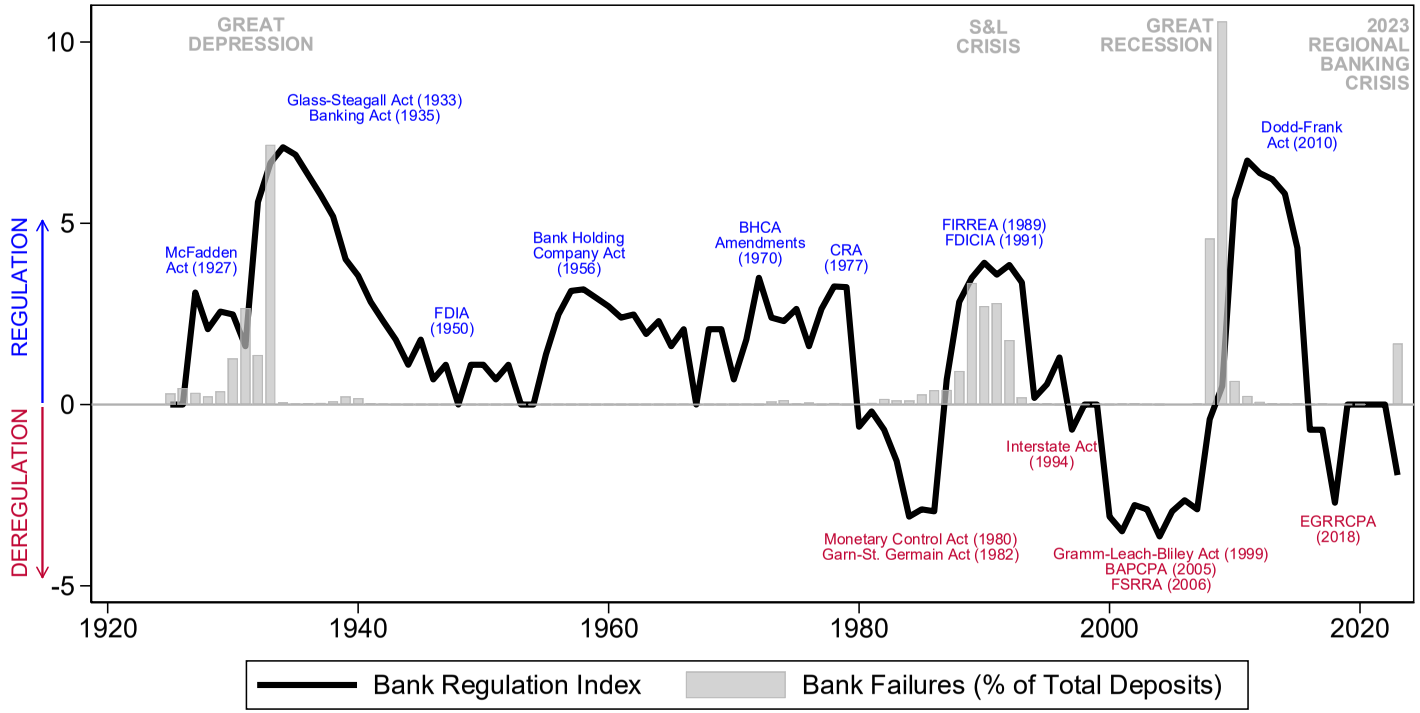


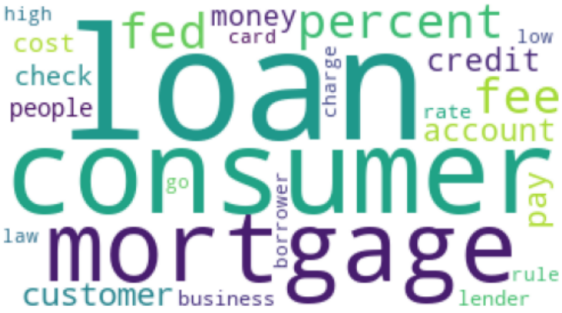
Figure 1. Bank Regulation Index and Banking Crises

The figure illustrates the relationship between *BRI* and *BankFailures*, where *BankFailures* represents the proportion of deposits in failed banks to total deposits. The plot highlights four distinct peaks in *BankFailures* since 1926. The initial surge corresponds to the Great Depression of the 1930s. In response, stringent banking regulations were instituted, leading to a subsequent decline in *BankFailures*. A period of deregulation during 1979-82 set the stage for the next notable peak, representing the Savings and Loans Crisis of the 1980s. Regulatory reforms between 1989 and 1991 then followed. A subsequent deregulatory phase from the late 1990s to early 2000s paved the way for the Great Recession of 2007-09, which triggered the enactment of the Dodd-Frank Act. This recurring pattern of regulatory interventions post-crisis, followed by deregulatory periods, with the Economic Growth, Regulatory Relief, and Consumer Protection Act of 2018 leading to the most recent increase in *BankFailures* by 2023. The cyclical trajectory of *BRI* against *BankFailures* underscores a consistent narrative: post-crisis regulatory measures often transition into deregulatory phases, resulting in bank failures. The 2023 decline in *BRI* is attributed to mentions of the EGRRCPA following the recent bank failures.

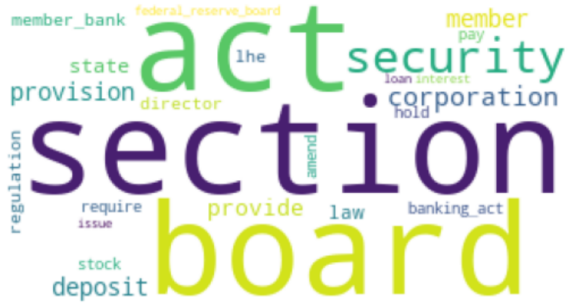
(A) Bank Activities

(B) Dodd Frank

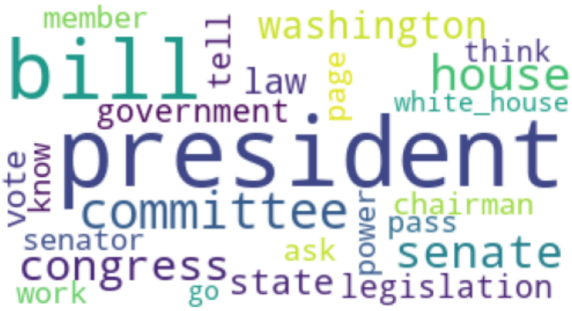
(C) Lending



(D) Legalese



(E) Government



(F) Monetary

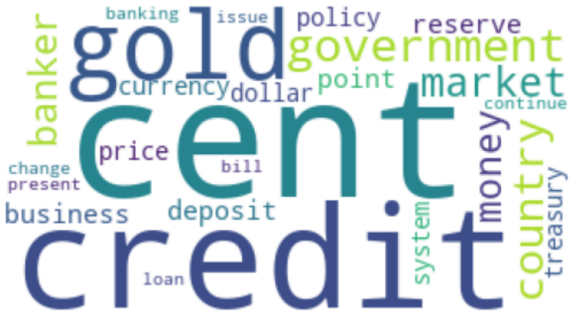
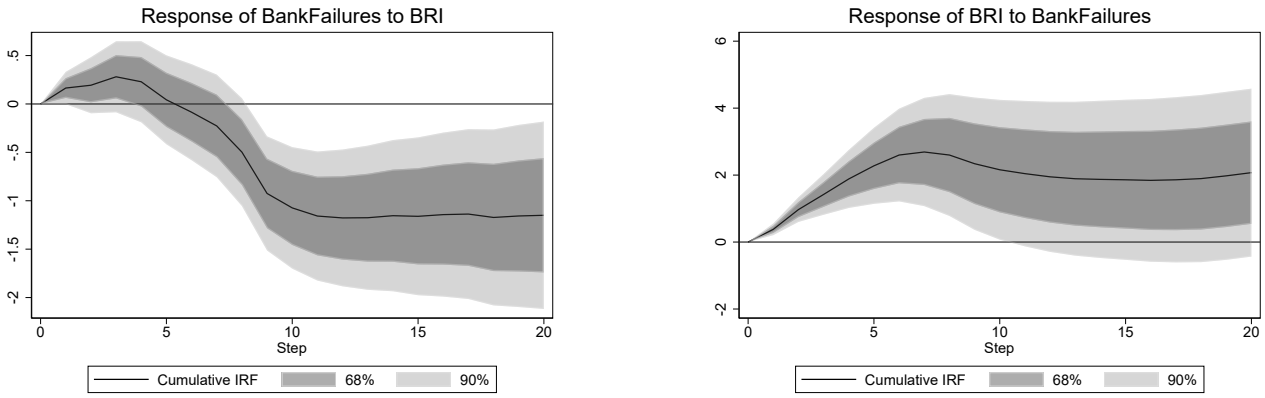


Figure 2. Word Clouds for LDA Topics

LDA provides two different distributions: a distribution of each document on the topics and a distribution of each topic on a set of words or terms. See Appendix B for details on the LDA procedure. The term distribution of each topic can be used to create the word clouds associated with the topic. The size of each term in the image is proportionate to the score it receives in the LDA distribution.

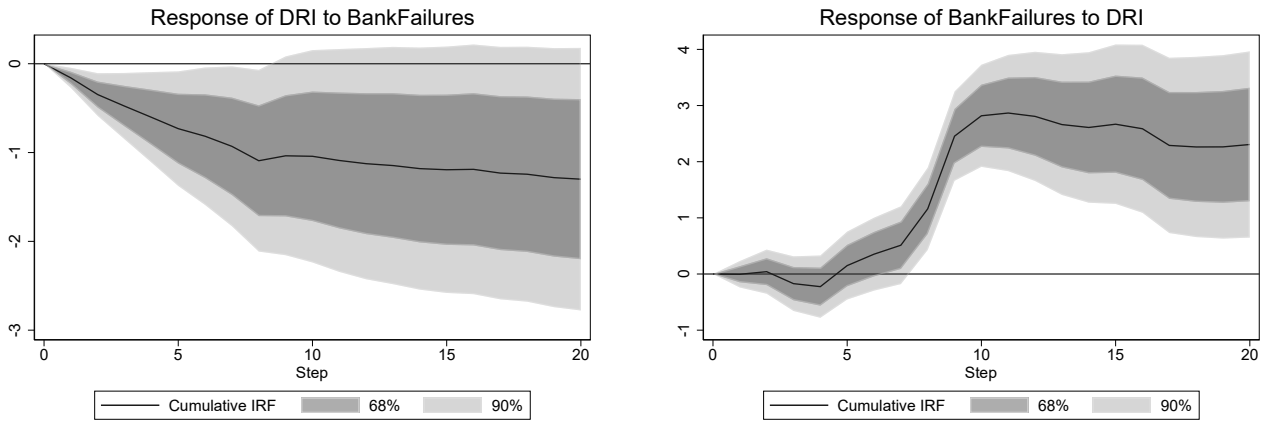


(A) Impulse (BRI)

(B) Impulse (BankFailures)

Figure 3. IRF of BRI and Bank Failures

VAR-estimated impulse-response functions for BankFailures and Bank Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006). Identification is based on ten lags. The VAR model is specified in Section 4.2.



(A) Impulse (BankFailures)

(B) Impulse (DRI)

Figure 4. IRF of DRI and Bank Failures

VAR-estimated impulse-response functions for BankFailures and Decreased Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006). Identification is based on ten lags. The VAR model is specified in Section 4.2.

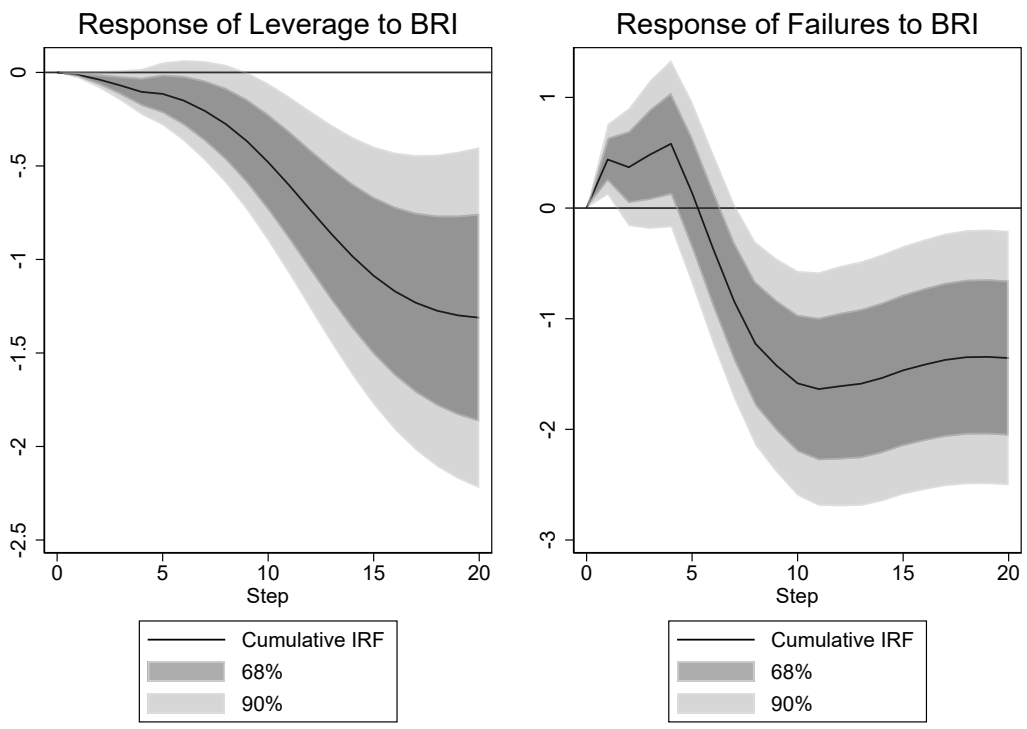


Figure 5. IRFs of Leverage

Augmented VAR-estimated impulse-response functions for Leverage and Bank Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006). Data for Leverage is from Jordà et al. (2017). Identification is based on ten lags.

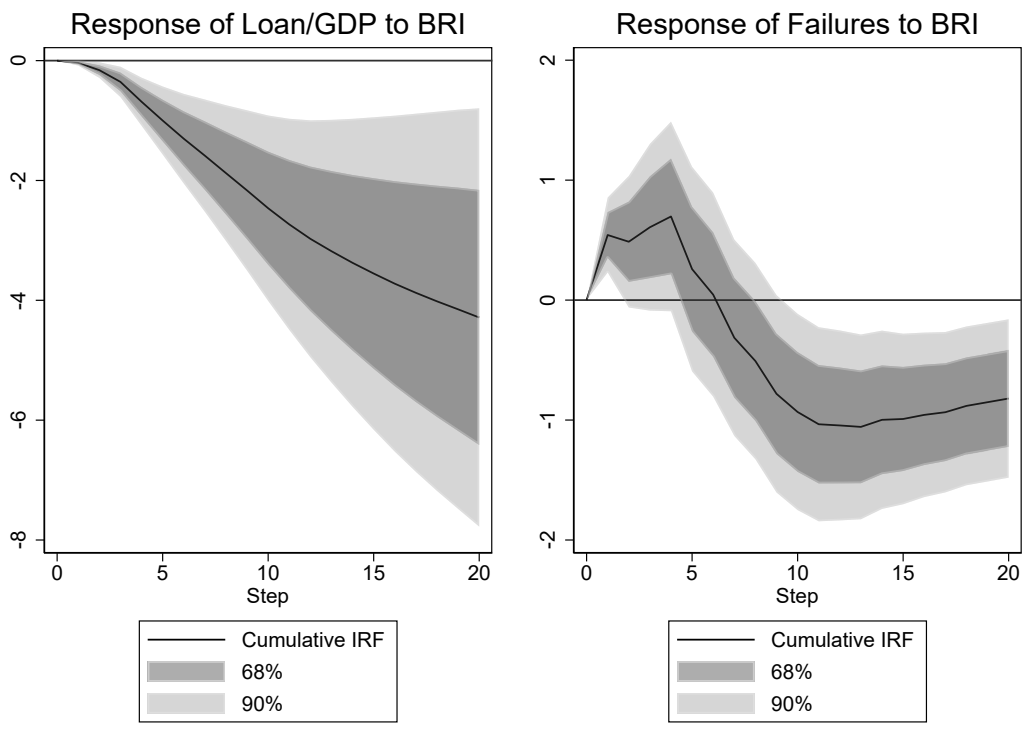


Figure 6. IRFs of Loan/GDP

Augmented VAR-estimated impulse-response functions for Loan/GDP and Bank Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006). Data for Leverage is from Jordà et al. (2017). Identification is based on ten lags.

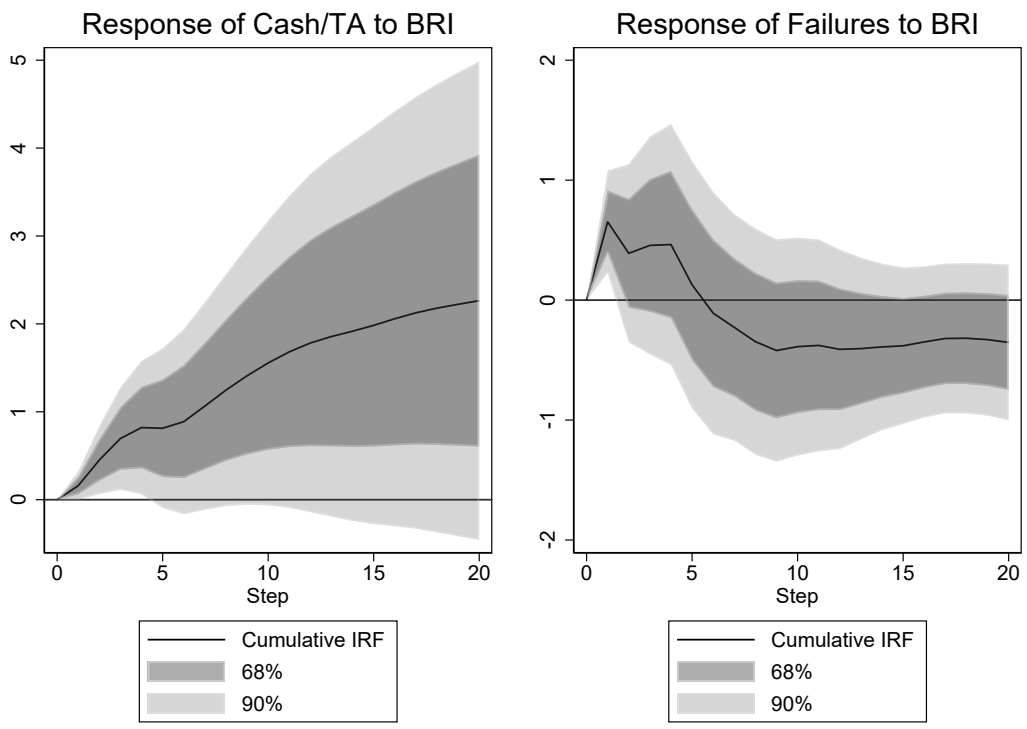


Figure 7. IRFs of Cash/TA

Augmented VAR-estimated impulse-response functions for Cash Ratio and Bank Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006). Data for Leverage is from Jordà et al. (2017). Identification is based on ten lags.

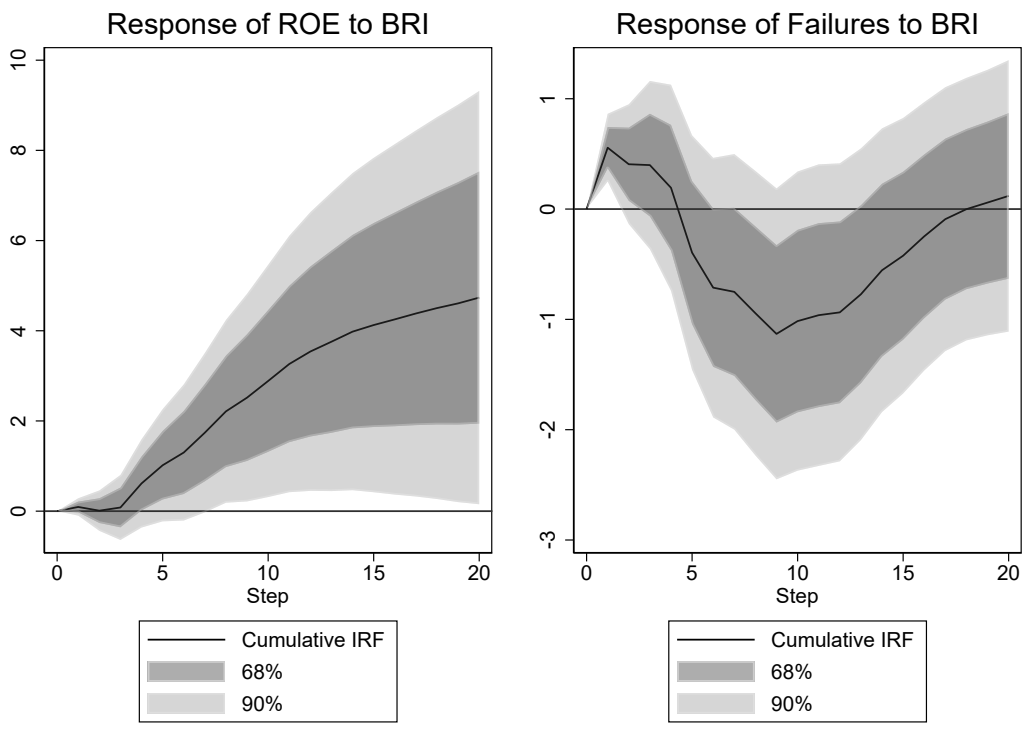


Figure 8. IRFs of ROE

Augmented VAR-estimated impulse-response functions for Return-on-Equity and Bank Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006).

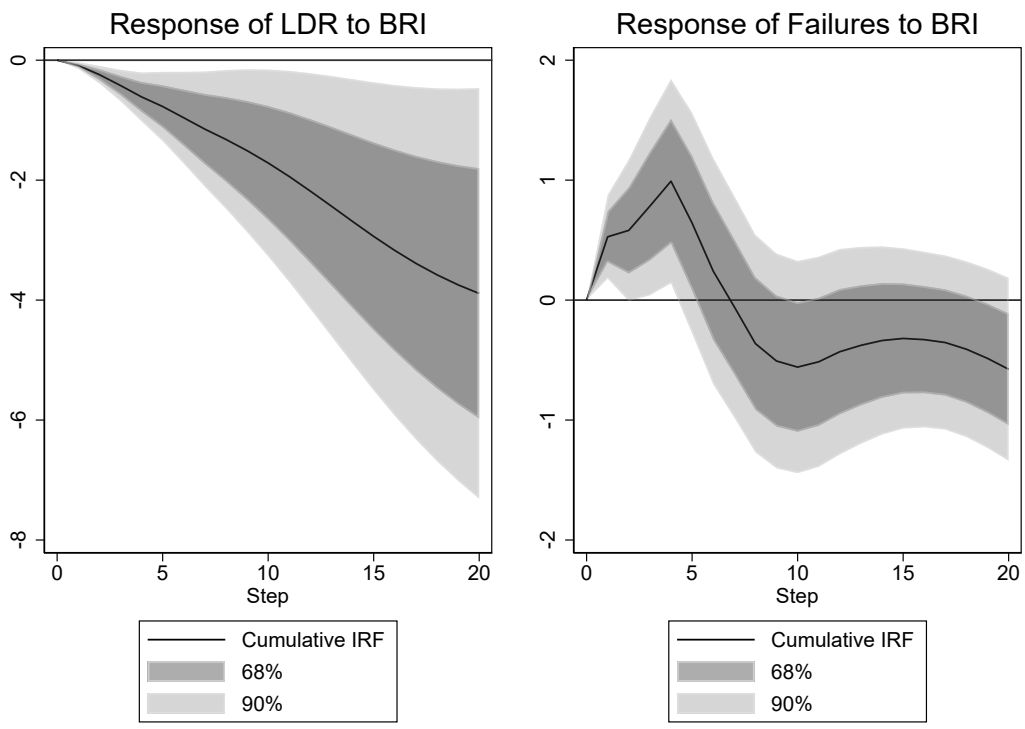


Figure 9. IRFs of LDR

Augmented VAR-estimated impulse-response functions for IRF Augmented by Loan-to-Deposit Ratio and Bank Regulation Index. 68% and 90% confidence bands are used following Sims and Zha (1999) and Sims and Zha (2006). Data for Leverage is from Jordà et al. (2017). Identification is based on ten lags.

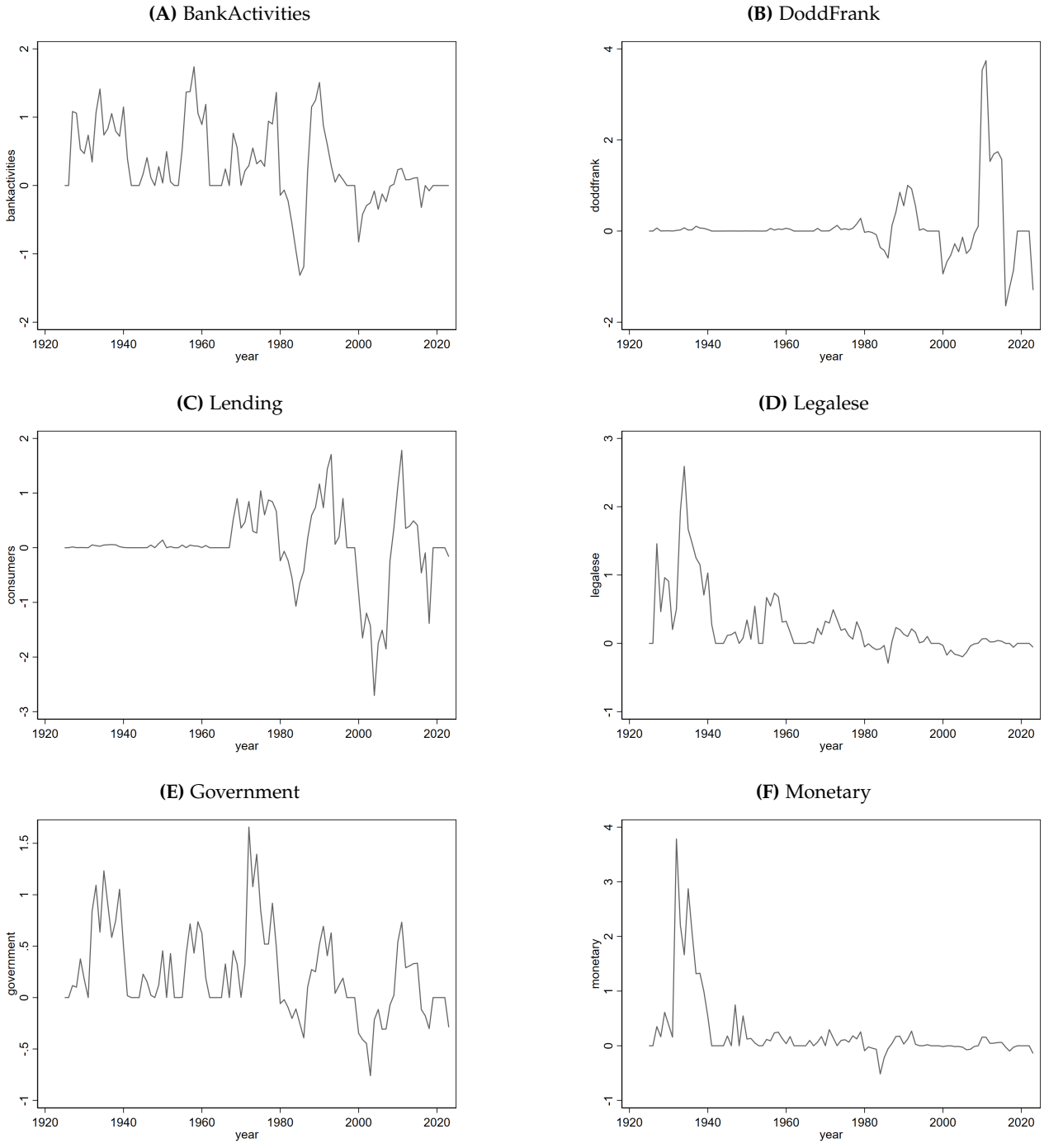
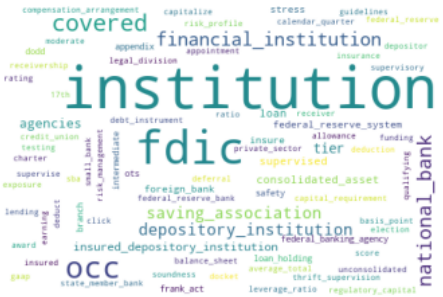


Figure 10. Time-Series Plots for LDA Topics

This figure plots the sub-index associated with each topic obtained from LDA (Appendix B).

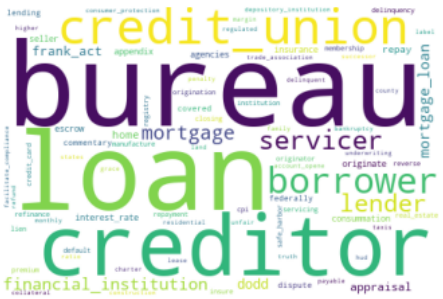
(A) Exposure

(B) Institutions



(C) Banking Entity

(D) Lending



(E) Credit Rating

(F) Accounting

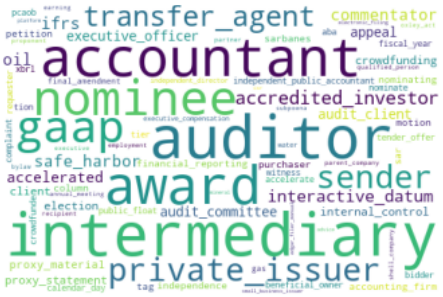


Figure 11. Federal Register: Word Clouds for LDA Topics

LDA provides two different distributions: a distribution of each document on the topics and a distribution of each topic on a set of words or terms. These word clouds, associated with each topic, are obtained by applying Latent Dirichlet Allocation (LDA) on Federal Register text. See Appendix B for details on the LDA procedure. The term distribution of each topic can be used to create the word clouds associated with the topic. The size of each term in the image is proportionate to the score it receives in the LDA distribution.

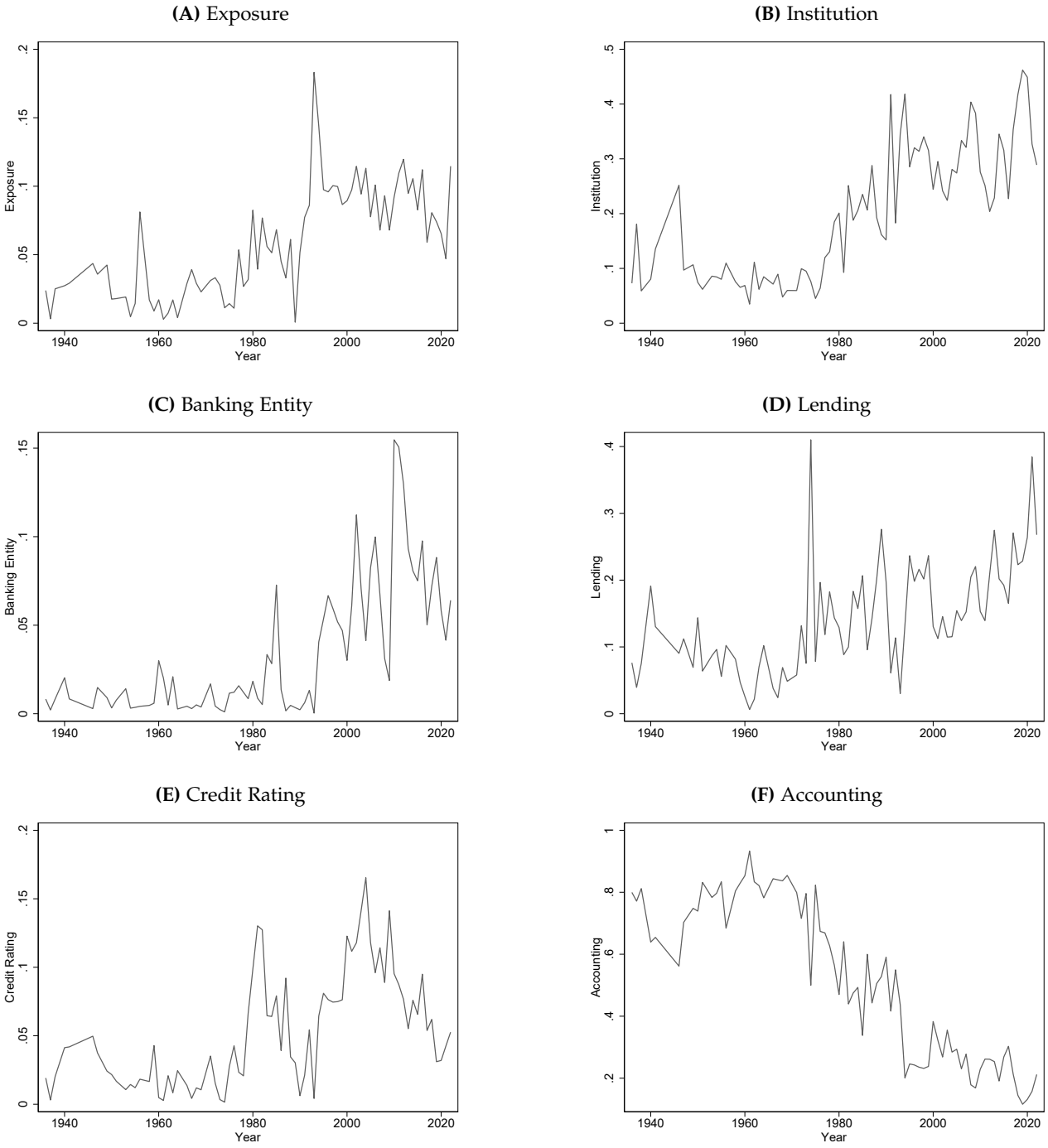


Figure 12. Time-Series Plots for Federal Register LDA Topics

This figure plots the weight associated with each topic obtained from LDA (Section 6).

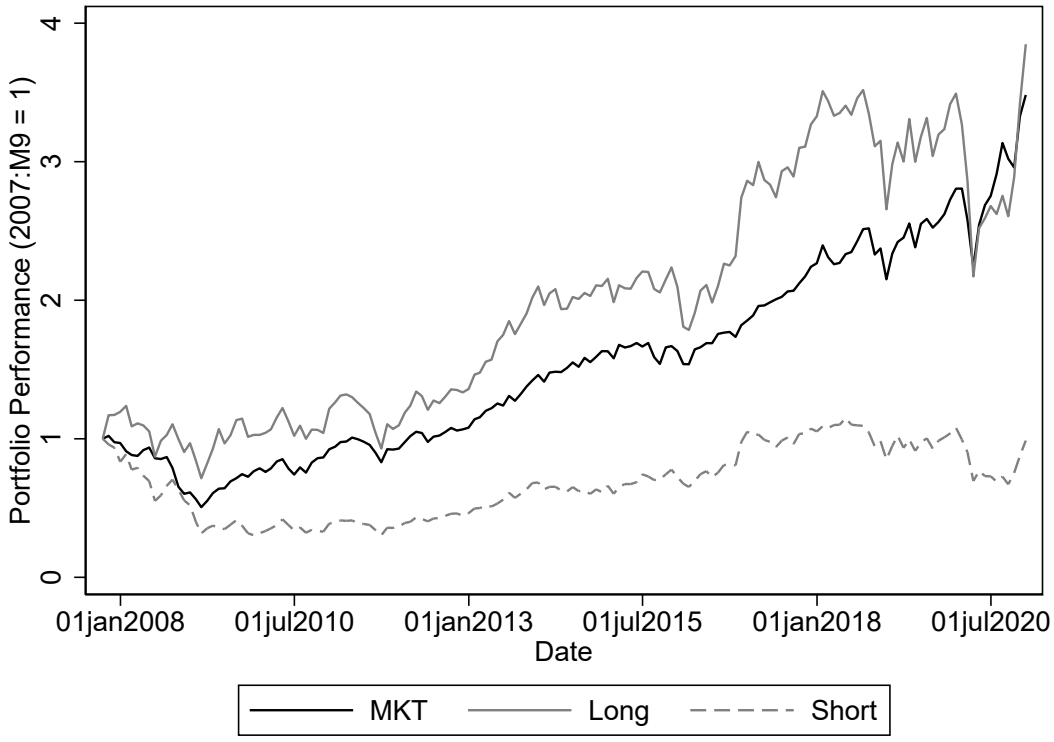


Figure 13. Performance of Long and Short Portfolios

This figure plots the performance of $1 invested in the long, market, and short portfolios. LDA model, trained on the Federal Register, is applied to the earnings call text, yielding weight for the *Lending* topic. Bank stocks in the following quarter are sorted into deciles by weight of the *Lending* topic (Section 6). *Long* and *Short* show the performance of a portfolio of bank stocks in the lowest and highest decile, respectively.

Table 1. Summary Statistics

BRI_t, the value of annual Bank Regulation Index. Similarly, IRI_t and DRI_t are values of Increased Regulation Index and Decreased Regulation Index, respectively. ΔGDP_t is GDP growth, π_t is inflation and r_{t-1} is short-term interest rate from Jordà et al. (2017). $LDR_{i,t}$, $Cash/TA_{i,t}$, Lev_t, $\ln(TA_{i,t})$ are Loan-to-Deposit Ratio, Cash Ratio, Leverage and log of Total Assets (in $ Millions) for year t. $ROE_{i,t}$ and $ROA_{i,t}$ is Net Income as a % of Equity and Total Assets, respectively. R_t and AR_t are annual stock return and annual abnormal return of year t, respectively. $\sigma(R_{i,t})$ and $\sigma(\epsilon_{i,t})$ volatility of stock return and the idiosyncratic volatility for bank i in year t, respectively.

Variable	Mean	Std. Dev.	P25	P50	P75
Year-Level					
BRI_t	1.61	2.64	0.00	1.79	3.18
IRI_t	1.98	1.97	0.00	1.61	3.26
DRI_t	0.80	1.30	0.00	0.00	1.61
ΔGDP_t	6.44	6.15	3.97	5.96	9.04
π_t	2.96	3.43	1.35	2.49	4.16
r_t	3.92	3.32	1.16	3.11	5.66
Bank-Year Level					
$LDR_{i,t}$	85.51	26.59	72.52	85.20	97.21
$Cash/TA_{i,t}$	5.89	6.82	2.10	3.26	5.99
$\ln(TA_{i,t})$	7.70	1.65	6.51	7.39	8.63
Lev_t	12.13	4.56	9.43	11.33	13.69
$ROE_{i,t}$	8.57	11.05	6.84	10.24	13.41
$ROA_{i,t}$	0.79	0.81	0.61	0.92	1.18
$R_{i,t}$	11.10	31.74	-8.45	8.34	29.39
$AR_{i,t}$	4.36	25.91	-9.39	0.58	17.79
$\sigma(\epsilon_{i,t})$	7.15	3.73	4.54	6.24	8.74
$\sigma(R_{i,t})$	7.54	3.79	4.88	6.72	9.32
$DD_{i,t}$	3.25	4.44	0.20	2.98	6.12

Table 2. LDA: Topic and Laws

Latent Dirichlet Allocation (LDA) is a machine-learning technique that analyzes sets of documents — in this case, a corpus of newspaper articles — to provide a distribution of each document over a specified number of topics, which in our study is set to six. It also determines how frequently certain words are associated with these topics, as illustrated in the accompanying word cloud visualizations (Figure 2). Given that LDA assigns a distribution of topics to each article, we can calculate the mean topic distribution for each piece of legislation mentioned within these articles. The table resulting from this analysis categorizes each topic and provides examples of laws. These examples are accompanied by the proportion (in third column) that denotes the extent to which a particular law is representated by a given topic.

Topic	Laws	Share
BankActivities	Bank Holding Company Act of 1956	0.51
	Community Reinvestment Act of 1977	0.42
	Garn-St Germain Depository Institutions Act of 1982	0.42
	Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994	0.36
DoddFrank	Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010	0.63
	Jumpstart Our Business Startups of 2012	0.65
	Economic Growth, Regulatory Relief, and Consumer Protection Act of 2018	0.47
	Financial Services Regulatory Relief Act of 2006	0.52
Lending	Credit CARD Act of 2009	0.87
	Bankruptcy Abuse Prevention and Consumer Protection Act of 2005	0.65
	Federal Deposit Insurance Corporation Improvement Act of 1991	0.56
	Monetary Control Act of 1980	0.49
Legalese	McFadden Act of 1927	0.43
	Federal Deposit Insurance Act of 1950	0.37
Government	Bank Secrecy Act of 1970	0.53
	Bank Holding Company Act Amendments of 1970	0.47
	Federal Credit Union Act of 1934	0.47
	Bank Protection Act of 1968	0.32
Monetary	Banking Act of 1933	0.41
	Federal Financing Bank Act of 1973	0.36
	Banking Act of 1935	0.34
	Export-Import Bank Act of 1945	0.33

Table 3. News about Bank Regulations, Topics and Sentiment

The dependent variable is the Sentiment of the news text calculated using FinBERT. *dereg* is a dummy variable indicating the mentioned Law is of deregulatory nature. Sentiment is calculated using a 3-sentence window around the mention of the Law Name, nickname or (4-letter or more) abbreviation in the news text. *BankActivities, DoddFrank, Lending, Legalese, Government, and Monetary* are assigned labels of six topics obtained from LDA. Standard Errors reported in brackets are double-clustered by Year and by Law.

	FinBERT Sentiment						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
dereg	0.056***	0.055***	0.055***	0.055***	0.055***	0.055***	0.055***
	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)
BankActivities	0.025***						0.005
	(0.009)						(0.012)
DoddFrank		0.011					0.016
		(0.007)					(0.021)
Lending			-0.003				0.010
			(0.012)				(0.025)
Legalese				-0.071***			-0.071***
				(0.009)			(0.014)
Government					-0.004		-0.011
					(0.021)		(0.020)
Monetary						0.037***	
						(0.011)	
Observations	6,261	6,261	6,261	6,261	6,261	6,261	6,261
R-squared	0.112	0.111	0.111	0.115	0.111	0.112	0.115
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Clusters	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Law Clusters	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Determinants of Bank Regulation

The dependent variable is BRI_t, the value of annual Bank Regulation Index. Similarly, IRI_t and DRI_t are values of Increased Regulation Index and Decreased Regulation Index, respectively. ΔGDP_{t-1} is last year's GDP growth. π_{t-1} is last year's inflation. r_{t-1} is last year's short-term interest rate. $BankFailures_{t-1}$ are defined as Deposits of failed banks as a percentage of total deposits in year $t-1$. $Republican_t$ is a dummy variable indicating Government being held by the Republican Party. Newey-West Standard Errors with 12 lags are reported in brackets.

	(1) BRI_t	(2) BRI_t	(3) BRI_t	(4) BRI_t	(5) IRI_t	(6) DRI_t
$BankFailures_{t-1}$	0.597***			0.488**	0.417**	-0.071
	(0.174)			(0.238)	(0.174)	(0.080)
$Republican_t$			-1.853**	-1.453	-1.148	0.305
			(0.940)	(0.949)	(0.701)	(0.473)
ΔGDP_{t-1}		0.012		0.025	0.010	-0.015
		(0.041)		(0.034)	(0.036)	(0.011)
π_{t-1}		-0.144		-0.148	-0.115	0.033
		(0.135)		(0.124)	(0.105)	(0.047)
r_{t-1}		-0.184**		-0.072	0.044	0.117
		(0.076)		(0.086)	(0.076)	(0.084)
Constant	1.331**	2.654***	2.523***	2.660***	2.733***	0.073
	(0.543)	(0.707)	(0.701)	(0.784)	(0.722)	(0.260)
Observations	95	96	96	95	95	95
R-squared	0.111	0.121	0.125	0.264	0.238	0.201
Std. Err.	NW	NW	NW	NW	NW	NW

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Regulations' Short-term Impact on Banks

The dependent variable is shown in column title for year t. BRI_{t-1} is the lagged value of annual Bank Regulation Index. Similarly, IRI_{t-1} and DRI_{t-1} are lagged values of Increased Regulation Index and Decreased Regulation Index, respectively. $LDR_{i,t}$, $DD_{i,t}$ and $Cash/TA_{i,t}$ are Loan-to-Deposit Ratio, Distance-to-Default and Cash Ratio for year t. $ROE_{i,t}$ is Net Income as a % of Equity. AR_t and $\sigma(\epsilon_{i,t})$ are annual abnormal return and the idiosyncratic volatility for bank i in year t, respectively. Bank-level *Controls* are LDR, $\ln(Total\_Assets)$, Leverage and Cash Ratio of year $t-1$. Newey-West Standard Errors with 12 lags are reported in brackets.

Panel A. Short-term Impact of BRI_{t-1}

	$ROE_{i,t}$	$\sigma(\epsilon_{i,t})$	$\Delta LDR_{i,t}$	ΔCash/TA$_{i,t}$	$DD_{i,t}$	$AR_{i,t}$
BRI_{t-1}	-0.321***	0.118***	-0.141***	0.019*	-0.018	-0.229**
	(0.045)	(0.015)	(0.042)	(0.011)	(0.017)	(0.104)
Observations	8,317	8,668	8,576	8,556	8,535	8,535
R-squared	0.109	0.136	0.161	0.169	0.088	0.088
Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Std. Err.	NW	NW	NW	NW	NW	NW

Panel B. Separating BRI_{t-1} into IRI_{t-1} and DRI_{t-1}

	$ROE_{i,t}$	$\sigma(\epsilon_{i,t})$	$\Delta LDR_{i,t}$	ΔCash/TA$_{i,t}$	$DD_{i,t}$	$AR_{i,t}$
IRI_{t-1}	-0.242***	0.030	-0.017	-0.010	0.091***	0.858***
	(0.057)	(0.020)	(0.055)	(0.018)	(0.024)	(0.144)
DRI_{t-1}	0.928***	-0.570***	0.814***	-0.227***	0.625***	3.975***
	(0.105)	(0.029)	(0.096)	(0.030)	(0.040)	(0.266)
Observations	8,317	8,668	8,576	8,556	8,535	8,535
R-squared	0.109	0.136	0.161	0.169	0.088	0.088
Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 6. Regulations' Short-term Impact on Banks

The dependent variable is shown in column title for year t. BRI_{t-1} is the lagged value of annual Bank Regulation Index. Similarly, IRI_{t-1} and DRI_{t-1} are lagged values of Increased Regulation Index and Decreased Regulation Index, respectively. *Large Bank* is a dummy indicating that bank is in top tercile by size for year t. $LDR_{i,t}$, $DD_{i,t}$ and $Cash/TA_{i,t}$ are Loan-to-Deposit Ratio, Distance-to-Default and Cash Ratio for year t. $ROE_{i,t}$ is Net Income as a % of Equity. AR_t and $\sigma(\epsilon_{i,t})$ are annual abnormal return and the idiosyncratic volatility for bank i in year t, respectively. Bank-level *Controls* are LDR, $\ln(Total\_Assets)$, Leverage and Cash Ratio of year $t-1$. Newey-West Standard Errors with 12 lags are reported in brackets.

	$AR_{i,t}$	$ROE_{i,t}$	$\sigma(\epsilon_{i,t})$	$\Delta LDR_{i,t}$	ΔCash/TA$_{i,t}$	$DD_{i,t}$
$IRI_{t-1} \times$ *Large Bank*	-1.494***	0.206**	-0.200***	-0.043	-0.017	0.049
	(0.226)	(0.095)	(0.029)	(0.095)	(0.026)	(0.031)
$DRI_{t-1} \times$ *Large Bank*	-2.601***	-0.574***	0.076	-0.311	0.089**	-0.229***
	(0.405)	(0.153)	(0.049)	(0.190)	(0.042)	(0.058)
Large Bank	5.948***	1.584***	0.045	0.547	-0.090	0.545***
	(1.381)	(0.589)	(0.210)	(0.594)	(0.145)	(0.180)
Observations	9,247	8,798	9,096	9,080	9,059	8,963
R-squared	0.020	0.018	0.048	0.163	0.189	0.019
Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 7. Short- and Long-Term Dichotomy in Impact of Regulations

The dependent variable is shown in row title i in year t. $\beta_{i,t}^{reg}$ are the winsorized (1% and 99%) and standardised values of -β_1 from Section 4.3. $ROE_{i,t}$, AR_t, DD_t, $\sigma(\epsilon_{i,t})$, $Cash/TA_{i,t}$ and $LDR_{i,t}$, are Net Income as a % of Equity, Annual Abnormal Return, Distance-to-Default, Cash Ratio and Loan-to-Deposit Ratio bank i for year t with lead order shown in column number. NW Standard Errors with 12 lags are reported in brackets.

	(t)	$(t+2)$	$(t+4)$	$(t+6)$	$(t+8)$	$(t+10)$
			$ROE_{i,t}$			
$\beta_{i,t}^{reg}$	-1.072***	-0.973***	0.090	0.444*	0.737***	0.354
	(0.197)	(0.174)	(0.188)	(0.244)	(0.227)	(0.259)
Observations	8,239	6,768	5,499	4,483	3,673	2,979
R-squared	0.085	0.097	0.099	0.071	0.063	0.048
			$AR_{i,t}$			
$\beta_{i,t}^{reg}$	-1.290***	-0.190	-0.248	1.058**	1.660***	0.500
	(0.415)	(0.363)	(0.400)	(0.419)	(0.467)	(0.437)
Observations	8,585	7,259	6,002	4,970	4,090	3,344
R-squared	0.037	0.041	0.017	0.011	0.016	0.014
			$DD_{i,t}$			
$\beta_{i,t}^{reg}$	-0.079	-0.385***	0.141**	0.163***	0.222***	-0.049
	(0.065)	(0.058)	(0.062)	(0.063)	(0.072)	(0.075)
Observations	8,425	6,938	5,654	4,625	3,801	3,097
R-squared	0.067	0.103	0.053	0.136	0.128	0.046
			$\sigma(\epsilon_{i,t})$			
$\beta_{i,t}^{reg}$	0.203***	0.344***	-0.004	-0.160*	-0.265***	0.010
	(0.062)	(0.066)	(0.062)	(0.083)	(0.089)	(0.090)
Observations	8,561	7,242	5,980	4,948	4,068	3,325
R-squared	0.100	0.119	0.109	0.123	0.137	0.040
			$LDR_{i,t}$			
$\beta_{i,t}^{reg}$	-0.388***	-0.818***	-1.458***	-1.006***	-0.143	0.382
	(0.137)	(0.220)	(0.258)	(0.290)	(0.258)	(0.288)
Observations	8,424	6,939	5,661	4,633	3,808	3,105
R-squared	0.545	0.211	0.091	0.067	0.067	0.059
			$Cash/TA_{i,t}$			
$\beta_{i,t}^{reg}$	0.046	0.213***	0.299***	0.083	-0.063	-0.170
	(0.033)	(0.055)	(0.077)	(0.083)	(0.081)	(0.104)
Observations	8,404	6,905	5,622	4,587	3,763	3,070
R-squared	0.374	0.106	0.053	0.049	0.026	0.014
Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macro Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 8. Predicting Bank Failures with the Bank Regulation Index

This table shows a probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 0.5$ and 0 otherwise. $BankFailures_t$ is the percentage of deposits in failed banks over total deposits in year t. Following Jordà et al. (2021), $\Delta_5 Loans/GDP$ is the average change in $Loans/GDP$ ratio over the last 5 years. $\Delta_{t-10 \to t-5} BRI$ is the average change in BRI_t over the years $t-10$ to $t-5$. $CapitalRatio_{t-1}$ is the Capital Ratio lagged by one year. LDR_{t-1} is the Loans-to-Deposits ratio lagged by one year. $Mortgages/GDP_{t-1}$ is the ratio of Mortgages to GDP lagged by one year. These variables are calculated from Jordà et al. (2017) data.

	$Crisis_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta_{t-10 \to t-5} BRI$		-0.927***		-0.917***		-0.462**
		(0.299)		(0.331)		(0.208)
Known Predictors						
$\Delta_5 Loans/GDP$	0.304**	-0.593*	0.384**	-0.992*		
	(0.152)	(0.332)	(0.195)	(0.552)		
$CapitalRatio_{t-1}$	-4.442	8.627	-1.181	13.795*		
	(3.446)	(6.216)	(3.847)	(7.795)		
LDR_{t-1}			0.043**	0.055		
			(0.018)	(0.049)		
$Mortgages/GDP_{t-1}$					0.049**	0.030
					(0.021)	(0.047)
Constant	-0.644	-3.243**	-4.279***	-8.014*	-2.565***	-2.492*
	(0.579)	(1.299)	(1.621)	(4.617)	(0.655)	(1.497)
Observations	96	87	95	86	95	87
Pseudo-R^2	0.0963	0.550	0.220	0.591	0.0940	0.464

Table 9. Predicting Bank Failures: Decomposing BRI into Regulatory vs. Deregulatory Indices

This table shows probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 0.5$ and 0 otherwise. $BankFailures_t$ is the percentage of deposits in failed banks over total deposits in year t. Following Jordà et al. (2021), $\Delta_5 Loans/GDP$ is the average change in $Loans/GDP$ ratio over the last 5 years (year $t-6$ to year $t-1$). $\Delta_{10\to5} IRI$ and $\Delta_{10\to5} DRI$ is the average change in Increasing and Decreasing Regulation Index over the years $t-10$ to $t-5$. $CapitalRatio_{t-1}$ is the Capital Ratio lagged by one year. LDR_{t-1} is the Loans-to-Deposits ratio lagged by one year. $Mortgages/GDP_{t-1}$ is the ratio of Mortgages to GDP lagged by one year. These variables are calculated from Jordà et al. (2017) data.

	$Crisis_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta_{t-10\to t-5} IRI$		-0.852**		-1.133**		-0.329
		(0.377)		(0.536)		(0.254)
$\Delta_{t-10\to t-5} DRI$		0.964***		0.854**		0.622**
		(0.330)		(0.343)		(0.276)
Known Predictors						
$\Delta_5 Loans/GDP$	0.301**	-0.581*	0.381*	-1.158*		
	(0.152)	(0.342)	(0.196)	(0.665)		
$CapitalRatio_{t-1}$	-4.175	7.964	-0.888	17.709		
	(3.413)	(6.527)	(3.817)	(11.169)		
LDR_{t-1}			0.044**	0.069		
			(0.018)	(0.056)		
$Mortgages/GDP_{t-1}$					0.049**	0.023
					(0.021)	(0.052)
Constant	-0.694	-3.259**	-4.358***	-9.358*	-2.565***	-2.630
	(0.574)	(1.303)	(1.627)	(5.448)	(0.655)	(1.668)
Observations	97	87	96	86	95	87
Pseudo-R^2	0.0963	0.552	0.220	0.598	0.0940	0.481

Table 10. Predicting Bank Failures: Decomposing BRI into Topics through LDA Methods

This table shows probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 0.5$ and 0 otherwise. $BankFailures_t$ is the percentage of deposits in failed banks over total deposits in year t. This table uses Latent Dirichlet Allocation (LDA) to decompose BRI into different topics, as explained in Appendix B. LDA provides a distribution of each news article over the 6 topics. The BRI of year t is then decomposed into six different topics for year t using distribution provided by LDA for news articles of year t. $\Delta_{t-10\to t-5}$ shows the average change of the subindex from year $t-10$ to year $t-5$. Figure 2 shows Word Cloud associated with each topic.

	$Crisis_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta_{t-10\to t-5}BankActivities$	-2.001***						-1.347*
	(-3.341)						(-1.711)
$\Delta_{t-10\to t-5}DoddFrank$		-0.791**					0.370
		(-2.234)					(0.664)
$\Delta_{t-10\to t-5}Lending$			-1.286***				-0.974**
			(-3.911)				(-2.116)
$\Delta_{t-10\to t-5}Legalese$				-6.929***			-1.289
				(-3.089)			(-0.394)
$\Delta_{t-10\to t-5}Government$					-3.401***		0.011
					(-3.036)		(0.007)
$\Delta_{t-10\to t-5}Monetary$						-3.489**	-0.138
						(-2.450)	(-0.096)
Constant	-1.431***	-1.365***	-1.722***	-1.444***	-1.460***	-1.278***	-1.819***
	(-5.813)	(-6.703)	(-6.108)	(-5.175)	(-5.165)	(-6.213)	(-4.980)
Observations	89	89	89	89	89	89	89
Pseudo-R^2	0.337	0.0954	0.369	0.311	0.320	0.158	0.498

Table 11. Performance of Lending-Exposure Bank Stock Portfolios

The table shows the alphas, t-statistics, and R^2 of *Lending* exposure bank stock portfolios from time-series regressions of monthly returns with different factors (Section 6). Sample is Sep-2007 to Dec-2020. Four different models are considered: the CAPM model, Four-Factor model (market ($MKTRF$), size (SMB), value (HML) and momentum (MOM)), Five-Factor model (market ($MKTRF$), size (SMB), value (HML), robust-minus-weak (RMW) and conservative-minus-aggressive (CMA)) and Four-Factor model with robust-minus-weak (RMW) and conservative-minus-aggressive (CMA). All alphas are expressed in percentages. t-statistics are shown in brackets.

	Long-Short			
	(1)	(2)	(3)	(4)
ALPHA	0.75	0.61	0.65	0.64
	(2.37)	(2.01)	(2.03)	(2.07)
MKTRF	0.10	0.15	0.18	0.13
	(1.55)	(2.05)	(2.46)	(1.71)
SMB		-0.34	-0.30	-0.33
		(-2.59)	(-2.18)	(-2.48)
HML		-0.35	-0.07	-0.23
		(-2.97)	(-0.51)	(-1.64)
MOM		-0.24		-0.23
		(-3.34)		(-3.18)
RMW			0.09	0.09
			(0.44)	(0.45)
CMA			-0.46	-0.39
			(-1.89)	(-1.62)
Observations	159	159	159	159
R-squared	0.06	0.18	0.14	0.20

A Data Appendix

Table A.1. Complete List of Laws

Law Date	Law Name	Other Names	Reg
2/25/1927	McFadden Act		1
7/22/1932	Federal Home Loan Bank Act	FHLBA	1
3/9/1933	Emergency Banking Relief Act		1
3/24/1933	State Bank Aid Act		1
6/16/1933	Banking Act of 1933	Glass-Steagall	1
6/26/1934	Federal Credit Union Act		1
8/23/1935	Banking Act of 1935		1
3/4/1939	Export-Import Bank Extension Act		1
6/30/1939	Glass Federal Reserve Note Act		1
7/31/1945	Export-Import Bank Act of 1945		1
9/21/1950	Federal Deposit Insurance Act	FDIA	1
5/9/1956	Bank Holding Company Act of 1956	BHCA, BHC Act	1
9/23/1959	Spence Act (Savings and Loan Holding Companies)	Spence Act	1
10/23/1962	Bank Service Company Act	BSCA	1
10/16/1966	Financial Institutions Supervisory Act of 1966	FISA	1
5/29/1968	Truth in Lending Act	TILA	1
7/7/1968	Bank Protection Act of 1968		1
10/26/1970	Bank Secrecy Act of 1970		1
12/31/1970	Bank Holding Company Act Amendments of 1970	BHCA	1
12/29/1973	Federal Financing Bank Act of 1973		1
12/22/1974	Real Estate Settlement Procedures Act of 1974	RESPA	1
10/12/1977	Community Reinvestment Act	Housing and Community Development Act of 1977	1
11/16/1977	Federal Reserve Reform Act of 1977	FRRA	1
9/17/1978	International Banking Act of 1978		1
11/10/1978	Financial Institutions Regulatory and Interest Rate Control Act of 1978	FIRA	1
3/31/1980	Monetary Control Act of 1980	DIDMCA, Depository Institutions Deregulation Act of 1980	-1
7/27/1981	Cash Discount Act		-1
12/26/1981	International Banking Facility Deposit Insurance Act		-1
10/8/1982	Export Trading Company Act of 1982		-1
10/15/1982	Garn-St Germain Depository Institutions Act of 1982	Garn-St Germain Act, Garn Act	-1
8/10/1987	Competitive Equality Banking Act of 1987	CEBA	1
8/9/1989	Financial Institutions Reform, Recovery, and Enforcement Act of 1989	FIRREA	1
12/12/1991	Resolution Trust Corporation Refinancing, Restructuring, and Improvement Act of 1991		1
12/19/1991	Federal Deposit Insurance Corporation Improvement Act of 1991	FDICIA, Truth in Savings Act	1
9/23/1994	Riegle Community Development and Regulatory Improvement Act of 1994		-1
9/29/1994	Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994	Interstate Act, Riegle-Neal	-1
9/30/1996	Deposit Insurance Funds Act of 1996		-1
7/3/1997	Riegle-Neal Amendments Act of 1997		1
11/12/1999	Gramm-Leach-Bliley Act	GLB Act, GLBA	-1
12/4/2002	FHA Downpayment Simplification Act of 2002		1
10/28/2003	Check Clearing for the 21st Century Act	Check 21 Act	-1
4/20/2005	Bankruptcy Abuse Prevention and Consumer Protection Act of 2005	BAPCPA	-1
10/13/2006	Financial Services Regulatory Relief Act of 2006	FSRRA	-1
5/22/2009	Credit CARD Act of 2009	CARD Act	1
7/21/2010	Dodd-Frank Wall Street Reform and Consumer Protection Act	Dodd-Frank	1
4/5/2012	Jumpstart Our Business Startups	JOBS Act	-1
12/18/2014	Insurance Capital Standards Clarification Act of 2014	Insurance Capital Standards Clarification Act	-1
12/18/2014	American Savings Promotion Act		-1
5/24/2018	Economic Growth, Regulatory Relief, and Consumer Protection Act	EGRRCPA	-1
1/3/2019	RBIC Advisers Relief Act of 2018	RBIC Advisers Relief Act	-1

Table A.2. Predicting Bank Failures: Robustness to Alternative Cutoff

This table shows an alternative cutoff of 1% in defining a banking crisis (i.e., when deposits of failed banks in year t amount to more than 1% of total deposits). It uses a probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 1$ and 0 otherwise. $failures_t$ is the percentage of deposits in failed banks over total deposits in year t. Following Jordà et al. (2021), $\Delta_5 Loans/GDP$ is the average change in $Loans/GDP$ ratio over the last 5 years (year $t-6$ to year $t-1$). $\Delta_{10\to5} BRI$ is the average change in BRI_t over the years $t-10$ to $t-5$. $CapitalRatio_{t-1}$ is the Capital Ratio lagged by one year. LDR_{t-1} is the Loans-to-Deposits ratio lagged by one year. $Mortages/GDP_{t-1}$ is the ratio of Mortgages to GDP lagged by one year. These variables are calculated from Jordà et al. (2017) data.

	$Crisis_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta_5 Loans/GDP$	0.278*	-0.687*	0.336*	-2.428		
	(0.155)	(0.375)	(0.198)	(1.545)		
$CapitalRatio_{t-1}$	-5.657	7.062	-2.790	19.465*		
	(3.739)	(6.393)	(4.123)	(10.648)		
LDR_{t-1}			0.039**	0.224		
			(0.018)	(0.163)		
$Mortages/GDP_{t-1}$					0.036*	0.015
					(0.020)	(0.046)
$\Delta_{10\to5} BRI$		-0.906***		-1.426*		-0.448**
		(0.332)		(0.781)		(0.217)
Constant	-0.554	-3.203**	-3.829**	-22.940	-2.281***	-2.189
	(0.612)	(1.349)	(1.680)	(14.651)	(0.631)	(1.414)
Observations	96	87	95	86	95	87
Pseudo-R^2	0.104	0.521	0.209	0.648	0.0559	0.407

Table A.3. Predicting Bank Failures: Robustness to Baron et al. (2021) crises years

This table shows probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 0.5$, or if the year is included in Baron et al. (2021) definition of banking panics, and 0 otherwise. $failures_t$ is the percentage of deposits in failed banks over total deposits in year t. Following Jordà et al. (2021), $\Delta_5 Loans/GDP$ is the average change in $Loans/GDP$ ratio over the last 5 years (year $t-6$ to year $t-1$). $\Delta_{10\to5} BRI$ is the average change in BRI_t over the years $t-10$ to $t-5$. $CapitalRatio_{t-1}$ is the Capital Ratio lagged by one year. LDR_{t-1} is the Loans-to-Deposits ratio lagged by one year. $Mortages/GDP_{t-1}$ is the ratio of Mortgages to GDP lagged by one year. These variables are calculated from Jordà et al. (2017) data.

	$Crisis_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta_5 Loans/GDP$	0.116	-0.348**	0.033	-0.490**		
	(0.095)	(0.156)	(0.109)	(0.215)		
$CapitalRatio_{t-1}$	-3.069	3.135	-1.412	6.144		
	(2.941)	(3.948)	(3.097)	(4.609)		
LDR_{t-1}			0.027**	0.027		
			(0.012)	(0.016)		
$Mortages/GDP_{t-1}$					0.044**	0.037
					(0.018)	(0.025)
$\Delta_{10\to5} BRI$		-0.420***		-0.385***		-0.169**
		(0.121)		(0.128)		(0.086)
Constant	-0.519	-1.414*	-2.525**	-3.731**	-2.200***	-2.081**
	(0.490)	(0.723)	(1.012)	(1.685)	(0.553)	(0.813)
Observations	96	87	95	86	95	87
Pseudo-R^2	0.0247	0.258	0.0593	0.310	0.0785	0.215

Table A.4. Predicting Bank Failures: Lag Length Robustness

This table shows probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 0.5$ and 0 otherwise. $failures_t$ is the percentage of deposits in failed banks over total deposits in year t. This table uses different lags of the changes in BRI as a robustness check.

	$Crisis_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta_{6\to1}BRI$	0.078						
	(1.212)						
$\Delta_{7\to2}BRI$		-0.030					
		(-0.467)					
$\Delta_{8\to3}BRI$			-0.131*				
			(-1.833)				
$\Delta_{9\to4}BRI$				-0.295***			
				(-3.100)			
$\Delta_{10\to5}BRI$					-0.585***		
					(-3.747)		
$\Delta_{11\to6}BRI$						-0.807***	
						(-3.371)	
$\Delta_{12\to7}BRI$							-0.640***
							(-3.604)
Constant	-1.251***	-1.144***	-1.137***	-1.261***	-1.654***	-2.035***	-1.651***
	(-6.236)	(-6.259)	(-6.240)	(-5.909)	(-4.637)	(-3.589)	(-4.235)
Observations	93	92	91	90	89	88	87
Pseudo-R^2	0.0208	0.00326	0.0581	0.221	0.478	0.575	0.494

Table A.5. Decomposing Federal Register text into Topics through LDA Methods

This table shows probit classification model, where the dependent variable is a dummy that takes the value 1 when $BankFailures_t > 0.5$ and 0 otherwise. $BankFailures_t$ is the percentage of deposits in failed banks over total deposits in year t. This table uses Latent Dirichlet Allocation (LDA) to decompose Federal Register text into different topics, as explained in Appendix B. LDA provides a distribution of each Federal Register document over the 6 topics. The LDA distribution of year t is then used to make six different topic values for year t. $\Delta_{t-10\to t-5}$ shows the average change of the subindex from year $t-10$ to year $t-5$. Figure 11 shows Word Cloud associated with each topic.

	$Crisis_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta_{t-10\to t-5}$*Exposure*	-0.007						-0.049
	(0.016)						(0.065)
$\Delta_{t-10\to t-5}$*Institution*		-0.004					0.019
		(0.008)					(0.020)
$\Delta_{t-10\to t-5}$*Banking Entity*			0.000				0.051
			(0.014)				(0.033)
$\Delta_{t-10\to t-5}$*Lending*				-0.023**			-0.056*
				(0.012)			(0.030)
$\Delta_{t-10\to t-5}$*Credit Rating*					0.016		0.011
					(0.015)		(0.040)
$\Delta_{t-10\to t-5}$*Accounting*						-0.000	0.003
						(0.002)	(0.003)
Constant	-1.057***	-1.045***	-1.069***	-1.120***	-1.109***	-1.070***	-1.311***
	(0.197)	(0.199)	(0.198)	(0.217)	(0.203)	(0.197)	(0.271)
Observations	63	63	63	63	63	63	63
Pseudo-R^2	0.004	0.005	0.000	0.0957	0.020	0.000	0.217

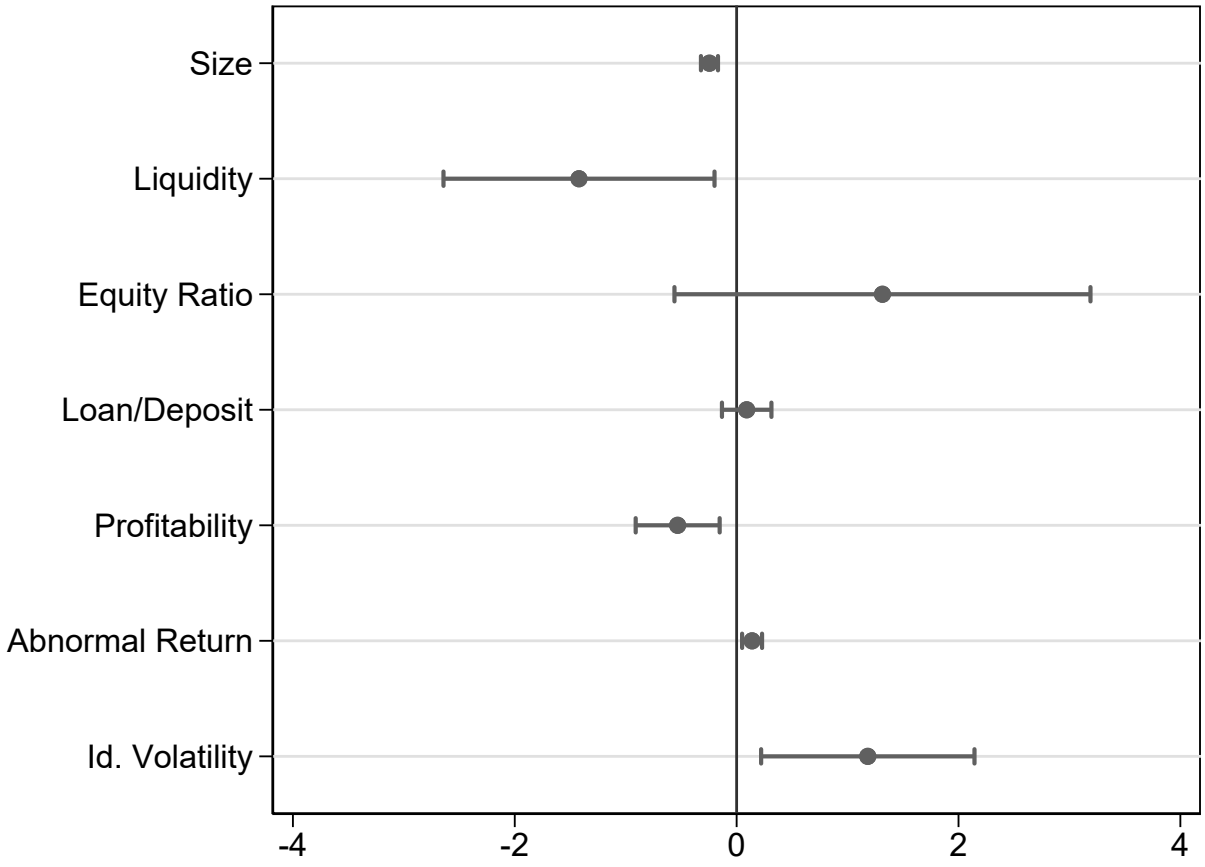


Figure A.1. Regulatory Exposure: Coefficient Plot

This figure plots the coefficients that explain the *Regulatory Exposure*. These estimates are obtained from the following regression:

$$\beta_{i,t}^{reg} = \gamma_0 + BankControls_{t-1} + MacroControls_{t-1} + \delta_i + \gamma_t + \epsilon_{i,t} \tag{12}$$

Bank *Controls* are LDR, $\ln(Total\_Assets)$, Leverage and Cash Ratio of year $t-1$. Macro *Controls* are ΔGDP_{t-1} (last year's GDP growth), π_{t-1} (last year's inflation), r_{t-1} (last year's short-term interest rate). δ_i are bank fixed effects. γ_t are decade fixed effects.

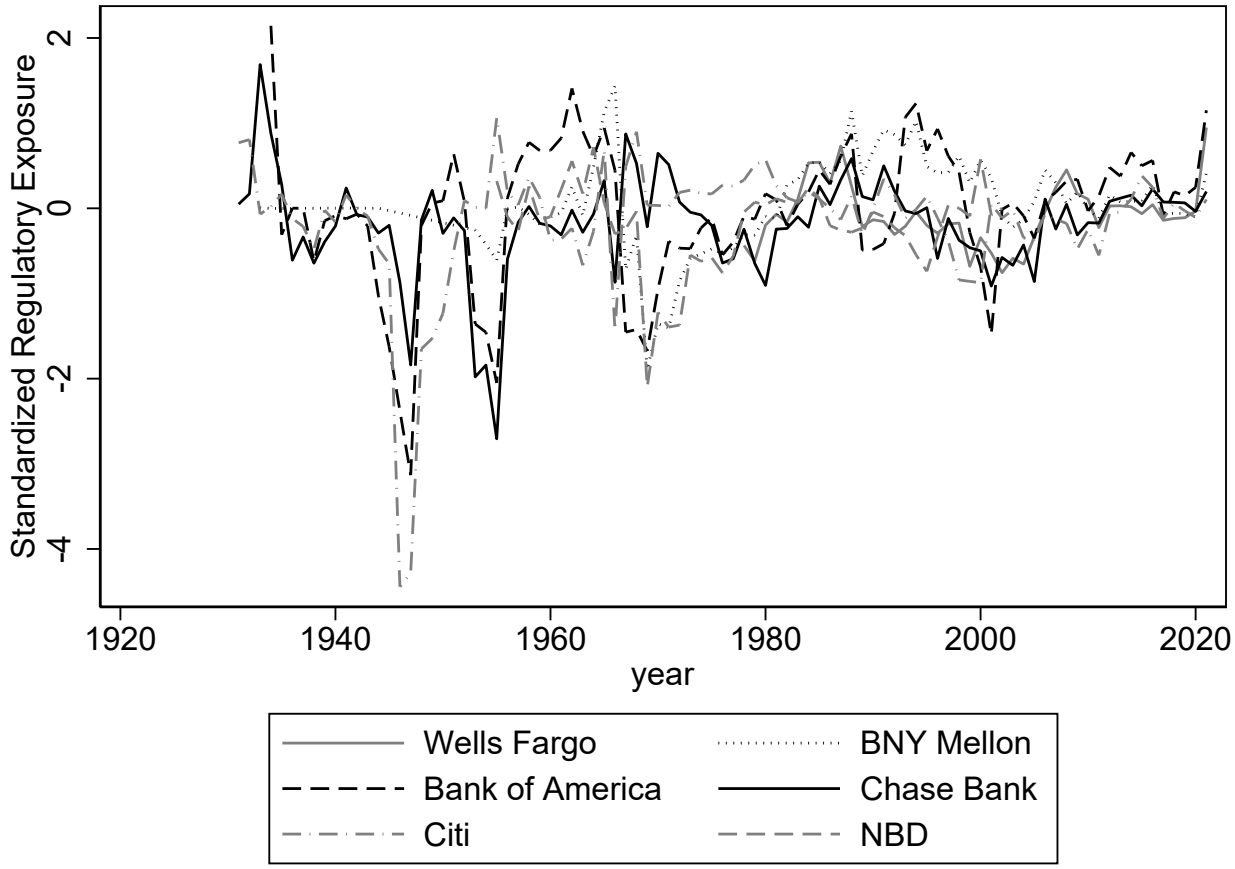
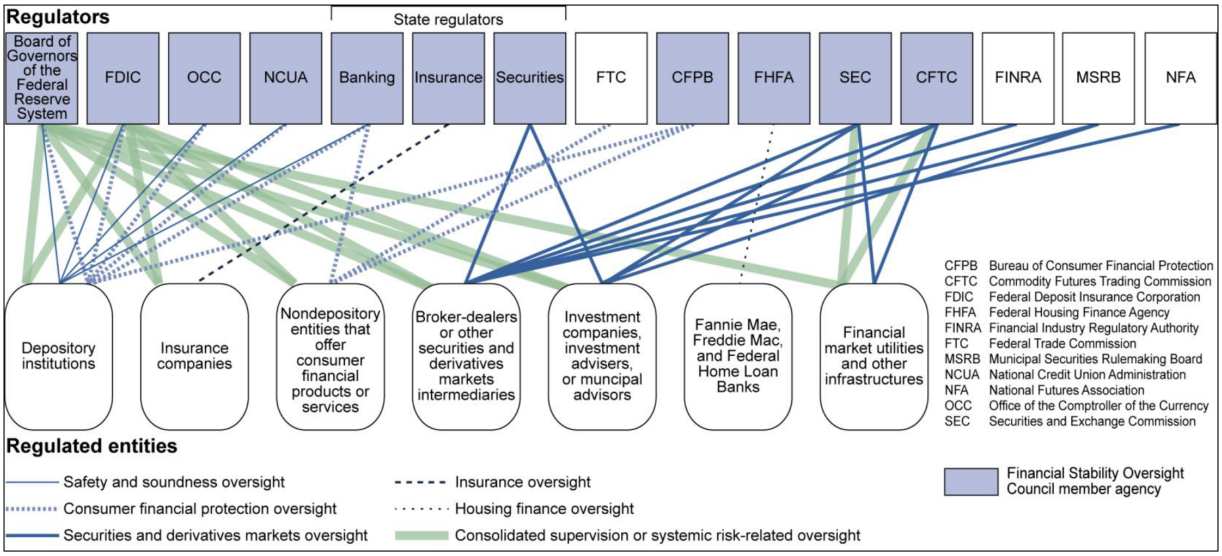


Figure A.2. Regulatory Exposure for Large Banks

This figure plots the *Regulatory Exposure* obtained for a selected sample of 6 large banks.



Source: Government Accountability Office (GAO), *Financial Regulation*, GAO-16-175, February 2016, Figure 2.

Figure A.3. FSOC-Member Agencies

A financial entity can fall under the purview of multiple regulatory bodies due to its involvement in various financial operations, as depicted here (Figure 1 from Labonte (2017)). For instance, a firm could be simultaneously regulated by an institutional overseer and an activity-specific regulator when it partakes in a regulated financial activity, and additionally by a market regulator during its participation in a regulated market. This intricate setup, as demonstrated in this figure, highlights the multifaceted regulatory roles and responsibilities assigned to different overseeing authorities.

B Latent Dirichlet Allocation: Implementation Details

LDA is an unsupervised machine-learning method. A challenge in implementing LDA is to decide the number of topics, as there is no optimal number from an interpretation standpoint. There is always a trade-off between fit statistics and substansive information fit. Following Calomiris et al. (2020), I decide this number to be six. The following steps are then taken in implementing the LDA model.

First, I convert the text to lowercase and use Natural Language Toolkit (NLTK) to *tokenize* the corpus. Then I remove line, paragraph and page breaks. The second step is to remove words that are related to days (Monday, Tuesday, etc), time (month, year etc), distance (miles etc) or numbers (two, thousand, million etc). This list is augmented by stopword list by gensim. Words of length 3 letters or larger are kept and special characters (@, *, etc.) are removed. Words are tagged for their part of speech and I keep adjectives, adverbs, nouns, proper nouns, and verbs. Third, bigrams are created using the NLTK library. Fourth is *lemmtization*, where a word is converted to its root word using spaCy.

The next step is TF-IDF (Term Frequency - Inverse Document Frequency). It is a procedure that scales the frequency of a term in a document by frequency of that term in documents across the corpus. For example, since the word "bank" appears in all news articles (by selection), it gets a low TF-IDF score and I keep terms with scores above a threshold. I keep only those terms that appear in at least 25 documents. Then I use gensim.corpora to create the dictionary and doc2bow to convert documents to vectors. Lastly, I use gensim.Ldamodel to conduct the LDA analysis.

The output of LDA is a distribution of each article i over each topic t. This weight is defined as $w_{i,t}$. Let w^r and w^d denote if the article is about a regulatory or deregulatory law, respectively. For each topic t, for articles dated in year T, the value of the time series plot is calculated as $BRI_{t,T}$:

$$BRI_{t,T} = \ln \quad \frac{\sum_{i \in T} w^r_{i,t} + 1}{\sum_{i \in T} w^d_{i,t} + 1} \Bigg) \Bigg(\tag{13}$$

Figure 10 shows the time-series plot for each subindex.