

## Borrower vulnerabilities, their distribution and credit losses<sup>1</sup>

*Central banks and other supervisory authorities have made significant efforts to collect borrower-level data on debt vulnerabilities. Do such data add to the information in aggregate measures? We find that statistics about household and non-financial corporate borrowers with low repayment capacity help to explain changes in non-performing loans and bankruptcies that aggregate measures would have missed. Borrower-level data could thus be useful for financial stability assessments.*

*JEL classification: E32, E44, G32, G33, G51.*

Private non-financial sector debt has increased strongly in recent years (CGFS (2022)). As monetary policy tightens and economic conditions worsen, weaker borrowers could be pushed to the cusp of default or beyond. Thus, in addition to aggregate measures that represent the typical borrower, information about the distribution of vulnerabilities across borrowers may be informative about future credit losses.

To analyse the distribution of borrower vulnerabilities, granular information on individual borrowers is required. However, such data are often unavailable or, when they can be obtained, provide only limited historical information for many countries. Hence, the existing literature has focused mainly on how measures of aggregate repayment capacity can explain or predict credit losses or financial crises (eg Claessens et al (2012), Drehmann et al (2012), Schularick and Taylor (2012), Juselius and Tarashev (2022)). Much less is known about how borrower-level vulnerabilities might be used to inform financial stability assessments.

This special feature uses a unique cross-country data set to analyse whether information about the distribution of borrower-level repayment capacity helps to explain future credit losses. Collected from 20 central banks for the CGFS report on private sector debt and financial stability (CGFS (2022)), this data set contains summary statistics on the distribution of two common measures of repayment capacity: the debt-to-income and debt service-to-income ratios. These are available separately for the household and non-financial corporate (NFC) sectors. The summary

<sup>1</sup> The authors thank Claudio Borio, Stijn Claessens, Alessio de Vincenzo, Egemen Eren, Andreas Fuster, Ulf Lewrick, Benoît Mojon, Hyun Song Shin, Nikola Tarashev, Phillip Wooldridge and Egon Zakrajšek for helpful comments and discussions, and Daniel Heimgartner, Nicolas Lemerrier and Oliver Surbek for excellent research assistance. The views expressed in this article are those of the authors and do not necessarily reflect those of the Bank for International Settlements, the Bank of Italy and the Swiss National Bank.

### *Key takeaways*

- Data collected from 20 central banks reveal information about differences in borrowers' capacity to repay debt, and how these have evolved from 2006 to 2021, within both the household and corporate sectors.
- The repayment capacity of the most vulnerable borrowers can evolve quite differently from trends in aggregate measures that represent the typical borrower.
- Data on borrowers with the lowest repayment capacity are useful for financial stability assessments because they help to predict future credit losses that would be missed by aggregate data.

statistics capture information about borrowers in the “tail of concern” – ie those with the lowest repayment capacity.

Can information about the tail of concern be used to complement aggregate measures for the prediction of credit losses? To answer this question, we start by constructing two metrics of borrower vulnerabilities in this tail. The first reflects the extent to which the vulnerability of the typical borrower within the tail of concern is greater than that of the typical borrower overall (ie the median borrower) – the *tail length*. The second conveys information about the extent to which the group of borrowers with the lowest repayment capacity is more vulnerable than the typical borrower in the tail of concern – the *tail thickness*. We then assess whether these two metrics help to explain credit losses, after accounting for the aggregate repayment capacity as well as that of the median borrower.

Our analysis produces two main findings.

First, the two metrics of borrowers' repayment capacity from the tail of concern – length and thickness – have evolved quite differently from the corresponding aggregate measure over our sample (2006–21). This provides a first indication that borrower-level data – be it about households or NFCs – could contribute information that is not contained in aggregate data.

Second, information about the tail of concern helps to explain future credit losses. After taking aggregate vulnerabilities into account, both the tail length and tail thickness appear useful in explaining banking sector non-performing loans (NPLs) and business bankruptcies over the next three years. Thus, despite being parsimonious, the two metrics appear useful for financial stability analysis.

The remainder of this special feature is organised as follows. The first section briefly surveys the literature on how the distribution of borrower vulnerabilities can help to inform financial stability assessments. The second describes the data set and explains how we capture information about the tail of concern. The third examines whether shifts in this tail reflect changes in borrower vulnerabilities that do not surface in aggregate statistics. The fourth analyses the extent to which the tail of concern helps predict future credit losses. Finally, we conclude with policy takeaways.

## Financial stability implications of debt distribution: the literature

There is a growing awareness that the distribution of borrower vulnerabilities matters for credit losses. Central banks and banking supervisors use granular data on the repayment capacity of vulnerable borrowers in their stress tests (eg Anderson et al (2014), Bilston et al (2015), Finansinspektionen (2022)). Fuster et al (2018) discuss the importance of accounting for the vulnerabilities of the riskiest borrowers when assessing potential losses under adverse shocks. Adelino et al (2016) use FICO scores to show that the share of delinquencies among high-income borrowers increased particularly sharply after the Great Financial Crisis (GFC) in areas that experienced the strongest run-up in house prices before the GFC.

Yet data scarcity has constrained research linking the distribution of borrower vulnerabilities to actual credit losses or financial crises. The lack of long and comprehensive historical data series has limited analysis largely to specific episodes or big firms. For example, Greenwood and Hanson (2013) show that, among listed firms, a higher concentration of debt growth in the riskier firms is associated with weaker GDP growth, downside risks to economic activity and a higher probability of financial crises. Brandao-Marques et al (2019) find similar results using a cross-country panel of listed firms. Gourinchas et al (2020) project non-performing loan (NPLs) that could materialise in the wake of the Covid-19 shock using firm-level data on debt in a sample of large and small firms. However, their NPL projections are based not on any historical pattern but on an assumed relationship between stressed cash flows and subsequent defaults.

## Characterising the distribution of borrower vulnerabilities

This special feature uses a unique cross-country data set on borrower-level vulnerabilities at an annual frequency over the period 2006–21. These data, provided by 20 central banks, consist of summary statistics about the distribution of common measures of repayment capacity, separately for the household and NFC sectors.<sup>2</sup> The specific statistics collected were the median, and the 75th and 90th percentiles of the debt-to-income and debt service-to-income distributions across borrowers.<sup>3</sup> To take the 90th percentile as an example, it is equal to a ratio that is exceeded by exactly 10% of the borrowers. Thus, the summary statistics in the data set capture information about the vulnerabilities of the median borrower as well as those in the “tail of concern” from a financial stability perspective.

We complement this data set in two ways. First, we include aggregate – ie country-level debt-to-income and debt service-to-income ratios for the same time period. Second, we also add country-level data on bank NPLs, business bankruptcies and GDP growth.

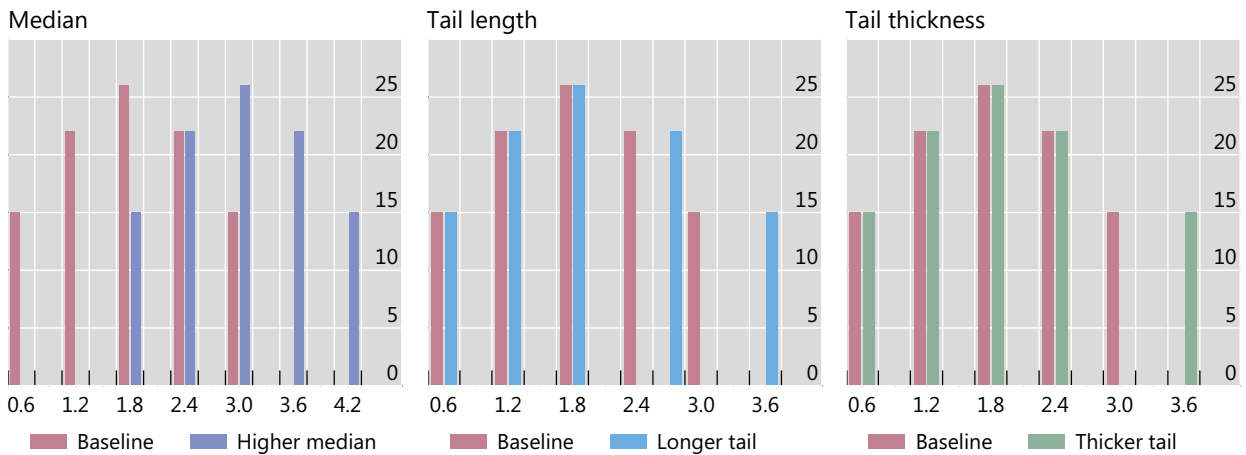
<sup>2</sup> The data set consists of data from central banks in AU, BE, BR, CH, DE, ECB, ES, FR, GB, HK, IN, IT, KR, LU, MX, NL, RU, SA, TH and US.

<sup>3</sup> For the NFC sector, the summary statistics are based on transformations of raw data on the income-to-debt and interest rate coverage ratios.

## Illustrative changes to the tails and centre of the distribution

Density in percent

Graph 1



Source: Authors' calculations.

Our analysis starts by characterising the distribution of borrower vulnerabilities with three metrics. The first captures the repayment capacity of the “average” (or typical) borrower with the median of the distribution. An increase in the median corresponds to a rightwards shift of the whole distribution, all else equal (Graph 1, left-hand panel). We expect this metric to capture information similar to that contained in the corresponding aggregate measure of vulnerability.

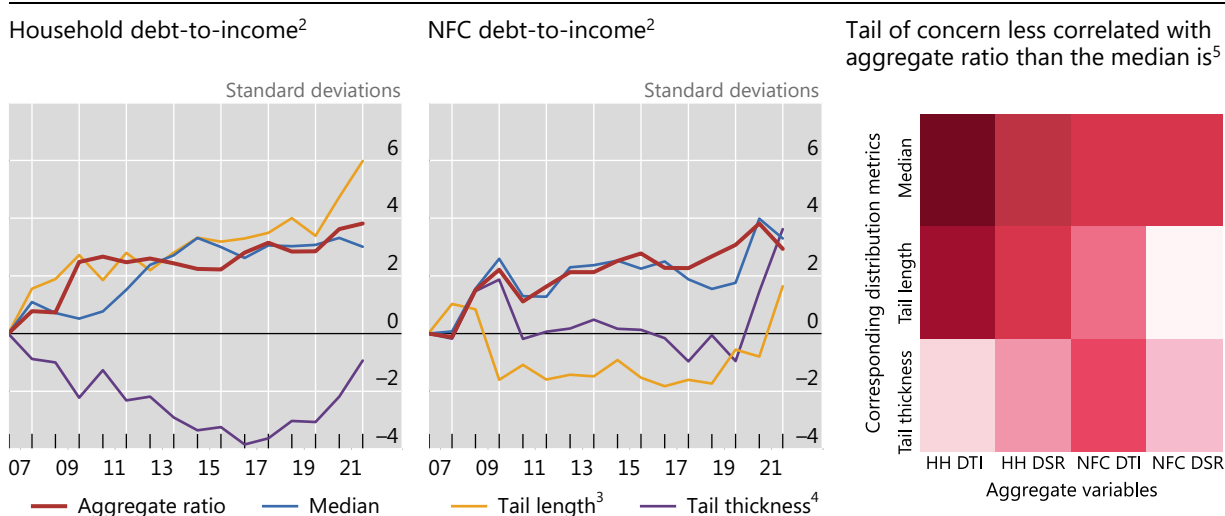
The second metric captures the extent to which the vulnerability of the typical borrower within the tail of concern differs from that of the median borrower. It is equal to the distance between the 75th percentile and the median – the *tail length*. An increase in the tail length indicates that repayment capacity has deteriorated particularly strongly in the tail of concern (Graph 1, centre panel).

The third metric conveys information about the extent to which the riskiest group of borrowers is more vulnerable than the typical borrower in the tail of concern. It is equal to the ratio of the distance between the 90th and 75th percentiles to that between the 75th percentile and the median – the *tail thickness*. An increase in the tail thickness indicates that repayment capacity has deteriorated in particular among the riskiest borrowers (Graph 1, right-hand panel).

## Information in the distribution vs the aggregate

Do the metrics capturing the tail of concern contain information that is different from that in the corresponding aggregate measure? If so, one would expect that the tail will evolve differently from the rest of the distribution.

First, we record that, as expected, the medians tend to co-move closely with aggregate ratios that capture repayment capacities at the country level. For example, averaging across countries, we see that the median household debt-to-income ratio increased steadily alongside the aggregate ratio between 2006 and 2021 (Graph 2, left-hand panel). A similar picture is evident in terms of NFC debt-to-income (centre



HH DTI = household debt-to-income ratio; HH DSR = household debt-service ratio; NFC DTI = non-financial corporate debt-to-income ratio; NFC DSR = non-financial corporate debt-service ratio.

<sup>1</sup> Higher values indicate greater debt vulnerabilities. <sup>2</sup> Average across countries. Each series is standardised to start at zero in 2006 and have a unit standard deviation. <sup>3</sup> Difference between the 75th percentile and the median. <sup>4</sup> Ratio of the difference between the 90th and the 75th percentiles and the difference between the 75th percentile and the median. <sup>5</sup> Darker colours indicate higher correlations. Correlations computed on the pooled sample after subtracting the within-country means of each series.

Sources: CGFS Working Group on Private Sector Debt and Financial Stability; Fitch; national data; authors' calculations.

panel). Correlation coefficients, based on within country variation, point in the same direction (right-hand panel, top row).

By contrast, the time series profiles (Graph 2, left-hand and centre panels) and correlations (right-hand panel) reveal substantial differences between the tail length and thickness, on the one hand, and the aggregate measures on the other. For example, even though the aggregate household debt-to-income ratio rose across countries after the GFC, the corresponding tail thickness decreased as the most vulnerable households repaired their balance sheets more than others and financial intermediaries cut back on their riskier lending. After 2017, the tail thickened again, alongside a general rise in mortgage growth and rising house prices. Turning to the Covid-19 shock, its impact on the most vulnerable NFCs was reflected in a surge in both tail length and thickness, as firms in customer-facing sectors borrowed to fill a cashflow gap (Banerjee et al (2021)). In parallel, aggregate and median NFC vulnerabilities declined, possibly on the back of blanket fiscal support schemes.

## Distribution of borrower vulnerabilities and credit losses

We next investigate the information value of the tail metrics for explaining future credit losses. We do so by estimating the following predictive regressions for the household and NFC sector separately, using one measure of repayment capacity at a time – debt-to-income or debt service-to-income:

$$Outcome_{it+h} = \beta_1^h Tail\ length_{it} + \beta_2^h Tail\ fatness_{it} + \beta_3^h Median_{it} + \beta_4^h Aggregate\ ratio_{it} + \gamma^h X_{it} + \alpha_i + \varepsilon_{it}, \quad h=1,\dots,3 \quad (1)$$

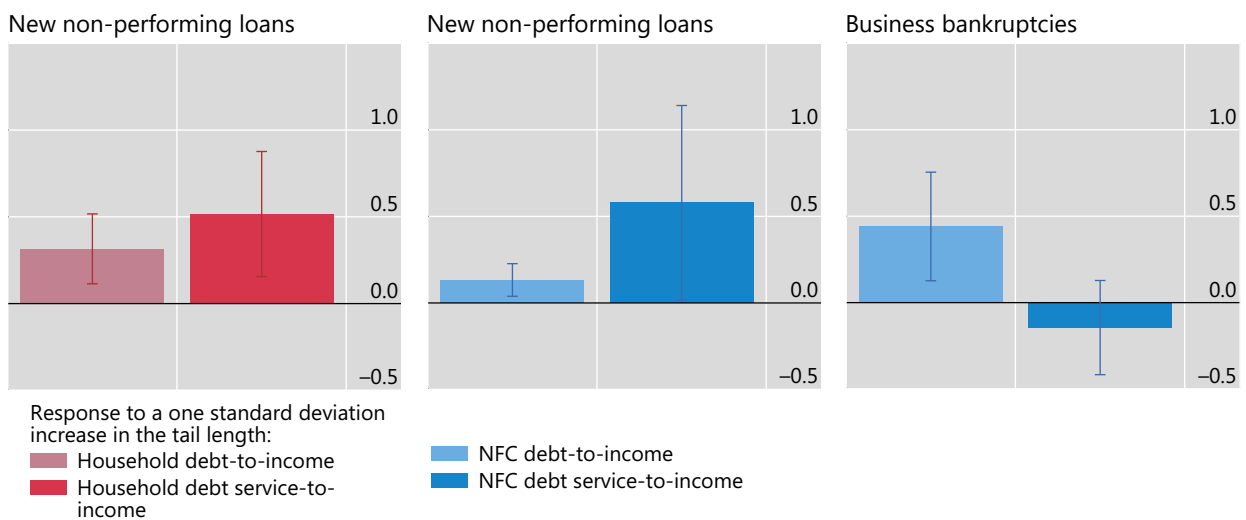
where  $i$ , is the country and  $t$  stands for the year. The longest horizon,  $h$  is equal to three years, a time span that is often used to test the signalling quality of early warning indicators for financial distress.<sup>4</sup> The variable  $Outcome_{it+h}$  for country  $i$  in year  $t+h$  is equal to the change in total new NPLs in the banking sector between period  $t-1$  and  $t+h$  as a share of total loans in year  $t-1$ <sup>5</sup> or the log change in business bankruptcies between  $t+h$  and  $t-1$ . Our key explanatory variables are  $Tail\ length_{it}$ , and  $Tail\ fatness_{it}$ , which seek to establish if information about the tail of concern helps to predict credit losses. We include the corresponding  $Aggregate\ ratio_{it}$  and  $Median_{it}$ , to control for their influence on the outcome variable. In addition, we include controls dated in period  $t$  or earlier. These include real GDP growth in period  $t$  to control for general economic activity, an indicator variable taking on the value of one during 2020 and 2021 to account for the Covid-19 shock, and country fixed effects,  $\alpha_i$ , to control for time-invariant unobserved country characteristics. It is important to note that, due to the short time series of our data, the analysis is in-sample. This means that our estimations capture the co-movement of variables but does not necessary imply that they also forecast credit losses out-of-sample.

Information about the tail length does appear to help explain credit losses that aggregate data and the median metric alone would miss. Our estimates indicate that a one standard deviation increase in the tail length of either the household debt-to-income or the debt service-to-income distribution is associated with a 0.6 to 1 percentage point (0.3 to 0.5 standard deviation) rise in the share of new NPLs in total loans after two to three years (Graph 3, left-hand panel). NPLs increase by between

### Peak response of NPLs and bankruptcies to shifts in the tail length<sup>1</sup>

Standard deviations

Graph 3



<sup>1</sup> Over a three-year horizon. Based on a panel regression of the variable in the panel title on the tail length, tail thickness, median and, aggregate measure of the repayment capacity metric in the legend. Tail length is defined as the difference between the 75th percentile and the median of the distribution. Tail thickness is defined as the ratio of the difference between the 90th and 75th percentiles and the difference between the 75th percentile and the median. Additional control variables include GDP growth and a Covid-19 indicator variable taking on the value of one in 2020 and 2021. Whiskers indicate 90% confidence intervals. Driscoll-Kraay standard errors.

Sources: CGFS Working Group on Private Sector Debt and Financial Stability; Fitch; national data; authors' calculations.

<sup>4</sup> For example, Drehmann and Tsatsaronis (2014).

<sup>5</sup> New NPLs are defined as total NPLs at the end of the period minus total NPLs at the start of the period plus net charge-offs over the period of observation.

0.3 to 1.2 percentage points (0.1 to 0.6 standard deviations) following similar-sized increases in NFC tail length (centre panel), while the annual growth rate in business bankruptcies would go from -8% (sample average) to +16% (a 0.4 standard deviation rise) (right-hand panel).

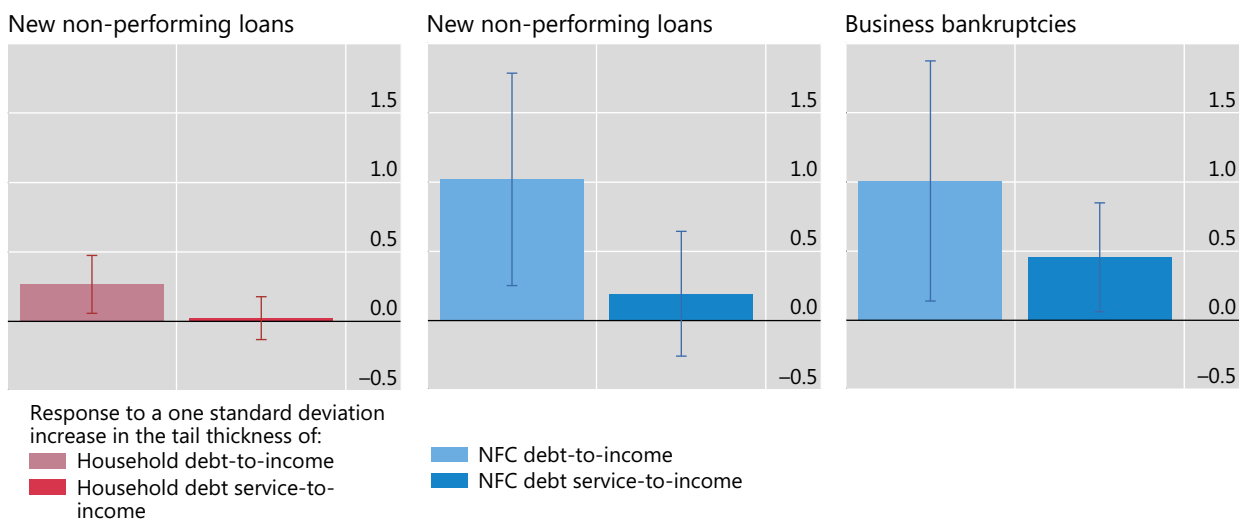
An increase in tail thickness is also associated with a rise in credit losses. For example, a one standard deviation increase in the tail thickness of the household debt-to-income distribution is associated with a 0.5 percentage point (0.3 standard deviation) increase in NPLs (Graph 4, left-hand panel). A similar-sized increase in the tail thickness of the NFC debt-to-income distribution is associated with a one standard deviation increase in NPLs and business bankruptcies (centre and right-hand panels). Turning to the debt service-to-income ratio, an increase in tail thickness has significant predictive power only for business bankruptcies.<sup>6</sup>

Accounting for the tail of concern helps to explain a material portion of the variability in credit losses. For example, including such metrics from the household debt-to-income distribution in regression specification (1) raises the explained portion of new NPL variability (adjusted R<sup>2</sup>) three years in the future by around 8 percentage points. Similarly, including tail metrics stemming from NFC debt-to-income distributions raises the explained portion of new NPL variability by around 6 percentage points. The increase in explanatory power for business bankruptcies rises from around 3% to 13%. Overall, our results suggest that information about borrowers in the tail of concern adds value in explaining future credit losses.

### Peak response of NPLs and bankruptcies to shifts in the tail thickness<sup>1</sup>

Standard deviations

Graph 4



<sup>1</sup> Over a three-year horizon. Based on a panel regression of the variable in the panel title on the tail length, tail thickness, median and, aggregate measure of the repayment capacity metric in the legend. Tail length is defined as the difference between the 75th percentile and the median of the distribution. Tail thickness is defined as the ratio of the difference between the 90th and 75th percentiles and the difference between the 75th percentile and the median. Additional control variables include GDP growth and a Covid-19 indicator variable taking on the value of one during 2020 and 2021. Whiskers indicate 90% confidence intervals. Driscoll-Kraay standard errors.

Sources: CGFS Working Group on Private Sector Debt and Financial Stability; Fitch; national data; authors' calculations.

<sup>6</sup> We ran several additional tests to confirm the robustness of our results. Our broad findings do not materially change if we include lags of the dependent variable as additional control variables in the regression. Similarly, our results are robust to excluding the indicator variable for the Covid-19 shock and either the median or aggregate ratio from our regression specification.

## Takeaways for policymakers

The added informational value of granular data should be set against the costs of collecting it. CGFS (2022) highlights the significant efforts central banks have made to gather such data as well as the related privacy issues. This special feature has underscored that, when it comes to predicting credit losses, it is important to track the repayment capacity of the most vulnerable borrowers. In addition, it has shown that parsimonious summary statistics – which may help to overcome privacy concerns – can complement aggregate measures in explaining credit losses.

Our results should be interpreted with caution because of the short underlying time series. Although the coverage of 20 countries helps to mitigate this deficiency by increasing the number of observations, it is important to note that our sample covers only one financial cycle, with one episode of distress. Future research based on longer time series could shed useful additional light on the value of information about the most vulnerable borrowers for predicting credit losses. Importantly, it should also seek to clarify whether such information improves the real-time (ie “out of sample”) forecasts on which policymakers would need to rely when taking action.



## References

- Adelino, M, A Schoar and F Severino (2016): "Loan originations and defaults in the mortgage crisis: the role of the middle class", *Review of Financial Studies* vol 29, no 7, July, pp 1635–70.
- Anderson, G, P Bunn, A Pugh and A Uluc (2014): "The potential impact of higher interest rates on the household sector: evidence from the 2014 NMG Consulting Survey", Bank of England, *Quarterly Bulletin* (fourth quarter), pp 358–68.
- Banerjee, R, J Noss and J Vidal-Pastor (2021): "Liquidity to solvency: transition cancelled or postponed?", *BIS Bulletin*, no 40.
- Bilston, T, R Johnson and M Read (2015): "Stress-testing the Australian household sector using the HILDA Survey", *Reserve Bank of Australia Research Discussion Paper*, no 2015-01, March.
- Brandao-Marques, L, Q Chen, C Raddatz, J Vandenbussche and P Xie (2019): "The riskiness of credit allocation and financial stability", *IMF Working Papers*, no 2019-207.
- Claessens, S, A Kose and M Terrones (2012): "How do business and financial cycles interact?", *Journal of International Economics*, vol 87, no 1, pp 178–190.
- Committee on the Global Financial System (2022): "Private sector debt and financial stability", *CGFS Papers*, no 67.
- Drehmann, M, C Borio and K Tsatsaronis (2012): "Characterising the financial cycle: don't lose sight of the medium term!", *BIS Working Papers*, no 380.
- Drehmann, M and K Tsatsaronis (2014): "The credit-to-GDP gap and countercyclical capital buffers: questions and answers", *BIS Quarterly Review*, March, pp 55–73.
- Finansinspektionen (2022): "The Swedish Mortgage Market 2015", [www.fi.se/en/published/reports/swedish-mortgage-reports/the-swedish-mortgage-market-2015/](http://www.fi.se/en/published/reports/swedish-mortgage-reports/the-swedish-mortgage-market-2015/)
- Fuster, A, B Guttman-Kenney and A Haughwout (2018): "Tracking and stress-testing U.S. household leverage", Federal Reserve Bank of New York, *Economic Policy Review*, September.
- Gourinchas, P-O, Ş Kalemli-Özcan, V Penciakova and N Sander (2020): "Estimating SME failures in real time: an application to the COVID-19 crisis", *NBER Working Papers*, no 27877.
- Greenwood, R and S Hanson (2013): "Issuer quality and corporate bond returns", *Review of Financial Studies*, vol 26, no 6, pp 1483–525.
- Juselius, M and N Tarashev (2022): "When uncertainty decouples expected and unexpected losses", *BIS Working Papers*, no 995.
- Schularick, M and A Taylor (2012): "Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008", *American Economic Review*, vol 102, no 2, pp 1029–61.

