Macrofinancial Feedback, Bank Stress Testing and Capital Surcharges

Tobias Adrian[†], Jose Berrospide[§], and Romain Lafarguette[†]

[†]International Monetary Fund
[§]Federal Reserve Board

This version: October 11, 2021

Abstract

We develop a framework to assess macrofinancial vulnerabilities across the business and financial cycles, and to calibrate a countercyclical capital buffer (CCyB) to account for those vulnerabilities in the context of bank stress tests. A parsimonious model captures the dynamic relationship between bank capital, financial conditions, and GDP growth, quantifies the causal impact of shocks to bank capital on the future distribution of financial conditions and future downside risks to GDP growth, and estimates macrofinancial feedback effects. Our model calibrates a bank capital surcharge, the additional bank capital that accounts for the macrofinancial feedback—typically overlooked in bank stress tests—as a macroprudential tool to avert banks' amplification of shocks to the economy. Using a Growth-at-Risk based metric as a measure of financial stability risks, we calibrate the size and the timing of the CCyB, as the capital needed to offset the macrofinancial feedback across the business cycle.

Keywords: Stress-testing, macrofinancial feedback, macroprudential policy, quantile regressions, density forecasting, granular instrumental variables, causal inference

JEL classification: G20, G28

Contacts: tadrian@imf.org, jose.m.berrospide@frb.gov and rlafarguette@imf.org. We thank participants at the 2020 Stress Testing Research Conference, and the Financial Stability Workshop at the Federal Reserve Board for valuable comments. We also thank Lawrence Schmidt, Yinchu Zhu and Brice Green for sharing with us their quantile spacing codes. The views expressed in this paper are those of the authors and do not necessarily represent those of the International Monetary Fund, its Management, its Executive Directors, or the Board of Governors of the Federal Reserve, the Federal Reserve System or its staff.

1 Introduction

Supervisory stress tests have played a central role since the 2008 crisis and have contributed to improve the resilience and safety of banking sectors worldwide. Supervisors use stress tests to assess whether banks operate with sufficient capital to withstand the potential effects of severe recessions or stressful financial conditions, and thus to ensure that they continue supporting the economy by providing credit to households and businesses. For that purpose, supervisors design macroeconomic scenarios using adverse macroeconomic assumptions such as significant declines in GDP growth and increases in unemployment rates, and stressful financial conditions such as large drops in asset prices. Internal models are then used to estimate banks' losses, revenues and balance sheets, and thus to determine capital shortfalls under baseline and adverse scenarios. Banks are required to hold sufficient capital under such scenarios, and may face restrictions to their payout policies if they fail the stress tests, that is, if their post-stress capital ratios fall below minimum capital requirements.

Despite their widespread use by supervisors in many countries, current bank stress tests remain primarily microprudential in some of these countries, with an emphasis on the solvency of individual banks and their ability "to pass or fail the test" and the subsequent supervisory measures that may be needed to beef up capital cushions when a bank does not pass the test. From a macroprudential perspective, however, the supervisory stress tests emphasize the adequacy of aggregate bank capitalization and the financial health of the entire financial system, and in particular, on the ability of the banking sector to remain robust to systemic risks and thus to support economic growth.

In achieving its macroprudential goal, supervisory stress tests focus on feedback effects that could exacerbate potential vulnerabilities, trigger systemic risk, and thus compromise financial stability. An important goal of macroprudential stress testing is the correct calibration of bank capital, accounting for both the direct risks to bank capital from adverse macroeconomic and financial conditions, and the indirect risks emanating from feedback effects resulting from the interactions of banks and the real economy. Such indirect risks give rise to important amplification effects that the current approach may overlook. Yet to date, no mainstream approach to macroprudential stress testing has been developed.[1] This paper fills this gap and proposes a tractable econometric model that can provide a basis for modeling macrofinancial feedback effects in supervisory stress tests, and that can be used to analyze the impact of increased bank capital on macroeconomic tail risks.

Beyond our proposed standard, top-down stress-testing model that calibrates macrofinancial feedback effects, our paper contributions include three additional innovations. First, we consider a parsimonious, reduced-form model with contemporaneous and lagged interactions that captures

[1]The lack of a widely adopted approach to macroprudential stress testing can be traced back to the difficulty of modeling macrofinancial feedback effects in a parsimonious manner. As we show, macrofinancial feedback effects can be easily modeled in a top-down framework, which can then become a powerful tool to provide supervisors with valuable information on both the direct and indirect effects of stressful conditions.

the dynamic relationship between bank capital, an index of financial conditions, and GDP growth, and is able to replicate remarkably well the projected path of bank capital in the most severe scenario of the U.S. supervisory stress tests. Second, bank capital in our model is a flow variable measured by pre-tax net income (PTNI) as percent of risk-weighted assets (RWA). Multiple drivers behind the changes in bank capital ratios, such as bank profits, capital distribution assumptions and the banks' own deleveraging decisions, can have different impacts on financial conditions and future GDP growth. By measuring the change in bank capital through PTNI as percent of RWA, we account for capital additions resulting from bank profits, and deleveraging decisions occurring through RWA. Our measure also excludes banks' capital distributions (e.g., dividend payments and share repurchases) and thus removes a source of potential endogeneity as these decisions may be conditioned on the macro and financial environment. Further, this approach overcomes non-stationarity problems associated with the upward trend in capital ratios, after the implementation of the Basel III standards, and thus improves upon former approaches using changes in bank capital, or deviations of the capital ratio relative to its long-run trend. Third, we propose a methodology to use the stress-testing framework to calibrate the countercyclical capital buffer (CCyB), and account for vulnerabilities through the business and financial cycles.

Our parsimonious model accounts for endogenous and nonlinear relationships, and thus quantifies the causal impact of shocks to bank capital on the future distribution of financial conditions and future downside risks to GDP growth. More importantly, our model calibrates a bank capital surcharge, the additional bank capital that accounts for the macrofinancial feedback between the financial sector and the real economy, as a macroprudential tool to avert banks' amplification of real and financial shocks to the economy. More specifically, the macroprudential feedback is constructed as the difference in the projected path of GDP growth, in the unrestricted model and a restricted model that shuts down responses of both GDP growth and the index of financial conditions to the changes in equity capital. A positive macrofinancial feedback indicates that the projected path of GDP growth exhibits a slower recovery when all dynamic interactions are accounted for than the GDP growth path in the model that overlooks the feedback effect, which in turn, suggests that banks need to operate with higher capital to avoid such costly slower recovery. Given its countercyclical nature (e.g., the macrofinancial feedback increases when financial vulnerabilities rise), the capital surcharge can guide policy makers in setting the CCyB.

We calibrate the macrofinancial feedback for the U.S. using information from the Federal Reserve's Comprehensive Capital Analysis and Review (CCAR) exercise. Based on macroeconomic and financial variables in the 2020 CCAR scenarios, we construct a parsimonious Financial Conditions Index (FCI). Our FCI is significantly correlated with the Chicago Fed's National Financial Conditions Index (NFCI), previously used in the literature, but unlike the NFCI, we can use it to forecast financial conditions under the severely adverse scenario in CCAR. Finally, we evaluate the impact of macrofinancial shocks to economic activity and financial markets comparable in severity with the extreme macroeconomic conditions featured by the severely adverse scenario in CCAR.

To account for endogeneity concerns, we instrument bank PTNI, bank capital and FCI. We use the granular instrumental variables (GIV) of Gabaix & Koijen (2020) for bank PTNI and bank capital, from the cross section of banks in CCAR. For the FCI, we use two instruments. One is a GIV from the cross section of the Capital Asset Pricing Model (CAPM) cost of capital for banks in CCAR, the other are monetary policy shocks from Cieslak & Schrimpf (2019). We show that all instruments are valid and reject that they are weak. This instrumental variable strategy allows us to make causal statements, i.e. go beyond Granger causality, and apply our findings directly to policy questions.

Using our parsimonious model and the assumptions in the severely adverse scenario of the stress tests, we are able to produce projected paths of bank capital similar to aggregate capital projections in the CCAR exercise. Our median path for the Tier1 Capital/RWA ratio implies a start-to-minimum decline of 3.2 percentage points, comparable to the 2.7 percentage point average decline in CCAR between 2013 and 2020. The projected path of the capital ratio for the 5th percentile implies a larger decline (up to 8 percentage points), a very extreme tail risk considering the macroeconomic shocks implicit in CCAR.

We find that the macroeconomic shocks implied by the 2020 CCAR severely adverse scenario produce a macrofinancial feedback of about 2 percentage points for the median of the distribution of future GDP growth in 2019. The feedback effect is larger, 7.5 percentage points, for the 5th percentile on aggregate GDP growth, suggesting a significantly large impact one year after the initial shock. In terms of financial conditions, the feedback effect is about 2.1 and 3.2 standard deviations for the median and the 5th percentile of the FCI distribution, repectively.

We link the macrofinancial feedback to vulnerabilities through the business cycle using the Growth at Risk (GaR) framework. GaR is defined as the conditional growþof GDP at the lower 5th percentile, and thus estimates downside risk to GDP growth. Using our model, we calculate both a GaR estimate conditional on the state of the economy (conditional GaR) and the unconditional GaR, based on historic averages of our variables. We then propose the GaR gap, the difference between the conditional GaR and unconditional GaR as our metric of countercyclical vulnerabilities. The difference between conditional and unconditional GaR captures the difference between cyclical and structural vulnerabilities, and thus the calculation of the GaR gap summarizes different vulnerabilities into one consistent metric, which we view as superior to alternative measures of financial vulnerabilities such as the credit-to-GDP gap.

Conceptually, countercyclical vulnerabilities include the macrofinancial feedback. Therefore, we decompose the GaR gap, at each period, into a direct effect of the shock on future GDP and financial conditions, and the macrofinancial feedback effect of bank capital. This way, we obtain the capital surcharge needed to offset the macrofinancial feedback across the business cycle and thus a measure of the CCyB. As a measure of countercyclical vulnerabilities, the GaR gap helps with the calibration of the size and the timing of the CCyB.

Using the CCAR scenario, the GaR gap informs the setting of the CCyB to offset the macrofinancial feedback through the business cycle. Our estimates indicate that the CCyB for the median in the pre-crisis should have been on average at 190 basis points (near the upper bound of the Basel III CCyB), and 460 basis points for the 5th percentile. The CCyB for the median in the post-crisis should be around 210 basis points on average, and about 510 basis points for the 5th percentile.

Our results have important policy implications for the design of stress tests as we provide a methodology to assess banks' resilience to the macrofinancial feedback. Our paper contributes to assess the extent to which bank capital regulation can be used as a macroprudential tool to help ease financial stability concerns. In particular, we illustrate how dynamic capital requirements (time-varying capital surcharges) can be used during financial stress to reduce a model-based measure of downside risk. Moreover, our results offer a menu of policy options by quantifying the magnitude of the capital surcharge at different points in the distribution of our variables (e.g., how much capital ratios would need to be raised through a capital surcharge) given the assessment of financial conditions and to reduce downside risks to economic growth. In practice, we provide a quantification of the CCyB from macrofinancial dynamics.

Our calibration of the CCyB builds on earlier literature. Drehmann et al. (2011) uses the credit-to-GDP gap (credit gap), the gap between the ratio of credit to GDP and its long-term backward-looking trend to calibrate the CCyB, arguing that it captures the build-up of systemic vulnerabilities that can lead to banking crises. The approach employs forecasting regressions for banking crises, but does not explicitly separate out the macrofinancial feedback, which is the goal of our paper. Furthermore, Wezel (2019) argues that the credit gap does not work well in assessing bank loan losses, and Edge & Meisenzahl (2011) show that the credit gap is an unreliable indicator in real time. In contrast, our estimator performs well in real time. In assessing the effects of countercyclical capital buffers, Jimenez et al. (2017) use Spain's experience with dynamic provisioning and show that capital buffers smooth credit supply cycles. Auer & Ongena (2019) find that the additional equity capital required as part of the introduction of the CCyB in Switzerland led to higher growth in commercial lending. Our model provides an alternative approach to metrics such as the credit gap or dynamic provisions to inform the CCyB calibration. In addition, Benes & Kumhof (2015) study a dynamic stochastic general equilibrium (DSGE) model to quantify a macrofinancial feedback and calibrate a CCyB, but the approach is very stylized, model specific, and difficult to use within the context of supervisory stress tests. In contrast, our methodology is straightforward to add onto existing supervisory stress tests and our results are not model specific.

Our paper also builds on previous work studying the effects of financial conditions on downside risks to GDP growth using quantile regressions (Adrian et al. (2019), Prasad et al. (2019)). Our metric of countercyclical vulnerabilities (GaR gap) is an extension of the GaR measure developed by Adrian et al. (2018). They find differences in the GaR term structure depending on whether loose financial conditions would be amplified by rapid credit growth, consistent with previous work showing that high credit growth helps predict the severity and duration of recessions (Jorda et al. (2013)). Boyarchenko et al. (2020) show that the role of bank capital in predicting future GDP growth goes over and above the predictive information in credit growth. Like these papers, we also allow for nonlinear effects of financial conditions and bank capital in predicting the GDP growth distribution but rather than looking at high credit growth, we focus on insufficient bank capital as the amplifying factor of a negative shock, in the context of stress tests.

Our paper is related to recent work that studies downsize risks to the economy stemming from financial stress and vulnerabilities. Chavleishvili et al. (2020) use a quantile vector autoregressive (QVAR) model relating GDP growth, financial stress and economic vulnerabilities, and show dynamic relationships among these variables across quantiles to conclude that policymakers should consider not only downside risks to the economy but also upside potential. Our paper also uses a reduced-form, dynamic model that relates downside risk with financial conditions and vulnerabilities but focuses on quantifying the causal impact of bank capital shocks on financial conditions and downside risk, and calibrating the extra capital needed to offset the macrofinancial feedback across the business cycle. Although we use quantile regressions, we do not employ a QVAR methodology and rely on sampling methods instead, which allows us to use a new type of probabilistic scenario to stress our model.

Our paper is also related to the emerging empirical literature on the lending implications of stress tests. Recent work shows that stress tests operate as higher capital requirements that reduce jumbo-mortgage loans (Calem et al. (2020)), consumer loans (Agarwal et al. (2020)), commercial and industrial loans (Acharya et al. (2018), Berrospide & Edge (2019)) and small business loans (Cortes et al. (2020)). Our model does not explicitly account for the impact of stress tests on bank lending, but consistent with this strand of the literature, it implicitly assumes a bank-lending channel through which changes in bank capital resulting from stress tests affect economic activity.

Finally, our paper contributes to the literature that evaluates the effectiveness of stress tests, for example by providing valuable information on banks' financial conditions to market participants (Flannery et al. (2017), and Fernandez et al. (2020)). Our goal is not to evaluate the CCAR stress testing approach, but, rather, to expand the existing machinery in a way that allows the quantitative assessment of macroprudential tools such as the Basel III CCyB, from a forward-looking stress-testing exercise.[2] The CCyB is a time-varying macroprudential tool that bank regulators can

[2]In the U.K., stress tests are used to inform the decision on the CCyB. See, Kohn (2019)

use to mitigate procyclicality in the financial sector, that is, to increase the resilience of banks by raising capital requirements (between 0 and 2.5 percent), when the risk of above-normal losses is elevated.[3] The CCyB is intended to create a capital buffer that helps banks absorb shocks associated with declining credit conditions. It can be raised gradually as financial vulnerabilities increase, and be reduced or removed when those vulnerabilities abate, to prevent a severe cutback in bank lending.

The rest of the paper is organized as follows: Section 2 provides background on the U.S. supervisory stress tests. Section 4 presents the empirical approach model. Section 5 discusses the results of the estimation. Section 6 provides a policy discussion. Finally, section 7 concludes.

2 A Primer on the U.S. Stress Tests

In the U.S., the Federal Reserve conducts annual stress tests as part of its Comprehensive Capital Analysis and Review (CCAR) exercise, established in 2011. The exercise includes both a qualitative assessment of banks' capital plans, including their risk management and governance practices, and a quantitative evaluation of bank capital positions under severe macroeconomic and financial conditions. In the quantitative exercise, known as the Dodd Frank Act Stress Tests (DFAST), the Federal Reserve uses its own independent empirical models to project bank revenues, losses, and capital, over a nine-quarter planning horizon, and under three hypothetical macro scenarios.[4] The severely adverse scenario features a deep recession, characterized by a substantial reduction in economic activity and a sharp increase in unemployment, large declines in asset prices, and increases in risk premia. More specifically, the 2020 DFAST severely adverse scenario features a severe recession (e.g., 8.5 percent real GDP contraction from its pre-recession peak and the unemployment rate reaching a peak of 10 percent), and the collapse of broad financial markets (e.g., 50 percent drop in stock prices, large declines in house prices and CRE prices of 28 and 35 percent, respectively, and a widening of corporate spreads of about 400 basis points).

One important difference between DFAST and CCAR relates to assumptions about capital distributions in the calculation of post-stress capital ratios. In DFAST, bank capital projections assume constant dividends equal to their average over the previous year (equity issuance and share repurchases are assumed to be zero). In CCAR, capital projections use the BHCs' reported

[3]In the U.S., the CCyB applies to banks with more than $250 billion in total assets or more than $75 billion in either short-term wholesale funding or international exposures, and to any depository institution subsidiary of those banks. Regulators are expected to activate the CCyB when systemic vulnerabilities are meaningfully above normal. For details, see the Federal Reserve Board's framework for setting the CCyB, published in September 2016, https://www.federalreserve.gov/newsevents/pressreleases/bcreg20160908b.htm.

[4]The Federal Reserve started DFAST in 2013 on the 19 largest banks that were subject to the 2009 Supervisory Capital Assessment Program (SCAP). In 2017, 35 banks with total assets of $50 billion or more were subject to DFAST, though the number of participating banks have changed as intermediate holding companies (IHCs) of foreign banking organizations that operate in the U.S. have been incorporated. The Economic Growth, Regulatory Relief, and Consumer Protection Act (EGRRCPA) of 2018 exempted institutions with less than $100 billion in assets from DFAST. Starting in 2019, banks with assets between $100 billion and $250 billion are subject to supervisory stress tests every other year.

planned capital distributions. Another difference, at least until 2019, referred to supervisory actions following the disclosure of results. No supervisory actions are associated with DFAST other than the requirement that the banks consider the results in their capital planning. In CCAR, however, the Federal Reserve could object a bank's capital plan on quantitative or qualitative grounds, and thus may require changes in planned capital distributions, or impose restrictions on those distributions if banks' post-stress capital ratios fall below minimum requirements (e.g., if the bank fails the test). Qualitative objections occur if the Federal Reserve evaluation concludes that the bank's governance or risk management systems are inadequate.

In 2019, the Federal Reserve modified the use of the qualitative objections. Under new capital rules, banks that participated in four consecutive exercises and successfully passed the qualitative evaluation in the fourth year are no longer subject to a potential qualitative objection. Finally, under new capital rules, in March 2020, the Federal Reserve integrated stress-testing results with point-in-time capital requirements by replacing the capital conservation buffer (flat at 2.5 percent) with a stress capital buffer (SCB). The SCB is calculated as the sum of the capital decline in the severely adverse scenario of the stress tests from start to minimum and four quarter of dividend payments, floored at 2.5 percent. With this final requirement, banks need to operate with capital ratios above minimum requirements that include the SCB to avoid restrictions on their capital distributions and compensation. Although DFAST and CCAR are distinct testing exercises, both efforts are complementary as they rely on similar processes, data, supervisory exercises, and requirements, and coordination exists to avoid duplicate requirements and to minimize regulatory burden.

Although IMF staff had been stress testing banking systems since the late 1990s (see Adrian et al. (2020b) for a comprehensive review of the IMF's stress tests), it wasn't until the aftermath of the 2008 financial crisis that the U.S. adopted supervisory stress tests (see Federal Reserve (2009a,b)). Hirtle & Lehnert (2015) provide an overview of the U.S. CCAR and DFAST, arguing that stress tests substantially change the nature of the supervisory process. The initial stress test in 2009 was designed to restore confidence in the U.S. banking system by recapitalizing banks to a degree sufficient to withstand even more adverse future shocks, but, subsequently, the CCAR and DFAST processes have changed the very notion of supervision. Our methodology builds directly on top down stress testing comparable the to DFAST exercise, and points towards a simple extension of the current approach to calculate the CCyB requirement in the aggregate.[5]

3 Data and Stress Scenarios

We use four aggregate macrofinancial variables: Gross Domestic Product (GDP) growth, financial conditions, bank profits, and bank capital. The motivation for such a parsimonious setup is to

[5] Since the stress test scenarios and the quantitative exercise before capital distributions, which we exclude in our analysis, are the same in DFAST and CCAR, we use DFAST and CCAR scenarios and exercises indistinctly.

illustrate the first order economic forces that are giving rise to macrofinancial feedback loops. Of course, in general, one can embed our proposed methodology in a richer setting with additional macroeconomic variables, more financial variables, and more granular supervisory bank data. However, our parsimonious analysis focuses on the basic insights, isolating the macrofinancial feedback. In this section, we describe the data and the stress scenarios.

Consistent with typical stress-testing frameworks that focus on banks' projected revenues and losses, we model changes in bank capital that originate from retained earnings, that is, from pre-tax net income (PTNI) and before capital distributions (e.g., dividend payments and stock repurchases).[6] More specifically, pre-tax net income is measured as the difference between pre-provision net revenue (PPNR) and total losses, which include loan losses and valuation losses in securities holdings. Pre-provision net revenue (PPNR) is defined as net interest income (interest income earned minus interest expense) plus non- interest income (including trading income, as well as non-trading noninterest income earned from fees and other sources), minus non-interest expense (compensation, expenses related to premises and fixed assets, and other non-credit-related expenses). Losses from different loan categories (residential mortgage, commercial real estate, commercial and industrial, auto, and credit card loans) are proxied by loan loss provisions. PTNI, PPNR, losses and risk-weighted assets (RWA) are aggregated across all banks participating in the stress tests exercise (DFAST) to obtain the time series of changes in bank capital.

Focusing on PTNI, and excluding both regulatory items and bank capital distributions, has several advantages. First, regulatory items such as minority interests, deferred tax assets, and certain intangible assets consistent with capital rules change over time with the implementation of different regulatory regimes. Second, bank decisions on capital distribution are endogenous to the macro and financial environment, both modeled through GDP growth and an index of financial conditions in our framework. Thus, by excluding regulatory items and capital distributions, our measure is more consistent and comparable over time, and helps alleviate some endogeneity concerns. Finally, using pre-tax net income before capital distribution avoids unnecessary distinctions between DFAST and CCAR exercises, as part of the U.S. supervisory stress tests.

PTNI is a flow concept that we use as the main state variable for bank capital. From PTNI, the evolution of bank capital ratios (a stock concept) can be calculated as follows:

$$Capital\ \ Ratio_{i,t} = Capital\ Ratio_{i,t-1} + PTNI_{i,t} - Taxes_{i,t} - Capital\ Distribution_{i,t} \quad (1)$$

For Capital Ratio, we use Tier 1 capital/RWA, consistent with the U.S. stress testing exercise and project it using equation (1), assuming a simple no corporate tax rate and no capital distributions.

[6]See, Overview of Modeling Framework in "Dodd-Frank Act Stress Test 2019: Supervisory Stress Test Methodology" at https://www.federalreserve.gov/publications/files/2019-march-supervisory-stress-test-methodology.pdf

Figure 1 shows the evolution of bank capital in our sample, as measured by PTNI before capital distributions, that is, PPNR minus losses, and Tier 1 capital, both as ratios to RWAs. Clearly, the Tier 1 capital ratio exhibits strong non-stationary, as it reflects the change in regulatory capital requirements in the decade after the global financial crisis with the phase in of the Basel III banking regulations. PTNI, on the other hand, is well behaved from an econometric point of view, reflecting the net profits of the banking system, as a major source of bank capital accumulation. The extent to which such profits are paid out or retained is largely a function of banking regulations and intertemporal considerations, which are endogenous bank decisions.

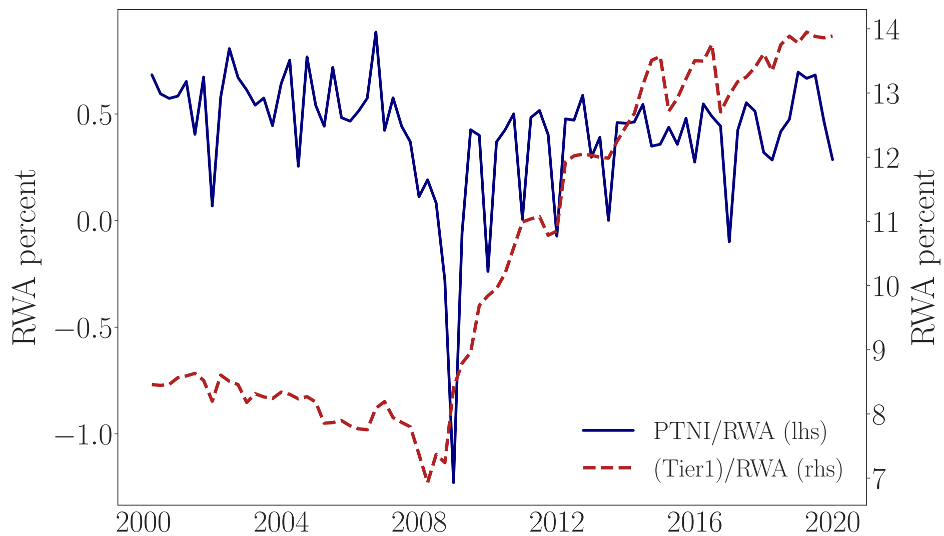

Figure 1: US banks' average PTNI/RWA and Tier1 Capital/RWA

Another notable feature of Figure 1 is the decline of average net profits after the global financial crisis (about 1.2 percent of RWA). There is much debate as to the reasons for the decline, including lower economic growth, lower interest rates, tighter regulation, and the impact of competition. PTNI is negatively skewed, with large losses in times of crises, recessions, and banking sector specific events. For macrofinancial modeling, the "flow" measure PTNI is certainly more appropriate than the Tier 1 capital ratio, though they are of course tightly linked from accounting identities.

One potential caveat in our approach is that by using aggregate data on bank capital, we do not distinguish between the impact of actual capital ratios and the impact of capital regulation on future GDP growth and financial conditions. Further, as has been already documented in the banking literature, banks tend to operate with capital levels above minimum requirements as they may fear regulatory pressures preventing them from breaching those requirements (e.g., Milne & Whalley (2001)). Therefore, changes in bank capital that lead to a level of capital that remains well above the requirements may trigger less of a procyclical impact on the economy relative to changes that bring the level of capital closer to the minimum.

We mitigate that concern, as discussed below, by including the level of capital (i.e., the capital ratio) in our model. Measuring the level of capital that accounts for the macrofinancial feedback helps us understand whether the banking sector operates with "enough" capital (given the real and financial shocks in the stress tests) to account for bank's potential amplification of shocks to the economy. By comparing the level of bank capital that accounts for the macrofinancial feedback with the current level of capital in the banking sector gives us a sense of a capital shortfall (and thus the need to raise capital requirements through the calibrated CCyB) or a capital surplus (e.g., banks operating with possibly inefficient excess capital), and their cost in terms of future GDP growth.

As a measure of real economic activity, we use the quarterly growth rate of real Gross Domestic Product from the Bureau of Economic Analysis.

Finally, we construct a financial conditions index using a dimension-reduction method based on Partial Least Square (PLS) regression on seven macro scenario variables from CCAR: the corporate credit spread (Merrill Lynch 10-year BBB corporate bond yield minus the 10-year Treasury yield), equity implied volatility "VIX" on the S&P500, house price index (year-over-year difference), the Dow-Jones equity market year-on-year return, Commercial Real Estate Price Index (year-over-year difference), mortgage spread defined as the 30-year fixed rate on conventional and conforming mortgages minus the 10-year Treasury yield, and the term spread defined as the difference between the 10-year and the 3-month Treasury yields. The PLS approach is a supervised dimension-reduction method (see Götz et al. (2010) for an introductory exposition): it uses a variable (or a set of variables) as anchor to compute the weights of the seven macro variables, so that the correlation in the projected subspace between the supervisors and the reduced data is maximal. The PLS is anchored to the Chicago Fed NFCI, which is computed over a hundred of variables (Brave et al., 2017). Figure 2, left panel, compares the Chicago Fed NFCI and the PLS FCI. The correlation between the PLS FCI from seven variables and the Chicago NFCI, is around 90 percent. The PLS FCI is, therefore, a parsimonious approach to estimate a financial conditions index from a subset of CCAR-stressed variables. Figure 2, right panel, shows the loadings of the seven underlying variables used in the PLS FCI.

Unlike the Chicago Fed NFCI, our financial conditions index (PLS FCI) is based on the CCAR exercise. Therefore, and as also shown in figure 2, we use the stress testing scenarios to construct a forecast of financial conditions between 2020 and 2022, that capture financial shocks consistent with the extreme macroeconomic conditions featured by the 2020 severely adverse scenario in CCAR.

Figure 3 shows the FCI together with the PTNI in the left panel, and with U.S. real GDP on the right panel. There is a visible negative correlation between the FCI and the PTNI. Intuitively, tight financial conditions correspond to an increase in the cost of funding of banks, and thus associated

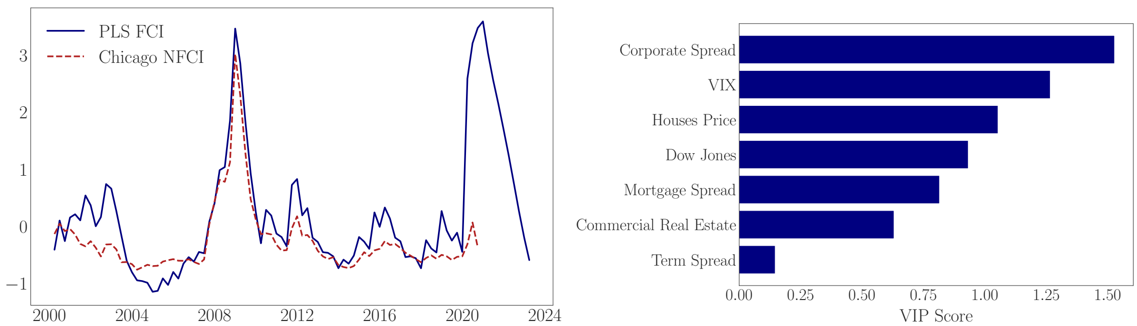

Figure 2: PLS FCI and Chicago Fed FCI (left panel) and FCI loadings (right panel)

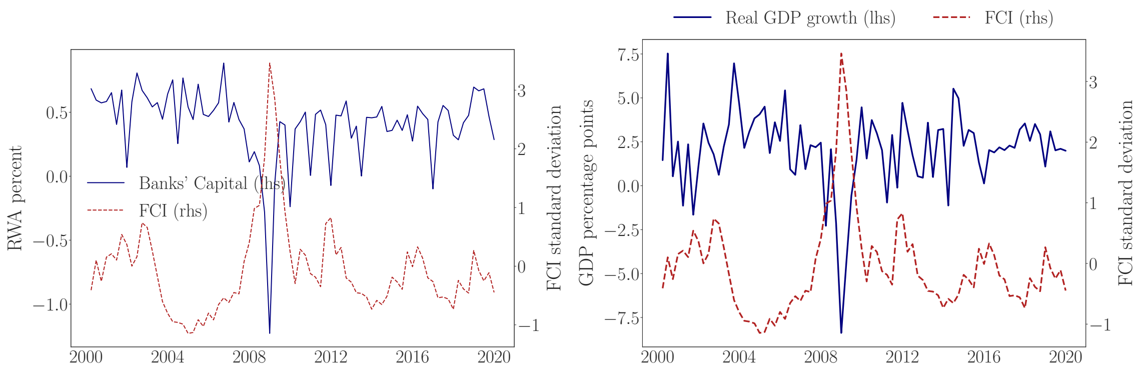

Figure 3: Financial Conditions and PTNI (left panel) and FCI with GDP (right panel)

with a drop in their profitability. Likewise, the correlation between FCI and GDP growth is negative, as loose financial conditions support lending and investment.

Table 1 presents the summary statistics of the main variables. All four endogenous variables have 80 quarters of data, from 2000Q1 to 2019Q4. On average, during our sample period, the U.S. economy exhibits positive economic growth (about 2 percent) and loose financial conditions (negative FCI of 0.05); and banks operate with about 10 percent capital ratios. There is also significant variation in the main variables: output volatility of about 2 percent, FCI that moved from about -1 in 2005 to about 350 during the global financial crisis, and the Tier1 Capital/RWA ratio that increased from its trough of 7 percent during the crisis to about 14 percent by the end of 2019.

	GDP growth	Financial conditions	Tier1/RWA pp	PTNI/RWA pp
Mean	2.05	-0.05	10.42	0.41
Standard deviation	2.29	0.79	2.38	0.29
Min	-8.38	-1.15	6.93	-1.23
25%	1.09	-0.53	8.26	0.36
50%	2.22	-0.22	9.895	0.46
75%	3.21	0.24	12.92	0.57
Max	7.53	3.46	13.95	0.88
Num obs	80	80	80	80

Table 1: Summary statistics

4 Empirical Framework

4.1 Data Generating Process and Modeling Assumptions

One important contribution of our paper is to formalize bank stress tests into a stochastic framework and to offer a practical estimation strategy using relatively simple and standard econometric tools. This section presents the formalization and explains the modeling assumptions.[7]

Notations

We are interested in modeling the endogenous dynamic between bank capital, financial conditions (FCI), and GDP growth in a stochastic framework. We represent these three variables as stochastic processes $(\mathcal{Y}, \mathcal{FCI}, \mathcal{C})$, for GDP, FCI and bank capital respectively,[8] defined over a discrete time-set $[0, \cdots, T]$. We distinguish between:

- The stochastic process $\mathcal{X}$, which is a collection of random variables defined over a discrete time-set $[0, \cdots, T]$: $\mathcal{X} = \{X(t) : t \in T\}$
- For a given period t, the random variable X_t, defined over the real support $\mathbb{R}$. Its cumulative distribution function (CDF) is written as: $P[X \leq x_0] = \int_{-\infty}^{x0} f_x(u)du$, where $f_x(u)$ is the density function of X_t
- A realization x_t of the random variable X_t for a given period t
- $S(X_t)$ is a sample drawn from the distribution of X_t. In practice, there are no observational samples in a time-series context (as we only observe one realization of the random variable at each period) and we therefore bootstrap the sample.

This means that, for instance, the observed GDP values in 2018 Q1 and 2019 Q1, y_{2018Q1} and y_{2019Q1}, are realizations of the random variables Y_{2018Q1}, Y_{2019Q1}, which can be different in terms of moments, family distribution etc. Although this level of generality increases the complexity of the model, it allows us to incorporate change in risk using an evolving stochastic process.

Conditional Joint Distribution

The joint distribution $F(\mathcal{Y}, \mathcal{FCI}, \mathcal{C}) = \{F(Y_t, FCI_t, C_t) : t \in T\}$ informs about the co-movement between GDP, financial conditions and capital at any point of the distribution, for instance, on the left tail of GDP and capital and the upper tail of financial conditions (tightening). More precisely, in the context of stress-testing, we are interested in the **conditional** joint distribution, for instance

[7] Our model is in essence a frequentist approach, but could also be solved in a Bayesian framework. Although we present a solution via standard quantile regressions, we use sampling techniques to estimate the model recursively, making it close in spirit to a Bayesian approach.

[8] With standard assumptions: defined over standard probability spaces $(\Omega_i, \mathcal{A}_i, \mathcal{P}_i)_{i \in y,f,c}$, where Ω, $\mathcal{A}, \mathcal{P}$ are the sampling space, a σ-algebra and a probability measure, respectively

based on an information set Ω_{t-1} and a series of shocks S_t on the conditioning information set: $\{F(Y_t, FCI_t, C_t \mid \Omega_{t-1}, S_t) : t \in T\}$

In practice, however, we only observe one realization of each random variable at each period. Therefore, we don't have enough degrees of freedom to estimate the joint distribution at every period: the model is by essence under-identified.[9]

From joint distribution to conditional densities using triangular ordering

We circumvent the issue of estimating joint densities by focusing instead on the conditional univariate densities, in a similar spirit as in Caselli et al. (2020). We estimate a recursive three equations system of conditional univariate densities with **contemporaneous feedback**, in a triangular ordering shape. We use a two-by two evaluation of the Granger-causality test of the three variables to determine the ordering of the system.

We present the p-values of the SSR F-test[10] of the Granger causality test in Table 2. For each cell, the table presents the p-value of the Granger test null hypothesis: the variable in index does NOT Granger-cause the variable in column. For instance, the p-value of the Granger test of FCI "Granger causing" GDP is 53.9 percent, meaning that we can **not reject** the null hypothesis that FCI does not Granger-cause GDP at a high degree of confidence. The results show that:

- We cannot reject at more than 95 percent of confidence that FCI and ΔCapital do not Granger-cause GDP: GDP is therefore the most exogeneous variable in the system
- We can reject at more than 99 percent of confidence that GDP and ΔCapital do not Granger-cause FCI: FCI is therefore the most endogeneous variable in the system
- We can reject at more than 99 percent of confidence that GDP does not Granger-cause ΔCapital. However, we cannot reject the hypothesis that FCI does not Granger-cause ΔCapital. Therefore, ΔCapital should contemporaneously depend on GDP but not on FCI.

Based on the results of the Granger test, we impose a fixed sequencing in our system, to break down the contemporaneous relationship between the variables. We assume that GDP in period $t + 1$ only depends on variables at time t; banks' variation in capital contemporaneously depends on GDP and variables at time t. Financial conditions are contemporaneously determined by GDP and banks' capital, and depends on variables at time t. Hence FCI is the most endogenous variable in

[9] One possible solution would be to parametrize the conditional joint distribution via copulas, as done for instance in Brechmann et al. (2013) with an application of Archimedean and vine copulas for systemic stress testing. Another solution would be to follow the approach used in Adrian et al. (2020a) or in Caselli et al. (2020), and estimate a large joint conditional density via a conditional kernel estimation. However, these two approaches rely on advanced mathematical tools, are complex to explain and difficult to use in a policy context

[10] Alternative tests based on the Chi-Square give similar results; for a textbook presentation of Granger causality tests, see Greene (2003). The p-values are estimated with the Python statsmodels package (Seabold & Perktold (2020))

	GDP	FCI	Delta Capital
GDP		0.000	0.000
FCI	0.539		0.324
Delta Capital	0.057	0.010	

Table 2: Two-by-two P-values of the Granger causality test of the system variables
H_0: the variable in index does not Granger-cause the variable in column

our system. These assumptions form a triangular ordering, and allow us to estimate the system of univariate conditioning densities recursively. Therefore, rather than estimating the three-dimensional joint density

$$F(\mathbf{Y_{t+1}}, \mathbf{FCI_{t+1}}, \mathbf{C_{t+1}} \mid Y_t, FCI_t, C_t) \quad \forall\ t\ \in\ [0, \cdots, T]$$

we estimate instead the following recursive three-equation system:

$$\begin{aligned} F(\mathbf{Y_{t+1}} \mid & \quad\quad\quad\quad\quad\quad Y_t, FCI_t, C_t) \\ F(\mathbf{C_{t+1}} \mid & \quad\quad\quad \mathbf{Y_{t+1}}, Y_t, FCI_t, C_t) \\ F(\mathbf{FCI_{t+1}} \mid & \ \mathbf{Y_{t+1}}, \mathbf{C_{t+1}}, Y_t, FCI_t, C_t) \end{aligned} \tag{2}$$

This approach is very convenient, as it simplifies the problem of estimating a joint-density into the estimation of three univariate conditional densities, done recursively line by line. Importantly, it avoids the problem of functional compatibility in the estimation of the joint density[11] and allows us to use relatively simple sampling methods - the quantile spacing from Schmidt & Zhu (2016) and the inverse probability integral transform - instead of bi-directional sampler such as the Gibbs sampler.

4.2 Semi-Parametric Quantile Recursive System

We now present a practical solution to estimate the recursive system of conditional densities (2) in a quantile regression framework. We formulate a parsimonious, semi-parametric model that captures the dynamic relationship between bank capital, financial conditions, and GDP growth. We particularly focus on the causal impact of shocks to bank capital on the future distribution of financial conditions and future downside risks to GDP growth.[12]

[11] Namely, that the estimate $\hat{F}(X,Y)$ should be equal to $\hat{F}(Y,X)$

[12] The estimation has been done using Python 3.8.3 (Anaconda, 2020) and R 3.6.3 (R Core Team, 2020). We used the R `quantreg` package of Koenker (2020) for the estimation of quantile regressions with inequality constraints. We used the Just-in-Time Python compiler `Numba` to optimize the estimation process via low-level functions. The full reproducible code is available on Github at https://romainlafarguette.github.io/software/

The model features a three-variable dynamic recursive system with one lag and the accounting identity relating PTNI to bank Tier 1 capital. Rather than modeling directly the level of bank's Tier 1 capital as an endogeneous variable, we instead focus on the change in capital $\Delta\mathcal{C}$ captured here by PTNI. We infer the level of bank capital $\mathcal{C}$ via a simple accounting equality. We prefer modeling the flow rather than the stock, as the stock of capital is highly persistent, and presents structural breaks in the data, making it difficult to model in practice.

As explained above, we impose a certain ordering to estimate the model. The recursive system also features a simple law of motion to update capital, where the PTNI shock depends on the PTNI stock of the previous period and the variation in capital at the current period. U.S. Real GDP growth is denoted y_t, U.S. Financial conditions are fci_t, the flow variable PTNI/RWA is Δc_t, and Tier 1 Capital/RWA is c_t. Our recursive formulation of macrofinancial dynamics is thus:

$$
\begin{aligned}
y_{t+1} &= \alpha_y^q + \beta_{y,y}^q y_t + \beta_{\Delta c,y}^q \Delta c_t + \beta_{f,y}^q fci_t + \beta_{c,y}^q c_t \ + \epsilon_{y,t+1}^q \\
\Delta c_{t+1} &= \alpha_c^q + \beta_{y1,\Delta c}^q \mathbf{y_{t+1}} + \beta_{y,\Delta c}^q y_t + \beta_{\Delta c,\Delta c}^q \Delta c_t + \beta_{f,\Delta c}^q fci_t + \beta_{c,\Delta c}^q c_t + \epsilon_{\Delta c,t+1}^q \\
fci_{t+1} &= \alpha_f^q + \beta_{y1,f}^q \mathbf{y_{t+1}} + \beta_{\Delta c1,f}^q \mathbf{\Delta c_{t+1}} + \beta_{c,f}^q \mathbf{c_{t+1}} + \underbrace{\beta_{y,f}^q y_t + \beta_{\Delta c,f}^q \Delta c_t + \beta_{f,f}^q fci_t + \beta_{c,f}^q c_t}_{\Gamma_t} + \epsilon_{f,t+1}^q \\
c_{t+1} &= c_t + \Delta c_{t+1} \quad \text{(Stochastic law of motion)}
\end{aligned}
\tag{3}
$$

We present in Appendix A our framework to estimate a conditional density from quantile regressions, by using the quantile spacing interpolation of Schmidt & Zhu (2016) and the inverse probability integral transform sampling approach. We outline here the estimation procedure, explained in detailed in the appendix:

4.3 Summary of the estimation process

Assuming a random variable Y_{t+1} of dimension $[T \times 1]$ and a set of regressors X_t of dimension $[T \times K]$, where T represents the number of observations in the time series and K the length of regressor, we generate a random sample $S(y_{t+1}|S(x_t))$ of length $[L \times 1]$ distributed as $Y_{t+1}|X_t$ via:

1. Estimate the set of quantile regressions over the historical time-series $Y_{t+1} = \alpha^\tau + \beta^\tau X_t + \epsilon^\tau$ for a set $\Theta = \{\tau_1, \cdots, \tau_n\} \in (0,1)$ of length N
2. Retrieve the stacked matrix of quantiles coefficients $\underbrace{\widehat{\beta}^\Theta}_{[N\times K]}$, where the rows corresponds to the quantile and the columns to the regressors

3. Consider a given conditioning sampling vector $S(x_t) \sim X_t$ to project the quantiles of Y_t. Note that $S(x_t)$ can be observed, in which case it will be of dimension $[1 \times K]$ or sampled of dimension $[L \times K]$. We take the most general case and assume it is of dimension $[L \times K]$

4. Use $\widehat{\beta}^{\Theta}$ to project the quantiles $\underbrace{Q(Y_{t+1}|S(x_t), \Theta)}_{[N \times L]} = \underbrace{\widehat{\beta}^{\Theta}}_{[N \times K]} \underbrace{S(x_t')}_{[K \times L]}$

5. Use the quantile spacing interpolation method of Schmidt & Zhu (2016) and the probability integral transform to sample $\underbrace{S(y_{t+1}|S(x_t))}_{[L \times 1]}$ from $\underbrace{Q(Y_{t+1}|S(x_t), \Theta)}_{[N \times L]}$

6. Asymptotically, $S(y_{t+1}|S(x_t)) \sim Y_{t+1}|X_t$ for a given conditioning vector $S(x_t) \sim X_t$.

Therefore, at the end of the process, from a conditioning vector $\{y_t, fci_t, \Delta c_t, c_t\}$ we have bootsrapped the multivariate sample of $S(y_{t+1}, fci_{t+1}, \Delta c_{t+1}, c_{t+1})$ which follows the distribution of each random variable in $t+1$.

4.4 Instruments for Causal Identification

Sources of Endogeneity

While the recursive specification provides a way to determine Granger causality (i.e. the identification of "temporal causality"), instruments are needed to estimate causal effects. More specifically, simultaneous endogeneity bias between financial conditions and regulatory capital occurs in two equations of the recursive system. Because financial conditions are forward-looking, they anticipate potential capital shortfalls of banks, which could be reflected in tighter financial conditions, and which could bias the true impact of financial conditions to banks' capital. Likewise, in the third equation, Tier1/RWA capital and PTNI/RWA are contemporaneously correlated with financial conditions, and thus simultaneity bias could occur. Financial conditions impact banks' funding costs, and therefore, their ability to raise capital.

- $\Delta c_{t+1} = \beta^q_{y1} y_{t+1} + \beta^q_y y_t + \beta^q_c \Delta c_t + \beta^q_f fci_t \;+\; \epsilon^q_{\Delta c,t+1} \;\rightarrow\; fci_t \not\perp \epsilon^q_{c,t+1}$

- $fci_{t+1} = \beta^q_{y1} y_{t+1} + \beta^q_{\Delta c1} \Delta c_{t+1} + \beta^q_{c1} c_{t+1} + \Gamma_t \;+\; \epsilon^q_{f,t+1} \;\rightarrow\; c_{t+1}, \Delta c_{t+1} \not\perp \epsilon^q_{f,t+1}$

The exogeneity assumption is therefore violated, for instance:

$$\mathbb{P}\left[\Delta C_{t+1} \leq \widehat{\beta^q_{y1}} y_{t+1} + \widehat{\beta^q_y} y_t + \widehat{\beta^q_c} \Delta c_t + \widehat{\beta^q_f} fci_t \middle| \{Y_{t+1}, Y_t, \Delta C_t, FCI_t\}\right] \neq \tau$$

Instruments

To tackle the endogeneity issue in the second and third equations of the system, we use a combination of instruments to estimate the causal coefficients. Because the endogeneous variables are system-wide aggregates ((total PTNI)/(total RWA), (total Tier1 capital)/(total RWA), and financial conditions which aggregate many variables), we use the Granular Instrument Variables (GIV) framework of Gabaix & Koijen (2020) to construct exogeneous instruments for each of these endogeneous variables. The system-wide PTNI/RWA is instrumented via the GIV applied to bank level PTNI/RWA. The system-wide Tier1 Capital/RWA is instrumented via the GIV applied to banks' level Tier1 Capital/RWA. The financial conditions index is instrumented via the GIV applied to banks' expected default probabilities (Moody's EDFs) and banks' CAPM cost of capital (Kovner & Van Tassel (2018)), the latter being a proxy for the banks' costs of equity funding.[13] As a robustness check, we also include the U.S. monetary policy shocks from Cieslak & Schrimpf (2019) as an instrument for FCI. However, because the monetary policy shocks are not available after 2018, it prevents us to use the full historical sample for estimation. This is why we only include this extra instrument in the robustness check.

Following Gabaix & Koijen (2020), the estimation procedure to construct the GIVs follows three steps:

1. **Panel regression** with time and fixed effects at the granular level: $c_{i,t} = \alpha_i + \lambda_t + \epsilon_{i,t}$. By construction $\epsilon_{i,t}$ is orthogonal to the cross-sectional and time heterogeneity captured by the time and fixed effects

2. **Principal component analysis** with K components on the panel residuals: $\epsilon_{i,t} = \sum_{k \in K} \Lambda_k + \nu_{i,t}$. By construction, $\nu_{i,t}$ is orthogonal to the co-movements across banks' residuals (the shocks). $\nu_{i,t}$ therefore represent the idiosyncratic shocks for each bank.

3. The granular instrument is the **average of the largest banks' idiosyncratic shocks** $\nu_{i,t}$: $I_t = \sum_{l \in L} w_{l,t} \nu_{l,t}$ where $w_{l,t}$ is the share of bank l assets into the banking system total assets. This average of the largest banks' idiosyncratic shocks represents the unexpected shocks which can be used to estimate the transmission of financial conditions to the real economy through the banking sector.

Appendix B has a detailed presentation of the construction of the granular IV and the estimation of the quantile coefficients using a two-step process.

4.5 Signs restrictions

Estimation of the recursive system is done line by line, via quantile regressions. First, instrumental variables are constructed, and the FCI and PTNI are projected onto the instrumental variables via

[13] Souza et al. (2016) provide evidence of the link between systemic risk and banks' default probabilities in financial networks. Kovner & Van Tassel (2018) find an empirical relationship between bank-level CAPM cost of capital and credit supply measures such as bank lending standards and loan spreads.

OLS. To make sure that the dynamic system is stable and accords with economic priors, we impose inequality constraints on the quantile coefficients, for all quantiles.

In particular, we impose that:

- The impact of GDP on PTNI (Δc) is positive: when GDP goes down, banks' losses increase and capital goes down.
- The impact of financial conditions on capital is negative: when FCI tightens, banks' have more difficulties to raise capital.
- The impact of capital on financial conditions is negative: lower average banks' capital tighten financial conditions.

We use the interior point algorithm approach of Koenker & Ng (2005) to compute the quantile regressions with inequality constraints. Appendix C presents the coefficients with the signs restrictions.

4.6 Probabilistic Stress-Scenarios and Realized Value-at-Risk

Our recursive model can simulate the conditional densities of GDP, financial conditions, and banks' capital, considering the interaction among these variables at different points of the distribution over a given horizon. However, in the context of stress testing, one needs to go beyond the system's endogenous dynamics and impose a specific stress scenario. A stress scenario assumes that the economy is hit by a series of shocks and the goal is to assess whether banks remain solvent under the scenario. For instance, stress testers can assume a "V-shaped" scenario, where the economy collapses for the first 3-4 quarters and rebounds quickly within a year, or a "U-shaped" scenario, where the recovery is slower.

Generating scenarios in a recursive system is not straightforward. To replicate a given shape with only one initial shock – as in an impulse response function – the fitted system should feature both short-term hysteresis and long-term mean-reverting properties, so that the economy can amplify the first shock for three to four periods, and then rebound. The magnitude of the coefficients will determine a "V"- or a "U"-shape recovery, for instance. In practice, this is extremely difficult to achieve with a recursive model. One would need to fit a large number of lags with significant coefficients to achieve each variable's desired shape. Over-fitting the model makes it intractable and with insufficient explanatory power.

One solution consists of imposing a time-series of shocks in absolute value to the system, negative for the three or four quarters, up to the trough, and positive afterwards. However, this approach is undesirable because it imposes an ad-hoc absolute path for the endogeneous variables, conflicting with the system's own dynamics. In the end, the stress-tester is more or less imposing the results he or she would like to estimate.

We develop a new probabilistic-scenario concept to solve this problem, distinguishing between a shock – a probability – and its impact: the realized Value at Risk. We define a probabilistic-scenario as a time series of probabilities – a sequence of risk levels. For instance, rather than assuming that GDP collapses by 8 percent during the first period, then 4 percent during the second period, etc. we assume that the GDP value during the first period will be the realized VaR at the 5th percentile of the conditional distribution, then the 10th percentile, etc.

Formally, a probabilistic-scenario for an endogenous variable is defined as:

$$\{\theta^X : t \in T\} \quad \text{such that} \quad \mathbb{P}[X_t \leq Q(X_t, \theta_t^X)] = \theta_t \tag{4}$$

To simulate a given probabilistic scenario in the recursive model (3), we condition the projection of the distribution to the realized Value at Risk (the conditional quantile at a given risk level). For instance, in the case of PTNI/RWA:

$$F(\mathbf{\Delta C_{t+1}} \mid \mathbf{Q}(\mathbf{Y_{t+1}}, \theta^{\mathbf{Y}}_{\mathbf{t+1}}, Q(Y_t, \theta_t^Y), Q(FCI_t, \theta_t^F), \Delta C_t) \tag{5}$$

Note that in this case, we condition the projection of the density of $\mathbf{\Delta C_{t+1}}$ on the quantiles of Y_{t+1}, Y_t, FCI_t and on the full density of ΔC_t, without restrictions. In stress testing, the scenarios are often imposed on the macrofinancial conditions – GDP growth and FCI – and the stress-tester studies the reaction of banks' capital to a particular scenario.

Our approach defines shocks as the realization of a given risk and leaves the model determine the Value at Risk. Therefore, the probabilistic scenarios are automatically incorporating the evolving conditions of the economy, based on the conditional densities. The same risk (e.g., 5th percentile of GDP and 95th of FCI) might have a different impact at different times, as the conditioning vector (GDP, FCI, and Capital) is also changing. For instance, the economy might be more resilient, which means that the Value at Risk will be lower for the same level of risk (typically with lower variance or skinnier tails).

Empirically, there are three ways to determine a probabilistic scenario:

1. **Ad-hoc**: assuming a given risk path, such that a series of lower percentiles for the first year to plunge the economy to the trough, then the median (the time to stabilize in the trough), then upper quantiles (to generate a rebound)

2. **Historically-based**: looking at the GFC for instance, we can use our model to estimate the corresponding quantiles of the 2008 Q4 observed realizations of GDP and FCI, based on the conditioning vector in 2008 Q3. And iterate forward.

3. **Model-based**: we translate the shocks in absolute value derived from a model, into the corresponding percentiles, using our model to generate the conditional distributions.

To align our stress-testing approach with the Federal Reserve Board's exercise, we have translated the CCAR scenarios in absolute value into their corresponding percentiles. We start the simulation of our model in 2019:Q4, where the conditioning vector is observed and we project the distribution of GDP, FCI, and capital in 2020: Q1. We then infer the corresponding risk-scenario of GDP and FCI as the percentile of the CCAR absolute values for GDP and FCI in the conditional distribution estimated by the model. And we iterate forward for the entire horizon. The conversion from CCAR absolute values to risk levels is presented in Figure 4.[14]

Mathematically, the conversion is done using the empirical cumulative distribution function:

$$\theta_t^Y : \mathbb{P}[Y_t \leq Y_t^{CCAR}] = \theta_t^Y$$

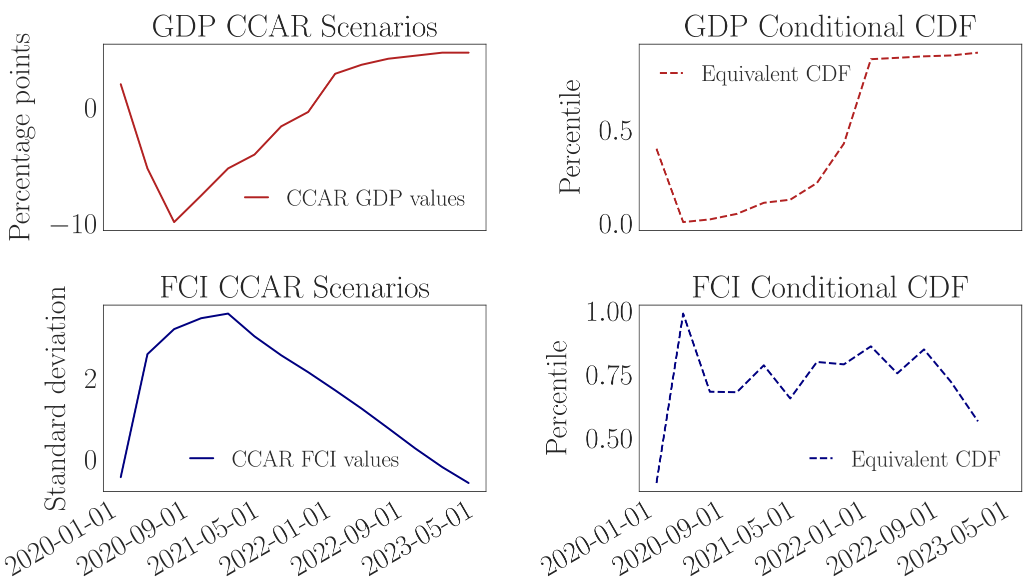

Figure 4: CCAR-implied probabilistic scenarios

4.7 Density Goodness of Fit Test

Finally, we propose to estimate the goodness of fit of our recursive density model using a probability integral transform (PIT) test, with the confidence interval given by Rossi & Sekhposyan (2019). Our testing approach is not truly an out-of-sample exercise: because of the limited original sample size

[14] In practice, when simulating our model, we need to repeat the first shock to align the time dimension with CCAR, because in our recursive model, GDP growth reacts with a lag of one period.

(80 observations), we cannot estimate the quantile regressions coefficients on a smaller sub-sample without losing all accuracy. Therefore, we propose a PIT test where the future density density in $t + 1$ is estimated on the conditioning vector in t, respecting the recursive structure presented in section 4.2, but with the same in-sample quantile coefficients at each periods. The results are presented in figure 5. The figure shows that the goodness of fit is excellent for the three endogeneous variables in our system, GDP, FCI and capital flows (PTNI/RWA), but not on capital in level (Tier1 Capital/RWA), consistent with the large structural break in capital discussed in section 3. These results confirm the validity of our modeling approach and demonstrate that our stylized, parsimonious three-variable model has reasonable goodness of fit.

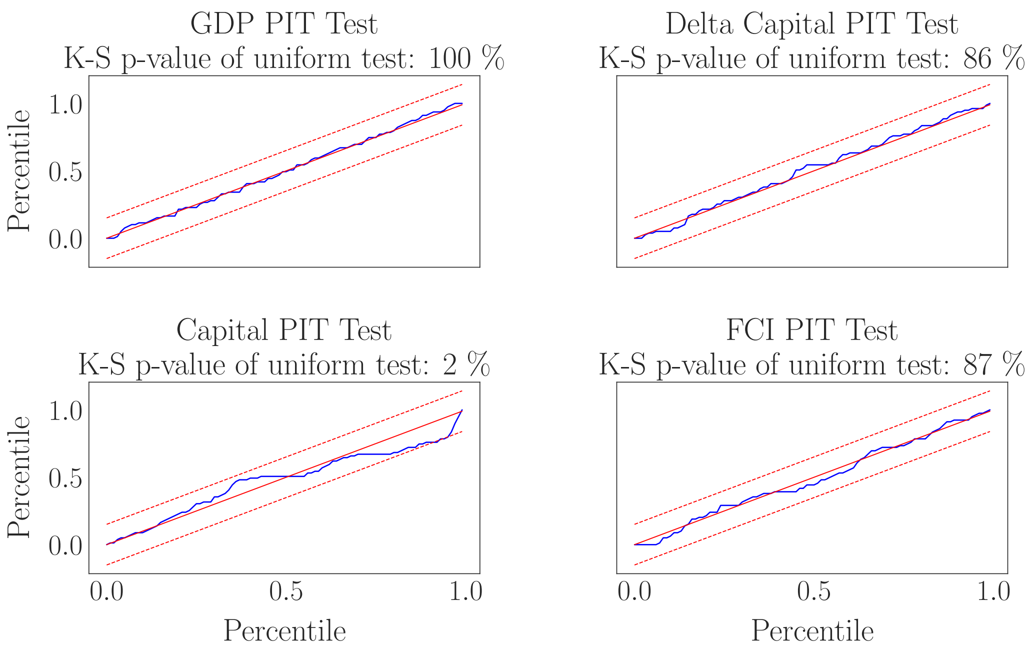

Figure 5: Probability Integral Transform Test of our Recursive Model

5 Results

5.1 CCAR Capital Fan Chart

We start replicating the severely adverse scenario of the CCAR 2020 exercise with our macrofinancial state variables – GDP growth and financial conditions – and consider only bank capital as endogenous variable.[15]. While the CCAR assumptions cover projections of 28 macrofinancial variables (including economic activity, asset prices, and interest rates in the U.S. economy),

[15]The 2020 CCAR assumptions are provided by the Federal Reserve website at `https://www.federalreserve.gov/supervisionreg/ccar-2020.htm`

we focus on a parsimonious stress scenario with just two macrofinancial state variables. Figures 6 show the path of GDP and FCI under the Federal Reserve's 2020 CCAR severely adverse scenario.[16]

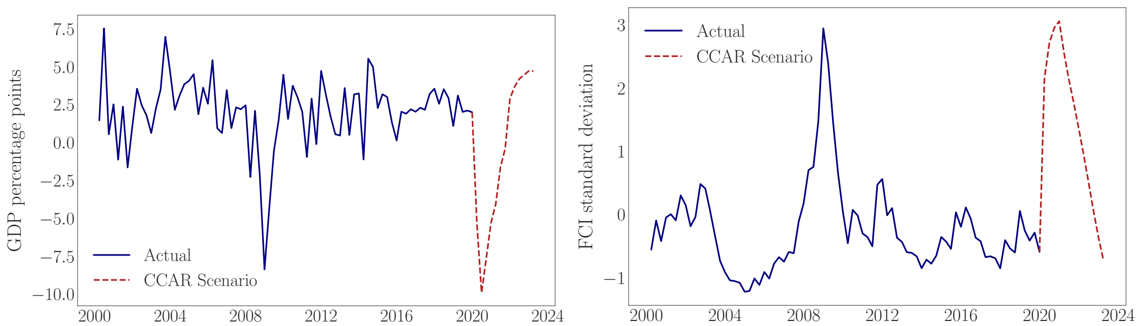

Figure 6: Historical GDP data and CCAR stressed projections (left panel) and VIP scores of the CCAR-stressed Financial Conditions Indicator (right panel)

The CCAR severely adverse scenario features a severe recession, with a GDP contraction of about 10 percent in 2020 (more severe than the 8.4 percent contraction in 2008), though GDP growth is expected to rebound in 2021 and 2022. Financial conditions are modeled as tightening sharply in 2020, with the FCI reaching a peak similar to that in 2008, followed by a pronounced easing. Both GDP and FCI dynamics implicitly incorporate policy support from expansionary monetary and fiscal policies.

Turning to the results, we first discuss the projected distribution of capital under the CCAR scenario, without taking into account the macroprudential feedback, via direct projection. Then, we move to the full-fledged model, where GDP growth and financial conditions are impacted by banks' capital dynamics. We then turn to the macrofinancial feedback effect, quantifying the extent to which banks can be amplifying shocks to economic activity and financial conditions. In turn, we can compute the counterfactual path of the endogenous variables by estimating a counterfactual model where the macrofinancial feedback, i.e. the impact of banks' capital on GDP and FCI, is shut down. Finally, we study the conditional forecast distribution of GDP growth and the FCI, focusing particularly on skewness and bi-modality.

Figure 7 shows the direct distribution projection, the "fan chart", of average Tier 1 Capital/RWA, based on the CCAR 2020 severely adverse scenario for GDP growth and financial conditions. These projections only use the second equation of the recursive density model and the law of motion in the model, where GDP growth and financial conditions shocked paths are exogenously given by the CCAR scenarios. Under this restricted, "pure-CCAR" approach, Tier 1 capital/RWA is endogeneously determined by the (i) CCAR-based GDP growth, (ii) the CCAR-based FCI and the

[16]We do not consider the risk factor shocks to financial variables in the global market shock applied to banks with the largest trading portfolios, which is an add-on exercise assumed to occur instantaneously, that is, in the first quarter of the projection period.

(iii) evolution of pre-tax net income, PTNI/RWA, as modeled below:

$$\Delta c_{t+1} = \beta^q_{y1} y^{CCAR}_{t+1} + \beta^q_y y^{CCAR}_t + \beta^q_{\Delta c} \Delta c_t + \beta^q_f fci^{CCAR}_t + \beta^q_c c_t + \epsilon^q_c$$
$$c_{t+1} = c_t + \Delta c_{t+1} \quad \text{(Deterministic law of motion)}$$

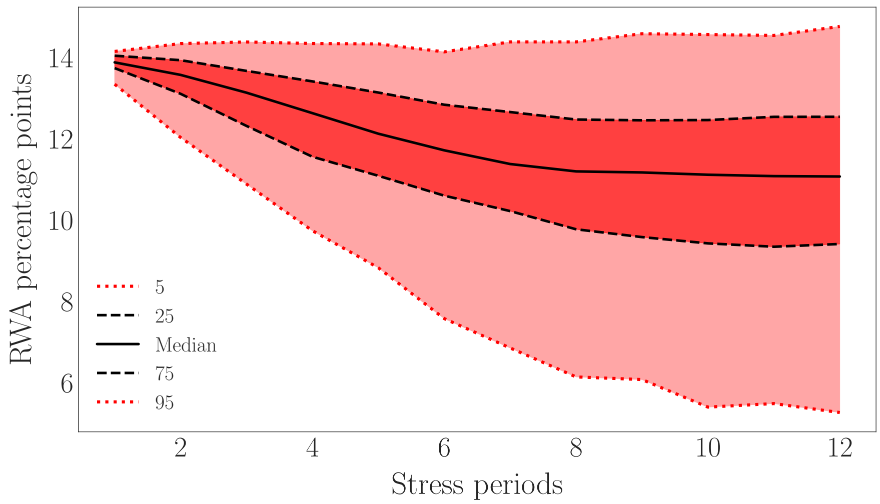

Figure 7: Direct projection of bank capital based on the CCAR scenario

Under this direct approach, the median path for the change in bank capital, conditional on CCAR shocks, is 3.2 percentage points from peak to trough, or equivalently, from start to minimum. In other words, the median Tier 1 Capital/RWA ratio declines from about 13.9 percent at the start of the projection period to a minimum of 10.7 percent after 12 quarters. The projected path of capital and the implied decline from start to minimum, using our parsimonious model, are remarkably similar to the paths and average declines in projected capital ratios in CCAR exercises. The average decline in post-stress capital ratios in CCAR, between 2013 and 2020, is approximately 2.7 percentage points (excluding the global market shock impact). The projected path for the Tier 1 Capital/RWA ratio for the fifth percentile could go up to 8 percentage points lower, from start to minimum, a very extreme tail risk, considering macroeconomic shocks implicit in CCAR assumptions.

In short, this section shows that the capital path in our model, derived from the CCAR scenarios, is well aligned with supervisory stress-test expectations about the impact of a large shock to bank capital from severe macrofinancial conditions.

5.2 Macrofinancial Feedback

This section presents one of the main results of the paper, the estimated magnitude of the macrofinancial feedback. First, the full dynamic feedback is estimated. Then, the restricted system that shuts off the feedback between PTNI and the macrofinancial variables is computed. The difference between the restricted and the unrestricted path of bank capital is our proposed estimate of the macrofinancial feedback, expressed in bank capital terms (e.g., as percent of RWA). Of course, we also calculate the wedge of GDP growth and the FCI with and without the macrofinancial feedback.

The formalization of the restricted model is presented below. The model is simulated, and the distributions projected using the exact same quantile coefficients as with the unrestricted model, so that the counterfactual path we estimate only differs because of the macroprudential feedback, and not because the other coefficients have changed.

$$F(Y^r_{t+1} \mid Y^r_t, FCI^r_t, \Delta C_t = 0, C_t = \bar{c})$$

$$F(\Delta c^r_{t+1} \mid Y^r_{t+1}, Y^r_t, \Delta C^r_t, FCI^r_t, C^r_t)$$

$$F(FCI^r_{t+1} \mid Y^r_{t+1}, Y^r_t, FCI^r_t, \Delta C_{t+1} = 0, C_{t+1} = \bar{c}, \Delta C_t = 0, C_t = \bar{c})$$

To avoid the possibility of a misspecified intercept, we keep the stochastic process $\mathcal{C}$ at is pre-crisis level $\bar{c}$[17] in the specification of GDP and FCI. However, we do force the endogeneous PTNI/RWA variable to be a constant equals to 0 in the specification of GDP and FCI at 0, as it is the endogeneous flow variable in our system. The only difference between the restricted path and the CCAR paths for GDP and FCI is that the restricted paths take into account the distributional effects, while the CCAR paths are exogeneously given.

Although we don't incorporate the feedback of capital on GDP growth and FCI, we estimate PTNI/RWA as an endogeneous variable and derive Tier 1 capital/RWA from the law of motion, based on the restricted paths of GDP growth and FCI. These restricted capital paths are free from the impact of banks' on the economy: banks' are only impacted by the direct GDP growth and FCI shocks, and not by the second round effects they, themselves, generate.

Paths for the median, 5th percentile, and other percentiles are defined as the time series of the median, 5th percentile, and other percentiles of the conditional distributions for GDP growth, FCI, and capital. Note that they are conditional on the probabilistic stress-scenario, as defined in Section 4.6. The advantage of the probabilistic stress-scenarios is that they are the same for both

[17] Putting the capital ratio at 0 instead of the 14 percent pre-crisis level would generate a spurious shock on GDP and FCI.

the unrestricted and the restricted models. However, the realized Value-at-Risks between the two models are different because the restricted model does not incorporate the feedback loop. We then study the projected distribution under the restricted and unrestricted models, conditional to the probabilistic scenarios. For instance, in the case of GDP, we study the properties of the unrestricted stochastic process defined as:

$$\left\{F(Y_t^u t+1 \mid Q(Y_t^u, \theta_t^{Y_{CCAR}}), \Delta C_t^u, C_t^u, Q(FCI_t^u, \theta_t^{FCI_{CCAR}}) : t \in T\right\}$$

The restricted stochastic process is defined as:

$$\left\{F(Y_t^r t+1 \mid Q(Y_t^r, \theta_t^{Y_{CCAR}}), \Delta C_t = 0, C_t = \bar{c}, Q(FCI_t^r, \theta_t^{FCI_{CCAR}}) : t \in T\right\}$$

For instance, the time series of a given τ-quantile for the unrestricted GDP stochastic process is:

$$\left\{Q\left(F(Y_t^u t+1 \mid Q(Y_t^u, \theta_t^{Y_{CCAR}}), \Delta c_t^u, C_t^u, Q(FCI_t^u, \theta_t^{FCI_{CCAR}}), \tau\right) : t \in T\right\}$$

Figure 8 shows the endogenous path of GDP growth with and without feedback, for the median and for the 5th percentile, which is our proxy for downside risk. For the median, the difference in the endogenous path of GDP growth (red area) is around 2 percentage points. Put differently, without accounting for the feedback effect of bank capital on economic activity and financial conditions, banks appear to operate with insufficient capital and thus adversely amplify adverse shocks to GDP growth in crisis time by 2 percentage points. The effect in the 5th quantile is much larger, of the order of 7.5 percentage points. This reflects strong downside risk in our model (the non-linearity in GDP dynamics), similar to the findings of Adrian et al. (2019) in the case of a univariate density GDP growth model for the US. This is a large magnitude by any measure, and the chart shows that this large impact is already reached one year after the initial shock.

Figure 9 shows the endogenous evolution of the change in bank capital (e.g., projected Tier 1 Capital/RWA) for the median and the 5th percentiles with and without the macrofinancial feedback (red area). While the feedback effect for the median change in bank capital is of the order of 1.5 percentage points, the magnitude is larger in the 5th percentile, at 3.4 percentage points. Hence the adverse feedback loop from the macrofinancial interaction has a quantitatively large impact on capital depletion in times of adverse shocks. This result suggests that the second-round effects in bank capital (first bank capital impacts GDP growth and FCI, then GDP growth and FCI impact bank capital) are large and of concern for policymakers.

Finally, Figure 10 shows the endogenous response of the FCI with and without the macrofinancial feedback. The difference between the median and the 5th percentile is relatively substantial, at 2.1 and 3.2 standard deviations, respectively. However, while we observe a strong non-linearity between the median and the 5th percentile for GDP growth, the difference between the median and the 5th percentile is smaller for the conditional FCI distribution. This results suggests that FCI reacts in a

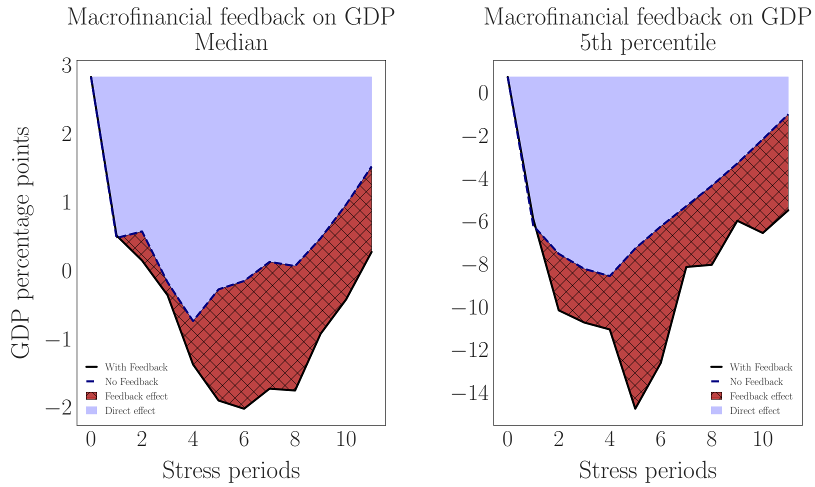

Figure 8: Impact of the macrofinancial feedback on GDP path

more or less linear fashion to shocks.

5.3 Multi-modality and skewness

This subsection investigates the importance of non-linearities and multi-modality in stress-testing exercises. The quantile recursive model estimates the full fledged conditional distributions of the endogeneous regressors across time, subject to the probabilistic scenarios. Figure 11 presents the evolution of the conditional distributions of GDP growth across time, from the first shock to the rest of the CCAR scenario time frame.

The distribution evolves with the dynamics of the CCAR shocks: initially, the distribution deteriorates in the first four quarters, and then slowly shifts back towards the post-crisis regime.

A first striking feature is that the distribution is multi-modal when the first shock hits the economy. At this stage, there is some room for the economy to either absorb the shock smoothly, as a very temporary shock, or on the contrary, to shift towards a crisis regime. As the crisis materializes with the subsequent series of CCAR shocks, the probability of avoiding a crisis regime disappears, and the distribution becomes uni-modal (quarters three to five).

The multi-modality and skewness features illustrate the advantage of our non-parametric approach: although the conditional quantiles are derived from relatively simple quantile regressions, the quantile spacing sampling approach delivers interesting insights. Despite the relative parsimonious nature of our empirical approach, our model allows us to capture rich and complex dynamics, in

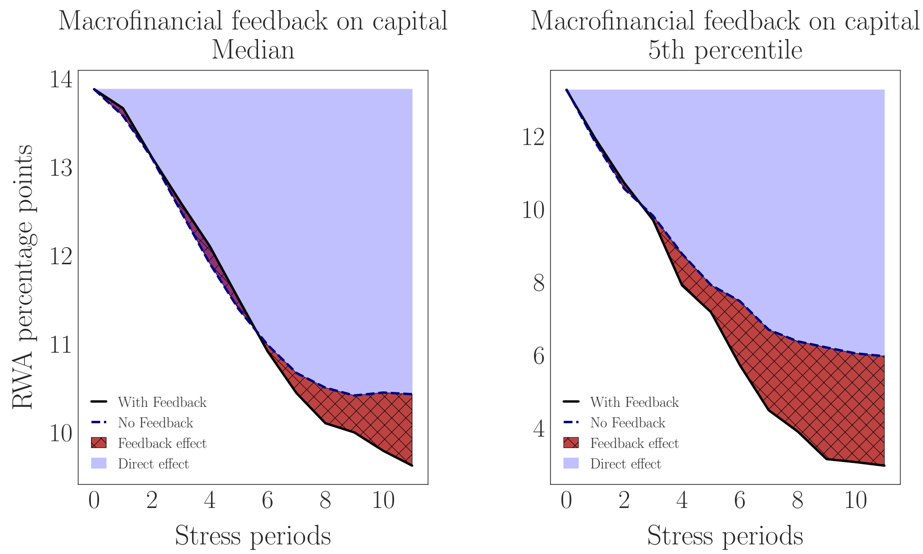

Figure 9: Impact of the macrofinancial feedback on the change in bank capital: PTNI/RWA

particular during crisis times.

In sum, this section has shown that our model captures quantitatively important macrofinancial feedback effects, especially on the downside risk (the 5th percentile). This result has two crucial implications: (i) it stresses the importance of capturing the macrofinancial feedback in stress-test design; (ii) there are strong non-linear effects in the endogeneous dynamic of GDP growth, FCI and bank capital, which are of primary importance for policymakers, who are designing macroprudential policies to mitigate downside risk. Estimating mean-models without macroprudential feedback largely underestimates macrofinancial risks.

5.4 Cyclical Vulnerabilities and the Growth-at-Risk Gap

We link the macrofinancial feedback to cyclical vulnerabilities using the Growth-at-Risk (GaR) framework, developed by Adrian et al. (2019). GaR is defined as the conditional growth of GDP at the lower 5th percentile, and thus estimates downside risk to GDP over the next four quarters. By putting the emphasis on the lower tail of the distribution of forecasted GDP growth, GaR identifies the likelihood of potential bad economic outcomes, such as an economic recession.

We use our parsimonious model (equation 3) and calculate a one-year ahead GaR estimate conditional on the current state of the economy, described by current GDP growth, financial conditions, and bank capitalization. We call this measure, conditional GaR, at a one-year horizon. We also calculate the unconditional GaR, namely, the Value-at-Risk of the GDP unconditional distribution.

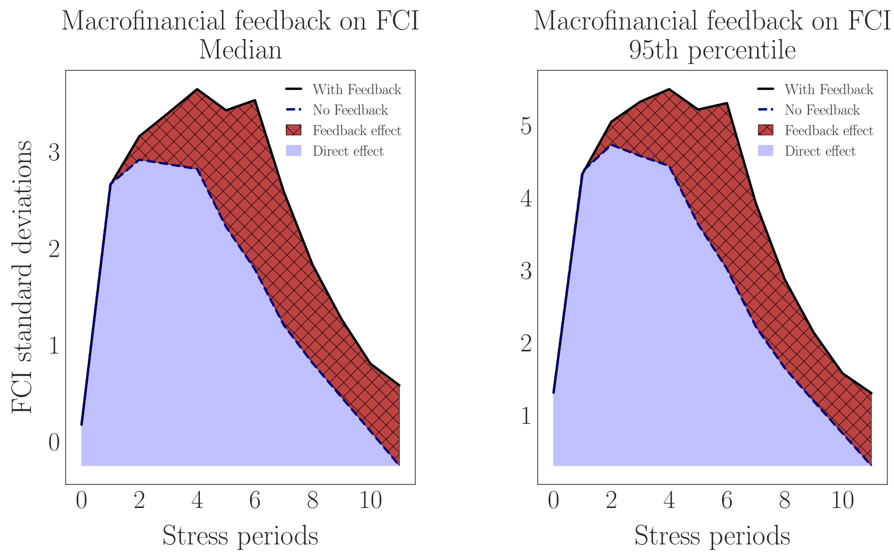

Figure 10: Impact of the macrofinancial feedback on Financial Conditions

To cancel out parametric noise in the GaR gap metric, we approximate the GDP unconditional distribution based on historic averages of the variables in our model, using a real-time expanding windows.[18]

We then propose the GaR gap, the difference between the conditional GaR and unconditional GaR as a metric of countercyclical vulnerabilities. Intuitively, the difference between conditional and unconditional GaR captures the difference between cyclical and structural vulnerabilities, and thus the GaR gap summarizes the extent of cyclical vulnerabilities in the economy by assessing downside risks to GDP growth.

$$\text{Gap}(\tau) \;=\; Q(y_{t+1year}|y_t, fci_t, \Delta c_t, \tau) - Q(y_{t+1year}|\overline{y_t}^m, \overline{fci_t}^m, \overline{\Delta c_t}^m, \tau)$$

Figure 12 shows the GaR gap at the 5th percentile (left panel) and the median (right panel) of the distribution, calculated at each point in time between 2000:Q1 and 2019:Q4.[19] The conditional GaR (red dashed lines) captures vulnerabilities at a high frequency relative to the unconditional GaR (solid black line), which evolves more slowly over time, but exhibits a hump-shaped path during the 2007-09 financial crisis in both the 5th percentile and the median. The GaR gap, shown as green and pink areas for positive and negative values, respectively, is a time-varying metric that captures vulnerabilities at a high frequency, with more pronounced swings at the 5th percentile than at the

[18] The unconditional GaR at time t is computed based on historical averages of the regressors up to time t: $\frac{1}{t}\sum_0^t X_j$

[19] For graphical reasons, we have increased the frequency of the GaR metric from quarterly to monthly by interpolation, and we also smooth both the unconditional and conditional GaR with a four-period moving average

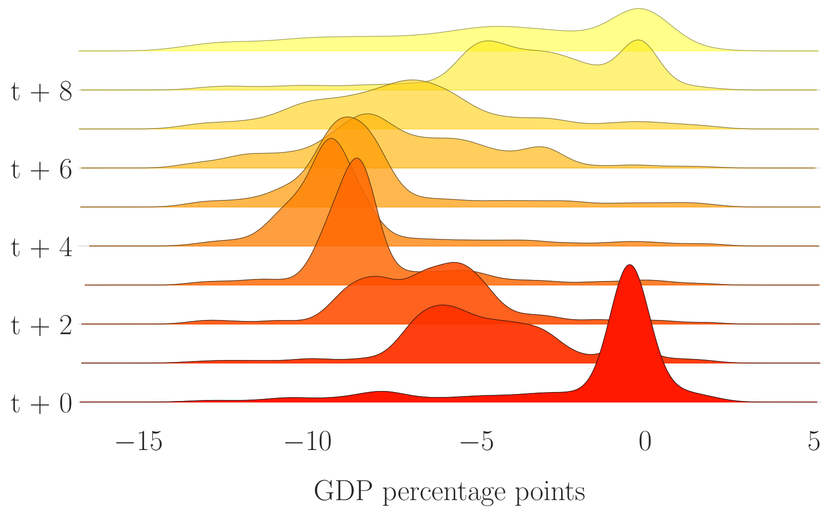

Figure 11: GDP conditional distributions over time

median. This feature of the GaR gap is consistent with the higher sensitivity of the left tail of the distribution of forecasted GDP growth relative to the median as previously shown by Adrian et al. (2019).

Figure 12 shows positive GaR estimates (green areas) before the 2007-09 financial crisis, and in two periods after the crisis, 2013 through 2015, and 2017 through 2019, suggesting times of economic expansion with the possible build-up of vulnerabilities. Large negative GaR estimates occurred during the global financial crisis, whereas small negative estimates are observed around the sovereign debt crisis in Europe in 2012, and the oil price shock in 2016. Within each of the green areas, the GaR gap increases up to a peak, and then starts declining and turns negative during times of low or negative economic growth. For example, the GaR gap increased by the end of 2003 and reached a peak in mid-2005. This is consistent with the idea of an economic expansion and thus low vulnerabilities (e.g., lower probability of a recession). After the peak in 2005, vulnerabilities start to increase and the GaR gap starts to decline, consistent with a higher probability of a recession -as the GaR is forecasted one-year ahead, which occurred in the midst of the 2007-2009 financial crisis.

5.5 Growth-at-Risk Gap and the Credit Gap

We compare our GaR metric of countercyclical vulnerability with alternative measures of financial vulnerability such as the Credit-to-GDP Gap (credit gap), the gap between the ratio of credit to GDP and its long-term backward-looking trend, a metric that has received a prominent role in

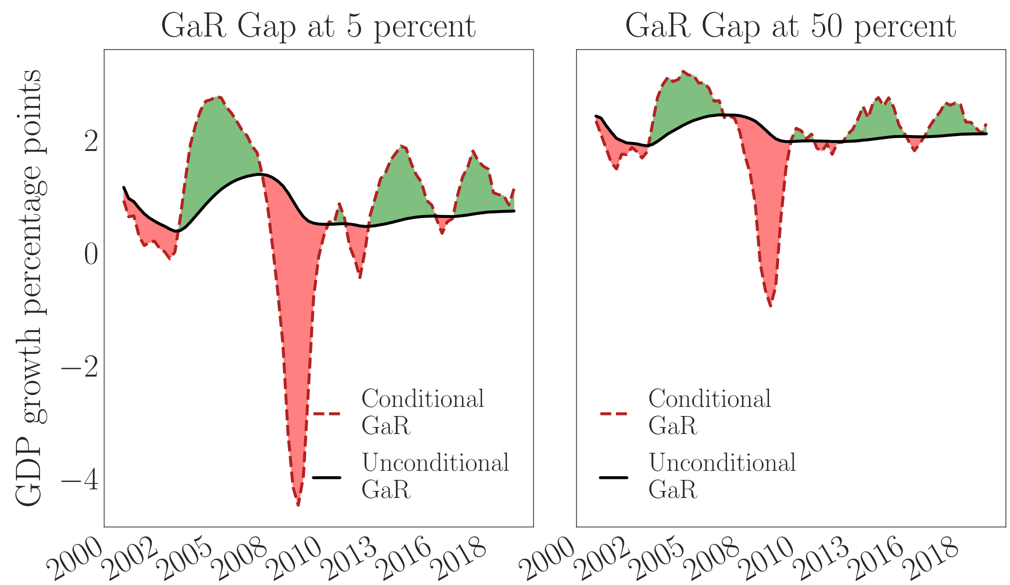

Figure 12: The Growth-at-Risk Gap metric, 5th percentile and median risk tolerance

the Basel III guides for policymakers to calibrate the countercyclical capital buffer (CCyB).[20] As Drehmann et al. (2011) argues, the credit-to-GDP gap captures the build-up of systemic vulnerability in the economy that can lead to banking crises.

Figure 13 compares the credit-to-GDP gap (left panel) with the GaR gap at the 5th percentile of the distribution (right panel), showing positive (green) and negative (pink) gap values. Compared with the credit gap, the GaR gap shows more time variation, with cycles at higher frequency than the business cycle. In contrast, the credit gap exhibits variation at much lower frequency than the business cycle. Both measures surge in the period leading to the global financial crisis, pointing to a significant increase in vulnerabilities and thus suggesting positive values for the CCyB in the runup to the global financial crisis. Both measures drop abruptly in 2008 as the crisis unfolds, suggesting a deactivation of the CCyB to prevent a severe cutback in bank lending. However, unlike the GaR gap, which signals periods of increased vulnerabilities and thus positive CCyB values during the post-crisis era, the credit gap remains at negative values, suggesting a zero CCyB for the U.S. throughout the 2009-2019 post-crisis decade.

The literature has criticized the credit gap as a basis for setting the CCyB on several grounds. For example, empirical evidence suggests that the credit gap is a poor countercyclical indicator, that it does not capture well bank loan losses (Wezel (2019)), and that it does not perform well in real time (Edge & Meisenzahl (2011)). Compared with the GaR gap, which performs well in real

[20]The Basel III rule for setting the CCyB is based on a linear relationship between the credit-to-GDP gap and the CCyB level. Accordingly, the CCyB is set to zero for values of the credit gap below 2 percentage points and capped at 2.5 percent for values of the gap above 10.

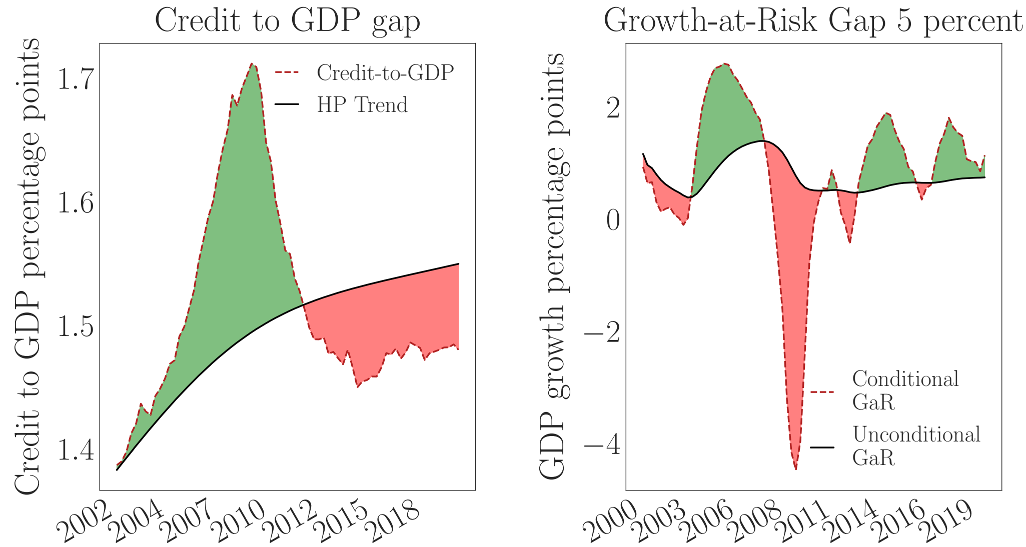

Figure 13: Credit-Gap versus Growth-at-Risk Gap

time, the credit gap is not risk based and does not capture amplification in the tails.[21]. Of course, our set-up can be easily augmented by adding the credit gap as additional state variable, an exercise that we leave for future work.

5.6 Macrofinancial Vulnerabilities and Countery-Cylical Capital Buffer (CCyB)

In this section, we use our GaR measure of financial stability risks at the one-year horizon to guide the calibration of the CCyB. Conceptually, countercyclical vulnerabilities through the business and financial cycles include the macrofinancial feedback, that is, the size of the banks' amplification of shocks to the economy resulting from insufficient bank capital, recursively estimated with instrumented quantile regressions, as explained in section 4. As discussed above, the macrofinancial feedback is larger at the tails of the GDP growth distribution. Our GaR gap metric informs the setting of the CCyB as the additional bank capital or capital surcharge required to offset the macrofinancial feedback through the business cycle.

We infer the capital surcharge needed to offset the macrofinancial feedback, and thus derive a measure of the countercyclical capital buffer (CCyB). To do so, we compute the difference between the unrestricted and restricted macro-financial dynamics, in the peak-to-trough difference, for a set of probabilities (risk levels). We repeat the computation at each point in time and infer the capital surcharge needed, at different risk levels. Calling C_t^{PTT} the peak-to-trough difference for capital at time t, for a given risk τ:

[21]Regarding its calculation, additional criticism points to many statistical shortcomings such as an end-point problem, sensitivity to the choice of the smoothing parameter in the filtering procedure (e.g., the lambda parameter in the HP filter) and a time-varying trend, all of which makes it difficult for policy use

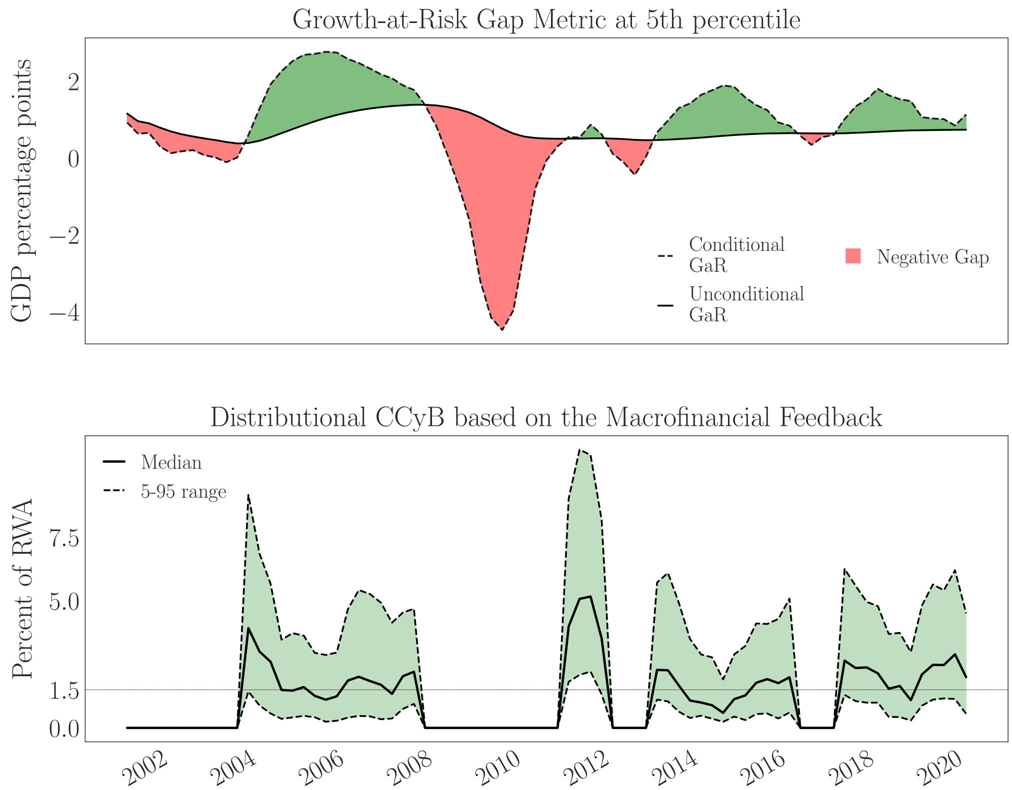

Figure 14: From macrofinancial vulnerabilities to the calibration of the CCyB

$$\text{CCyB}(\tau)_t \;=\; \mathbb{1}_{\text{Gap}(\tau)>0}\left[Q(C_j^{R,PTT} \mid \theta_j^{CCAR}\}, \tau)_t^{t+H} - Q(C_j^{U,PTT} \mid \theta_j^{CCAR}, \tau)_t^{t+H}\right]$$

The lower panel of figure 14 shows the range of CCyB estimates, given by the 5th and 95th percentiles, as well as the median in our sample, based on the macrofinancial feedback associated with the decomposition of the GaR gap in the upper panel. Notice that the CCyB is truncated at zero when the gap is negative.

Our estimates indicate that the CCyB for the median during the pre-crisis period (between September 2003 and June 2007) should have been on average around 190 basis points (near the upper bound of the Basel III CCyB) and about 460 basis points for the 5th percentile. During the post-crisis (between December 2010 and December 2019), our estimates suggest that the positive CCyB for the median should be around 210 basis points on average, and about 510 basis points on average for the 5th percentile.

Our measure of countercyclical vulnerability shows that periods of economic expansion, when GaR is high and points to low downside risk, are followed by periods with high downsize risk resulting from a build-up of systemic vulnerabilities, and vice versa, consistent with the term structure of GaR (Adrian et al. (2018)). Clearly, compared with alternative vulnerabilities metrics such as the credit gap, the GaR gap exhibits higher variations (e.g., more volatility) through the cycle, which we view as a desirable feature that helps inform the setting of the CCyB.

By connecting the evolution of cyclical vulnerability to different levels of the capital surcharge, our methodology suggests a higher capital surcharge (CCyB) when the GaR gap declines. As figure 14 suggests, the rule for deactivation of the CCyB is clear and easy to implement (e.g., when the GaR gap turns negative). However, activation or rise of the CCyB has to be gradual and may need a transition when vulnerabilities start to rise (when the GaR gap decreases), which as the lower panel shows, occurs quarters before the GaR gap turns negative. Consequently, we can argue that policymakers may announce the need to raise the CCyB in advance and during "good times" (when the GaR gap increases), and use the capital surcharge to guide the setting of the CCyB level when the GaR gap starts decreasing, which would be the period when systemic vulnerabilities are meaningfully above its normal level (unconditional GaR). For example, a policymaker desiring to hedge against the 5 percent probability of a recession could have announced an increase of the CCyB by year-end 2004, and could have set it in by year-end 2005 at about 460 basis points (increasing the aggregate Tier1/RWA ratio from 9.1 percent to 13.7 percent). Of course, the policymaker would have reduced the CCyB to zero around September 2007 (e.g., during the collapse of short-term funding markets), immediately before the GaR gap turned negative.

6 Policy Implications

Our results have important policy implications for the design of bank stress testing. Our proposed methodology allows supervisors to expand the classic microprudential bank stress tests to account for macrofinanical feedback, thus providing a rigorous basis for macroprudential policy. The methodology accounts for the indirect risks emanating from macrofinancial feedback that give rise to amplification effects of adverse credit supply effects via bank balance sheet management, which the current microprudential stress test approach cannot gauge. Our model provides an estimate of the macrofinancial feedback effects using the CCAR severely adverse scenario. More importantly, it calibrates a bank capital surcharge, which we design as the additional bank capital that accounts for the macrofinancial feedback between the financial sector and the real economy. The capital surcharge can be thought of as a macroprudential tool to mitigate banks' amplification of real and financial shocks to the economy.

Our results offer a menu of policy options to implement the capital surcharge at different points in the GDP growth and financial conditions distributions. More specifically, for 2019:Q4 (the most recent observation in our sample), our findings suggest a median capital surcharge of about 1.5 percentage points to avoid banks' amplification effect of 2 percentage points in the median GDP growth. The magnitude of the capital surcharge in the 5th percentile increases to 3.4 percentage points, to avoid a longer and deeper recession (7.5 percentage point decline on aggregate GDP growth). Therefore, the capital surcharge increases with financial vulnerabilities at the tails of the distribution of future GDP growth.

Our proposed measure of cyclical vulnerabilities, the GaR gap, can guide policymakers in setting the CCyB that offsets the macrofinancial feedback at different points in the business cycle. Our estimates suggest that the capital surcharge on the median before the global financial crisis should have been at 190 basis points, close to the upper value of the Basel III CCyB. The capital surcharge on the median is around 210 basis points on average, and could get up to 510 basis points for the 5th percentile, post crisis. From a policy perspective, the level of the capital surcharge at different points in time depends on the risk tolerance of the policymaker, that is, how much risk the policymaker would like to hedge against will determine how much extra capital is required from the banks. Our model shows a very strong non-linear relationship between countercyclical vulnerability and the macrofinancial feedback. Therefore, policymakers will require higher bank capital ratios to hedge the 5th percentile than to hedge the median.

Our framework expands existing rules in the calibration of the CCyB, using a forward-looking stress-testing perspective. Some jurisdictions have already moved towards integrating the CCyB calibration into stress testing. For example, in the U.K., the Bank of England uses the bank stress test results to help calibrate the CCyB. Further, between 2015 and 2019, 15 countries used the CCyB, and many of them raised it multiple times. As of March 2020, 13 of these countries had reduced their CCyB.[22] In the U.S., policymakers have never activated the CCyB as of the time of writing.

7 Conclusion

We propose a methodology for modeling and calibrating macrofinancial feedback effects in the context of bank stress tests. We propose three key innovations relative to the existing literature. First, we focus on PTNI as the relevant state variable, and show that that is closely related to both financial conditions and GDP growth. Second, we introduce valid instruments for bank capital and FCI to make causal statements. Third, we impose economically motivated sign restrictions in the estimation of our reduced-form macrofinancial model. We compute the macrofinancial feedback from a counterfactual exercise, where the feedback between bank capital and the state variables is shut off. We show that the counterfactual path of GDP growth and bank capital are quantitatively different with large magnitude during stress, implying a large macrofinancial feedback from undercapitalized banks.

We then propose a methodology to calibrate bank capital surcharges through the business cycle, such as the CCyB. Our findings for 2019:Q4 suggest a median capital surcharge of about

[22] For details, see Edge & Liang (2020) and Benediktsdottir et al. (2020).

1.5 percentage points to avoid banks' amplification effect of 2 percentage points in the median GDP growth. The capital surcharge for the 5th percentile increases to 3.4 percentage points, to avoid a longer and deeper recession (7.5 percentage point decline on aggregate GDP growth). The bank capital surcharge—the additional bank capital that offsets the macrofinancial feedback—is typically overlooked in bank stress tests. We then propose a measure of countercyclical vulnerabilities, the GaR gap, which we argue improves upon existing metrics such as the credit gap to inform policymakers about the size and the timing of a countercyclical capital buffer that offsets the macrofinancial feedback at different points in the business cycle.

Our methodology then allows bank supervisors to calibrate the CCyB within the regular supervisory stress testing exercise, a major advance over current practice. Our model suggests that the capital surcharge on the median before the global financial crisis should have been at 190 basis points, near the upper value of Basel III CCyB, and 460 basis points for the 5th percentile. The capital surcharge on the median should be around 210 basis points on average, and could go up to 510 basis points for the 5th percentile, during the post crisis. Our methodology can easily augment the current stress testing machinery to include the calculation of the macrofinancial feedback and the capital surcharge, and this can be done though simple auxiliary equations relative to models currently estimated. Finally, our methodology is easily applicable to any stress testing approach (e.g., macro scenarios of different severity, different planning horizons) and thus can be easily adopted by supervisors in different jurisdictions.

One caveat in our analysis based on a reduced-from model is the possibility that the estimation of the model parameters is distorted by policy changes (e.g., the Lucas critique). In the absence of a structural model, our empirical approach mitigates this concern by using the GIV method to deal with endogeneity concerns and to produce unbiased estimates, which we interpret as providing a causal impact of bank capital on downside risk of future GDP growth and financial conditions.

References

Acharya, V. V., Berger, A. N., & Roman, R. A. (2018). Lending implications of us bank stress tests: Costs or benefits? *Journal of Financial Intermediation*, 34, 58–90.

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, 109(4), 1263–89.

Adrian, T., Boyarchenko, N., & Giannone, D. (2020a). Multimodality in macro-financial dynamics. *Federal Reserve Bank of New York Staff Report*, (903).

Adrian, T., Grinberg, F., Liang, N., Malik, S., & Zhu, J. (2018). The term structure of growth at risk. *American Economic Journal: Macroeconomics*, (forthcoming).

Adrian, T., Morsink, J., & Schumacher, L. (2020b). Stress testing at the IMF. *IMF Departmental Paper*.

Agarwal, S., An, X., Cordell, L., & Roman, R. (2020). Banks' stress test results and their impact on consumer credit markets. *Working Paper. Federal Reserve Bank of Philadelphia*.

Auer, R. & Ongena, S. (2019). The countercyclical capital buffer and the composition of bank lending. *CEPR Discussion Paper*, 13942.

Benediktsdottir, S., Feldberg, G., & Liang, N. (2020). What macroprudential policies are countries using to help their economies through the covid-19 crisis? *The Brookings Institute*.

Benes, J. & Kumhof, M. (2015). Risky bank lending and countercyclical capital buffers. *Journal of Economic Dynamics and Control*, 58, 58–80.

Berrospide, J. & Edge, R. (2019). The effects of bank capital buffers on bank lending and firm activity: What can we learn from five years of stress-test results? *Working Paper. Board of Governors of the Federal Reserve System*.

Boyarchenko, N., Giannone, D., & Kovner, A. (2020). Bank capital and real gdp growth. *Federal Reserve Bank of New York Staff Report*, (950).

Brave, S. A., Kelly, D. L., et al. (2017). Introducing the chicago fed's new adjusted national financial conditions index. *Chicago Fed Letter*, 386.

Brechmann, E. C., Hendrich, K., & Czado, C. (2013). Conditional copula simulation for systemic risk stress testing. *Insurance: Mathematics and Economics*, 53(3), 722–732.

Calem, P., Correa, R., & Lee, S. J. (2020). Prudential policies and their impact on credit in the united states. *Journal of Financial Intermediation*, 42, 100–126.

Caselli, F., Grigoli, F., Lafarguette, R., & Wang, C. (2020). Predictive density aggregation: A model for global gdp growth. *IMF working paper*, 1(20/78).

Chavleishvili, S., Engle, R., Fahr, S., Kremer, M., Manganelli, S., & Schwaab, B. (2020). The risk management approach to macro-prudential policy. *ECB and NYU Stern School of Business working paper*.

Chen, L.-Y. & Lee, S. (2018). Exact computation of gmm estimators for instrumental variable quantile regression models. *Journal of Applied Econometrics*, 33(4), 553–567.

Chernozhukov, V. & Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1), 379–398.

Cieslak, A. & Schrimpf, A. (2019). Non-monetary news in central bank communication. *Journal of International Economics*, 118, 293–315.

Cortes, K., Demyanyk, Y., Li, L., Loutskine, E., & Strahan, P. (2020). Stress tests and small business lending. *Journal of Financial Economics*, 136, 260–279.

Drehmann, M., Borio, C., & Tsatsaronis, K. (2011). Anchoring countercyclical capital buffers: The role of credit aggregates. *International Journal of Central Banking*, 7(4), 189–240.

Edge, R. & Liang, N. (2020). Stronger financial stability governance leads to greater use of the countercyclical capital buffer. *The Brookings Institute*.

Edge, R. & Meisenzahl, R. (2011). The unreliability of credit-to-gdp ratio gaps in real-time: Implications for countercyclical capital buffers. *International Journal of Central Banking*, 7(4), 261–298.

Federal Reserve (2009a). The supervisory capital assessment program: Design and implementation. *http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090424a1.pdf*, April 24.

Federal Reserve (2009b). The supervisory capital assessment program: Overview of results. *http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090507a1.pdf*, May 7.

Fernandez, M., Igan, D., & Pinheiro, M. (2020). March madness in wall street: (what) does the market learn from stress tests? *Journal of Banking and Finance*, 112.

Flannery, M., Hirtle, B., & Kovner, A. (2017). Evaluating the information in the federal reserve stress tests. *Journal of Financial Intermediation*, 29, 1–18.

Gabaix, X. & Koijen, R. S. (2020). Granular instrumental variables. *SSRN paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3368612*.

Götz, O., Liehr-Gobbers, K., & Krafft, M. (2010). Evaluation of structural equation models using the partial least squares (pls) approach. In *Handbook of partial least squares* (pp. 691–711). Springer.

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Hirtle, B. & Lehnert, A. (2015). Supervisory stress tests. *Annual Review of Financial Economics*, 7, 339–355.

Jimenez, G., Ongena, S., Peydro, J., & Saurina, J. (2017). Macroprudential policy, countercyclical bank capital buffers and credit supply: Evidence from the spanish dynamic provisioning experiments. *Journal of Political Economy*, 125, 2126–2177.

Jorda, O., Schularick, M., & Taylor, A. (2013). When credit bites back. *Journal of Money, Credit and Banking*, 45(2), 3–28.

Koenker, R. (2020). *quantreg: Quantile Regression*. R package version 5.61.

Koenker, R. & Ng, P. (2005). Inequality constrained quantile regression. *Sankhya: The Indian Journal of Statistics*, (pp. 418–440).

Kohn, D. (2019). Stress tests and the countercyclical capital buffer: the uk experience. *The Brookings Institution*.

Kovner, A. & Van Tassel, P. (2018). Evaluating regulatory reform: Banks' cost of capital and lending. *Federal Reserve Bank of New York, Staff Reports*, 1(854).

Milne, A. K. L. & Whalley, A. E. (2001). Bank capital and incentives for risk-taking. *Cass Business School Research Paper*.

Prasad, A., Elekdag, S., Jeasakul, P., Lafarguette, R., Alter, A., Alan, X., & Wang, C. (2019). Growth at risk: Concept and applications in IMF country surveillance. *IMF working paper*, 1(19/36).

Rossi, B. & Sekhposyan, T. (2019). Alternative tests for correct specification of conditional predictive densities. *Journal of Econometrics*, 208(2), 638–657.

Schmidt, L. & Zhu, Y. (2016). Quantile spacings: A simple method for the joint estimation of multiple quantiles without crossing. *Available at SSRN 2220901*.

Seabold, S. & Perktold, J. (2020). Statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

Souza, S., Silva, T., Tabak, B., & Guerra, S. (2016). Evaluating systemic risk using bank default probabilities in financial networks. *Journal of Economic Dynamics and Control*, 66, 54–75.

Wasserman, L. (2013). *All of statistics: a concise course in statistical inference*. Springer Science & Business Media.

Wezel, T. (2019). Conceptual issues in calibrating the basel III countercyclical capital buffer. *IMF Working Paper*, 19/86.

A Estimate Conditional Densities via Quantile Regressions and Quantile Spacing Sampling

We present in this section our approach to estimate the recursive set of conditional densities presented above using quantile regressions and the quantile spacing approach proposed by Schmidt & Zhu (2016). To keep the notation simple, we use a generic example of a single random variable Y_{t+1} conditioned on a set of regressors $\{X_{1,t}, \cdots, X_{K,t}\}$.[23] We then discuss how to generalize the estimation when the regressors are not scalar but samples of random distributions themselves, and how we deal with the curse of dimensionality.

A.1 Quantile Regressions and Conditional Quantiles

Quantile regressions are a standard and robust tool to estimate conditional quantiles via a simple specification; in our case, we use a linear specification. Quantile regressions offer simple asymptotic inference and hypothesis testing. They are also appealing to practitioners and policymakers as their coefficients can be interpreted in a straightforward manner - the marginal impact of the regressor on the conditional quantile of the dependent variable.

We want to estimate the conditional distribution of a random variable Y_{t+1} from a set of conditioning variables $\{X_{1,t}, \cdots, X_{K,t}\}$. We first estimate a set of quantile regressions for different quantile values, $\tau \in (0, 1)$:

$$Y_{t+1} = \alpha^\tau + \sum_{k=1}^{K} \beta_k^\tau X_{k,t} + \epsilon_{t+1}^\tau \quad \text{for } \tau \ \in \ (0,1) \tag{6}$$

Assuming that the exogeneity assumption[24] holds $\mathbb{P}[\epsilon_{t+1}^\tau \leq 0 | \{X_{1,t}, \cdots, X_{K,t}\}] = \tau$, then:

$$\mathbb{P}\left[Y_{t+1} \leq \sum_{k=1}^{K} \hat{\beta}^\tau X_{k,t} \middle| \{X_{1,t}, \cdots, X_{K,t}\} \right] = \tau$$

Which means, by definition, that $\sum_{k=1}^{K} \hat{\beta}^\tau X_{k,t}$ is the conditional τ-quantile of Y_{t+1}:

$$Q(Y_{t+1} | \{X_{1,t}, \cdots, X_{K,t}\}, \tau) = \sum_{k=1}^{K} \hat{\beta}^\tau X_{k,t}$$

Where $Q(Y|X,\tau)$ is the τ-quantile of Y conditional on X ; for simplicity, we call X_t the vector of regressors $\{X_{1,t}, \cdots, X_{K,t}\}$. Hence, estimating a collection of quantile regressions over the set of quantile $\Theta = \{\tau_1, \cdots, \tau_Q\}$ allows us to recover a set of conditional quantiles $\hat{Q}(Y_{t+1}|X_t, \Theta)$. This is a point-wise quantile function defined over $\{\tau_1, \cdots, \tau_Q\}$, a subset of $(0, 1)$

A.2 Quantiles Uncrossing via Quantile Spacing

The point-wise quantile function $\widehat{Q}_{Y_{t+1}}(\Theta|X_t)$ should be monotonically increasing,[25] as quantiles are by construction increasing. However, because the conditional quantiles are estimated over a finite distance as $\sum_{k=1}^{K} \hat{\beta}^\tau X_{k,t}$, they might be crossing due to parametric noise, i.e:

[23] All variables are defined over the real line $\mathbb{R}$ for simplicity

[24] This is the quantile equivalent of the exogeneity assumption for OLS $\mathbb{E}[\epsilon|X] = 0$

[25] Throughout the paper, by "monotone" we mean (weakly) increasing

$$\exists\ \tau_i, \tau_j \in (0,1):\ \widehat{Q}_{Y_{t+1}}(\tau_i|X_t) > \widehat{Q}_{Y_{t+1}}(\tau_j|X_t) \text{ with } \tau_i\ \leq \tau_j$$

We follow the quantile spacing approach, suggested by Schmidt & Zhu (2016). The idea is to interpolate any conditional quantiles using the set of the estimated conditional quantiles set from the quantile regression using a base function (for instance, the quantile function of a $N(0,1)$). Therefore, from the quantile regressions fitted values (for instance 5%, 25%, 50% and 75%, 95%), we can estimate any quantile between 0 and 1.

1. From the point-wise quantile function $\widehat{Q}_{Y_{t+1}}(\Theta|X_t)$ estimated above estimate any conditional quantile $Q(\alpha, X),\ \forall\ \alpha\ \in\ (0,1)$ interpolate differently on whether the quantiles are outside or inside the estimated range:
    - $Q(\alpha, X) = a_{out}(X) + b_{out}(X)\Psi(\alpha)$ when α is below the minimal estimated conditional quantile via the quantile regression (e.g. 5%) or above the maximal one (95%). For instance when estimating the quantile at 1% when the quantile regression is done at 5%. The values for a_{out} and b_{out} are provided in Schmidt & Zhu (2016)
    - $Q(\alpha, X) = a_{in}(X) + b_{in}(X)\Psi(\alpha)$ when α is within the estimated range, in our case, between 5% and 95%
    - The function $\Psi(\alpha)$ is the base function, a quantile function of a known distirbution, for instance the $N(0,1)$
2. The quantile function of $\tilde{Y}_X$, defined as: $\widehat{Q}^*_{\tilde{Y}_X}(u|X_t) \equiv \widehat{F}^{-1}(u|X_t) = \inf\{y_{t+1} : \widehat{F}(y_{t+1}|x_t) \geq u\}$ is by construction monotonically increasing. This function coincides with $\widehat{Q}_{Y_{t+1}}(\Theta|X_t)$ when $\widehat{Q}_{Y_{t+1}}(\Theta|X_t)$ is monotonically increasing, i.e. when the quantiles are not crossing. The process is therefore consistent both for crossing and non-crossing conditional quantiles.

A.3 Inverse Probability Integral Transform, Quantiles Conditioning on Samples and the Curse of Dimensionality

The inverse probability integral transform - also called the inverse transform sampling - is a simple method to generate sample from any probability distribution given its quantile function.[26] The probability integral transform states that if X is a random variable with a cumulative distribution function F_X, then the random variable $Z \equiv F_X(X)$ has a uniform distribution over $(0,1)$. Therefore, if we have an estimator for $F_X^{-1}(.)$, we can generate a random sample $S(x)$ from a uniform distribution $\mathcal{U}$, so that $S(x) \sim X$. In practice:

1. Draw a uniform random sample U of length L over $(0,1)$ using a random number generator
2. Using $\widehat{F}^{-1}_{Y_{t+1}}(.|X_t)$ from the quantile spacing procedure explained above, generate an inverse transform sample $S(Y_{t+1}) \equiv \widehat{F}^{-1}_{Y_{t+1}}(U|X_t)$ of length L
3. Under the assumption that $\widehat{F}^{-1}_{Y_{t+1}}(.|X_t)$ converges in distribution to the true value $F^{-1}_{Y_{t+1}}(.|X_t)$, then we know that the sample drawn in step 2. follows the true distribution: $S(Y_{t+1}) \sim Y_{t+1}|X_t$

[26] For a textbook exposition of the inverse probability integral transform, please refer to Wasserman (2013). We have been using the basic version of the inverse probability integral transform but faster algorithm exist, such as the ziggurat algorithm and other rejection sampling approaches

So far, we have described a simple way to bootstrap a sample of Y_{t+1}, conditional on a set of explanatory variables X_t.

Generally, when fitting quantile regressions, we only observe one realization of X_t, from the time series, with one observation $x_{i,t}$ per regressor X_i per period t. Hence, from a vector x_t of size $[1 \times K]$, where K is the number of regressors, the procedure outlined above generate a sample $S(y_{t+1})$ of length $[L \times 1]$ for Y_{t+1}. The length of the set Θ, i.e. the number of conditional estimated via quantile regressions has no impact on the sample length of $S(y_{t+1})$, which is only determined by the length of the uniform distribution drawn in the first step of the inverse probability integral transform. The choice of the conditional quantiles to estimate is driven only by accuracy estimation - for instance, avoiding extreme quantiles when the number of observations is limited. In our case, we estimate the quantiles from 5% to 95% by increment of 5 percentage points: $5\%, 10\%, 15\%, \cdots, 85\%, 90\%, 95\%$ and we interpolate the rest.

Now, we consider the case where $x_{i,t}$ $\forall i \in [1, \cdots, K]$ is itself a sample of length L,[27] so that the sample $S(x_t)$ of regressors is now of length $[L \times K]$. This situation occurs if, for instance, the set of regressors X_t is endogeneously estimated in a recursive system, from a previous step. In this case, for each vector $x_{l,k} \in S(x_t)$ of length $[1 \times K]$ (one observation per regressor in the regressors set) corresponds a sample $S(y_{l,t+1})$ of length $[L \times 1]$. Hence, we can map the full sample $S(x_t)$ into a $[L * L, 1]$ sample-length of $S(y_{t+1})$

$$\underbrace{S(x_{l,t})}_{[1\times K]} \rightarrow \underbrace{S(y_{l,t+1})|x_{l,t}}_{[L\times 1]}$$

$$\underbrace{S(x_t)}_{[L\times K]} \rightarrow \underbrace{S(y_{t+1})|x_{l,t}}_{[L*L\times 1]}$$

Therefore, if $S(y_t)$ is now used as a conditioning vector in the next recursion, the sample length of the dependent variable will be L^3, and will grow exponentially with the number of recursions. To avoid this curse of dimensionality, we need to introduce another sampling step, to keep the sample length of the generated sample $S(y_t)$ constant equal to L. Rather than drawing directly a random sample of length L from $\underbrace{S(y_t)}_{[L*L\times 1]}$, we draw a singleton from each conditioning sample $S(y_{l,t+1})$:

$$S(y_{l,t+1}|x_{l,t}) \rightarrow y'_{l,t+1}|x_{l,t} \quad \forall\, l \in [0, L] \rightarrow \underbrace{S(y_{t+1})|S(x_t)}_{[L\times 1]}$$

This approach ensures that the resulting sample $S(y_{t+1})$ will always be of length L and not of length $L * L$. Also, because we sample from each conditional sample $S(y_{l,t+1}|x_{l,t})$, we make sure that full conditioning sample vector $S(x_t)$ is impacting $S(y_t)$, respecting the conditioning structure and avoiding to neglect some conditioning vectors $S(x_{l,t})$ from $S(x_l)$

[27] To simplify and without loss of generality, we assume that we draw the same sample length L for all variables

A.4 Summary of the estimation process

Therefore, assuming a random variable Y_{t+1} of dimension $[T \times 1]$ and a set of regressors X_t of dimension $[T \times K]$, where T represents the number of observations in the time series and K the length of regressor, we generate a random sample $S(y_{t+1}|S(x_t))$ of length $[L \times 1]$ distributed as $Y_{t+1}|X_t$ via:

1. Estimate the set of quantile regressions over the historical time-series $Y_{t+1} = \alpha^\tau + \beta^\tau X_t + \epsilon^\tau$ for a set $\Theta = \{\tau_1, \cdots, \tau_n\} \in (0,1)$ of length N
2. Retrieve the stacked matrix of quantiles coefficients $\underbrace{\widehat{\beta}^\Theta}_{[N \times K]}$, where the rows corresponds to the quantile and the columns to the regressors
3. Consider a given conditioning sampling vector $S(x_t) \sim X_t$ to project the quantiles of Y_t. Note that $S(x_t)$ can be observed, in which case it will be of dimension $[1 \times K]$ or sampled of dimension $[L \times K]$. We take the most general case and assume it is of dimension $[L \times K]$
4. Use $\widehat{\beta}^\Theta$ to project the quantiles $\underbrace{Q(Y_{t+1}|S(x_t),\Theta)}_{[N \times L]} = \underbrace{\widehat{\beta}^\Theta}_{[N \times K]} \underbrace{S(x_t')}_{[K \times L]}$
5. Use the quantile spacing interpolation and probability integral transform to sample $\underbrace{S(y_{t+1}|S(x_t))}_{[L \times 1]}$ from $\underbrace{Q(Y_{t+1}|S(x_t),\Theta)}_{[N \times L]}$
6. Asymptotically, $S(y_{t+1}|S(x_t)) \sim Y_{t+1}|X_t$ for a given conditioning vector $S(x_t) \sim X_t$.

B Construction of the Granular Instruments

Following Gabaix & Koijen (2020), the estimation procedure to construct the GIVs is done in three steps:

1. **Panel regression** with time and fixed effects at the granular level: $c_{i,t} = \alpha_i + \lambda_t + \epsilon_{i,t}$. By construction $\epsilon_{i,t}$ is orthogonal to the cross-sectional and time heterogeneity captured by the time and fixed effects
2. **Principal component analysis** with K components on the panel residuals: $\epsilon_{i,t} = \sum_{k \in K} \Lambda_k + \nu_{i,t}$. By construction, $\nu_{i,t}$ is orthogonal to the co-movements across banks' residuals (the shocks). $\nu_{i,t}$ therefore represent the idiosyncratic shocks for each bank.
3. The granular instrument is the **average of largest banks' idiosyncratic shocks** $\nu_{i,t}$: $I_t = \sum_{l \in L} w_{l,t} \nu_{l,t}$ where $w_{l,t}$ is the share of bank l assets into the banking system total assets. This average of largest banks' idiosyncratic shocks represents the unexpected shocks which can be used to estimate the transmission of financial conditions to the real economy through the banking sector.

The Granular Instrument Variable framework of Gabaix & Koijen (2020) ensures that the two main conditions for the validity of an instrument are respected. First, the **exclusion restriction**, i.e. the orthogonality of the instrument with the residuals of the regressions, is obtained by the double orthogonalization process: (i) the panel regression orthogonalizes banks' shocks with time and cross-sectional heterogeneity; (ii) the PCA of these banks' shocks remove the co-movement among shocks, hence guaranteeing that these shocks are, indeed, idiosyncratic. Second, the **relevance condition**, i.e. the explanatory power of the instrument to the endogeneous regressor, is achieved via the averaging over the largest banks. Idiosyncratic shocks to large banks have a system wide impact, hence also over averaged metrics, such as average Tier 1 capital/RWA or PTNI/RWA.

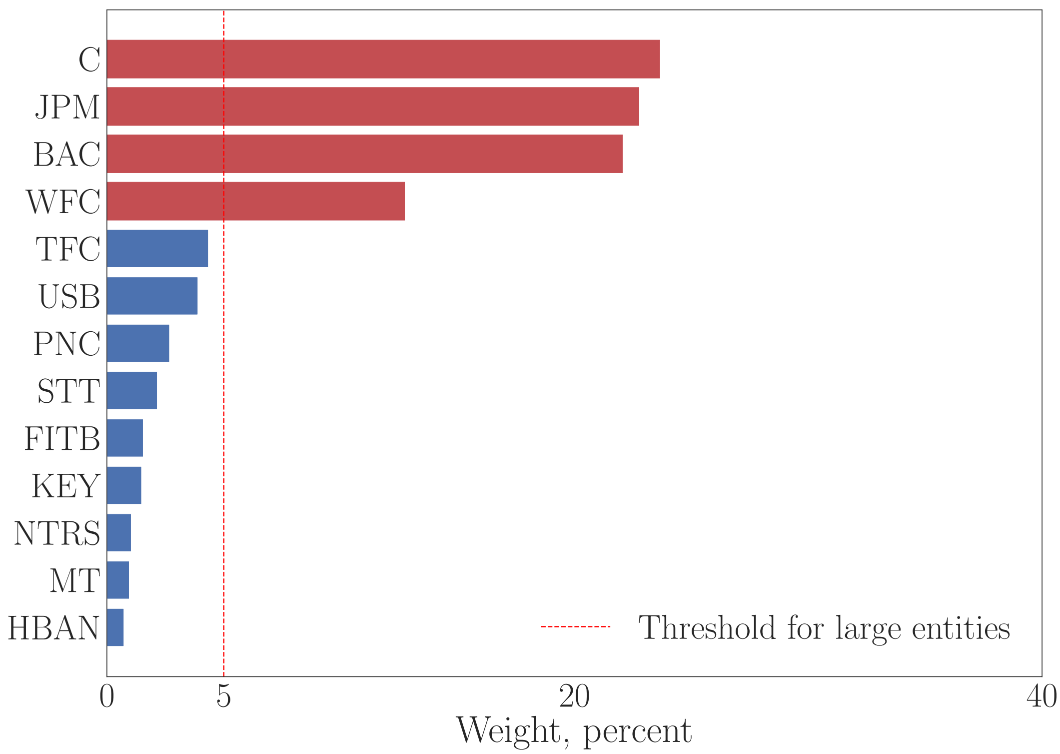

Figure 15: Market share by banks and selection threshold

Figure 15 shows the market share of US banks, together with the selection threshold for the granular instrumental variable construction. As presented, the method extracts the idiosyncratic shocks of the fourth largest banks. The relative concentrated nature of the US banking system is particularly well suited for Granular IV estimation: it reinforces the strength of the argument, as the largest banks are more likely to have an impact when they dominate the others by a large margin, rather than in a relatively uniformly distributed banking system.

Figure 16 shows the fraction of variance explained by a PCA factor model applied to the banks' CAPM cost of capital. With five principal components, which we use for the PCA specification, 90 percent of variance is explained, which suggests that we manage to extract most of the comovement among banks' shocks to construct the idiosyncratic shocks to the CAPM cost of capital.

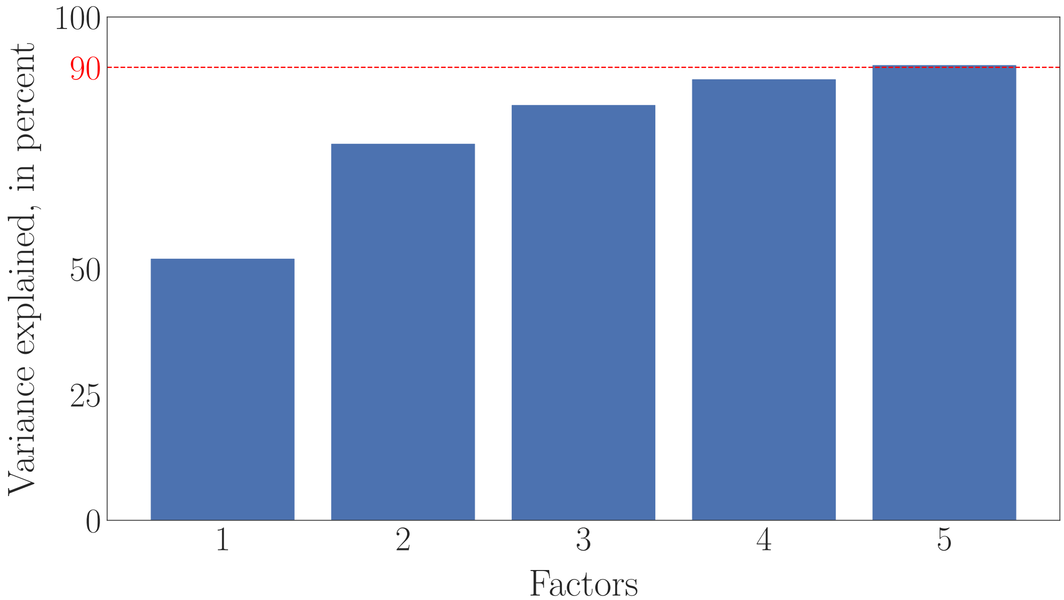

Figure 16: Cumulated variance explained by the PCA factors

B.1 Two-stage estimation

We assume that the endogeneity is linear and doesn't vary by quantiles.[28] This assumption allows to apply a two-step approach in the estimation of the instrumented beta coefficients:

1. In the first stage, the endogeneous variables are regressed on the instruments and the other exogeneous regressors, using a simple heteroscedastic-robust OLS estimator.

2. In the second-stage, the fitted values of the first step are used in the quantile regressions.

The first stage OLS specifications incorporate both the instruments and the other exogeneous variables, following the literature on instrumental variables two-stage least squares:

$$\Delta c_{t+1} = \alpha + \beta_{\Delta c,GIV}\mathrm{GIV}(\Delta c_{t+1}) + \beta_{y1}y_{t+1} + \beta_y y_t + \beta_f fci_t + \beta_{\Delta c}\Delta c_t + \beta_c c_t + \epsilon_{\Delta c,t+1}$$

$$c_{t+1} = \alpha + \beta_{c,GIV}\mathrm{GIV}(c_{t+1}) + \beta_y y_t + \beta_f fci_t + \beta_{\Delta c}\Delta c_t + \beta_c c_t + \epsilon_{c,t+1}$$

$$fci_t = \alpha + \beta_{e,GIV}\mathrm{GIV}(\mathrm{EDF}_t) + \beta_{cp,GIV}\mathrm{GIV}(\mathrm{CAPM}_t) + \beta_{y1}y_{t+1} + \beta_y y_t + \beta_{\Delta c}\Delta c_t + \beta_c c_t + \epsilon_{f,t}$$

The summary tables of the first-stage regressions are presented in Tables 3, 4 and 5. The F-statistics of each first stage is above 10, suggesting a relatively strong explanatory power of the instruments to the endogeneous variables (the so-called "strength" of the instrument). In the case of the FCI instruments, where the model is overidentified, the Sargan's test of overidentification returns a statistics of 0.0924 and a p-value of 76% of the null hypothesis H_0 that "H_0: overidentifying

[28] Although the literature does develop instrumental quantile regressions (Chernozhukov & Hansen (2008) or Chen & Lee (2018)), we are not aware of a derivation of instrumental quantile regressions with inequality constraints. We impose inequality constraints to ensure consistency with economic priors. We thus use a hybrid two-stage approach with OLS in the first stage and inequality constrained quantile regressions in the second stage

restrictions are valid". Hence, we cannot reject that the overidentifying restrictions are valid.

Dep. Variable:	PTNI/RWA in t1	**R-squared:**	0.3025
Estimator:	OLS	**Adj. R-squared:**	0.2444
No. Observations:	79	**F-statistic:**	10.951
Distribution:	chi2(6)	**P-value (F-stat)**	0.0899
Cov. Estimator:	robust		

	Parameter	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
Intercept	0.0025	0.0017	1.4593	0.1445	-0.0009	0.0059
GDP growth in t+1	0.0004	0.0003	1.2216	0.2219	-0.0002	0.0009
GDP growth in t	9.406e-05	0.0001	0.7027	0.4822	-0.0002	0.0004
FCI in t	-0.0007	0.0006	-1.1886	0.2346	-0.0019	0.0005
PTNI/RWA in t	0.1393	0.1100	1.2658	0.2056	-0.0764	0.3550
Tier1 capital/RWA in t	0.0002	0.0095	0.0176	0.9859	-0.0185	0.0188
Capital granular instrument in t	-0.4949	0.3326	-1.4882	0.1367	-1.1467	0.1569

Table 3: First stage summary table of the instrumented PTNI/RWA

Dep. Variable:	Tier1 capital/RWA in t1	**R-squared:**	0.1231
Estimator:	OLS	**Adj. R-squared:**	0.0757
No. Observations:	79	**F-statistic:**	22.148
Distribution:	chi2(4)	**P-value (F-stat)**	0.0002
Cov. Estimator:	robust		

	Parameter	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
Intercept	0.1069	0.0049	21.978	0.0000	0.0974	0.1164
GDP growth in t	0.0007	0.0013	0.5435	0.5868	-0.0018	0.0032
FCI in t	-0.0076	0.0041	-1.8584	0.0631	-0.0157	0.0004
PTNI/RWA in t	-1.0214	0.8501	-1.2015	0.2296	-2.6876	0.6448
Tier1 capital/RWA instrument in t+1	2.8746	0.7057	4.0734	0.0000	1.4915	4.2578

Table 4: First stage summary table of the instrumented Tier1 capital/RWA

Figure 17 shows the actual and the fitted values for PTNI and FCI from the estimation of the dynamic system. The fit is very close, in particular in crisis time, as crisis are usually large unexpected shocks and should be reflected both in the endogeneous and the fitted variable.

Second-stage ordinary least squares projection of the first-stage model provides fitted values of $\widehat{\Delta c_{t+1}}, \widehat{c_{t+1}}, \widehat{fci_t}$, which are then used in quantile regressions of the recursive model, instead of their endogeneous counterparts $\Delta c_{t+1}, c_{t+1}, fci_t$:

	Dep. Variable:	FCI in t	**R-squared:**	0.6780		
	Estimator:	OLS	**Adj. R-squared:**	0.6511		
	Distribution:	chi2(6)	**F-statistic:**	232.60		
	Cov. Estimator:	robust				
	Parameter	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
Intercept	0.9627	0.2056	4.6812	0.0000	0.5596	1.3657
GDP growth in t+1	-0.0557	0.0265	-2.1008	0.0357	-0.1077	-0.0037
GDP growth in t	-0.1640	0.0346	-4.7457	0.0000	-0.2318	-0.0963
PTNI/RWA in t	-65.315	18.110	-3.6065	0.0003	-100.81	-29.819
Tier1 capital/RWA in t	-2.7722	1.7774	-1.5597	0.1188	-6.2559	0.7115
FCI EDF instrument	13.962	2.7725	5.0360	0.0000	8.5282	19.396
FCI CAPM instrument	-21.125	48.717	-0.4336	0.6646	-116.61	74.360

Table 5: First stage summary table of the instrumented FCI

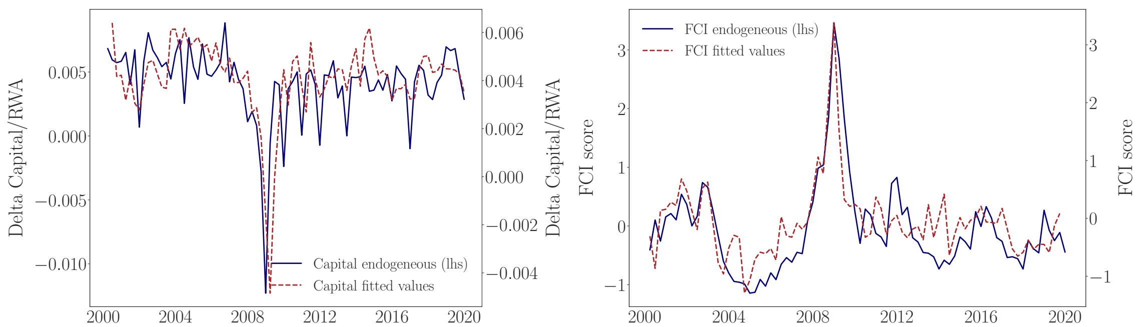

Figure 17: First stage fitted values for Δ capital (left panel) and FCI (right panel)

$$y_{t+1} = \alpha_y^q + \beta_y^q y_t + \beta_{\Delta c}^q \Delta c_t + \beta_f^q fci_t + \beta_c^q c_t \ + \epsilon_{y,t+1}^q$$

$$\Delta c_{t+1} = \alpha_c^q + \beta_{y1}^q \mathbf{y_{t+1}} + \beta_y^q y_t + \beta_{\Delta c}^q \Delta c_t + \beta_f^q \widehat{fci_t} + \beta_c^q c_t + \epsilon_{c,t+1}^q$$

$$fci_{t+1} = \alpha_f^q + \beta_{y1}^q \mathbf{y_{t+1}} + \beta_{\Delta c1}^q \widehat{\mathbf{\Delta c_{t+1}}} + \beta_c^q \widehat{\mathbf{c_{t+1}}} + \beta_y^q y_t + \beta_{\Delta c}^q \Delta c_t + \beta_f^q fci_t + \beta_c^q c_t + \ \epsilon_{f,t+1}^q$$

$$c_{t+1} = c_t + \Delta c_{t+1} \quad \text{(Deterministic law of motion)}$$

The quantile beta coefficients are estimated from the second-stage regression presented above, using inequality constrained quantile regressions, as explained below. So far, we have described how the recursive system is estimated, and how the instruments are constructed that allow us to make causal statements. We next turn to an explanation of the quantile regressions with sign restrictions.

C Sign Restrictions

The coefficients for the constrained quantile regressions are presented in Figures 18, 19, 20. As the constraint can be binding on some parts of the distribution, the coefficients are very close to the zero line. Note that the constraint is imposed at the regression level, so it has an impact on all coefficients, even for those that are left unconstrained.

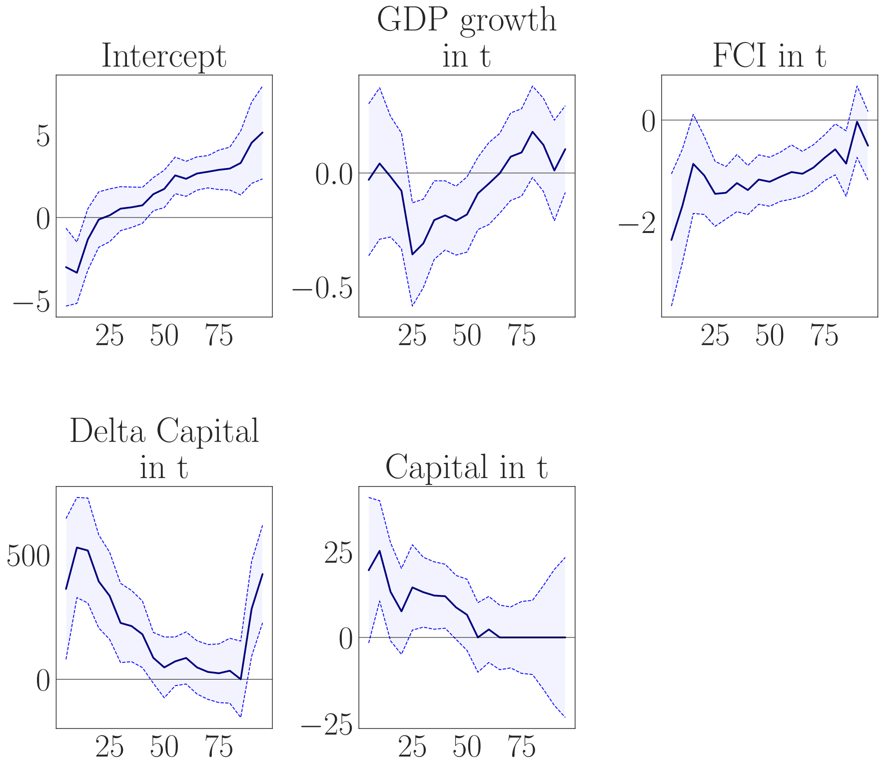

Figure 18: Constrained quantile coefficients for the GDP specification

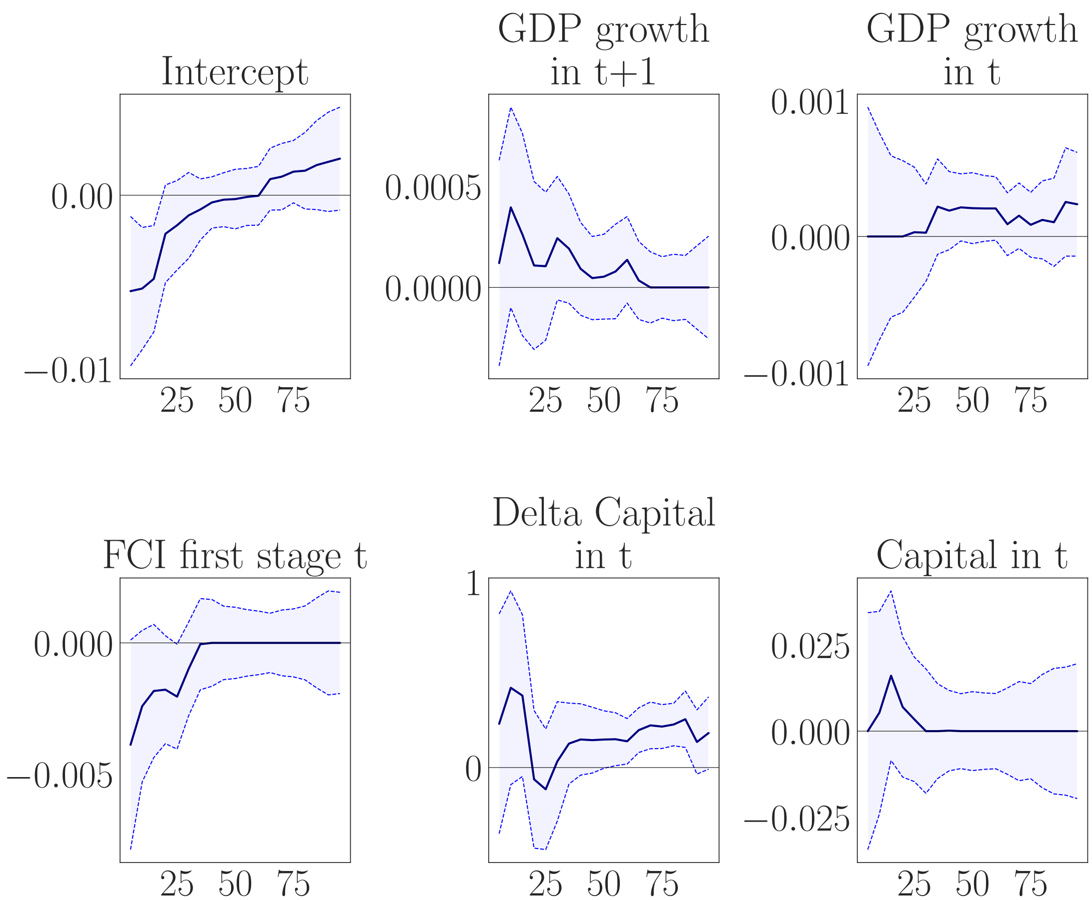

Figure 19: Constrained quantile coefficients for the PTNI/RWA specification

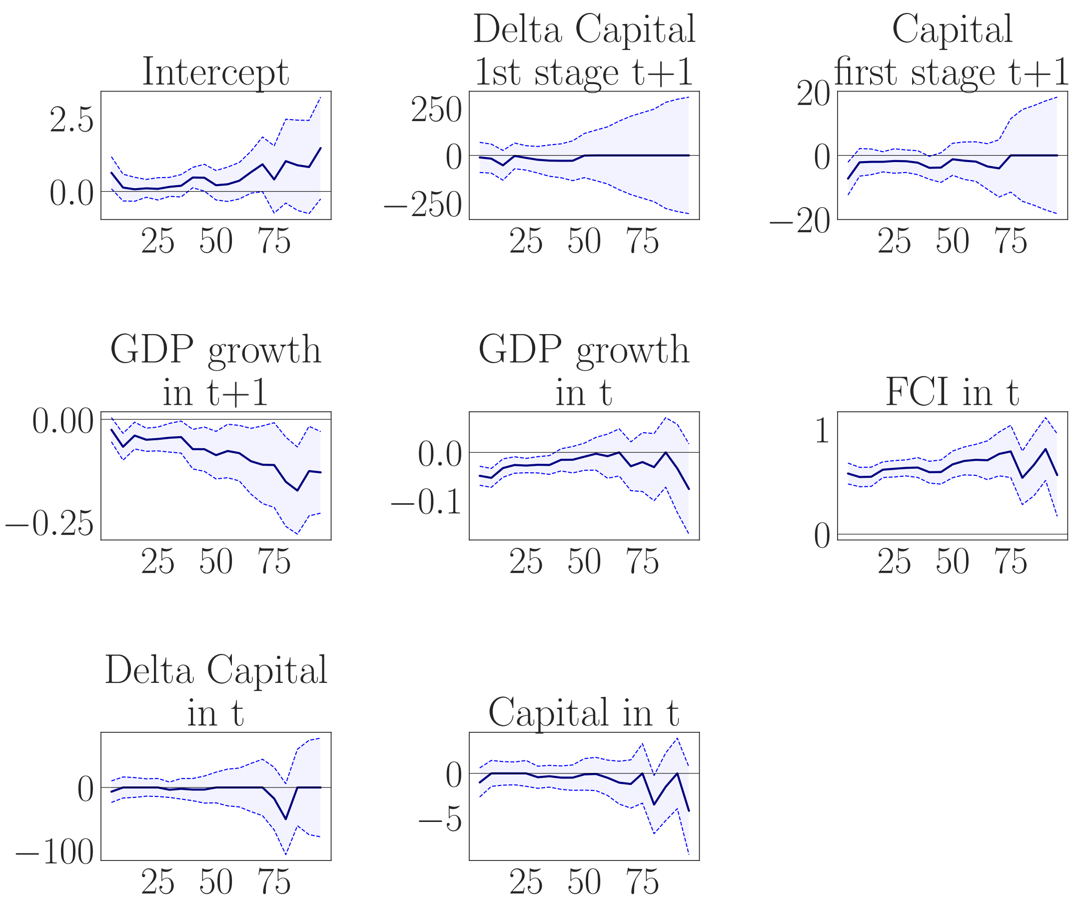

Figure 20: Constrained quantile coefficients for the FCI specification