Windsorization, Imputation, and Unraveling Leverage Pro-Cyclicality in Financial Institutions

Introduction

The aim of this study is to analyze whether US financial institutions show pro or counter-cyclical behaviour.

After pre-processing and inspecting the data in *Section 1*, *Section 2* and *3* show the results of the two-way regression analyses.

Section 2 analyzes the impact of the size growth rate of financial entities on their leverage growth rate over the entire time period, that covers years from 2005 to 2023.

Section 3 applies the same analysis to some selected subperiods and inspects the effect of different types of financial entities. Financial entities are distinguished into financial services (FSs), commercial banks (CBs) and real estate (REs).

Section 1

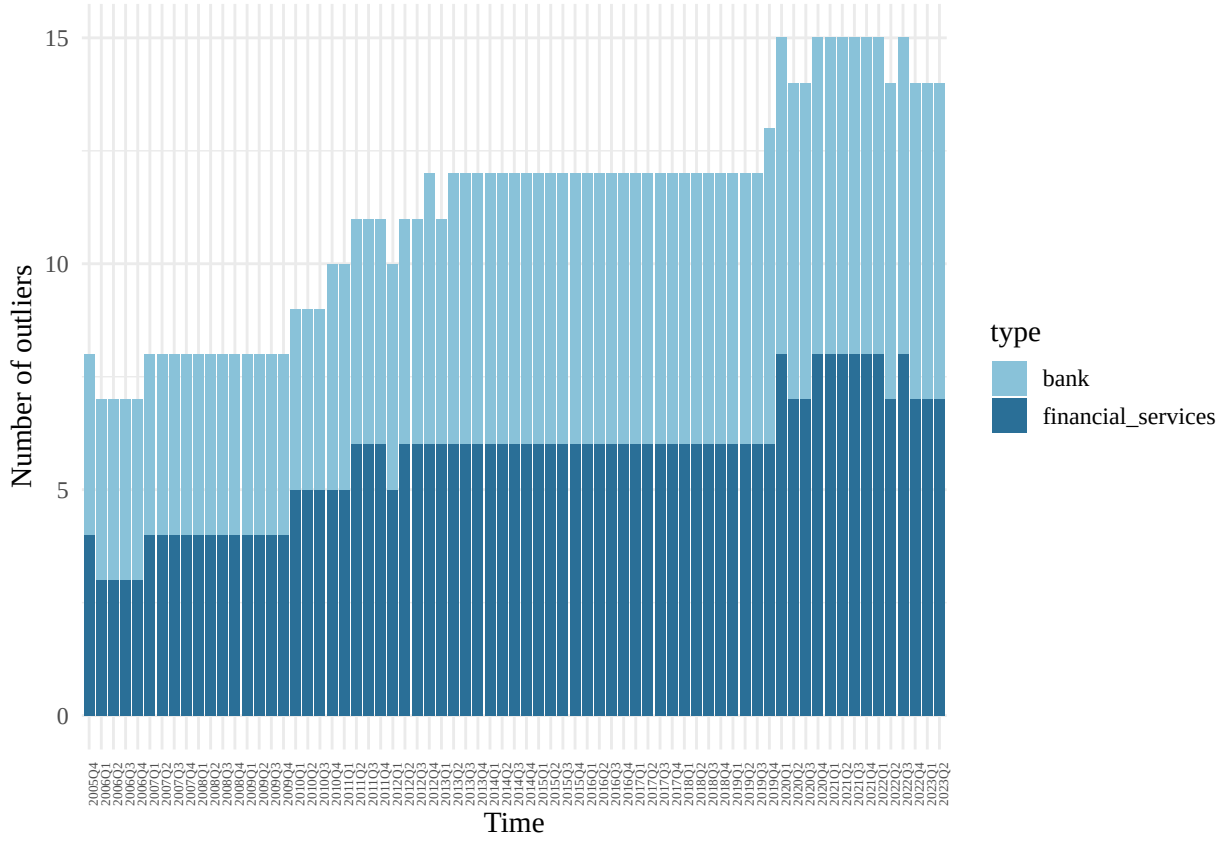
1.1. Preprocessing

Firstly, we create a complete dataset by removing the entities [1] that contain errors and rearranging the data in a panel structure, composed of 643 firms observed in 71 quarters from the 4th quarter of 2005 to the 2nd quarter of 2023.

After that, we inspect the data and we notice there are plenty of large outliers, which distort the distribution of our variables of interest. More specifically, as Figure 1 shows, extreme values are present in banks and to a lesser extent in financial services. Moreover, we can appreciate that their frequency is increasing over time and that it follows a similar pattern for the two categories.

[1] We will be referring to the distinction between banks, financial services and real estate as type of entity, where the entities are the individual firms.

Figure 1: Frequency counts of outliers by type and over time



To solve this problem, we winsorize them by type using the Z-Score method and considering 4 mean absolute deviations from the median as the threshold. In Figure 2, we can appreciate that the winsorization was able to effectively reduce the distortions in the distributions.

Secondly, we further clean the data by dropping 153 entities which have negative leverage values because, by construction, leverage must take only positive values and therefore we can assume that they are entry errors.

Then we inspect missing values and we adopt two different strategies:

- We remove the entities that have more than 22 (30%) missing values for each variable. They are 141.
- We impute missing values for the remaining entities.

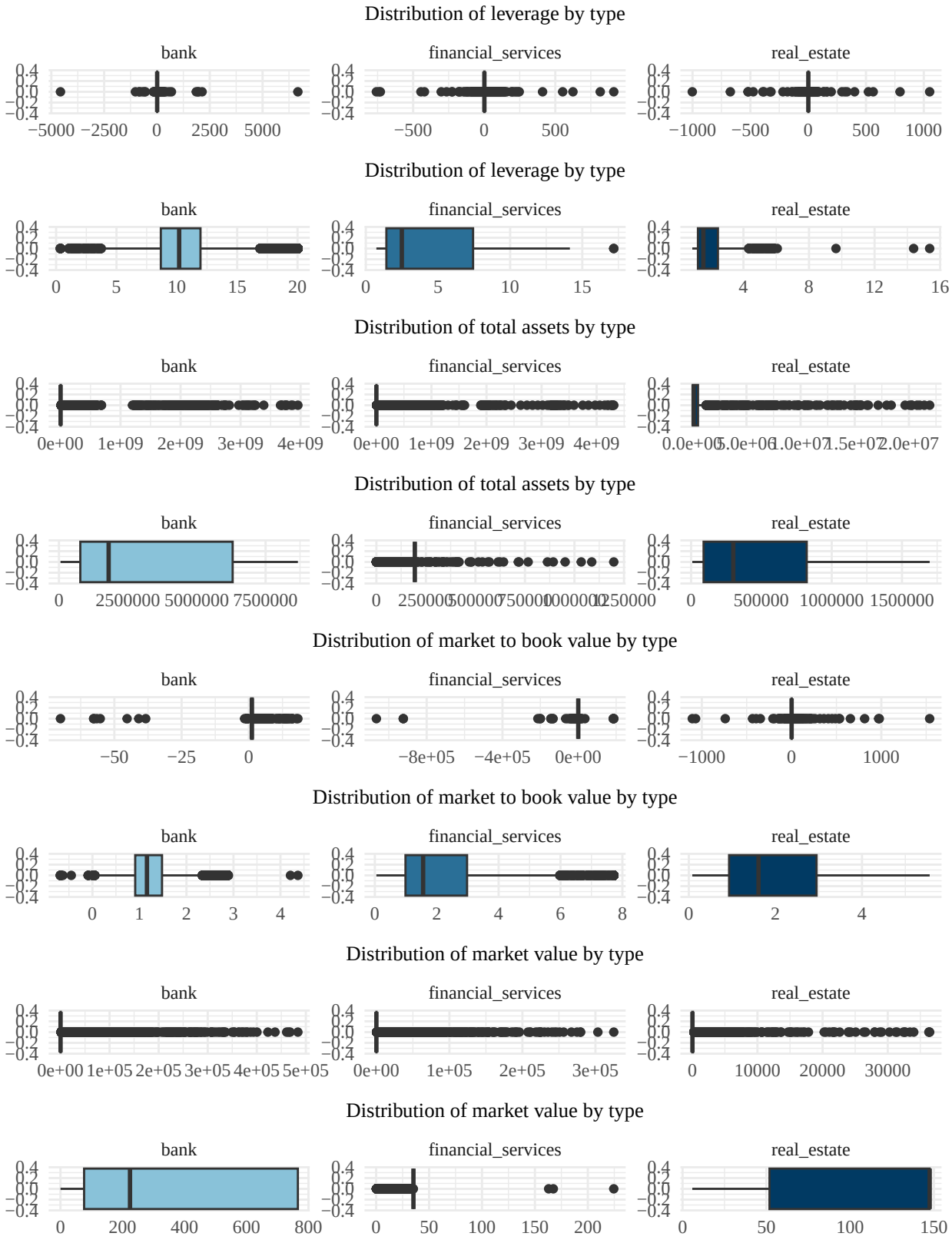
For what concerns the imputation, we use a combination of natural spline and KNN (considering 3 neighbours), to account for both the variability in time and in the cross-section dimension. In particular, we weight the two methods based on the relative position of the missing with respect to the observed values. This imputation should not be distorted by outliers, since we use winsorized data.

Finally, we generate the logarithms and the growth rates transformations for leverage, total assets and market value and we remove 2 entities that present infinite growth rates of market value.

At the end of the data pre-processing, we obtain a balanced panel dataset containing:

- 24637 observations
- 347 entities (273 banks, 59 financial services and 15 real estate)
- 71 quarters

Figure 2: Distributions of interest variables by type before and after windsorization



1.2 Summary statistics

We compute summary statistics on two different datasets, namely the original (not winsorized and with missing values) and the balanced one after winsorization. By comparing them we can confirm that winsorization worked: there are no more large outliers, the standard deviation is reduced, and the skewness is lower. In this report, we include only the summaries of the final dataset for matters of relevance and brevity (Table 1, Table 2, Table 3).

Looking at our statistics we can see that on average the leverage ratio, the market value and the amount of total assets are higher for banks. This means that CBs are on average bigger than REs and FSs and have a higher level of indebtness which can imply lower solvency with respect to the other financial entities. This is coherent with the fact that banks are among the most leveraged institutions in the US. In fact, since they lend out money that is callable at any time, it is safer for them to have a higher leverage ratio than it is for REs and FSs. Finally, REs are slightly bigger than FSs and have a higher level of solvency.

All minimum values are positive, except for the market to book value ratio for banks: this could indicate that, for a specific bank, liabilities are greater than assets. For what concerns the skewness, it is always positive, meaning that all the variables for the three financial entities have extreme values on the right (this is mainly true for the price distribution). The market value for real estate is the only variable with a negative skewness.

The differences just observed in the summary tables can be better appreciated in the bubble plot in Figure 3, which shows the relationship between mean and standard deviation by type for leverage and total assets (while the bubble size represents the numerosity of each group). In the case of leverage we can observe that, for a similar value of standard deviation, the mean for CBs is more than double with respect to the one of FSs.
For total assets there is a positive relationship between the mean and the standard deviation: as the mean increases the standard deviation gets bigger too.

Figure 3: Bubble plots of summary statistics by type of leverage and total assets

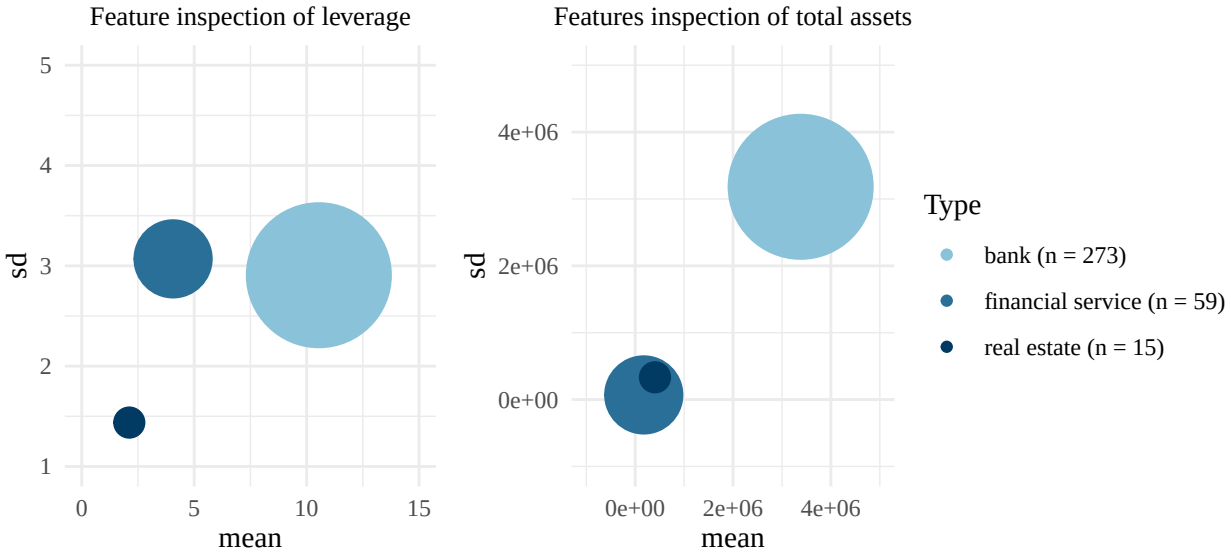


Table 1: Financial services - Summary statistics

name	mean	st_dev	min	q25	q50	q75	max	skewness	kurtosis
leverage	4.061	3.069	0.734	1.425	2.496	7.430	17.180	0.677	-1.070
mkt_2_BV	2.311	1.963	0.050	0.990	1.560	2.980	7.722	1.530	1.526
mkt_value	33.395	7.948	0.220	35.226	35.226	35.226	224.508	3.223	121.230
price	3805.997	32429.784	0.233	12.090	27.750	60.310	528921.000	10.342	119.194
total_assets	174696.623	69609.238	18.000	194120.504	194120.504	194120.504	1198569.667	2.261	37.968

Table 2: Banks - Summary statistics

name	mean	st_dev	min	q25	q50	q75	max	skewness	kurtosis
leverage	10.544	2.906	0.362	8.674	10.193	11.959	20.028	0.700	1.409
mkt_2_BV	1.243	0.519	-0.679	0.910	1.160	1.480	4.371	0.918	1.334
mkt_value	351.649	300.662	0.090	76.235	224.300	765.949	765.949	0.403	-1.559
price	30.462	63.031	0.075	11.751	19.160	30.840	1762.232	12.033	199.525
total_assets	3381974.013	3182301.070	57228.000	776397.500	1805175.000	6304974.500	8672367.449	0.756	-1.085

Table 3: Real estate - Summary statistics

name	mean	st_dev	min	q25	q50	q75	max	skewness	kurtosis
leverage	2.114	1.437	0.879	1.223	1.553	2.443	15.361	2.720	12.973
mkt_2_BV	2.118	1.589	0.080	0.930	1.610	2.950	5.563	0.958	-0.195
mkt_value	105.023	51.284	5.790	51.860	147.452	147.452	147.452	-0.597	-1.362
price	41.260	85.358	1.390	11.590	18.850	34.418	876.546	5.689	38.332
total_assets	404112.568	333011.495	8415.000	88738.000	299908.000	822536.333	1696922.667	0.287	-1.354

1.3 Stylized facts

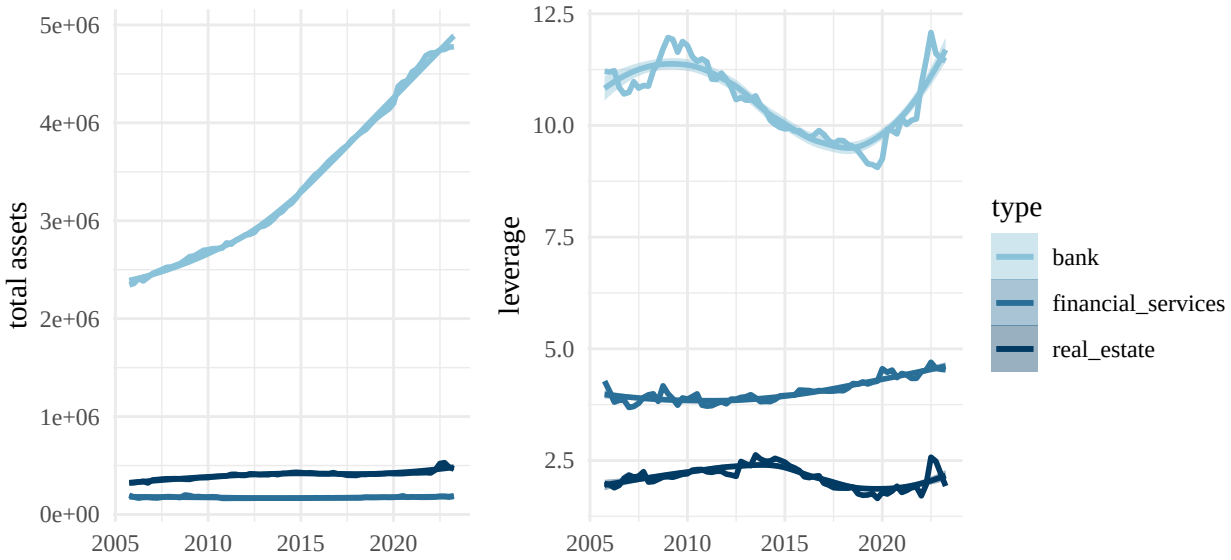
In Figure 4, we represent the evolution over time of the averages of total assets and leverage by type of firm. We also add a smoothing spline to visualize the trend better.

Since the means of both variables do not show interesting fluctuations for financial services and real estate, we focus our considerations on banks. As a matter of fact, banks have constantly increasing total assets, meaning that they are always expanding in size over time. On the other hand, leverage increases from 2005 to 2008's financial crisis, after which it plummets until about 2020. Afterwards, it grows up to even higher values than those of 2007.

We remind that: "leverage (L_t) defined as the ratio between total assets (A_t) over total equity (E_t), is pro-cyclical if $\Delta L_t = f(\Delta A_t)$, and $f' > 0$" (Beccalli et al., 2015). In this light, since leverage is a proxy for debt we can conclude that:

- In the first stage, debt increases more than proportionally with respect to equity and also total assets increase, thus suggesting a pro-cyclical behaviour.
- After the Financial Crisis, since total assets are still increasing, the decreasing leverage must be due to equity growing more than proportionally with respect to debt. This implies a counter-cyclical behaviour, that could be due to the regulations applied to prevent future crises. In this context, the Dodd Frank Act (2010) tried to achieve financial stability and consumer financial protection, while the BASEL III (2010-2018) introduced a minimum of 3% for what concerns the banks' leverage ratio [2]. Another explanation is that, after the Financial Crisis and the money market collapsed, banks stopped lending to each other due to a lack of mutual trust.
- Lastly, we can notice that in the last phase, the leverage becomes again pro-cyclical, as the first one. This may be due to the COVID-19 pandemic.

Figure 4: leverage and Total Asset's averages by type over time



We look further into the relationship of leverage versus total assets (Figure 5) and market value (Figure 6) through scatterplots. total assets growth rate and leverage growth rate show, in general, a positive relationship. The same relationship is analyzed for the three different types of entity and it is confirmed to be positive. On the other hand, the market value growth rate and leverage growth rate show no significant relationship neither in the general case nor in the three different types of entity.

[2] To note that the BASEL III defines this ratio as the ratio between Regulatory Capital and Average Total Assets

Figure 5: Scatterplots of total assets growth and leverage growth

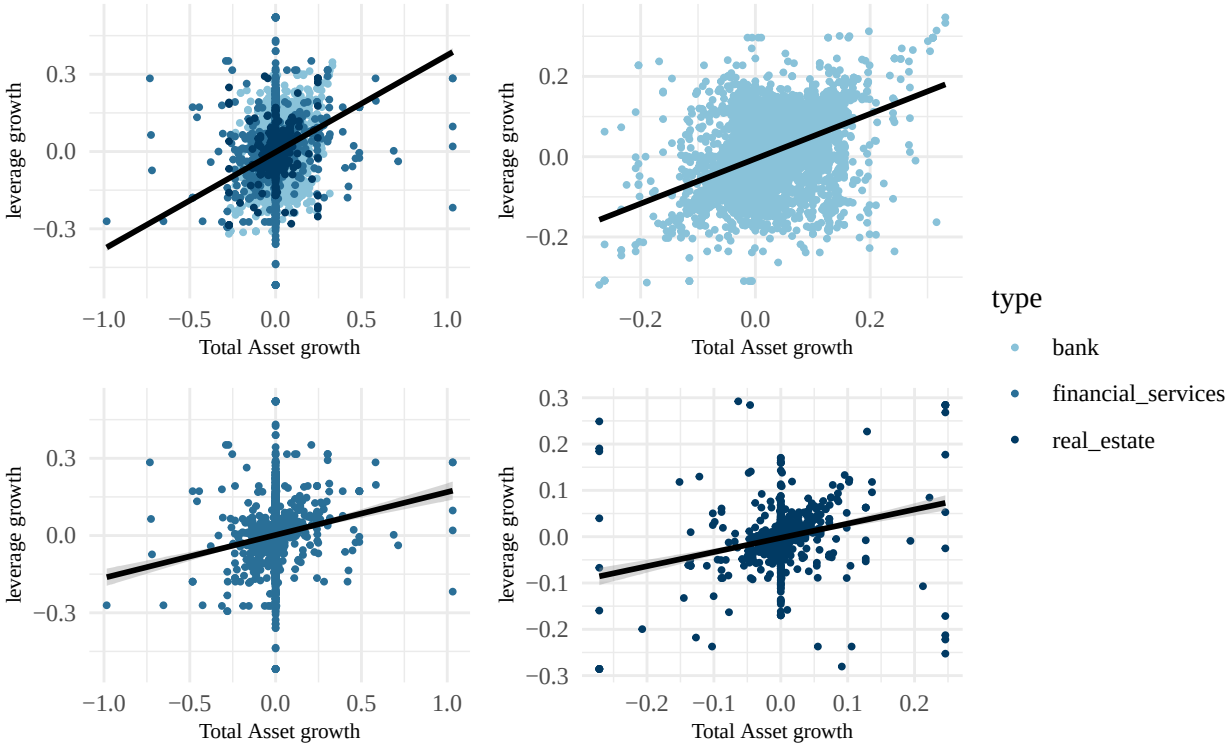
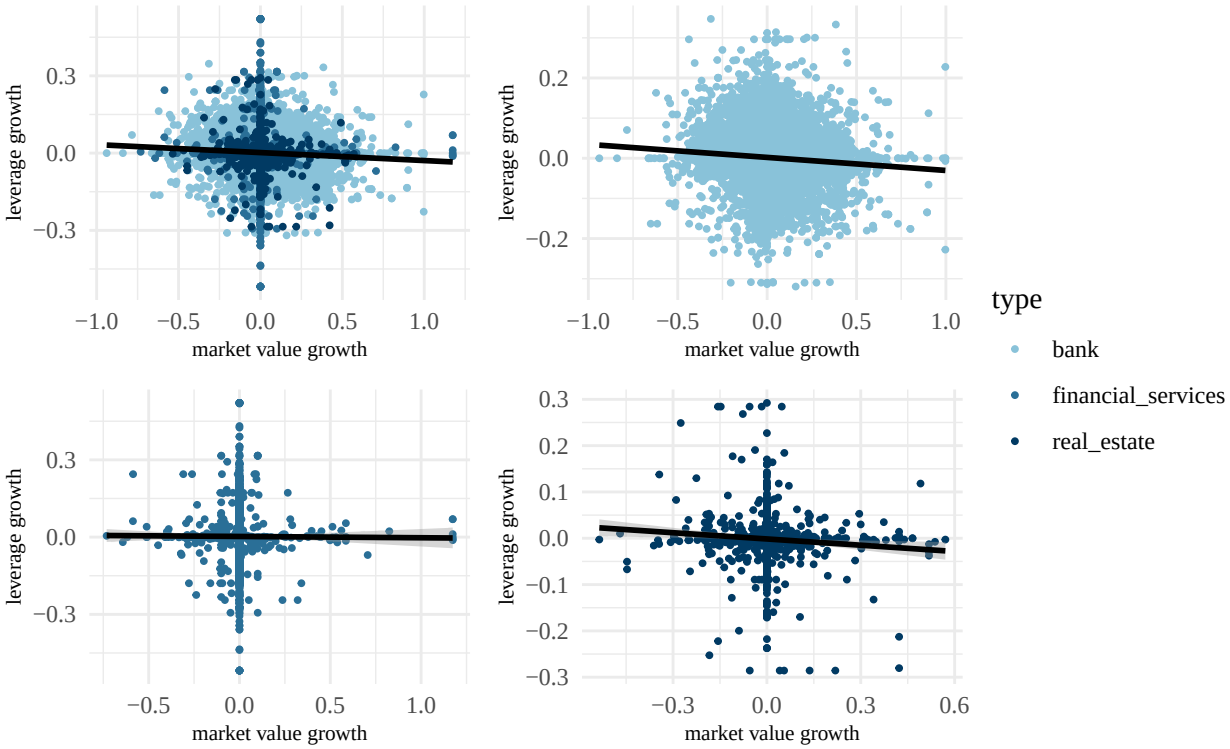


Figure 6: Scatterplots of market value growth and leverage growth



Section 2 - Two-way panel regressions on the long period

Regressions in the present sections use as the dependent variable the leverage growth rate and as the main independent variables:

- The growth rate of either total assets or market value, both considered as measures of size.
- The lag of the leverage in logarithmic form, which "captures financial institutions' reaction to the leverage level in the previous quarter" (Cincinelli et al., 2021).

As we know from the definition, a positive coefficient associated to the total assets growth rate would imply the presence of pro-cyclicality. In addition, we expect a negative coefficient for $(ln)Leverage_{i,t-1}$ as banks try to correct deviations from some target levels (Cincinelli et al., 2021). This is because if the leverage at time t-1 increases, entities aim to reduce it at time t.

The general specification we use in the present section is the following:

$$\begin{aligned}\Delta Leverage_{i,t} =& \alpha + \beta_1 \Delta Size_{i,t} + \beta_2 (ln)Leverage_{i,t-1} + \beta_3 MarketToBook_{i,t-1} \\ &+ \sum_{i=1}^{347} FinancialInstitutions_i + \sum_{t=2005:4}^{2023:2} Time_t + \epsilon_{i,t}\end{aligned}$$

We included entity-fixed effects to control for individual heterogeneity and time dummies to account for quarterly effects.

Results are reported in Table 4 (standard errors are adjusted for both heteroskedasticity and serial correlation).

Model (1) shows that the coefficient for the growth of the total assets is 0.37 and significant at 1%, suggesting leverage pro-cyclicality on the entire time period. More specifically, if the total assets growth rate increases by 1%, the growth rate of leverage increases by 0.37%[3]. The coefficient of the lag of the logarithm of leverage is statistically significant as well and has a negative sign as expected.

Model (2) adds the lag of the Market-to-Book Value Ratio to the previous regression, but the inclusion of this variable doesn't affect the estimates.

Model (3) and Model (4) do the same as Models (1) and (2), but using the growth rate of market value instead of the growth rate of total assets as a measure of size. The coefficient of size shows the opposite sign with respect to what observed in Models (1) and (2). The rationale behind this is that, keeping other things equal, as the market value of the entity increases, the leverage decreases by construction. Moreover, as in the previous regressions, the inclusion of the lag of the Market-to-Book Value Ratio doesn't affect the estimates. Overall, we find evidence of leverage pro-cyclicality.

Concerning the goodness-of-fit, in Models (1) and (2), Adjusted R^2 are around 0.11, while they are extremely lower for Models (3) and (4), where they are around 0.02. This means that despite the model fit being poor in both cases, using total assets as a measure of size better explains the growth rate of the leverage

Section 3 - Two-way panel regressions for subperiods and type of financial entity

In this section, we propose three different specifications. All specifications are written in complete form, even though some terms could be dropped afterwards due to collinearity issues.

With the first specification, we aim to study how leverage pro-cyclicality differs among the three financial entity types:

$$\begin{aligned}\Delta Leverage_{i,t} =& \alpha + \beta_1 \Delta TotalAssets_{i,t} + \beta_2 (ln)Leverage_{i,t-1} + \beta_3 MarketToBook_{i,t-1} \\ &+ \beta_4 \Delta TotalAssets_{i,t} * CBs_i + \beta_4 \Delta TotalAssets_{i,t} * FSs_i + \beta_5 \Delta TotalAssets_{i,t} * REs_i \\ &+ \sum_{j=1}^{3} FinancialEntity_j + \sum_{i=1}^{347} FinancialInstitutions_i + \sum_{t=2005:4}^{2023:2} Time_t + \epsilon_{i,t}\end{aligned}$$

[3]Which is larger than what found in Cincinelli et al. (2021) for the Chinese financial sector.

Table 4: Regression results for pro-cyclicality leverage

	Dependent variable:			
	$\Delta Leverage_{i,t}$			
	(1)	(2)	(3)	(4)
$\Delta TotalAssets_{i,t}$	0.371^{***}	0.370^{***}		
	(0.076)	(0.076)		
$\Delta MarketValue_{i,t}$			-0.025^{***}	-0.024^{***}
			(0.004)	(0.004)
$(ln)Leverage_{i,t-1}$	-0.036^{***}	-0.036^{***}	-0.042^{***}	-0.043^{***}
	(0.005)	(0.005)	(0.005)	(0.005)
$MarketToBook_{i,t-1}$		0.002^{**}		0.002^{***}
		(0.001)		(0.001)
Entity FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES
Observations	24,290	24,290	24,290	24,290
R^2	0.123	0.123	0.036	0.037
Adjusted R^2	0.108	0.108	0.019	0.020
Note:				*p<0.1; **p<0.05; ***p<0.01

Where CBs is a dummy equal to 1 if the entity type is "commercial bank" and 0 otherwise, FSs is a dummy equal to 1 if the entity type is "financial services" and 0 otherwise, and REs is a dummy equal to 1 if the entity type is "real estate" and 0 otherwise. $\sum_{j=1}^{3} FinancialEntity_j$ is the sum of those three dummies. Notice that, in the estimation, $\Delta TotalAssets_{i,t} * CBs$ will be dropped due to collinearity, thus, when interpreting the results, one should consider "commercial bank" as the baseline. In this specification, we expect the coefficient of total assets to be positive (since it indicates the pro-cyclicality of banks) and the coefficients associated with the two interaction terms to be negative but smaller in absolute value with respect to the one of total assets. This is because, as stated by Cincinelli et al. (2021) and Beccalli et al. (2015), pro-cyclicality in leverage characterizes financial institutions that are involved consistently in banking activity. This is a result we empirically obtained from the scatter plots in Figure 5 as well.

In the second specification, we focus on specific subperiods to study how pro-cyclicality varies in different contexts:

$$\begin{aligned}\Delta Leverage_{i,t} =& \alpha + \beta_1 \Delta TotalAssets_{i,t} + \beta_2 (ln) Leverage_{i,t-1} + \beta_3 MarketToBook_{i,t-1} + \\ &+ \beta_4 \Delta TotalAssets_{i,t} * GFC_t + \beta_5 \Delta TotalAssets_{i,t} * MPE_t + \beta_6 \Delta TotalAssets_{i,t} * COV_t \\ &+ \sum_{h=1}^{3} CrisisDummy_h + \sum_{i=1}^{347} FinancialInstitutions_i + \sum_{t=2005:4}^{2023:2} Time_t + \epsilon_{i,t}\end{aligned}$$

Where GFC is a dummy variable equal to 1 for the period of the global financial crisis[4] and 0 otherwise, MPE is a dummy variable equal to 1 for the period of the monetary policy expansion[5] (in which the FED reduced interest rates at zero) and 0 otherwise and COV is a dummy variable equal to 1 for the pandemic period[6] and 0 otherwise. $\sum_{h=1}^{3} CrisisDummy_h$ is the sum of those three dummies. In this case, we pursue our objective by comparing the coefficients of the interaction terms.

In the last specification, we disentangle the marginal effects of pro-cyclicality for both types of financial

[4]The Financial Crisis period includes observations from January 2007 to June 2009

[5]The Recovery measures period goes from July 2009 to December 2015 and includes a series of expansionary policies that have been implemented to overcome the crisis. Among these, the Dodd-Frank Act and BASEL III described in *Section 1* and the Zero-Interest-Rate-Policy (ZIRP) and the Open Market Operations (OMO) implemented by the FED

[6]The COVID-19 pandemic, from January 2020 to March 2022.

entities and subperiods:

$$\begin{aligned}\Delta Leverage_{i,t} =& \alpha + \beta_1 \Delta TotalAssets_{i,t} + \beta_2 (ln) Leverage_{i,t-1} + \beta_3 MarketToBook_{i,t-1} \\ &+ \beta_4 \Delta TotalAssets_{i,t} * FinancialEntity_j * CrisisDummy_h \\ &+ \sum_{h=1}^{3} CrisisDummy_h + \sum_{i=1}^{347} FinancialInstitutions_i + \sum_{t=2005:4}^{2023:2} Time_t + \epsilon_{i,t}\end{aligned}$$

Where we interact total assets with all the dummy variables we described before, obtaning a total of 9 interaction terms. With this specification, we can compare differences in pro-cyclicality between entity types for each specific time period and between time periods for a specific entity type. All specifications include two-way fixed effects.

Table 5 presents the results for the present section (standard errors are adjusted for both heteroskedasticity and serial correlation).

Model (1) highlights that all three entity types are leverage pro-cyclical. An increase of 1% in total assets growth rate reflects an increase in leverage of 0.56% for banks, an increase of 0.16%[7] for financial services and an increase of 0.28%[8] for real estate entities. All coefficients are statistically significant at the 1% level. Empirical evidence confirms that banks have larger leverage pro-cyclicality.

In Model (2), the coefficient of total assets is smaller with respect to Model (1) because here we are focusing on subperiods and not entities (however it is still statistically significant at 1%). Looking at the interaction terms, we don't find any significant difference in pro-cyclicality in the three subperiods we considered.

Model (3) gives some interesting insights. Looking at entity types, we see that banks have always been more pro-cyclical than the baseline period (especially during the global financial crisis), with all the coefficients that are positive and statistically significant at the 1% level. Instead, financial services have been less pro-cyclical during the global financial crisis and the period of monetary policy expansion than they were during the baseline period (coefficients are again statistically significant at 1%), but there seems not to be any difference during the pandemic crisis. Finally, real estate entities register a weakly different behaviour in pro-cyclicality only in the monetary policy expansion period. In addition, we see that banks have the most pronounced pro-cyclical behaviour among all types of financial entities. This is a result we already found previously in Model (1), and it is confirmed here for each subperiod as well.

For what concerns the models' fit, adjusted R^2s are very low for all the regressions but higher than the ones of the previous sections' regressions.

Conclusions and policy implications

In this work we aimed to study the presence of leverage pro-cyclicality in the US financial sector. Firstly, we found empirical evidence of pro-cyclicality behaviour for the entire period. When we moved on to analysing entity types, results showed that banks act more pro-cyclically than financial services and real estate entities, confirming what was suggested by the theory. We then focused our attention on different subperiods, i.e. the global financial crisis, the monetary policy expansion period and the recent pandemic. Here we did not find evidence of a variation in pro-cyclicality behaviours in the financial sector. However, when we disentangled how each entity type managed its balance sheet in these subperiods, we found again strong pro-cyclicality for banks, especially in the global financial crisis period. From these results some policy suggestions can be drawn. In particular, it is evident that the presence of pro-cyclicality constitutes a risk for the stability of the economic system, especially during downturns. This requires a regulatory framework with a special attention on the banking activity, since as we have seen those entities show the most prominent pro-cyclical behaviour in the US financial sector. A proper regulation should function as an automatic stabilizer in the business cycle, namely preventing excessive debt accumulation during expansions and avoiding credit crunches in

[7]Where 0.16 is obtained from 0.56, that is the coefficient for the growth rate of total assets, minus 0.40, that is the coefficient of the interaction term between the growth rate of total assets and the dummy for REs.

[8]Where 0.28 is obtained from 0.56, that is the coefficient for the growth rate of total assets, minus 0.28, that is the coefficient of the interaction term between the growth rate of total assets and the dummy for FSs.

Table 5: Regression results for pro-cyclicality leverage (marginal effects for financial entities and subperiods)

	Dependent variable:		
	$\Delta Leverage_{i,t}$		
	(1)	(2)	(3)
$\Delta TotalAssets_{i,t}$	0.560^{***}	0.368^{***}	0.372^{***}
	(0.034)	(0.048)	(0.047)
$(ln)Leverage_{i,t-1}$	-0.035^{***}	-0.036^{***}	-0.036^{***}
	(0.005)	(0.005)	(0.005)
$MarketToBook_{i,t-1}$	0.002^{***}	0.002^{**}	0.002^{***}
	(0.001)	(0.001)	(0.001)
$\Delta TotalAssets * FSs_{i,t}$	-0.400^{***}		
	(0.066)		
$\Delta TotalAssets * REs_{i,t}$	-0.278^{***}		
	(0.056)		
$\Delta TotalAssets * GFC_{i,t}$		0.041	
		(0.068)	
$\Delta TotalAssets * MPE_{i,t}$		-0.070	
		(0.099)	
$\Delta TotalAssets * COV_{i,t}$		0.142	
		(0.098)	
$\Delta TotalAssets * CBs * GFC_{i,t}$			0.282^{***}
			(0.074)
$\Delta TotalAssets * CBs * MPE_{i,t}$			0.165^{***}
			(0.058)
$\Delta TotalAssets * CBs * COV_{i,t}$			0.275^{***}
			(0.076)
$\Delta TotalAssets * FSs * GFC_{i,t}$			-0.124^{***}
			(0.047)
$\Delta TotalAssets * FSs * MPE_{i,t}$			-0.262^{***}
			(0.097)
$\Delta TotalAssets * FSs * COV_{i,t}$			-0.178
			(0.199)
$\Delta TotalAssets * REs * GFC_{i,t}$			-0.419^{*}
			(0.225)
$\Delta TotalAssets * REs * MPE_{i,t}$			-0.431
			(0.308)
$\Delta TotalAssets * REs * COV_{i,t}$			0.410^{**}
			(0.178)
GFC		0.006	0.005
		(0.005)	(0.005)
MPE		0.005	0.002
		(0.004)	(0.004)
COV		0.050^{***}	0.050^{***}
		(0.005)	(0.005)
Entity FE	YES	YES	YES
Time FE	YES	YES	YES
Observations	24,290	24,290	24,290
R^2	0.148	0.177	0.198
Adjusted R^2	0.133	0.162	0.184
Note:			*p<0.1; **p<0.05; ***p<0.01

contraction periods, thus imposing some capital constraints on financial firms. Some steps have been taken in this direction by the Basel III Committee on Banking Supervision, albeit as some authors argue (see Beccalli et al. (2015) and Baglioni et al. (2013)) the degree of indebtness allowed is still too high.

Bibliography

111th US Congress. (2010). Dodd-Frank Wall Street Reform and Consumer Protection Act (Public Law 111–203).

Basel Committee on Banking Supervision, & Bank for International Settlements. (2011). Basel III: A global regulatory framework for more resilient banks and banking systems. Bank for International Settlements.

Baglioni, A., Beccalli, E., Boitani, A. et al. (2013). Is the leverage of European banks procyclical?. Empir Econ 45, 1251–1266. https://doi.org/10.1007/s00181-012-0655-4

Beccalli, E., Boitani, A., & Di Giuliantonio, S. (2015). Leverage pro-cyclicality and securitization in US banking. Journal of Financial Intermediation, 24(2), 200–230. https://doi.org/10.1016/j.jfi.2014.04.005

Cincinelli, P., Pellini, E., & Urga, G. (2021). Leverage and systemic risk pro-cyclicality in the Chinese financial system. International Review of Financial Analysis, 78. https://doi.org/10.1016/j.irfa.2021.101895