproportionately larger share of flight-to-safety flows while maintaining a wider structural wedge in their cost of funds.

The volatility of the bank's underlying assets, σ, affects the spread through two opposing channels: the kinetic effect, which increases the probability of hitting the barrier, and the strategic effect, which lowers the optimal default barrier V_B as shareholders delay default to preserve option value. The net effect on the spread is positive

$$\frac{\partial S^*}{\partial \sigma}(V_t, V_B) > 0.$$

This static highlights how aggregate uncertainty exacerbates TBTF distortions. During periods of high market volatility, such as the Global Financial Crisis or the COVID-19 pandemic, the TBTF National Bank gains a competitive advantage even if the Local Bank's expected asset value remains constant. Depositors, fearing that higher variance increases the likelihood of a left-tail event, demand a higher risk premium from the unprotected Local Bank. Meanwhile, the TBTF bank, shielded by the implicit guarantee, sees its funding costs remain insensitive to the volatility shock. This suggests that TBTF policies distort competition most severely during turbulent market regimes.

The parameter α represents the Loss Given Default (LGD) faced by uninsured depositors at the Local Bank. The sensitivity is straightforward

$$\frac{\partial S^*}{\partial \alpha}(V_t, V_B) = \frac{2}{3}\Psi(V_t, V_B) > 0.$$

This static formalizes the direct link between regulatory credibility and market pricing. The magnitude of the TBTF distortion scales linearly with the expected loss severity. If regulators signal a credible commitment to impose losses on uninsured depositors at small banks, increasing α, the spread widens. Conversely, if a blanket guarantee is effectively extended to all institutions, as seen in the immediate aftermath of the Silicon Valley Bank failure FDIC (2023) $\alpha \to 0$, the

risk-driven component of the spread collapses. This implies that ambiguity regarding the scope of deposit insurance is a primary driver of the cost of capital differential between large and small banks.

This discussion elucidates the economic mechanisms driving the flight-to-safety. The model is calibrated using an insured deposit spread of 17 basis points and a historical uninsured recovery rate of $1 - \alpha = 91\%$ found in Perry (2026). In this analysis, the probability of default proxies for aggregate sector distress. Crucially, the model features asymmetric risk exposure: while the Local Bank's solvency is tethered to this systemic risk, the National Bank's TBTF designation effectively inoculates it from asset volatility.

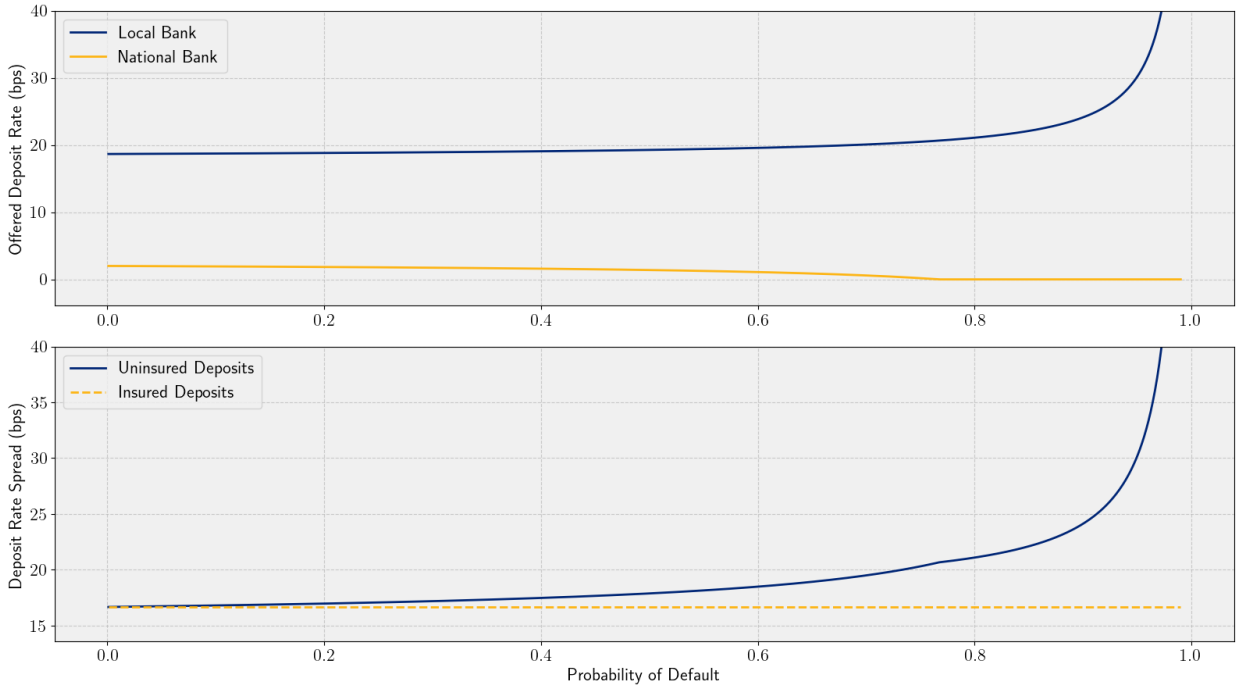


Figure 2. The top panel plots the equilibrium offered deposit rates in basis points for the Local Bank (blue line) and National Bank (yellow line) as a function of the Local Bank's probability of default. The bottom panel displays the resulting deposit rate spread ($r_L^* - r_N^*$) for uninsured deposits (blue line) and insured deposits (yellow dashed line) across the range of default probabilities.

The bifurcation in equilibrium pricing strategies is vividly illustrated in Figure 2. The data reveals a distinct "scissors" dynamic predicted by Proposition 1. As the probability of default rises, the Local Bank is compelled to aggressively raise its offered rate to compensate uninsured depositors for the escalating expected loss. Conversely, the National Bank responds to this same distress by lowering its deposit rate. This strategic markdown occurs because the Local Bank's increasing risk profile reduces its substitutability, softening the competitive constraint on the incumbent. Consequently, in the region of extreme distress where the default probability approaches unity, the National Bank can optimally effectively zero-out its deposit rate. In this regime, the safety attribute dominates the depositor's utility function, allowing the TBTF bank to retain funds while paying negligible pecuniary returns.

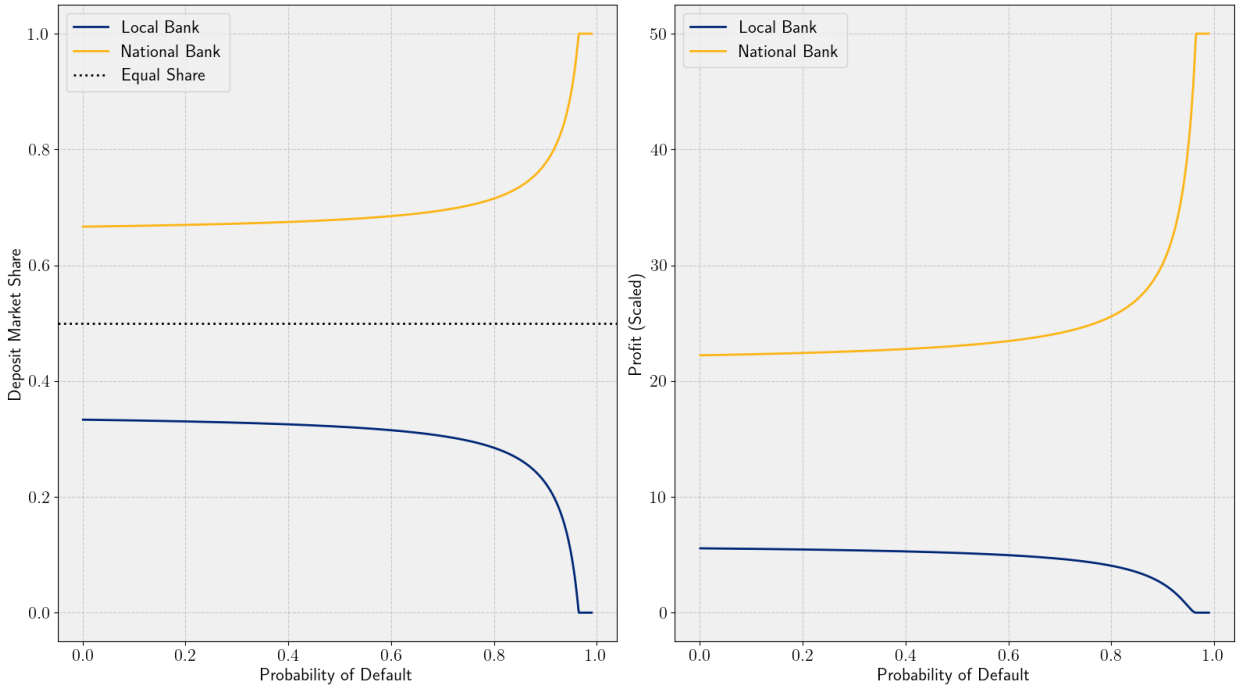


Figure 3. The left panel plots the equilibrium deposit market shares for the Local Bank and National Bank as a function of the probability of default, with the dotted line indicating an equal 50% share. The right panel plots the scaled equilibrium profit functions for both institutions against the same domain of default probabilities.

Figure 3 translates these pricing dynamics into equilibrium market shares and profitability, highlighting the quantity channel of the distortion. The left panel demonstrates that despite the Local Bank's efforts to defend its funding base through higher rates, it suffers a monotonic loss of market share as risk rises. This implies that demand elasticity with respect to safety overwhelms elasticity with respect to yield; during crises, price adjustments fail to clear the market for unprotected banks. The right panel underscores the destabilizing implications for profitability. While the Local Bank's profits collapse toward zero as it faces both volume outflows and margin compression, the National Bank's profits exhibit a convex, counter-cyclical profile. By capturing cheap flight-to-safety flows, the TBTF institution enjoys a profit boost precisely when the sector is failing, effectively transferring franchise value from the periphery to the core.

To examine these mechanics in a continuous setting, Figure 4 presents a single simulation path of the sectors asset value.

Figure 4 confirms that the safety premium is not static but behaves as a highly volatile, state-dependent variable. During periods of relative calm (e.g., years 0–1), the spread remains compressed, primarily reflecting the baseline convenience yield ϕ. However, as asset values deteriorate and the probability of default spikes, the spread widens non-linearly. This convexity highlights the option-like nature of the implicit guarantee; the cost of being small is negligible in benign economic environments but expands dramatically during distress episodes, forcing the spread to widen exactly when the Local Bank can least afford it.

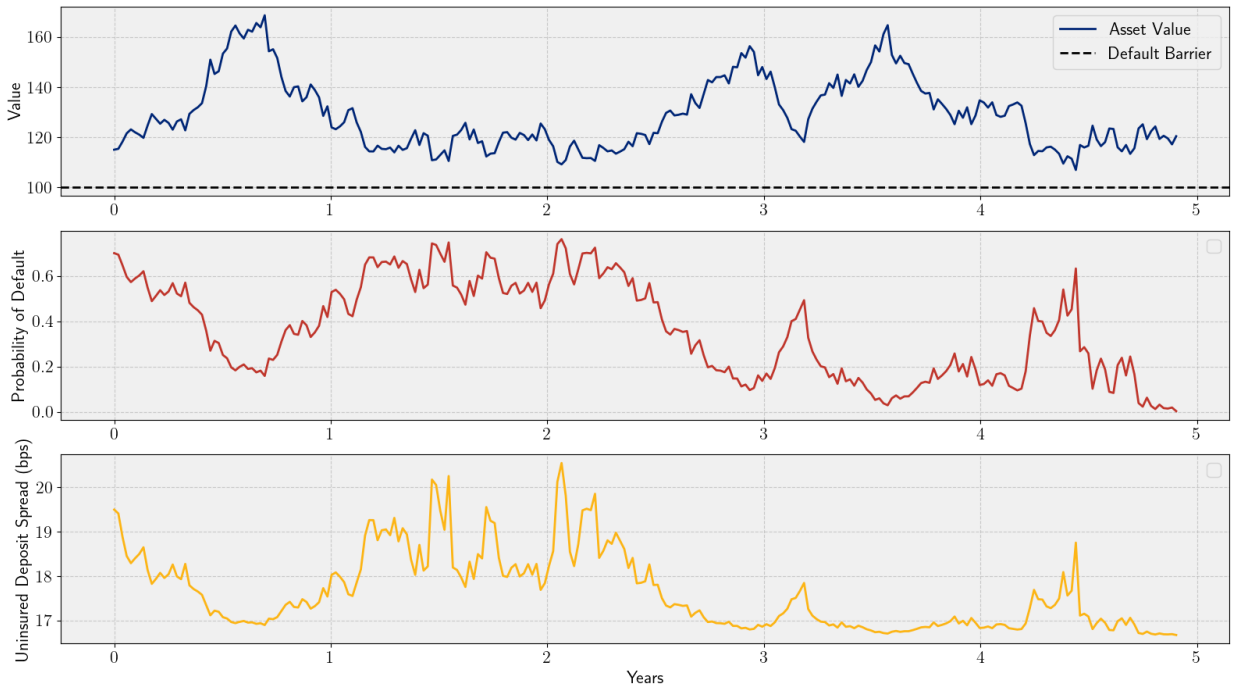


Figure 4. A single stochastic simulation of the economy over a five-year horizon. The top panel displays the path of the banking sector's asset value (V_t) relative to the constant default barrier (V_B, dashed line). The middle panel shows the corresponding instantaneous probability of default derived from the structural model. The bottom panel plots the resulting equilibrium spread for uninsured deposits in basis points.

The raw time-series dynamics presented in Figure 5 provide a granular view of the scissors mechanism in operation. As economic conditions deteriorate, evidenced by the spikes in default probability, we observe a simultaneous divergence in bank behavior: the National Bank lowers rates and gains deposit inflows, while the Local Bank raises rates and suffers outflows. This inverse correlation confirms that the theoretical predictions hold dynamically; the National Bank effectively monetizes the systemic instability by widening its net interest margin during the very periods where its competitor is fighting for liquidity.

Finally, Figure 6 scales profits to reveal a positive feedback loop. Because the National Bank generates windfall profits during "bad times" the TBTF incumbent accumulates capital at a faster rate over the business cycle. This dynamic suggests that the current regulatory architecture creates an evolutionary pressure toward industry consolidation. Over time, the repeated profit boosts enjoyed by TBTF banks during downturns could drive the market toward an equilibrium where all surviving banks are systemically important, not due to superior efficiency, but because the implicit guarantee allows them to cannibalize the profitability of smaller rivals during every contraction. This mechanism is exemplified by historical consolidations, such as Wells Fargo's acquisition of Wachovia during the 2008 Great Financial Crisis and J.P. Morgan Chase's purchase of First Republic Bank in 2023, where TBTF incumbents absorbed distressed competitors and expanded their market dominance amidst systemic fragility.

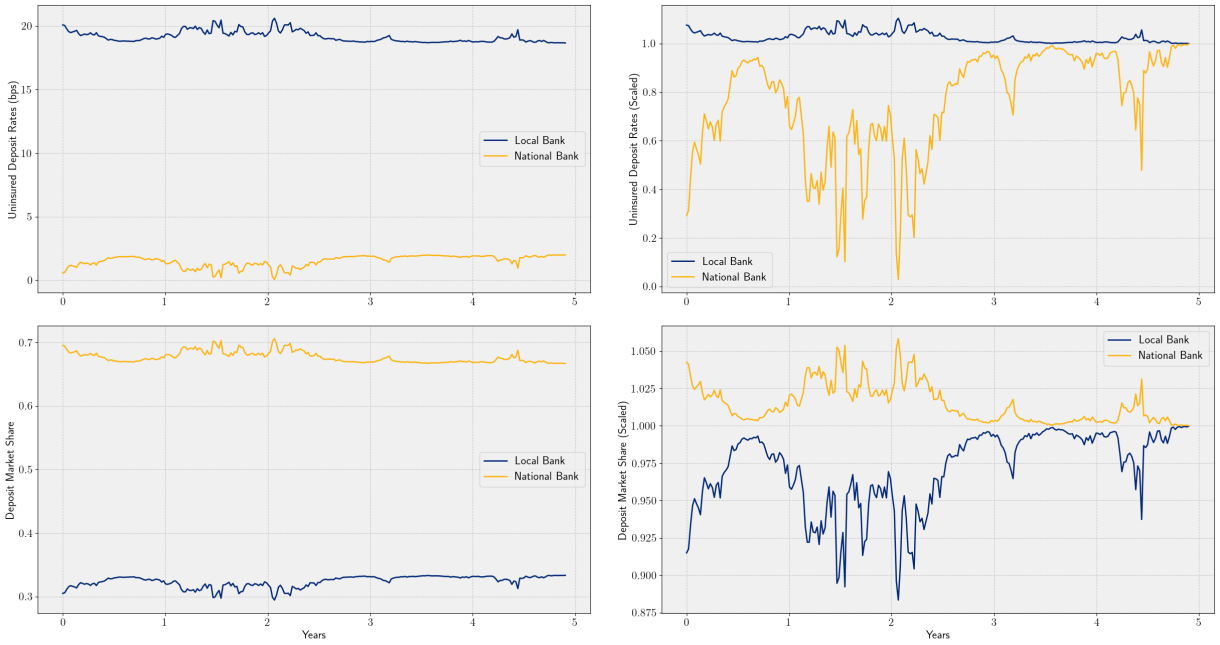


Figure 5. Time-series realizations of equilibrium pricing and quantity metrics corresponding to the simulation path in Figure 4. The top-left panel plots the uninsured deposit rates (bps) for the Local and National Banks. The bottom-left panel plots the deposit market shares for both banks. The right column displays the scaled versions of the deposit rates (top-right) and market shares (bottom-right) over the same five-year period. Scaled by the corresponding metric when the probability of default is zero.

IV. Empirics

To estimate the elasticity of deposit demand with respect to bank failure risk, we require a continuous measure of the probability of default for every commercial bank in the United States. While the default probabilities of Globally Systemically Important Banks (GSIBs) can be inferred directly from the market prices of Credit Default Swaps (CDS), no such market exists for the vast majority of regional and local banks. Consequently, we must construct a proxy for these unobservable default probabilities.

We approach this as a supervised learning problem. We operate under the structural assumption that the mapping between observable bank accounting variables, derived from regulatory filings, and default risk is consistent across bank sizes, conditional on a sufficiently rich feature set. The estimation procedure follows a multi-stage framework designed to balance the bias-variance trade-off inherent in high-dimensional financial data. First we filter the feature space using Elastic Net regularization, then we utilize a Weighted Ensemble model that aggregates predictions from a diverse set of base learners, including Random Forests and Gradient Boosted Decision Trees (GBM).

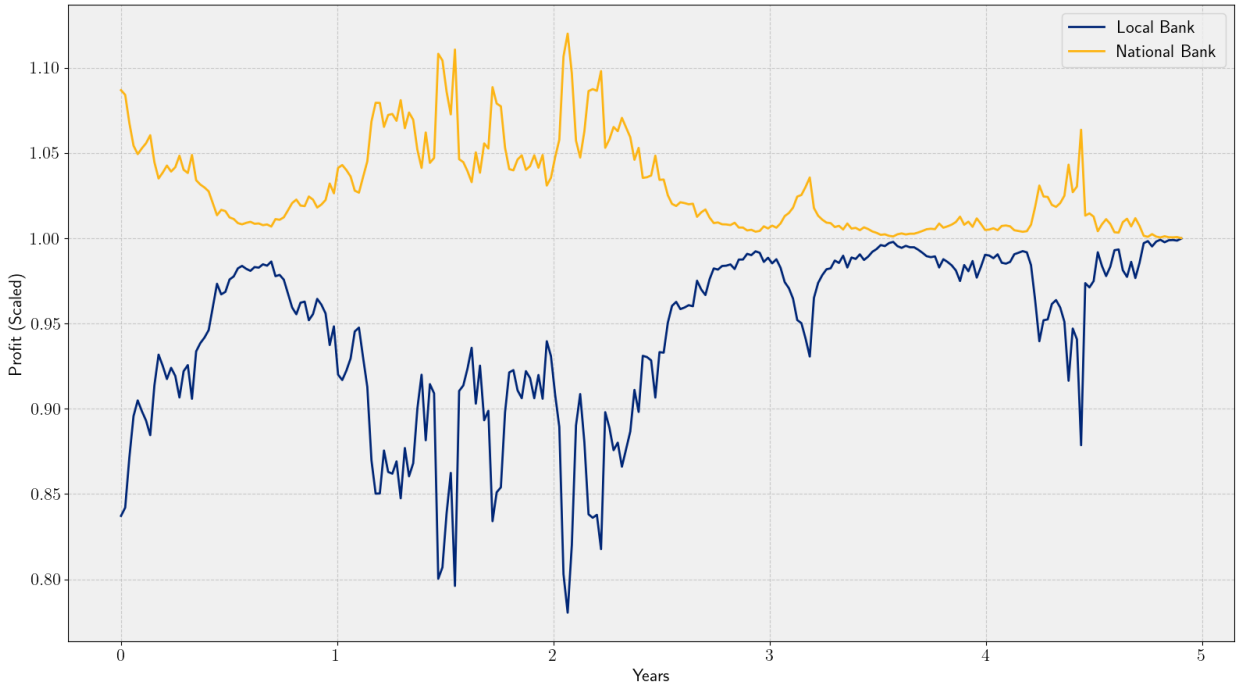


Figure 6. The plot displays the time-series evolution of scaled profits for the Local Bank (blue line) and National Bank (yellow line) over the five-year simulation period shown in Figure 4. Scaled by the profit when the probability of default is zero.

A. Data

The primary dataset is constructed at a quarterly frequency spanning from 2002 to 2023. This period encompasses several distinct phases of the banking cycle, including the pre-crisis expansion, the Global Financial Crisis (GFC), the zero-lower-bound (ZLB) era, and the post-COVID inflationary period. We integrate data from three primary sources to construct the training and prediction sets.

First, to proxy for default risk, we utilize 5-year CDS spreads sourced from Markit. The raw data consists of daily spread observations for the 17 largest U.S. commercial banks. To align this high-frequency market data with quarterly regulatory filings, we take the average CDS spread of the final month of each quarter. This "end-of-quarter" alignment ensures that the market signal reflects the information set available at the time of the regulatory report.

Second, granular balance sheet and income statement data are sourced from the Federal Financial Institutions Examination Council's (FFIEC) Reports of Condition and Income ("Call Reports") and the Uniform Bank Performance Report (UBPR). The UBPR serves as a cleaned and extended version of the Call Reports, providing pre-calculated regulatory ratios and peer group averages that are critical for benchmarking bank performance.

Third, to ensure our analysis covers the full spectrum of the U.S. banking system, we incorporate data from the Statistics on Depository Institutions (SDI), which covers 6,980 savings institutions. To ensure consistency in the unit of observation, all data—including the 2,464 commercial banks

in the training set—are aggregated to the highest Bank Holding Company (BHC) level. This aggregation aligns with the regulatory treatment of TBTF designations, which are applied at the holding company level.

Following Begenau and Stafford (2023), we define the transaction deposit rate as the interest expense on non-time domestic deposits divided by the quarterly average of non-time domestic deposits. This precise definition isolates the cost of funding for the most liquid, flight-prone liabilities, which are the primary focus of our elasticity analysis.

The target variable for our training set is the implied probability of default derived from the CDS spreads of the 17 largest banks. Following the methodology of Hull (2015), we extract the risk-neutral default probability from the credit spread.

A CDS is priced so that, at inception, the expected present value (EPV) of the premium leg equals the EPV of the protection leg under the risk-neutral measure. Assuming continuous discounting, exponential default density, and payment dates of $0 < t_1 < \cdots < t_n = T$ the pricing relationship is

$$S\left[\sum_{i=1}^{n}(t_i - t_{i-1})e^{-(r_f+\lambda)t_i} + \sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(t - t_{i-1})\lambda e^{-(r_f+\lambda)t}dt\right] = (1-R)\int_0^T \lambda e^{-(r_f+\lambda)t}dt$$

where S is the CDS spread and R is the recovery rate for the insurance contract. The bracketed term is the PV of contractual premium payments plus the PV of accrued premium paid at default; the right-hand side is the PV of the default-contingent protection payment $(1 - R)$.

We solve the equation above numerically for λ. The implied risk-neutral default probability over horizon T is then

$$\rho(T) = 1 - e^{-\lambda T}.$$

This delivers a market-implied, risk-neutral measure of default risk that we use as the supervised learning target. Over the sample the average 5-year market implied default probability is 1.4%.

Observations: 1124

Variable	Mean	Std. Dev.	Min	Median	Max
Total Deposits ($Bn)	334.792	439.185	2.684	153.289	2243.484
Insured Deposits ($Bn)	188.581	235.899	2.148	87.376	1250.422
Uninsured Deposits ($Bn)	146.21	215.058	0.536	63.502	1240.76
CDS Spread (bps)	89.19	90.008	4.63	66.885	1297.946
Probability of Default	1.481%	1.5%	0.075%	1.112%	21.707%
Deposit Rate (bps)	39.323	39.678	1.066	23.614	175.232

Table I: Summary statistics for the in-sample training data, covering the 17 largest U.S. commercial banks from 2002 to 2023. Variables reported are total deposits, insured and uninsured deposits (in billions of dollars), CDS spreads (basis points), implied probability of default, and the quarterly deposit rate (basis points).

Table I reports summary statistics for the 1,124 bank-quarter observations that comprise the in-sample training data. The deposit distribution is highly right-skewed: the median institution holds approximately $153 billion in total deposits, while the mean exceeds $335 billion, reflecting balance sheet concentration among the GSIBs. Insured deposits account for roughly 56% of total deposits on average, though this share varies across institutions and over time. The CDS spread and implied default probability distributions exhibit comparable skewness; the median default probability of 1.1% lies well below the sample mean of 1.5%, and the maximum of 21.7% reflects acute GFC-era stress. The average quarterly deposit rate of 39 basis points captures the full span of monetary policy regimes in the sample, from the zero lower bound through the post-COVID tightening cycle.

A critical contribution of this paper is the construction of a high-dimensional feature space capable of capturing the non-linear dynamics of bank failure. We begin with the universe of 2,464 commercial banks that have complete coverage of all 523 features available in the UBPR. While these level variables provide a static snapshot of bank health, they fail to capture the trajectory of financial distress. To address this, we augment the feature space by calculating first order lags and percent changes. First-order lags are included to capture the persistence and mean-reversion in bank performance; while percent changes capture rapid deteriorations in asset quality or liquidity that may not yet be reflected in static capital ratios. Multiple lag and percent change horizons were tested with one order being the most performant.

This augmentation expands the initial dimensionality to 1,569 features. Given the scale variance among banking metrics (e.g., Total Assets vs. Return on Equity), we standardize all features prior to training using Z-score normalization. For every feature j, the observation x_{ij} is transformed as

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j},$$

ensuring that the optimization algorithms in subsequent steps are not biased by the magnitude of the underlying variables.

B. Default Model

The initial feature space 1,569 suffers from severe multicollinearity, a pervasive issue in accounting data where stock and flow variables are mechanically linked. A standard OLS approach in this environment would yield high-variance estimators and poor generalization.

To address this, we employ Elastic Net regularization as a filtering mechanism. Elastic Net is chosen over Lasso or Ridge regression individually because it combines the sparsity-inducing properties of the ℓ_1 penalty (Lasso) with the grouping effect of the ℓ_2 penalty (Ridge). In banking data, where groups of liquidity and capital ratios are highly correlated, Lasso tends to select a single variable from the group arbitrarily. Elastic Net encourages a grouping effect, retaining correlated predictors that collectively inform default risk.

We model the relationship as a linear equation $\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$. The estimator $\hat{\boldsymbol{\beta}}$ is obtained by

minimizing the residual sum of squares subject to the weighted sum of the ℓ_1 and ℓ_2 norms

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2N} ||\boldsymbol{y} - X\boldsymbol{\beta}||_2^2 + \xi \left(\zeta ||\boldsymbol{\beta}||_1 + \frac{1-\zeta}{2} ||\boldsymbol{\beta}||_2^2 \right) \right\} \tag{15}$$

where the first term is the standard OLS goodness-of-fit, $||\boldsymbol{\beta}||_1$ is the ℓ_1 ("Taxi Cab") norm, and $||\boldsymbol{\beta}||_2^2$ is the ℓ_2 ("Euclidean") norm. The taxi cab norm enforces sparsity by driving coefficients of non-predictive features to exactly zero, while the Euclidean norm stabilizes the solution path and handles collinearity.

The hyperparameters ξ and ζ control the overall strength and mixing of the regularization. We tune these hyperparameters via 5-fold cross-validation with an 85/15 train test-split. Post-estimation, we perform feature selection by retaining only the subset of variables for which $\hat{\beta}_j \neq 0$. This selection process distills the feature space down to 170 distinct variables, composed of: 75 level variables, 73 lagged variables, and 22 percent change variables. The most predictive features identified by the model include Non-Current Loans to Tier 1 Capital, the Lag of Total Servicing with Recourse to Total Assets, and the Percent Change in Allowance for Credit Losses.

While the Elastic Net provides interpretability and efficient feature selection, it assumes a linear relationship between bank fundamentals and default risk. However, financial distress is inherently non-linear; the marginal impact of a liquidity shock increases exponentially as capital buffers deplete. To capture these non-linearities and high-order interactions, we deploy a Weighted Ensemble model using the AutoGluon framework.

The ensemble aggregates predictions from diverse base learners: Random Forests (RF) and Gradient Boosted Machines (GBM), specifically XGBoost and LightGBM. The core of our ensemble consists of decision tree-based models, which are uniquely suited to capturing threshold effects in financial data (e.g., a bank is safe until its capital ratio drops below a specific regulatory minimum).

The Random Forest component addresses the variance inherent in individual decision trees. We utilize a bagging approach. For a forest of B trees, we generate B bootstrap samples from the training data. A decision tree T_b is grown on each sample. Unlike standard CART trees, the Random Forest introduces additional randomness by considering only a random subset of features at each split candidate. The final prediction is the average of the individual tree predictions

$$\hat{f}_{RF}(\boldsymbol{x}) = \frac{1}{B} \sum_{b=1}^{B} T_b(\boldsymbol{x}; \Theta_b).$$

To reduce bias, we employ Gradient Boosted Decision Trees (GBDT). Unlike Random Forests, which train trees in parallel, GBDTs train trees sequentially. Each new tree attempts to correct the residuals of the previous ensemble.

We define the ensemble model at step K as the sum of K additive functions

$$\hat{y}_i = \eta(\boldsymbol{x}_i) = \sum_{k=1}^{K} g_k(\boldsymbol{x}_i)$$

for $g_k \in \mathbb{G}$ is the space of regression trees. The objective is to minimize a regularized loss function

$$L(\eta) := \sum_i \ell(\hat{y}_i, y_i) + \sum_k \Omega(g_k).$$

Here, ℓ is the mean-square-error loss function measuring the difference between the predicted default probability $\hat{y}_i$ and the target y_i. The term $\Omega(g_k)$ penalizes model complexity to prevent overfitting.

Both XGBoost and LightGBM optimize this objective using a second-order Taylor expansion of the loss function. At iteration t, we seek the function h_t that minimizes

$$\Xi^{(t)} \approx \sum_{i=1}^{N} \left(\ell(y_i, \hat{y}_i^{(t-1)}) + \partial_{\hat{y}_i^{(t-1)}} \ell(y_i, \hat{y}_i^{(t-1)}) h_t(\boldsymbol{x}_i) + \frac{1}{2} \partial^2_{\hat{y}_i^{(t-1)}} \ell(y_i, \hat{y}_i^{(t-1)}) h_t^2(\boldsymbol{x}_i) \right) + \Omega(h_t). \quad (16)$$

This formulation allows the model to utilize the curvature of the loss function, resulting in faster convergence and more precise estimates of the probability of default compared to standard gradient descent.

The final stage of the estimation is the aggregation of the base learners into a single predictive model. We employ a stacking strategy via the Weighted Ensemble algorithm. Let $\hat{g}_{ij}$ denote the prediction of base learner $j \in \{1, \ldots, M\}$ for observation i. The final predicted default probability $\hat{\rho}_i$ is a convex combination of the base learners

$$\hat{\rho}_i = \sum_{j=1}^{M} w_j \hat{g}_{ij}.$$

The optimal weights vector $\mathbf{w} \in \mathbb{R}^M$ is determined by solving a constrained least squares optimization problem on a hold-out validation set with N_v observations. We minimize the MSE

$$\min_{\mathbf{w} \in \mathbb{R}^M} \frac{1}{N_v} \sum_{i=1}^{N_v} \left(y_i - \sum_{j=1}^{M} w_j \hat{g}_{ij} \right)^2 \quad \text{subject to} \quad \sum_{j=1}^{M} w_j = 1, \quad w_j \geq 0 \quad \forall j.$$

The non-negativity constraint $w_j \geq 0$ and the summation constraint $\sum w_j = 1$) ensure models are weighted by their validation performance. This reduces the generalization error by averaging out the idiosyncratic errors of diverse models. For instance, if the Random Forest overestimates risk in high-volatility environments while the LightGBM underestimates it, the ensemble average will converge closer to the true probability.

To ensure the stability of the estimated default probabilities, we implement a rigorous validation protocol. The model training utilizes an 85/15 train/test split. Within the training phase, we employ 5-fold cross-validation to tune the hyperparameters of the Elastic Net and the base learners. Furthermore, we bootstrap the training data for the ensemble. By training on repeated bootstrap

samples, we reduce the variance of the final estimator and mitigate the risk of overfitting to specific idiosyncratic bank failures in the training set.

	N	Mean	Std. Dev	Min	p50	p99	Max
$\rho_{k,t}$	1,076	0.0141	0.0139	0.0008	0.0108	0.0659	0.2003
$\hat{\rho}_{k,t}$	39,527	0.0143	0.0122	0.0014	0.0116	0.0620	0.1282

Table II: Summary statistics for observed default probabilities $\rho_{k,t}$ for the in-sample banks and estimated default probabilities $\hat{\rho}_{k,t}$ for the out-of-sample local banks.

Table II compares the distribution of market-implied default probabilities for the in-sample banks with the model's out-of-sample predictions for the local bank universe. The central moments align closely, the means differ by less than two basis points and the medians and standard deviations are similarly matched. The primary discrepancy arises in the right tail, where the observed distribution extends to 20% while the estimated distribution reaches approximately 13%.

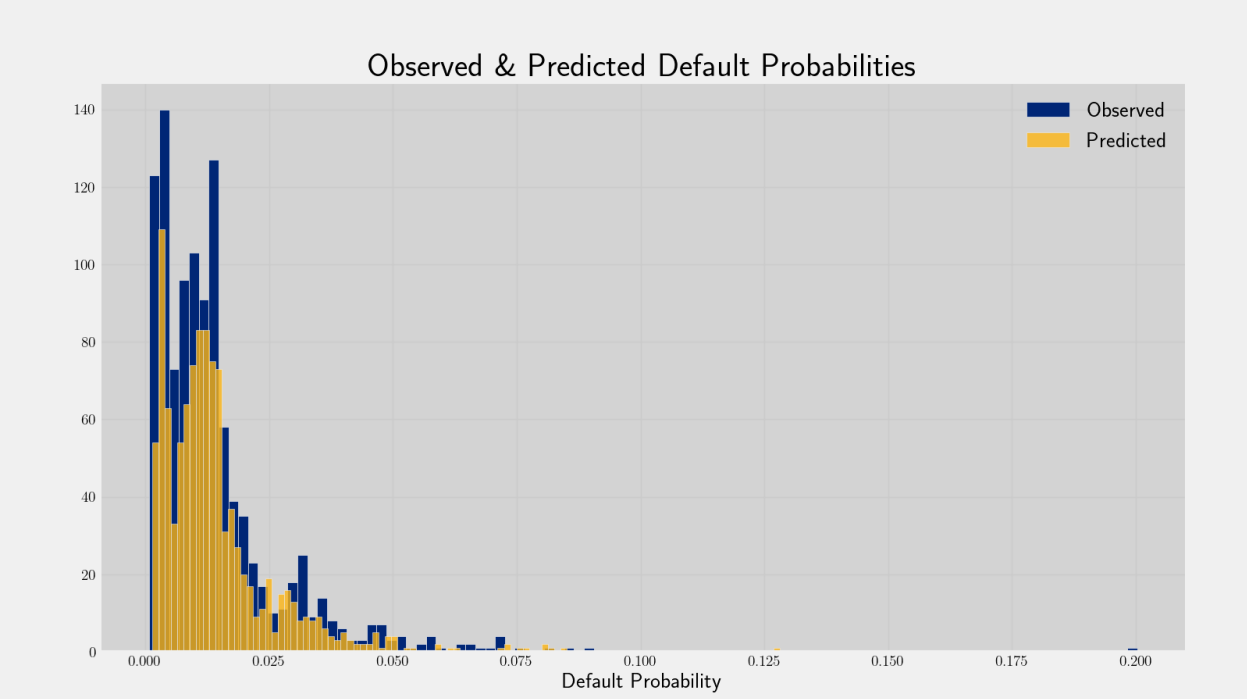


Figure 7. : Distribution of observed default probabilities for the in-sample large regional banks and GSIBs and estimated default probabilities for the out-of-sample local banks.

Figure 7 presents the distributional comparison as overlapping histograms. The two distributions share a common modal range and broadly similar shape, consistent with the summary statistics in Table II.

Figure 8 presents a hold-one-out validation for Wells Fargo. The model is re-estimated on all remaining banks and the resulting out-of-sample predictions are compared against the market-implied default probabilities over the full 2004–2024 period. The two series track each other closely,

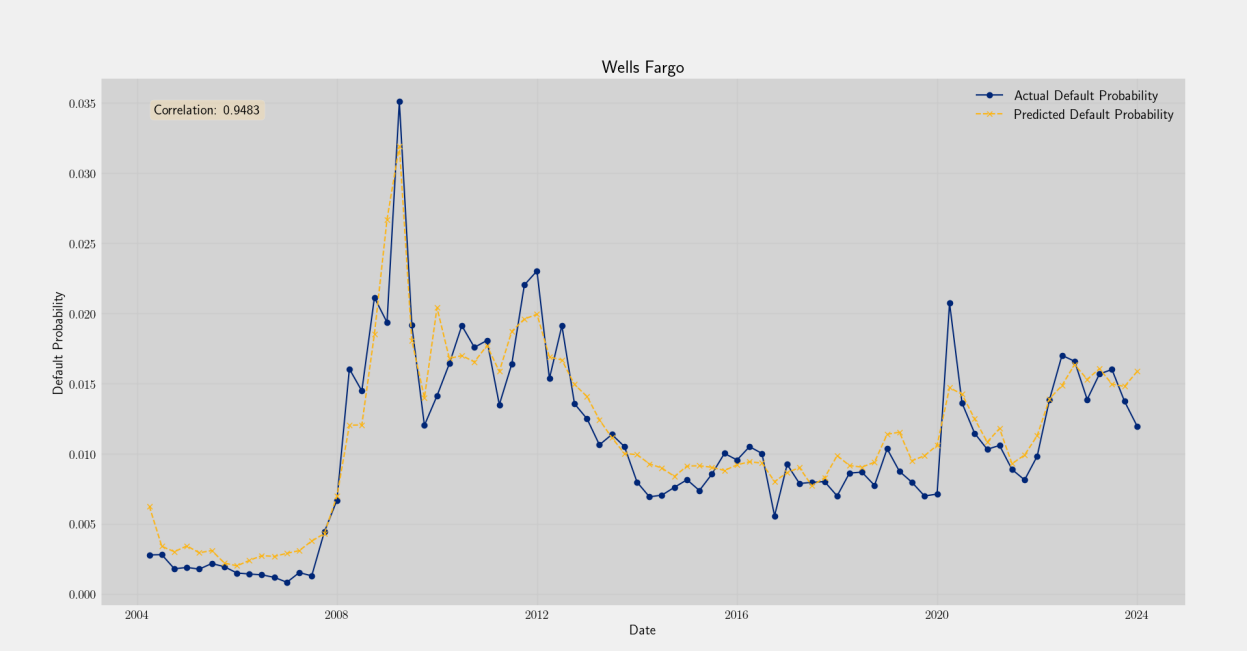


Figure 8. : Estimated and implied default probabilities for Wells Fargo under a hold-one-out specification. Out-of-sample predictions are compared against market-implied default probabilities from 2004 to 2024.

with a correlation of 0.95, and the model captures both the sharp increase in default risk during the GFC and the more moderate elevation observed around 2020. This performance across distinct stress episodes, estimated entirely without information from the held-out institution, supports the view that the model generalizes beyond the training sample.

The weighted ensemble achieves an R^2 of 0.8636 on the held-out test set, compared with 0.4164 for the baseline Elastic Net, and a root mean squared error of 0.0058, approximately half that of the linear benchmark. The ratio of validation RMSE to training RMSE is 23.5%, indicating some in-sample overfit that likely reflects the limited number of institutions available for training. The resulting $\hat{\rho}_{k,t}$ provides a continuous, risk-sensitive measure of default probability for the universe of small and regional banks, enabling the subsequent difference-in-differences analysis of deposit demand elasticity.

C. Causal Tests

To identify the causal mechanism through which TBTF guarantees distort deposit market competition, we must estimate the elasticity of deposit demand with respect to bank failure risk. Our approach follows the structural identification strategy of Egan et al. (2017), identifying demand parameters by exploiting cross-sectional variation in bank default probabilities. We extend their framework by introducing a difference-in-differences (DiD) specification that isolates the safety premium generated by government interventions following the Global Financial Crisis (GFC).

We model the deposit market as a discrete choice problem where depositors choose between depositing funds at a specific commercial bank k or retaining them in a risk-free outside option

(e.g., physical cash or government treasuries). We define the market share of bank k at time t for a given deposit type l (insured or uninsured) as

$$s_{k,t}^{l} := \frac{D_{k,t}^{l}}{\sum_{j=1}^{K} D_{j,t}^{l}} \tag{17}$$

where $D_{k,t}^{l}$ represents the total dollar amount of deposits of type l held by bank k. Crucially, the inclusion of the outside option allows for aggregate contractions in the banking sector's deposit base. Following Egan et al. (2017), we define the deposit share of the outside option, $s_{O,t}^{l}$, as the difference between the total deposits in the market and the sum of all in-sample bank deposits.

To disentangle the TBTF subsidy from fundamental quality differences, we employ a difference-in-differences methodology. We estimate the sensitivity of deposit market shares to the estimated probability of default $\hat{\rho}_{k,t}$, derived from our machine learning model described in Section IV.B, conditional on the bank's regulatory status, TBTF or not-TBTF, and the macroeconomic regime, pre/post GFC and between GSIBs and Local Banks.

The baseline specification for the quantity channel is given by the following linear regression:

$$\begin{aligned} \ln\left(\frac{s_{k,t}}{s_{O,t}}\right) = {} & \beta_0 + \beta_1 \hat{\rho}_{k,t} + \beta_2 \times \text{Regional} \times \hat{\rho}_{k,t} + \beta_3 \times \text{GSIB} \times \hat{\rho}_{k,t} \\ & + \beta_4 \times \text{Post GFC} \times \text{GSIB} \times \hat{\rho}_{k,t} + \beta_5 \times \text{Post GFC} \times \text{Regional} \times \hat{\rho}_{k,t} \\ & + \beta_6 \times \text{Post GFC} \times \hat{\rho}_{k,t} + \beta_7 \times \text{Post GFC} \times \text{GSIB} + \beta_8 \times \text{Post GFC} \times \text{Regional} \\ & + \delta_k + \delta_t + \Gamma X_{k,t} + \epsilon_{k,t} \end{aligned} \tag{18}$$

where $\hat{\rho}_{k,t}$ is the continuous measure of default probability for bank k at time t estimated via the Weighted Ensemble model. "Regional" and "GSIB" are indicator variables for large regional commercial banks and Globally Systemically Important Banks, respectively, with local banks serving as the omitted reference group. "Post GFC" is a time indicator set to one for periods following the implementation of the Dodd-Frank Act, post-2012. δ_k and δ_t represent bank and time fixed effects to control for unobserved heterogeneity and aggregate shocks. $X_{k,t}$ includes time-varying bank-level controls as in Egan et al. (2017), in particular the number of domestic branches is included.

The coefficients of primary interest are the triple interaction terms β_4 and β_5. These coefficients capture the differential change in depositor sensitivity to default risk for GSIBs and regional banks relative to local banks in the post-GFC period. A positive β_4 indicates that GSIBs gain deposit share as their measured risk rises in the post-reform era, consistent with the TBTF distortion. A positive β_5 indicates that regional banks similarly attract deposits as their risk rises relative to local banks, reflecting a partial flight-to-safety effect that extends across the size hierarchy.

We estimate a parallel specification for the pricing channel, where the dependent variable is the deposit spread, defined as the 3-month U.S. T-Bill rate minus the bank's offered deposit rate. This allows us to test the scissors dynamic predicted by our theoretical model: whether distressed local banks are forced to compress spreads to retain funds while protected GSIBs widen them.

D. Results

Table III presents estimates from the deposit share regression on an unbalanced panel of 2,464 commercial banks observed from 2003 to 2023. Throughout, local banks serve as the omitted reference category. All four specifications include bank and time fixed effects and control for the number of domestic branches; standard errors are clustered at the bank level.

	Total	Insured	Uninsured	Uninsured - Insured
	(1)	(2)	(3)	(4)
$\hat{\rho}_{k,t}$	4.048***	4.708***	2.972***	-1.736*
	(0.605)	(0.674)	(0.993)	(0.971)
Regional $\times \hat{\rho}_{k,t}$	1.012*	5.494	5.494	5.943***
	(3.465)	(3.765)	(3.765)	(1.737)
GSIB $\times \hat{\rho}_{k,t}$	-15.430*	-17.483**	-10.327	7.156*
	(8.523)	(7.685)	(9.794)	(3.802)
Post GFC $\times \hat{\rho}_{k,t}$	-3.400***	-4.091***	-3.762**	0.330
	(1.022)	(1.052)	(1.612)	(1.496)
Post GFC $\times$ Regional	-0.138	-0.281	-0.005	0.276*
	(0.143)	(0.180)	(0.120)	(0.145)
Post GFC $\times$ GSIB	-0.379	-0.443	-0.473	-0.029
	(0.348)	(0.407)	(0.368)	(0.286)
Post GFC $\times$ Regional $\times \hat{\rho}_{k,t}$	10.279**	8.229	14.549**	6.320
	(4.764)	(5.610)	(5.754)	(6.345)
Post GFC $\times$ GSIB $\times \hat{\rho}_{k,t}$	29.144*	17.422	50.590***	33.168*
	(16.485)	(20.698)	(16.911)	(18.931)
Bank FE	X	X	X	X
Time FE	X	X	X	X
N	40,603	40,603	40,603	40,603
R^2	0.985	0.980	0.960	0.727

Table III Deposit Share Regressions by Bank Type and Deposit Class. This table presents estimates from the regression of a bank's log deposit share relative to the outside option on its probability of default, estimated on an unbalanced panel of 2,464 commercial banks from 2003 to 2023. Local Banks serve as the omitted reference category. All specifications include bank and time fixed effects and control for the number of domestic branches. Standard errors are clustered at the bank level. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

I estimate four specifications that progressively isolate the credit-risk channel from alternative explanations. Column (1) reports the aggregate effect on total deposits. Columns (2) and (3) decompose this into insured and uninsured deposits, respectively. Column (4) estimates the within-bank difference between the uninsured and insured responses, a triple-differences specification following Egan et al. (2017) that removes the contribution of unobserved, time-varying bank quality from the identified effect.

Column (2) provides the critical identification test for the credit-risk mechanism. Because the FDIC guarantees deposits up to $250,000 per depositor per institution, insured depositors bear no credit exposure to the issuing bank. Their decisions should therefore be insensitive to the probability of default, and the post-GFC triple interaction terms for insured deposits should be

statistically indistinguishable from zero. If, instead, the documented post-GFC shifts in deposit flows were driven by changing bank quality or service provision, insured depositors would exhibit behavior symmetric to uninsured depositors, since quality affects both cohorts equally.

The data support the identification restriction. The coefficient on Post GFC × Regional $\times\hat{\rho}_{k,t}$ and the coefficient on Post GFC × GSIB $\times\hat{\rho}_{k,t}$ in Column (2) are both statistically insignificant. Insured depositors do not discriminate between bank types in response to rising default risk, consistent with the FDIC guarantee eliminating the incentive to monitor solvency. The large post-GFC effects documented in the uninsured specification cannot be attributed to a general shift in bank attractiveness or quality.

The pre-GFC coefficients establish the deposit-risk relationship before crisis-era government interventions created sharp asymmetries across bank types. Local and regional banks exhibit a directionally similar but statistically indistinguishable pre-GFC relationship. For GSIBs, the interaction GSIB $\times\hat{\rho}_{k,t}$ is -10.327 and marginally significant at the 10% level, suggesting that uninsured depositors at the largest banks were somewhat less likely to increase allocations in response to rising risk even in the pre-crisis period. The pre-GFC period therefore reveals no dramatic hierarchy in the deposit-risk relationship across bank types. The TBTF guarantee had not yet been made credible through the experience of a systemic bailout, and the structural break emerges sharply only in the post-GFC period.

The triple interaction terms with the Post GFC indicator reveal a fundamental reorganization of the relationship between bank default risk and deposit flows. For regional banks the coefficient on Post GFC × Regional $\times\hat{\rho}_{k,t}$ in Column (3) is positive and significant at the 5% level. As regional bank default risk rises in the post-reform era, uninsured depositors reallocate toward these institutions rather than away from them. This pattern is consistent with a partial flight-to-safety effect that extends beyond the four GSIBs: depositors withdrawing from local banks may perceive large regional institutions as relatively safe havens, even absent a formal systemic designation.

The effect is largest and most precisely estimated for GSIBs. The coefficient on Post GFC × GSIB $\times\hat{\rho}_{k,t}$ is 50.590, significant at the 1% level. The net post-GFC elasticity for GSIBs is 39.4 per unit of $\hat{\rho}_{k,t}$. Since $\hat{\rho}_{k,t}$ is measured as a decimal probability, a one-percentage-point increase in a GSIB's probability of default is associated with an approximately 48% increase in its uninsured deposit share in the post-GFC era. GSIBs have become liquidity sinks in the post-reform era, attracting rather than repelling uninsured capital as their measured default risk increases. The implicit government guarantee, reinforced by the systematic rescues during the GFC and the resolution framework established under Dodd-Frank, has decoupled credit risk from depositor response at the top of the banking hierarchy.

The flight-to-safety gradient across the size hierarchy is economically striking. Relative to local banks, GSIBs capture 50.590 additional units of log deposit share per unit increase in default probability in the post-GFC period, and 36.041 more than regional banks. The deposit reallocation is not simply a flow from small to any larger institution; it concentrates disproportionately in the four institutions perceived to carry the most credible government backstop.

Column (4) directly tests whether the uninsured deposit flows in Column (3) reflect credit-risk sensitivity or changes in unobserved bank quality. The triple-differences estimator computes the within-institution differential response of uninsured versus insured depositors to default probability. Because quality improvements generate symmetric inflows across both deposit types, any within-bank asymmetry between insured and uninsured responses must be attributed to credit risk rather than quality.

For local banks, the baseline coefficient on $\hat{\rho}_{k,t}$ in Column (4) is negative and significant at the 10% level. Higher default probability at local banks is associated with a greater decline in the uninsured deposit share than in the insured share. This within-institution differential is consistent with market discipline: uninsured depositors monitor and respond to solvency risk at local banks, while insured depositors do not.

The coefficient of primary interest is Post GFC × GSIB $\times\hat{\rho}_{k,t}$, significant at the 10% level. This estimate survives the within-bank quality control. Relative to insured depositors at the same GSIB, uninsured depositors disproportionately increase their allocations to GSIBs as default probability rises in the post-GFC era.

The analogous triple-difference coefficient for regional banks, Post GFC × Regional $\times\hat{\rho}_{k,t}$ is not statistically significant. This contrast with the significant GSIB result clarifies the scope of the TBTF mechanism. Regional banks attract deposit flows on the basis of general safety perceptions relative to local banks, but the clean credit-risk channel identified by the within-bank comparison is concentrated in the four institutions with a credible systemic backstop.

Time Period	Pre GFC 2003Q1-2007Q4	GFC 2008Q1-2009Q4	ZLB 2010Q1-2019Q4	Covid 2020Q1-2023Q1
$\hat{\rho}_{k,t}$	-0.064	0.001	0.026**	-0.036
	(0.072)	(0.007)	(0.011)	(0.023)
$\hat{\rho}_{k,t} \times$ GSIB	0.266***	0.008	-0.020	0.219***
	(0.048)	(0.015)	(0.018)	(0.041)
$\hat{\rho}_{k,t} \times$ Local	0.036	-0.010	-0.027**	0.032
	(0.073)	(0.008)	(0.011)	(0.024)
Bank FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
R^2	0.987	0.974	0.992	0.998
N	12,737	5,982	16,620	4,912
$\hat{\rho}_{k,t} + \hat{\rho}_{k,t} \times$ GSIB	0.202***	0.010	0.006	0.183***
	(0.051)	(0.014)	(0.016)	(0.035)
$\hat{\rho}_{k,t} + \hat{\rho}_{k,t} \times$ Local	-0.028***	-0.008*	-0.001	-0.004
	(0.009)	(0.004)	(0.002)	(0.005)

Table IV. This table shows the regression of a bank's deposit spread on its probability of default. A banks deposit spread is defined as the 3-month U.S. T-Bill minus the bank's offered deposit rate. Each regression is estimated using an unbalanced panel of 2,464 commercial banks in the United States over the period 2003-2023. All specifications control for the number of bank branches. Standard errors are clustered at the bank level.
Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table IV presents the results for the pricing channel, regressing the deposit spread on default

probabilities. The results corroborate the quantity channel but with weaker statistical significance. The results are presented with regional banks as the reference group.

A critical factor driving these results is the underlying interest rate environment. In the Pre-GFC period, when benchmark rates were well above zero, the pricing mechanism functioned consistently with the theory presented in Section III. The baseline coefficient for regional banks is negative, suggesting that riskier banks paid higher rates or that deposit rates were insensitive to default probabilities, while the GSIB interaction is significantly positive, suggesting an early presence of the safety subsidy. However, the lack of statistical significance in the interaction terms during the Post GFC and ZLB periods is likely a mechanical artifact of the extremely low interest rate environment that characterized the post-crisis era. When the Federal Funds Rate is pinned near zero, the variance in deposit rates collapses; banks cannot lower nominal deposit rates below zero to widen spreads, nor can riskier banks raise rates significantly without destroying net interest margins. This compression removes the variation required to identify the pricing coefficient using interaction terms alone.

To rigorously evaluate the net impact of these distortions across regimes, we calculate the total marginal effect of default risk for each bank type, presented in the bottom two rows of Table IV. These total effect rows sum the baseline coefficient with the respective interaction terms $\hat{\rho}_{k,t} + \hat{\rho}_{k,t} \times \text{Type}$, providing a clearer picture of the absolute elasticity. The results confirm a persistent divergence. For GSIBs, the total effect remains positive and statistically significant across the Pre-GFC and Covid periods. This indicates that even during the pandemic, GSIBs retained the ability to widen spreads as their risk profile deteriorated. In sharp contrast, the total effect for local banks is consistently negative or near zero in the Pre-GFC period and during the GFC. This negative total elasticity confirms that local banks are subject to market discipline, forcing them to compress their spreads and pay a premium to retain funding when their solvency is threatened, while their systemic competitors effectively monetize the distress.

V. Conclusion

This study provides a structural and empirical characterization of the safety premium in the U.S. deposit market, demonstrating that implicit government guarantees fundamentally distort the transmission of market discipline. The theoretical framework, combining Hotelling's spatial competition with Merton's credit risk mechanics, reveals that the elasticity of deposit demand with respect to safety overwhelms the elasticity with respect to yield during periods of stress. This generates a convex "scissors" dynamic in equilibrium pricing: unprotected local banks face a self-reinforcing death spiral where rising default risk necessitates higher deposit rates that further deplete capital, while TBTF institutions exploit the inelasticity of flight-to-safety flows to widen net interest margins. Crucially, the model demonstrates that this distortion is not static but highly counter-cyclical, expanding non-linearly when the banking sector is most fragile and the marginal utility of safety is highest.

Empirically, this paper overcomes the challenge of unobservable default risk for non-publicly traded institutions by deploying a Weighted Ensemble machine learning model, which maps the accounting fundamentals of the universe of commercial banks to market-implied probabilities of default. The subsequent difference-in-differences analysis validates the theoretical predictions via the quantity channel, confirming a regime change following the Global Financial Crisis. While the Zero Lower Bound environment constrained the identification of the pricing channel, the flow of funds data provides distinct evidence of bifurcation: uninsured depositors exert severe discipline on local banks while treating GSIBs as effectively risk-free substitutes for government obligations. This asymmetry confirms that TBTF status grants large banks a put option on the taxpayer that is valued most highly by the market during tail events.

The policy implications of these findings suggest that the current regulatory architecture inadvertently creates evolutionary pressure toward industry consolidation, not through efficiency gains, but through regulatory arbitrage. By shielding GSIBs from the market discipline that constrains their smaller rivals, the TBTF regime allows systemic institutions to accumulate capital and market share at an accelerated rate during downturns, effectively cannibalizing the periphery of the banking system. Furthermore, the sensitivity of the deposit spread to the loss given default parameter implies that ambiguity regarding the scope of deposit insurance remains a primary driver of cost-of-capital differentials. Consequently, unless regulatory frameworks can credibly commit to imposing losses on the unsecured creditors of systemic institutions, the flight-to-safety subsidy will continue to incentivize excessive scale and concentration, undermining the competitive neutrality of the financial system.

Appendix A. Model Derivations

Appendix A.A. First Passage Time ODE

We derive the differential equation governing the value of the claim $P(V_t)$ that pays \$1 at the moment of default τ and 0 otherwise. Default is defined as the first passage time $\tau := \inf\{t \geq 0 : V_t \leq V_B\}$.

Assumption A.A.1: *Asset values follow a Geometric Brownian Motion under the risk-neutral measure $\mathcal{Q}$: $dV_t/V_t = (r_f - \delta)dt + \sigma dZ_t$, with constant parameters risk-free rate r_f, payout rate δ, and volatility σ*

Let $P(V_t)$ be the price of the security at time t. Since the security pays no dividends before default, its discounted price process must be a martingale under the risk-neutral measure $\mathcal{Q}$. We define the discounted process

$$Y_t = e^{-r_f t} P(V_t).$$

For Y_t to be a martingale, its expected instantaneous drift must be zero $\mathbf{E_Q}[dY_t] = 0$.

We apply Ito's Lemma for the product of two processes to find the differential

$$dY_t = d(e^{-r_f t})P(V_t) + e^{-r_f t}dP(V_t) = -r_f e^{-r_f t}dt + e^{-r_f t}dP(V_t).$$

Next, we apply Ito's Lemma to $P(V_t)$. Since P is a function of V_t, we expand it to the second order:

$$dP(V_t) = P'(V_t)dV_t + \frac{1}{2}P''(V_t)(dV_t)^2.$$

Substitute the dynamics of V_t

$$dP(V_t) = P'(V_t)[(r_f - \delta)\, V_t dt + \sigma V_t dZ_t] + \frac{1}{2}P''(V_t)\sigma^2 V_t^2 dt.$$

Grouping the dt and dZ_t terms

$$dP(V_t) = \left[(r_f - \delta)\, V_t P'(V_t) + \frac{1}{2}\sigma^2 V_t^2 P''(V_t)\right] dt + \sigma V_t P'(V_t) dZ_t.$$

Now, substitute $dP(V_t)$ back into the expression for the discounted process

$$dY_t = -r_f e^{-r_f t} P(V_t)dt + e^{-rt}\left[\left((r_f - \delta)V_t P'(V_t) + \frac{1}{2}\sigma^2 V_t^2 P''(V_t)\right) dt + \sigma V_t P'(V_t) dZ_t\right].$$

Factor out $e^{-r_f t}$ and collect all dt terms

$$dY_t = e^{-r_f t} \underbrace{\left(\frac{1}{2}\sigma^2 V_t^2 P''(V_t) + (r_f - \delta) V_t P'(V_t) - r_f P(V_t) \right)}_{\text{Drift}} dt + e^{-rft} \sigma V_t P'(V_t) dZ_t.$$

By the Martingale Representation Theorem, the drift term must equal zero to prevent arbitrage.

$$\frac{1}{2}\sigma^2 V^2 P''(V) + (r_f - \delta) V P'(V) - r_f P(V) = 0.$$

Appendix A.B. First Passage Time Security

We derive the value of a state-contingent claim $P(V_t, V_B)$ that pays \$1 at the moment the Local Bank defaults. Default is defined as the first passage time $\tau := \inf\{t \geq 0 : V_t \leq V_B\}$.

Assumption A.B.1: *Asset values follow a Geometric Brownian Motion under the risk-neutral measure $\mathcal{Q}$: $dV_t/V_t = (r_f - \delta)dt + \sigma dZ_t$, with constant parameters risk-free rate r_f, payout rate δ, and volatility σ*

The claim pays \$1 if V_t hits V_B and 0 otherwise. Let $P(V)$ denote the expected present value of the claim

$$P(V) = \mathbf{E}_{\mathcal{Q}}[e^{-r_f \tau} | V_0 = V].$$

In the continuation region $V > V_B$, the discounted value of the claim is a martingale. In Section A.A the governing ODE for this system was shown to be

$$\frac{1}{2}\sigma^2 V^2 P''(V) + (r_f - \delta) V P'(V) - r P(V) = 0.$$

We guess a solution of the form $P(V) = AV^\beta$. Substituting this into the ODE yields the characteristic equation

$$\frac{1}{2}\sigma^2 \beta(\beta - 1) + (r_f - \delta)\beta - r_f = 0.$$

Rearranging terms gives the fundamental quadratic equation

$$\frac{1}{2}\sigma^2 \beta^2 + \left(r_f - \delta - \frac{1}{2}\sigma^2 \right) \beta - r = 0.$$

This quadratic has two roots. Since the value of the default claim must vanish as the bank becomes infinitely solvent, we select the negative root

$$-\gamma = \frac{-\left(r_f - \delta - \frac{1}{2}\sigma^2\right) - \sqrt{\left(r_f - \delta - \frac{1}{2}\sigma^2\right)^2 + 2 r_f \sigma^2}}{\sigma^2}.$$

Applying the boundary condition at default, $P(V_B) = 1$, we solve for the constant A

$$P(V_B) = AV_B^{-\gamma} = 1 \implies A = V_B^{\gamma}.$$

Thus, the value of the first-passage security is

$$P(V_t) = \left(\frac{V_t}{V_B}\right)^{-\gamma}.$$

Appendix A.C. Equity Value ODE

We derive the differential equation governing the value of bank equity $E(V_t)$, which pays a continuous dividend stream equal to asset income minus debt service $(\delta V_t - C)$ until default. Unlike the default security, equity pays a continuous flow of dividends. For the equity price to be arbitrage-free, the discounted value of the stock price plus the discounted value of cumulative dividends received must be a martingale under $\mathcal{Q}$. We define the gains process G_t

$$G_t = \underbrace{e^{-r_f t}E(V_t)}_{\text{Discounted Price}} + \underbrace{\int_0^t e^{-r_f s}(\delta V_s - C)ds}_{\text{Discounted Dividends}}.$$

The condition for no arbitrage is $\mathbf{E}_{\mathcal{Q}}[dG_t] = 0$.

Differentiate G_t with respect to time t and apply Ito's Lemma. The derivative of the integral term is simply the integrand evaluated at t

$$\begin{aligned} dG_t &= d(e^{-r_f t}E(V_t)) + e^{-r_f t}(\delta V_t - C)dt \\ &= -r_f e^{-r_f t}E(V_t)dt + e^{-r_f t}\left(E'(V_t)dV_t + \frac{1}{2}E''(V_t)(dV_t)^2\right) + e^{-r_f t}(\delta V_t - C)dt \end{aligned}$$

Plugging in asset dynamics

$$dE(V_t) = \left((r_f - \delta)V_t E'(V_t) + \frac{1}{2}\sigma^2 V_t^2 E''(V_t)\right)dt + \sigma V_t E'(V_t)dZ_t.$$

Substitute $dE(V_t)$ back into the expression for dG_t

$$dG_t = e^{-r_f t}\left(-r_f E(V_t)dt + \left[(r_f - \delta)V_t E'(V_t) + \frac{1}{2}\sigma^2 V_t^2 E''(V_t)\right]dt + \sigma V_t E'(V_t)dZ_t\right) + e^{-r_f t}(\delta V_t - C)dt.$$

Factor out $e^{-r_f t}$ and group all dt terms

$$dG_t = e^{-r_f t}\underbrace{\left(\frac{1}{2}\sigma^2 V_t^2 E''(V_t) + (r_f - \delta)V_t E'(V_t) - r_f E(V_t) + (\delta V_t - C)\right)}_{\text{Drift}}dt + e^{-r_f t}\sigma V_t E'(V_t)dZ_t.$$

Set the drift term to zero to satisfy the martingale condition, giving the ODE

$$\frac{1}{2}\sigma^2 V^2 E''(V) + (r_f - \delta)VE'(V) - r_f E(V) + (\delta V - C) = 0.$$

Appendix A.D. Equity Value and Optimal Default Bound

We derive the optimal default barrier V_B chosen by shareholders to maximize equity value.

Assumption A.D.1: *Equity holders receive a net flow of dividends $\delta V - C$ per instant, where C is the total interest expense.*

Assumption A.D.2: *Shareholders possess limited liability; equity value is zero at default.*

Assumption A.D.3: *Shareholders choose V_B optimally (Smooth Pasting condition).*

Section A.C derives the ODE for the equity value $E(V_t)$

$$\frac{1}{2}\sigma^2 V^2 E''(V) + (r_f - \delta)VE'(V) - r_f E(V) + (\delta V - C) = 0.$$

This is a non-homogenous Euler-Cauchy ODE with the known solution as the sum of the particular solution, value of solvent assets minus risk-free debt, and the homogeneous solution, the option value

$$E(V) = \left(V - \frac{C}{r}\right) + AV^{-\gamma}$$

Note: The positive root term BV^{η} is discarded because the default option value must vanish as $V \to \infty$ to satisfy the solvency constraint.

Now, we solve for the coefficient A and the optimal default bound V_B using the boundary conditions

- Value Matching: $E(V_B) = 0$

$$V_B - \frac{C}{r} + AV_B^{-\gamma} = 0 \tag{A.D.1}$$

- Smooth Pasting: $E'(V_B) = 0$

$$1 - \gamma A V_B^{-\gamma-1} = 0 \tag{A.D.2}$$

From A.D.2, we express A in terms of V_B

$$A = \frac{1}{\gamma} V_B^{\gamma+1}.$$

Substitute A back into A.D.1

$$V_B - \frac{C}{r} + \left(\frac{1}{\gamma} V_B^{\gamma+1}\right) V_B^{-\gamma} = 0$$
$$V_B - \frac{C}{r} + \frac{1}{\gamma} V_B = 0$$
$$V_B \left(1 + \frac{1}{\gamma}\right) = \frac{C}{r}.$$

Solving for V_B yields the closed-form endogenous barrier

$$V_B = \frac{C}{r}\left(\frac{\gamma}{\gamma+1}\right).$$

Appendix A.E. Proposition 1

We derive the equilibrium deposit rates r_N^* and r_L^* under the assumption of asymmetric transport costs.

Assumption A.E.1: *Depositors are uniformly distributed on* $[0, 1]$.

Assumption A.E.2: *Transport cost for Local Bank is normalized to 1; National Bank is* $\phi \in (0, 1]$.

Assumption A.E.3: *Depositors are risk-neutral. Local Bank depositors face expected credit loss flow* $\alpha\Psi$.

The indifferent depositor x^* equates net utility

$$r_N - \phi x^* = r_L - \alpha\Psi - (1 - x^*).$$

Solving for x^* gives the National Bank's market share M_N:

$$M_N = x^* = \frac{r_N - r_L + 1 + \alpha\Psi}{1 + \phi}. \tag{A.E.1}$$

The Local Bank's share is $M_L = 1 - M_N$.
Banks maximize profit $\Pi_k = (R - r_k)M_k$. The FOCs are

$$\frac{\partial \Pi_N}{\partial r_N} = -M_N + (R - r_N)\left(\frac{1}{1+\phi}\right) = 0 \implies R - r_N = (1+\phi)M_N \tag{A.E.2}$$
$$\frac{\partial \Pi_L}{\partial r_L} = -M_L + (R - r_L)\left(\frac{1}{1+\phi}\right) = 0 \implies R - r_L = (1+\phi)M_L \tag{A.E.3}$$

We solve the system for the spread $S = r_L - r_N$. Subtract A.E.2 from A.E.3

$$
\begin{aligned}
(R - r_L) - (R - r_N) &= (1+\phi)(M_L - M_N) \\
r_N - r_L &= (1+\phi)(1 - 2M_N) \\
-S &= (1+\phi) - 2(1+\phi)M_N \\
&= (1+\phi) - 2\left(r_N - r_L + 1 + \alpha\Psi\right) \\
&= -S + 1 + \alpha\Psi
\end{aligned}
$$

which we solve to find the equilibrium spread

$$
S = \frac{1-\phi}{3} + \frac{2}{3}\alpha\Psi \tag{A.E.4}
$$

Using the spread, we solve for levels

$$
\begin{aligned}
r_N^* &= R - (1+\phi)M_N^* = R - (-S^* + 1 + \alpha\Psi) \\
&= R - 1 - \alpha\Psi + \left(\frac{1-\phi}{3} + \frac{2}{3}\alpha\Psi\right) \\
&= R - \frac{2+\phi}{3} - \frac{\alpha}{3}\Psi
\end{aligned} \tag{A.E.5}
$$

$$
\begin{aligned}
r_L^* &= r_N^* + S^* \\
&= \left(R - \frac{2+\phi}{3} - \frac{\alpha}{3}\Psi\right) + \left(\frac{1-\phi}{3} + \frac{2}{3}\alpha\Psi\right) \\
&= R - \frac{1+2\phi}{3} + \frac{\alpha}{3}\Psi
\end{aligned} \tag{A.E.6}
$$

The term $\alpha\Psi$ represents the flow cost of credit risk. We explicitly derive this term, in terms of the underlying asset value V_t and the endogenous default boundary V_B, by requiring that, in equilibrium, the risky deposit rate must offer a arbitrage-free risk-adjusted return equal to the risk-free rate r_f.

Let $D(V_t)$ be the value of a unit deposit paying a continuous coupon yield $y(V_t, V_B) = r_f + \alpha\Psi(V_t, V_B)$ until default, with a recovery of $(1-\alpha)$ at the default moment τ. Under the risk-neutral measure $\mathcal{Q}$, the value of this claim is

$$
D(V_t, V_B) = \mathbf{E}_{\mathcal{Q}}\left[\int_0^\tau e^{-r_f s} y(V_t, V_B) ds + e^{-r_f \tau}(1-\alpha)\right].
$$

Using the closed-form solution for the First Passage Time security in Section A.B

$$
P(V_t, V_B) = \mathbf{E}_{\mathcal{Q}}[e^{-r_f \tau}] = \left(\frac{V_t}{V_B}\right)^{-\gamma}
$$

we rewrite the value of the unit deposit $D(V_t, V_B)$

$$D(V_t, V_B) = \frac{y(Vt; V_B)}{r_f}(1 - P(V_t, V_B)) + (1 - \alpha)P(V_t, V_B).$$

For the deposit to trade at par $D(V_t, V_B) = 1$, the offered yield $y(V_t, V_B)$ must satisfy

$$1 = \frac{r_f + \alpha\Psi(V_t, V_B)}{r_f}(1 - P(V_t, V_B)) + (1 - \alpha)P(V_t, V_B)$$

Solving this gives the functional form for the credit premium Ψ

$$\Psi(V_t, V_B) = r_f \left(\frac{P(V_t, V_B)}{1 - P(V_t, V_B)} \right) = r_f \left(\frac{(V_t/V_B)^{-\gamma}}{1 - (V_t/V_B)^{-\gamma}} \right). \tag{A.E.7}$$

Appendix A.F. Corollary 1

Substitute the equilibrium spread A.E.4 into the market share equation A.E.1

$$\begin{aligned} M_N^* &= \frac{-S^* + 1 + \alpha\Psi}{1 + \phi} \\ &= \frac{1}{1 + \phi} \left(1 + \alpha\Psi - \left(\frac{1 - \phi}{3} + \frac{2}{3}\alpha\Psi \right) \right) \\ &= \frac{1}{1 + \phi} \left(\frac{3 + 3\alpha\Psi - 1 + \phi - 2\alpha\Psi}{3} \right) \\ &= \frac{2 + \phi + \alpha\Psi}{3(1 + \phi)} \end{aligned}$$

The Local Bank share is

$$M_L^* = 1 - M_N^* = \frac{3(1 + \phi) - (2 + \phi + \alpha\Psi)}{3(1 + \phi)} = \frac{1 + 2\phi - \alpha\Psi}{3(1 + \phi)}.$$

For profits, recall from the FOCs that optimal profit is $\Pi_k^* = (1 + \phi)(M_k^*)^2$. Substituting the expressions for M_k^*

$$\begin{aligned} \Pi_N^* &= (1 + \phi) \left(\frac{2 + \phi + \alpha\Psi}{3(1 + \phi)} \right)^2 \\ \Pi_L^* &= (1 + \phi) \left(\frac{1 + 2\phi - \alpha\Psi}{3(1 + \phi)} \right)^2 \end{aligned}$$

REFERENCES

Begenau, Juliane, and Erik Stafford, 2023, Uniform rate setting, branch expansion, and the deposit channel, *Working Paper* .

Berndt, Antje, Darrell Duffie, and Yichao Zhu, 2025, The decline of too big to fail, *American Economic Review* 115, 945–974.

Black, Fischer, and John C. Cox, 1976, Valuing corporate securities: Some effects of bond indenture provisions, *The Journal of Finance* 31, 351–367.

Cookson, J. Anthony, Corbin Fox, Javier Gil-Bazo, Juan Felipe Imbet, and Christoph Schiller, 2023, Social media as a bank run catalyst, *Université Paris-Dauphine Research Papers* Available at SSRN: https://ssrn.com/abstract=4422754 or http://dx.doi.org/10.2139/ssrn.4422754.

d'Avernas, Adrien, Andrea L. Eisfeldt, Can Huang, Richard Stanton, and Nancy Wallace, 2023, The deposit business at large vs. small banks, *NBER Working Paper Series* .

Diamond, Douglas W., and Philip H. Dybvig, 1983, Bank runs, deposit insurance, and liquidity, *Journal of Political Economy* 91, 401–419.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2017, The deposit channel of monetary policy, *The Quarterly Journal of Economics* 132, 1819–1876.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, 2017, Deposit competition and financial fragility: Evidence from the us banking sector, *American Economic Review* 107, 169–216.

FDIC, 2023, Fdic acts to protect all depositors of the former silicon valley bank, santa clara, california, Online.

Federal Reserve System, 2026, U.s. domestically chartered commercial banks, `https://www.federalreserve.gov/releases/lbr/current/`, Accessed: April 20, 2026.

Financial Stability Board, 2025, 2025 list of global systemically important banks (g-sibs), `https://www.fsb.org/2025/11/2025-list-of-global-systemically-important-banks-g-sibs/`, Accessed: April 20, 2026.

Gandhi, Priyank, and Hanno Lustig, 2015, Size anomalies in u.s. bank stock returns, *The Journal of Finance* 70, 733–768.

Goldstein, Itay, and Ady Pauzner, 2005, Demand-deposit contracts and the probability of bank runs, *The Journal of Finance* 60, 1293–1327.

Hotelling, Harold, 1929, Stability in competition, *The Economic Journal* 39, 41–57.

Hull, John C., 2015, *Options, Futures, and Other Derivatives*, 9th edition (Pearson Education).

Iyer, Rajkamal, and Manju Puri, 2012, Understanding bank runs: The importance of depositor-bank relationships and networks, *The American Economic Review* 102, 1414–1445.

Kelly, Morgan, and Cormac Ó Gráda, 2000, Market contagion: Evidence from the panics of 1854 and 1857, *The American Economic Review* 90, 1110–1124.

Leland, Hayne E., 1994, Corporate debt value, bond covenants, and optimal capital structure, *The Journal of Finance* 49, 1213–1252.

Leland, Hayne E., and Klaus Bjerre Toft, 1996, Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads, *The Journal of Finance* 51, 987–1019.

Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *The Journal of Finance* 29, 449–470.

O'Hara, Maureen, and Wayne Shaw, 1990, Deposit insurance and wealth effects: The value of being "too big to fail", *The Journal of Finance* 45, 1587–1600.

Penas, María Fabiana, and Haluk Unal, 2004, Gains in bank mergers: Evidence from the bond markets, *Journal of Financial Economics* 74, 149–179.

Perry, Andrew, 2026, Selling distress: The economics of failed bank resolutions, *Working Paper* .

Santos, Joao A. C., 2014, Evidence from the bond market on banks' "too-big-to-fail" subsidy, *Economic Policy Review* 29–39.

Stern, Gary H., and Ron J. Feldman, 2004, *Too Big to Fail: The Hazards of Bank Bailouts* (Brookings Institution Press).

Notes