BANK OF ENGLAND

Staff Working Paper No. 738

Measuring risks to UK financial stability

David Aikman, Jonathan Bridges, Stephen Burgess, Richard Galletly, Iren Levina, Cian O'Neill and Alexandra Varadi

July 2018

Staff Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate. Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Bank of England or to state Bank of England policy. This paper should therefore not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee, Financial Policy Committee or Prudential Regulation Committee.

BANK OF ENGLAND

Staff Working Paper No. 738

Measuring risks to UK financial stability

David Aikman,[1] Jonathan Bridges,[2] Stephen Burgess,[3] Richard Galletly,[4] Iren Levina,[5] Cian O'Neill[6] and Alexandra Varadi[7]

Abstract

We present a framework for measuring the evolution of risks to financial stability over the financial cycle, which we apply to the United Kingdom. We identify 29 indicators of financial stability risk, drawing from the literature on early warning indicators of banking crises. We normalise and aggregate these indicators to produce three composite measures, capturing: *leverage in the private nonfinancial sector*, including the level and growth of household and corporate debt, as well as the United Kingdom's external debt; *asset valuations* in residential and commercial property markets, and in government and corporate bond and equity markets; and *credit terms* facing household and corporate borrowers. We assess these composite measures relative to their historical distributions. And we present preliminary evidence for how they influence downside risks to economic growth and different horizons. The measures provide an intuitive description of the evolution of the financial cycle of the past three decades. And they could lend themselves to simple communication, both with macroprudential policymakers and the wider public.

Key words: Macroprudential policy, financial crises, financial stability, early warning indicators, countercyclical capital buffers, data visualisation.

JEL classification: E44, G01, G10, G28.

(1) Bank of England. Email: david.aikman@bankofengland.co.uk (corresponding author)
(2) Bank of England. Email: jonathan.bridges@bankofengland.co.uk
(3) Bank of England. Email: stephen.burgess@bankofengland.co.uk
(4) Bank of England. Email: richard.galletly@bankofengland.co.uk
(5) Bank of England. Email: iren.levina@bankofengland.co.uk
(6) Bank of England. Email: cian.o'neill@bankofengland.co.uk
(7) Bank of England. Email: alexandra.varadi@bankofengland.co.uk

The views expressed in this paper are those of the authors and should not be taken to represent those of the Bank of England or its policy committees. For helpful discussions and comments, the authors would like to thank Kristina Bluwstein, Alex Brazier, Stephen Cecchetti, Matthieu Chavaz, Dimitri Demekas, Sinem Hacioglu, Anil Kashyap, Christina Segal-Knowles and Rhiannon Sowerbutts. Karam Shergill provided excellent research assistance.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Publications and Design Team, Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 7601 4030 email publications@bankofengland.co.uk

© Bank of England 2018
ISSN 1749-9135 (on-line)

'For historians each event is unique. Economics, however, maintains that forces in society and nature behave in repetitive ways. History is particular; economics is general.'

– Charles Kindleberger (1978)

1. Introduction

The experience of the 2007-8 global financial crisis has underlined the importance of monitoring the build-up of risks in the financial system – risks that can manifest themselves as elevated asset valuations, low volatility, deteriorating underwriting standards, and increasingly fragile balance sheets in the real economy and financial system. The conditions that foster an expansion in the financial cycle are liable to shift abruptly into reverse, tightening financial conditions and reducing economic activity, as financial market participants, borrowers and lenders alike reappraise their appetite for bearing risk (Minsky (1977), Kindleberger (1978)). In many advanced economies, new macroprudential policy regimes have been set up, tasked with monitoring threats to financial stability and with taking preventative policy action where required to shore up the resilience of the financial system.[1]

If these regimes are to be successful, practical tools are needed to inform policymakers of build-ups in vulnerabilities in real time. This paper presents an initial attempt at setting out such a risk monitoring framework. We identify 29 indicators of financial stability risk in the United Kingdom, drawing heavily from the literature on early warning indicators of banking crises.[2] We normalise and aggregate these indicators to produce three composite measures, capturing: (a) leverage in the private nonfinancial sector, including the level and growth of household and corporate debt, as well as the external debt of the United Kingdom, (b) asset valuation pressures in property and financial asset markets, and (c) credit terms facing household and corporate borrowers. We assess these composite measures relative to their historical distributions. And we present preliminary evidence for how they influence downside risks to economic growth at different horizons. The measures could lend themselves to simple communication, both with macroprudential policymakers and the

[1] In the United Kingdom, the Financial Policy Committee (FPC), set up on a statutory basis from April 2013, has a mandate from parliament to 'monitor, assess and take action to address potential systemic risks...with a view to protecting and enhancing the resilience of the U.K.'s financial system'.

[2] These indicators are presented for the purposes of this paper only; they are not an established set agreed upon by the Financial Policy Committee of the Bank of England.

wider public. To the best of our knowledge, this is the first paper to articulate a multivariate risk monitoring framework for the United Kingdom with this purpose.[3]

The approach we advocate provides an intuitive characterisation of U.K. financial cycle dynamics over the past 30 years. These cycles have 3 distinct phases: in the 'boom' phase, asset valuations and private nonfinancial sector leverage are elevated and credit terms are loose by historical standards; there follows a 'bust' phase, during which asset valuations crash and credit terms tighten abruptly; finally, the system goes through a 'repair' phase, where households and corporates gradually de-lever their balance sheets, working off debt overhangs built up in the boom years. There are two such cycles in our dataset. First, the run-up in credit and housing valuations of the late 1980s, which ended with the recession of the early 1990s and with sterling exiting the Exchange Rate Mechanism. Second, the long build-up in risks that began in the late 1990s, punctuated temporarily by the dot-com crash and the decline in equity valuations that followed. Thereafter, financial stability risks increased further: credit and property prices grew rapidly, the current account deficit widened substantially, and investors' compensation for risk across a range of asset classes declined significantly. A policymaker armed with this toolkit would have observed unprecedented elevated readings of our risk measures throughout the great moderation era. When the crisis struck in mid-2007, our measures pick up the large increases in risk premia and tightening in credit conditions that presaged the great recession.

The measures we construct have time series properties that are fundamentally distinct from financial conditions/stress indices (see Hatsius *et al.* (2010) for an overview of these measures). Such metrics are geared towards measuring the cost and ease with which borrowers can obtain credit. By design, they provide a *contemporaneous* read on market functioning, and their typical time profile is one in which realised stresses in the financial system are low and stable in the upswing of the financial cycle, but spike upwards during a stress event. By contrast, our measures are designed to give *advanced warning* to policymakers of build-ups in risk in the financial system, which, if left unchecked, may pose a threat to financial stability. Such risk build-ups often occur against a backdrop of low measured stress levels in the financial system, *i.e.* low readings of financial

[3]Previous contributions include Giese *et al.* (2014), who analysed the behaviour of alternative definitions of the credit-to-GDP gap for the United Kingdom; and, in an early contribution to the literature, Davis (1995), who described a set of potential key leading indicators that could be monitored to inform 'macroprudential surveillance'.

conditions/stress indices. Indeed, there is evidence to suggest that periods of low volatility provide the breeding ground for financial stability risks to develop (Brunnermeier and Sannikov (2014)), as complacency sets in (Gennalioli *et al.* (2015)), and value-at-risk constraints encourage leveraged intermediaries to expand their balance sheets (Adrian and Shin (2014)).

Following a methodology outlined in Adrian *et al.* (2016), we use quantile regressions to estimate the dynamic relationship between our risk measures and quantiles of the U.K. GDP growth distribution (see also Cecchetti (2006) and Cecchetti and Li (2008) for an earlier application of this approach). Our measures have strong predictive power for the left-hand tail of this distribution, while leaving its mode and right-hand tail relatively unchanged. In particular, we find that elevated private nonfinancial sector leverage imparts a strong negative skew to the GDP growth distribution 1-to-4 quarters ahead. At longer horizons, this effect vanishes. Elevated asset valuations, by contrast, create a more persistent fat left-hand tail in the GDP growth distribution, with statistically significant effects even at a horizon of 12 quarters. In quantitative terms, our estimates imply projected growth 8 (12) quarters ahead at the fifth quantile of -1.3% (-1.4%) as of 2002Q4. Given the build-up in leverage and asset prices that occurred, by 2004 this 'GDP-at-risk' had reached -3.7% (-4.1%). And by end-2006, it had reached -4.1% (-4.5%).

We also explore the lead-lag relationship between our three composite measures using vector autoregressions and Granger causality tests. We find that both the asset valuations and credit terms indices Granger-cause (*i.e.* have incremental forecasting power for) private nonfinancial sector leverage. Moreover, impulse responses suggest forecasts for private nonfinancial sector leverage should be revised upwards persistently following positive news in asset valuations or credit terms. This provides a rigorous justification for the inclusion of fast-and-early moving asset valuations and credit terms indicators in our framework.

The focus of our paper is on quantifying the build-up of risks to the U.K.'s financial system driven, in the main, by shifts in the quantity, quality and price of credit to the household and corporate sectors. We do not attempt to quantify the financial system's resilience to absorb such risks through, for example, capital and liquidity buffers. An overall assessment of potential tail risks to growth from financial instability would be conditional on such a resilience assessment. We leave consideration of these issues for future work.

Related literature: This paper builds on the large literature on early warning indicators, which seeks to find empirical regularities in the run-up to financial crises to enable officials and/or private market participants to diagnose building vulnerabilities in advance and take remedial actions. While efforts to detect such signals date back at least to the 1970s (see Bilson (1979)), this literature developed in earnest following the Mexican 'Tequila' crisis and the Asian financial crisis of the mid-1990s.[4]

One of the first attempts to specify and systematically evaluate a set of early warning indicators is by Honohan (1997). This paper considered the antecedents of three types of crises: macroeconomic crises driven by endogenous boom-bust financial cycles; microeconomic crises associated with weak management and fraud at individual banks; and crises in government-permeated banking systems. In a finding that accords remarkably closely with current perspectives on the drivers of financial stability risks, Honohan found that crises with macroeconomic origins tended to be associated with high rates of credit growth, elevated loan-to-deposit ratios, and high levels of foreign borrowing in advance.

Later seminal papers in this literature are by Kaminsky *et al.* (1998) and Kaminsky and Reinhart (1999), who introduced the signal-extraction approach to this literature. Under this approach, prediction thresholds are found that minimise each indicator's noise-to-signal ratio (*i.e.* the ratio of the probability of false alarms to one minus the probability of missing a crisis). The advantage of this approach is its scalability: a large number of potential early warning indicators can be analysed. The authors found that the indicators with the lowest noise-to-signal ratios were the real exchange rate, equity valuations, real interest rates and the money multiplier. Building on this approach, Kaminsky (1999) developed a set of composite early warning indices, each obtained by tallying the number of indicators that have crossed their prediction thresholds at any given time. Out-of-sample crisis probabilities calculated from this framework increased substantially before the 1997 Asian crisis in Thailand, Phillipines and Malaysia, but not in Indonesia. (South Korea was not considered.) We consider a variant of this approach in Appendix C.

Borio and Lowe (2002a) and (2002b) applied the Kaminsky *et al.* (1998) approach to banking crises. To do so, they focused on signals provided by cumulative increases in credit and equity prices,

[4]For surveys of this literature, see Bell and Pain (2000), Eichengreen and Arteta (2000), Demirgüç-Kunt and Detriagiache (2005), Davis and Karim (2008a), Frankel and Saravelos (2012), and Chamon and Crowe (2012).

and analysed whether composite measures outperformed individual indicators viewed in isolation. On an individual basis, the best performing indicator in their analysis was a 'credit gap' variable, defined as the percentage point deviation of the actual credit-to-GDP ratio from its trend, as given by a rolling Hodrick-Prescott filter. This indicator has subsequently received great prominence in the literature, including as an anchor variable for the countercyclical capital buffer (CCyB) (Drehmann *et al.* (2011)). This indicator was found to be particularly well-suited to forecasting crises at a 3-year horizon rather than near-term risks. The best performing composite measure in their analysis was found to weight together credit and asset price gaps, resulting in a noise-to-signal ratio that was almost 50% lower than that obtained for the credit gap alone – a benefit achieved by reducing the number of false positive signals issued.

Other papers in this literature employed more traditional regression-based techniques to examine the covariates of financial crises; papers that take this approach include Demirgüç-Kunt and Detriagiache (1998), Eichengreen *et al.* (1995) and Frankel and Rose (1996) – see also and Berg and Patillo (1999) for an application to currency crises. Demirgüç-Kunt and Detriagiache (2000) used this approach to develop tools for monitoring bank crisis risk. The first tool was an early warning system that issued a signal when the projected crisis probability exceeded a certain threshold, chosen to reflect policymakers' preferences over avoiding false alarms versus missing crises. The second was a rating system for bank fragility, which mapped crisis probability forecasts into different fragility classes. In an out-of-sample forecasting exercise, the authors showed that, while these tools idenfied signs of fragility in Thailand and the Philippines in the run-up to the Asian crisis, they provided a more sanguine picture of risks in other Asian economies.

Measured in terms of their ability to predict the *timing* of crises, these models were probably a failure.[5] This likely explains why interest in this approach began to fade in the mid-2000s. However, as Chamon and Crowe (2012) argue, these papers were on firmer ground identifying *underlying vulnerabilities*, which may persist for a substantial period before a crisis occurs. And this insight also sheds light on why there has been renewed interest in searching for indicators of financial stability

[5] As Demirgüç-Kunt and Detriagiache (2005) put it: 'So far, these tools have met with only limited success, as in-sample prediction accuracy cannot be replicated out-of-sample, a problem common to many areas of economics. One explanation may be that new crises are different from those experienced in the past, so that the coefficients derived from in-sample estimation are of limited use out of sample. Another problem may be that banking crises are rare events, so in-sample estimates are based on relatively few data points.'

risk over the past decade since the global financial crisis.

We survey the insights from this recent literature in terms of the specific indicators it suggests including in a risk monitoring system in Section 3. Here, we briefly review papers that adopt a multivariate approach to identifying build-ups in vulnerabilities for macroprudential policy purposes. The paper closest to ours in terms of methodology and aim is Aikman *et al.* (2016), which presents a set of indicators of financial stability risk in the U.S. financial system; Dattels *et al.* (2010) employ a closely-related approach for the purposes of constructing the IMF's risk monitoring system.[6] We adopt a similar methodology to these papers, which involves identifying indicators in distinct risk categories, assessing each indicator relative to its historical distribution, and aggregating into composite measures.

Aikman *et al.* (2016) demonstrate that a vulnerability index generated by this procedure Granger-causes the credit-to-GDP gap in the United States. Our contributions relative to this paper are as follows: first, we apply a modified version of the approach to assess build-ups of risk in the U.K.'s financial system; second, we show that the composite measures we produce have predictive power for the tails of the U.K. growth distribution at different horizons using techniques introduced by Adrian *et al.* (2016); and third, we explore whether slower-moving indicators of private nonfinancial sector leverage can be forecasted by developments in earlier, faster-moving indicators of asset valuations.

The remainder of this paper is divided into eight sections. Section 2 sets out the simple analytical framework that underpins our risk monitoring approach. Section 3 then reviews key insights from the literature in terms of which indicators contain the most robust information for predicting financial crises. Drawing on this discussion, Section 4 sets out the 29 specific indicators we employ and the criteria that led to their inclusion. In Section 5, we set out our approach to aggregating these indicators into composite measures. As a cross-check on the plausibility of the diagnosis they provide, we use these measures as a lens for assessing key macro-financial developments in the U.K. over the past 30 years. We also consider the performance of our framework in pseudo real time, where composite indices are calculated using only data that were available at the time. Section 6 reports the results of quantile regressions, which link our measures to tails of the distribution of

[6] See also Alessi and Detken (2011), who adopted the signalling approach to analyse the role of real versus financial, and domestic versus global indicators in an early warning system.

U.K. GDP growth. Section 7 examines the dynamic properties of the measures we construct and whether they can be forecasted in advance. Finally, Section 8 concludes. There are three appendices. Appendix A assesses the signalling performance of a subset of the 29 indicators for which we have cross-country time series data available. Appendix B presents time series plots of the 29 indicators used in the analysis. Appendix C presents results from two alternative methods of aggregating information from the underlying indicators: principal components analysis and what we dub an 'intensity-score' metric.

2. Analytical Framework: Risks, Resilience, Tail Risks to Economic Growth

In this section, we introduce a simple analytical framework to describe specifically the concepts used in the paper. Suppose there are two types of shocks that can hit the financial system: exogenous shocks that are orthogonal to indicators of the financial cycle (*e.g.* cyber threats), x; and shocks created within the system (*e.g.* a reappraisal of risk that leads to a large drop in asset prices), e. If s is the set of shocks, then $s_t = \{x_t, e_t\}$.

There are also factors we can identify that would amplify the effect of any given shock that hits the system. Following Adrian *et al.* (2013), call these factors 'vulnerabilities', V. In principle, these can be partitioned into indicators of *risk* facing the system, R, and indicators of the *resilience* of the financial system, K, *i.e.* its capacity to absorb these risks:

$$V_t = V(R_t, K_t) \tag{1}$$

R_t includes indicators such as the level and distribution of debt in the private nonfinancial sector; measures of the debt service burden facing borrowers; asset valuation pressures; and measures of the quality of credit extended by the financial system, including both price and non-price terms. And K_t includes measures of financial intermediary leverage, and maturity transformation, and interconnectedness across the financial system. The derivatives are such that $\frac{\partial V}{\partial R} > 0$; $\frac{\partial V}{\partial K} < 0$.

The defining characteristic of e shocks is that their likelihood depends on previously built-up vulnerabilities: $Pr(e_t) = e(V_{t-1})$, with $\frac{\partial e}{\partial V} > 0$. For example, the probability of an equity market crash will depend, amongst other things, on the overvaluation in the market.

Tail risks to economic growth depend upon s_t and V_t, but also ξ_t, which summarises the potential for shocks unrelated to the financial system to influence growth (*e.g.* the outbreak of war) :

$$VaR_q(GDP_t) = f(s_t, V_t, \xi_t) \tag{2}$$

where $VaR(.)_q$ is the worst possible fall in growth we would expect over the relevant policy horizon at probability q (Jorion (2001)), and the derivatives of f are all negative, *i.e.* greater vulnerabilities make the worst potential fall in GDP even worse. Cecchetti (2006) refers to $VaR_q(GDP)$ as 'GDP-at-Risk'.

In high-level terms, one might characterise the objectives of a macroprudential regime as seeking to ensure GDP-at-Risk – $E_t\{VaR(GDP_{t+h}|\xi = 0)\}$, where h is the relevant horizon – remains close to the desired risk tolerance level while maintaining a central outlook for growth close to trend (an objective it shares with monetary and fiscal policies). The macroprudential regulator might be considered to adjust its policy instruments – bank capital requirements, real economy borrowing limits and so on – to achieve a favourable balance between these objectives. The potential for disturbances to growth from non-financial shocks, ξ_t, are outside its remit.

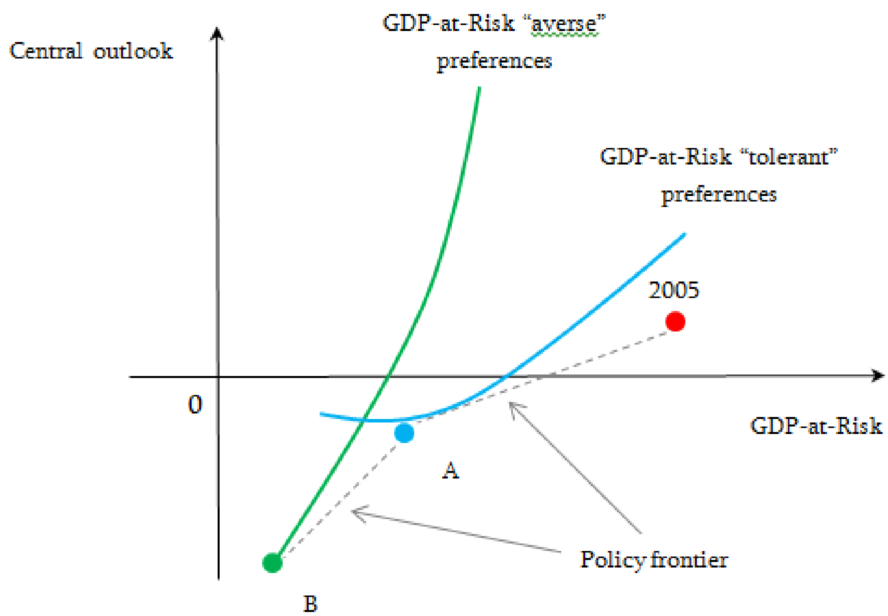
To illustrate this potential trade-off, Figure 1 shows a hypothetical problem in which a policymaker faces elevated GDP-at-Risk, yet a central outlook for the economy that is close to trend.[7] As indicated by the red dot, these are circumstances akin to those prevailing in many economies in the run-up to the global financial crisis (e.g. 2005). The policymaker has a choice of measures available to achieve a more favourable outlook. She can implement policies to enhance the resilience of the financial system (e.g. increase the CCyB). These measures generate a relatively flat trade-off – they reduce GDP-at-Risk at little cost to the central outlook – but can only bring GDP-at-Risk part way to its desired level, to point A in the diagram. Or she can implement policies to curb the debt build-up (e.g. tighten loan to income limits). These measures reduce GDP-at-Risk further, but at greater cost to the central outlook, as illustrated by point B on the chart. The locus between the green-blue-red dots is the policy frontier, which describes the set of achievable outcomes.[8]

The policymaker's preferences over these outcomes are summarised by a set of upward-sloping indifference curves: in order to be willing to tolerate higher GDP-at-risk, the central outlook for

[7]Thibaut Duprey and Alexander Ueberfeldt from the Bank of Canada present a closely related approach. See Duprey and Ueberfeldt (2018), 'How to manage macroeconomic and financial stability risks: a new framework', Bank of Canada Staff Analytical Note.

[8]This policy frontier is illustrative rather than being the output of a model; but it does reflect our priors as to the likely transmission mechanisms of these tools.

Figure 1: *Optimal macroprudential policy*



Notes: This figure illustrates an optimal policy problem using the GDP-at-Risk concept. The red dot is a stylised illustration of the outlook facing many economies in the run-up to the global financial crisis. The locus of points B-A-2005 defines the feasible policy frontier. The green and blue curves are indifference curves for highly risk-averse and less risk-averse policymakers respectively.

the economy must be more favourable. These curves are also convex (*i.e.* the slope is increasing), reflecting the notion that larger tail risks create disproportionately higher losses. Optimal policy is defined by the point of tangency between the indifference curves and the policy frontier. A policymaker who weights the central outlook relatively highly (blue indifference curve) might choose a point somewhere on the flat portion of the policy frontier, between the blue and red dots. By contrast, a GDP-at-Risk averse policymaker (green indifference curve) might choose a point on the more steeply-sloped section of the frontier, e.g. the corner solution, point B.

If macroprudential regimes are to be effective in fulfilling their objectives, a *programme* of research is required to guide appropriate policy responses. Elements of this programme include: (i) identifying a set of indicators, which provide reliable signals of prevailing risks and the resilience of the system (*i.e.* measuring V and e); (ii) generating insights on how to aggregate individual indicators to form an overall view of threats to growth, *i.e.* estimating the f function in equation

(2); (iii) identifying thresholds for these indicators at which vulnerabilities become unacceptable and require a policy response; (iv) forecasting the evolution of threats to growth at different horizons, $E_t\{VaR(GDP_{t+h}|\xi = 0)\}$; and (v) to facilitate cost benefit analysis, estimating policy impact multipliers, which describe both how threats to future growth and the central outlook for the economy, y_t, are likely to be affected by alternative policy measures, p – that is, $\frac{\partial f}{\partial p}$ and $\frac{\partial y}{\partial p}$, *i.e.* the slope of the policy frontier in Figure 1. This paper provides an initial attempt at implementing some aspects of this programme.

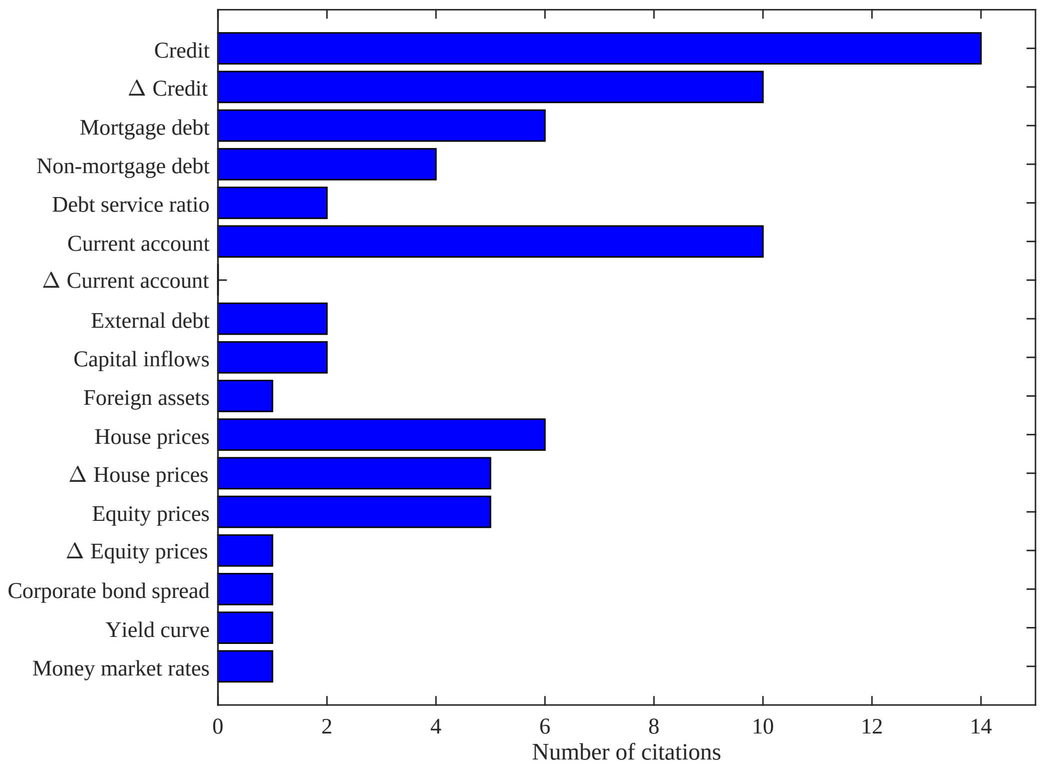
3. Which indicators predict financial crises?

To guide our search for specific indicators to include in our risk monitoring framework, we scanned GoogleScholar for empirical studies of the determinants of banking crises and the severity of the economic contraction that ensues in their aftermath. We filtered out studies that focused exclusively on emerging markets, given the structural differences between such economies and the United Kingdom. This procedure identified 37 distinct studies, most of which employed cross-country panel techniques. We then recorded the indicators that were identified as significant by these studies. Figure 2 presents the results of this meta-analysis.

The most commonly cited indicators are measures of the level and growth rate of credit (cited in 24 studies), the current account (cited in 10 studies), and house prices (cited in 11 studies). Most studies use broad measures of credit, although a number of studies emphasise the importance of run-ups in mortgage debt (cited in 6 studies). The most frequently cited asset price indicator, other than housing prices, are equity valuations (cited in 6 studies). Indicators of the quality of credit being extended, such as loan to value ratios, receive little attention, probably reflecting the limited data availability for such indicators.

Turning to a more detailed discussion of these indicators, perhaps the most robust result in the early warning literature is the importance of measures of credit-based variables as leading indicators of both the likelihood and severity of financial crises. Some of the most illuminating pieces of this research have drawn on evidence from long historical time series and across multiple countries: Schularick and Taylor (2012) report that a persistent one percentage point increase in credit-to-GDP on average across their sample of 14 developed countries from 1870 onwards raises the probability

Figure 2: *Number of citations in the banking crisis literature*



Notes: This chart presents a tally of citations of indicators that have been found to have statistically significant effects on crisis probabilities in the academic literature.

of a financial crisis from 4% to 4.3% per year; Jordá *et al.* (2013) report that it also raises the severity of a crisis, with real GDP per capita almost 1% lower after five years. This echoes and extends findings from earlier and subsequent research by numerous authors.[9] Not all credit booms end in financial crises, however. DellAriccia *et al.* (2012) report that only one-third of credit booms across their sample result in full-blown crises – albeit three-fifths of such episodes are followed by either crises or prolonged periods of subpar economic growth; Gourinchas *et al.* (2001) find that lending booms do not substantially increase the vulnerability of the banking system; and Boyd *et al.* (2000) find that in 10 out of the 21 countries they studied, there was at least one episode in which credit

[9] For research on the relationship between credit growth and financial crisis risk, see Gavin and Hausmann (1995); McKinnon and Pill (1996); Honohan (1997); Eichengreen and Arteta (2000); Bordo *et al.* (2001); Borio and Lowe (2002a) and (2002b); Borio and Lowe (2004); Borio and Drehmann (2009); Drehmann *et al.* (2011); Mendoza and Terrones (2011); Baron and Xiong (2017); Bridges *et al.* (2017). Drehmann and Juelius (2012) argue that the financial constraints associated with private sector indebtedness are well summarised by debt service ratios.

grew at an unusually rapid rate with no crisis occurring in the subsequent three years.

There is some evidence that it is growth in mortgage debt, rather than other forms of credit, which is the key determinant of crisis risk and severity. Mian and Sufi (2010) find that the increase in household leverage in the run-up to the global financial crisis was an important determinant of the severity of the recession across US counties; Mian, Sufi and Verner (2017) extend this analysis to an international context and show that increases in a country's household debt-to-GDP ratio predicts lower growth and higher unemployment in the medium run. Similarly, using a longer-run dataset, Jordá *et al.* (2013) document that growth in mortgage credit has been the key determinant of both crisis risk and crisis severity in the post-WW2 era. Korinek and Simsek (2016) and Bianchi and Mendoza (2016) present theoretical models of the aggregate demand externalities associated with excessive build-ups in household debt.

Turning to other indicators, a number of papers identify the current account deficit as a leading indicator of instability. For example, Catao and Milesi-Ferretti (2013) show that current account deficits are a powerful predictor of balance of payments crises (rather than banking crises) in both advanced and emerging economies. They emphasise the net international investment position as a particular source of vulnerability. But other research finds this indicator to be insignificant conditional on other relevant factors. For instance, Gourinchas and Obstfeld (2012) find that the current account is not a statistically significant predictor of banking crisis once they condition on credit growth. Jordá *et al.* (2013) report a similar finding. Other studies have found a role for the current account as an amplifier in conjunction with other vulnerabilities. For example, Davis *et al.* (2014) introduce an interaction term between the current account deficit and credit growth. And while credit growth is the superior crisis prediction indicator, the marginal effect of credit growth on the probability of crisis is higher at higher values of the current account deficit. Moreover, Whitaker *et al.* (2013) show that crisis severity is increasing in gross external debt – particularly short-term debt such as bank liabilities – even after controlling for domestic credit growth.

Yet another line of research emphasizes the role of inflated asset prices in generating financial market vulnerabilities. Boom-bust cycles in real estate prices, both residential and commercial, are viewed by many economists as key sources of financial fragility (see Cecchetti (2008) and Reinhart and Rogoff (2010)). Jordá *et al.* (2015) argue that synchronised booms in house prices and mortgage credit – which they refer to as 'leveraged bubbles' – create particularly strong vulnerabilities. Others have argued there may be complementary information in bond risk premia (Stein (2013), Mishkin (1990)). According to this view, when risk premia are unusually low, there is a greater probability of an upward spike, which may be associated with significant adverse economic effects. Lopez-Salido *et al.* (2015) find empirical support for this proposition: periods of aggressive risk pricing tend to be followed by increase in credit spreads, which are associated with persistent negative effects on output and employment; Gilchrist and Zakrajsek (2012) and Krishnamurthy and Muir (2017) report similar findings. Others have argued that low volatility may itself spur risk taking, with the potential for a destabilizing unraveling when volatility spikes. Brunnermeier and Sannikov (2014) term this the 'volatility paradox'. Danielsson *et al.* (2016) present cross-country empirical evidence on the relationship between volatility and financial crises, finding that periods of low volatility tend to be followed by banking crises.

4. Selecting risk indicators

In this section, we describe the specific indicators used, and the criteria that led to their inclusion. These indicators have been selected for the purposes of this paper only; they are not an established set agreed upon by the Financial Policy Committee of the Bank of England.

4.1. Selection criteria

Our most important criterion in deciding whether a particular indicator warrants inclusion in a risk assessment framework is whether it is likely to provide actionable, advance warning for policymakers of the build-up of risk in the financial system. Macroprudential policy responses will have implementation lags – for instance, decisions to increase the countercyclical capital buffer (CCyB) do not typically have binding effect on the capital banks are required to maintain for 12 months. Moreover, policymakers often choose to act gradually, which in part reflects uncertainty about the evolving state of the economy, heightening the need for advanced warning.

Given this, an ideal indicator would have the following characteristics. First, it would signal building vulnerabilities with potential threats to financial stability at least two to three years hence. The extensive academic literature surveyed in Section 3 has identified a number of indicators that provide informative signals in advance of past financial crises, and each of the indicators we adopt is drawn from this literature.

Second, an ideal indicator would provide reliable signals of building risks, with few erratic movements. In the jargon, this is known as a high 'signal-to-noise' ratio.[10] In weighting this criterion, we recognise that there is a potential trade-off with the desire for an early warning signal. In particular, indicators that provide a forward-looking perspective of building vulnerabilities, such as asset valuations, tend to also generate 'false positives' – as per Paul Samuelson's aphorism that the stock market has predicted nine out of the last five recessions. We opt to include some volatile indicators in our set to benefit from this forward-looking information. We examine this issue further in section 7.

Third, such an indicator would be available with a long time series, providing us with sufficient information to calibrate what constitutes elevated versus subdued readings. In general, we have good data availability for indicators of indebtedness in the private non-financial sector, external imbalances, and asset valuation pressures. We have granular information too about the terms and conditions on new loans, particularly residential mortgages for owner occupiers. But these data are only available with a relatively short time series (from the mid-2000s). But the U.K.'s financial sector accounts contain very little information about the maturity of debt instruments, and contain only partial information about interconnectedness between different parts of the financial system.[11]

We have sought to choose a broad, representative sample of indicators in our framework, recognising that no single variable will provide a perfect read on building vulnerabilities in all states. In particular, we explicitly aim to span different sources of vulnerability with our choice of indicators, capturing risks from high levels of indebtedness, rapid growth rates of credit and loose underwriting terms, which are indicative of deteriorating credit quality and hence future losses for financial intermediaries. And risks from elevated asset valuations, both in financial and real estate markets, which provide a gauge of financing conditions across the system and investors' appetite to bear risk; assessing valuations relative to fundamentals also helps size the potential shock to the

[10] As a cross-check on the indicators we include in the framework, Appendix A assesses the signalling performance of some commonly-cited macroprudential indicators in a cross-country panel.

[11] There are ambitious plans to improve the U.K.'s sectoral financial accounts over the next five years (see Nolan (2016)). This will provide policymakers with a wealth of additional risk measures, including improved data on which sectors hold outstanding debt claims, and on the maturity of debt claims.

system that would occur in the event of a correction.

We also seek a representative sample of indicators across sectors. We give particular prominence to the household and non-financial corporate sectors both because they represent a large share of total lending in the economy, particularly through their purchases of residential and commercial property, and because they provide a clear link to real activity. We also include measures of external sector vulnerability, because many crises are preceded by growing external leverage such as current account deficits and growing external debt.

Finally, we observe that there exists a trade-off between diversity and parsimony in the dimension of the indicator set. While a larger set of indicators is likely to better capture distinct sources of risk and provide a hedge against data mis-measurement, there are likely to be diminishing returns beyond some point. With too many indicators, the clarity of the overall message from the framework is likely to be reduced – unless the indicators in question are highly correlated, additional variables will tend to offset and reduce the signal provided. There are also practical constraints around communicating results to policymakers with a large number of indicators. We discuss methods for distilling the information contained in individual indicators in section 5, as well as Appendix C.

4.2. List of specific indicators

Table 1 presents the complete list of the 29 indicators we include in the framework, alongside the length of time series available (most series are available from 1987 onwards), the publication lag with which each is available, the nature of any transformations applied, and the source used.

The indicators have been grouped into three categories comprising: *private nonfinancial sector leverage*, which captures the scale of borrowing by households and companies, as well as external leverage; *asset valuations* in financial and property markets; and *credit terms and conditions*, to capture underwriting standards and credit quality on the flow of new lending. We work with quarterly transformations of all variables, taking averages over the quarter of monthly or daily data where required. Figure 3 presents a heat map showing the evolution of the 29 indicators since 1987; time series plots of raw series are presented in Appendix B.

We choose to smooth some indicators in order to reduce noise where the data are volatile. This reflects our focus on the build-up of vulnerabilities in the financial system, rather than providing a near-term indicator of stress, which is highly responsive to news. In particular, we take a three-year moving average when calculating growth rates of credit and property prices, and deflate both series by the equivalent growth rate of nominal GDP. Bridges *et al.* (2017) find that three-year growth rates of credit are the best predictors of the severity of recessions that follow financial crises. We also take a four-quarter moving average of the current account, gross external debt and gross capital inflow series. The academic literature suggests vulnerabilities in the financial system can build when some indicators are low relative to their historical averages – measures of volatility being the archetypal example. We invert the sign of such indicators in our framework, as indicated in the penultimate column of the table.

Table 1: *Indicators selected for our risk assessment*

Category	Individual series	Start date	Lag	Transformation	Inverted?	Source
Private nonfinancial sector leverage:						
Household	Mortgage credit growth	1988	1 quarter	12 quarter MA, less nGDP	–	ONS
	Consumer credit growth	1987	1 month	12 quarter MA, less nGDP	–	BoE
	Household debt to income ratio	1987	1 quarter	–	–	ONS
	Household debt service ratio	1988	1 quarter	–	–	ONS, BoE
PNFC	CRE credit growth	1988	1 month	12 quarter MA, less nGDP	–	BoE
	Non-CRE PNFC credit growth	1987	1 quarter	12 quarter MA, less nGDP	–	BoE
	Debt to profit ratio	1987	1 quarter	–	–	ONS
	PNFC debt service ratio	1990	1 quarter	–	–	ONS
External leverage	Current account	1987	1 quarter	4 quarter MA , % of nGDP	Yes	ONS
	Net foreign assets	1987	1 quarter	4 quarter MA, % of nGDP	Yes	ONS
	Gross external debt	1987	1 quarter	% of nGDP	–	ONS
	Gross capital inflows	1987	1 quarter	4 quarter MA, % of nGDP	–	ONS
Asset valuations:						
Financial	Term premium (10yr gilt)	1993	–	Quarterly average	Yes	BoE staff calculation
	Realised volatility (FTSE All Share)	1987	–	Quarterly average	Yes	Thomson Reuters
	Price/earnings ratio (FTSE All Share)	1987	–	Quarterly average	–	Thomson Reuters
	Equity risk premium (FTSE All Share)	2000	–	Quarterly average	Yes	Thomson Reuters, Bloomberg, IMF
	Investment grade (£) bond spread	1997	–	Quarterly average	Yes	Thomson Reuters, BAML
	High yield (£) bond spread	1998	–	Quarterly average	Yes	Thomson Reuters, BAML
Property	House price to income ratio	1990	1 quarter	–	–	ONS, DCLG, Halifax, Nationwide
	House price growth	1990	1 month	12 quarter MA, less nGDP	–	Halifax and Nationwide
	Housing risk premium	1987	1 month	–	Yes	BoE staff calculation
	CRE yield	1987	1 quarter	–	Yes	IPD
	CRE price growth	1988	1 month	12 quarter MA, less nGDP	–	IPD
	CRE risk premium	2001	1 month	–	Yes	BoE staff calculation
Terms of credit:						
Residential	Spread on new mortgage lending	1997	1 quarter	Quarterly average	Yes	CML, BoE, FCA
	LTV on new mortgages	2005	1 quarter	–	–	FCA
	LTI on new mortgages	2005	1 quarter	–	–	FCA
Commercial property	Spread on new CRE lending	2002	See notes	–	Yes	De Montfort and BoE
	LTV on new CRE lending	1999	See notes	See notes	–	De Montfort

Notes. 'Lag' refers to the reporting delay for the series. '12 quarter MA' indicates that the series has been smoothed by taking a moving average over the previous 12 quarters; 'less nGDP' indicates that we have subtracted nominal GDP growth over the equialent period. For some series, theory suggests lower values generate more risk; we multiply these series by -1, as indicated by the penultimate column.

The indicators span a range vulnerability types and sectors. For example, 12 of the 29 are drawn from the private nonfinancial sector leverage category; 12 are drawn from the asset valuations category; and 5 are drawn from the terms and conditions category. Roughly two-thirds of the list is comprised of indicators of credit, property prices and external imbalances – these are indicators that feature prominently in the early warning literature, surveyed in section 3. 10 of the 29 indicators reflect developments in the household sector and housing market, while 9 reflect the risks in the (nonfinancial) corporate sector, including commercial real estate. Bank credit features explicitly in 6 indicators; bond finance features in 5 indicators; and equity market finance features in 3 indicators.

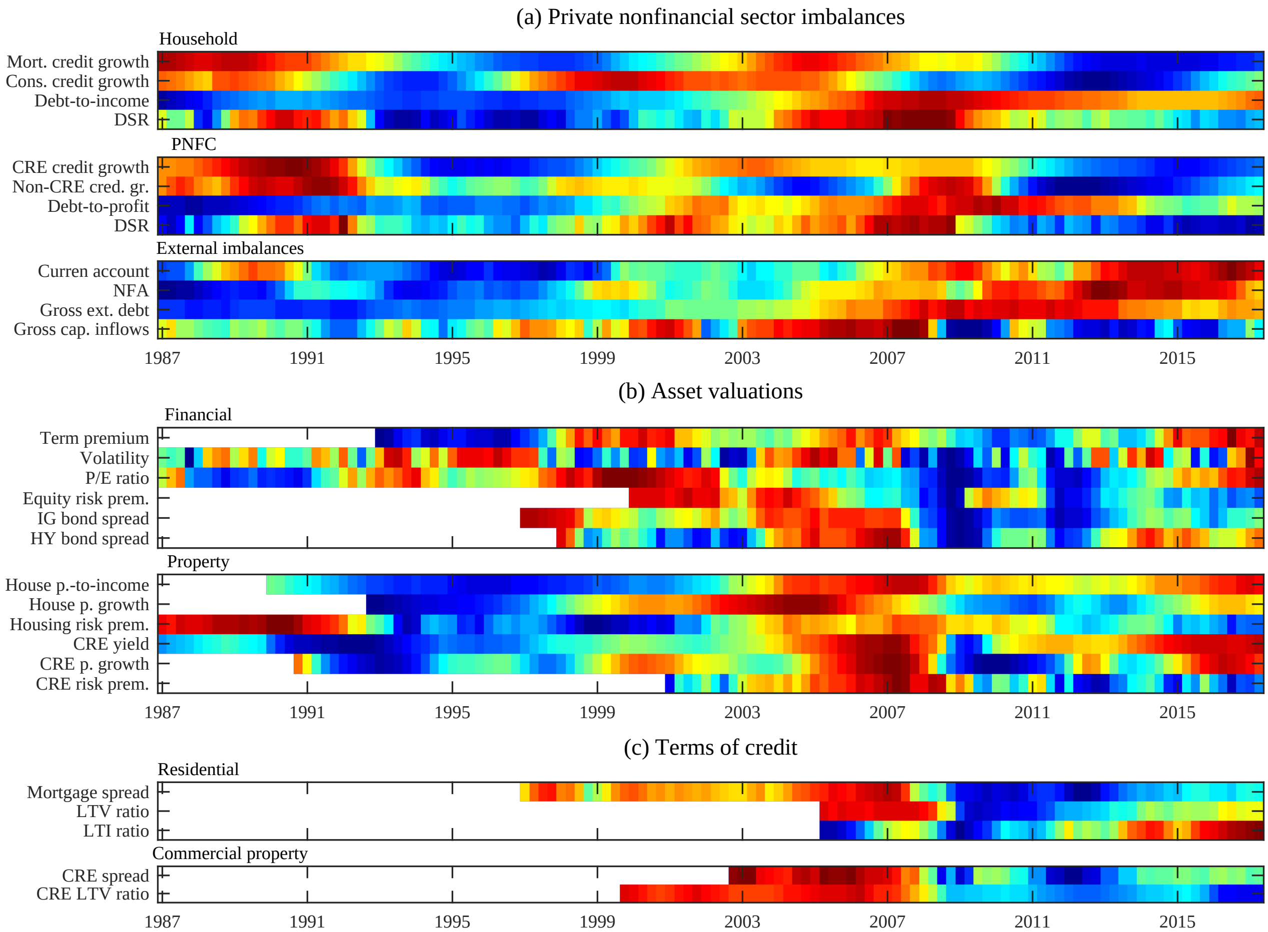
4.3. Discussion

We include three distinct measures of leverage in both the household and corporate sectors. These capture the *level* of indebtedness (*e.g.* the household debt to income ratio), the *growth rate* of indebtedness (*e.g.* mortgage credit growth), and the *interest burden* of debt (*e.g.* the household debt service ratio).[12] We do so because these measures capture distinct economic channels via which indebtedness can threaten financial stability. Episodes of rapid credit growth are often indicative of deteriorating lending standards and an increase in the financial system's exposure to less credit-worthy borrowers. High debt levels may amplify the impact of adverse shocks via debt-deleveraging mechanisms and associated aggregate demand externalities. And debt service burdens provide a gauge on the sustainability of the level of credit and its sensitivity to interest rate or income shocks.

Much work on the leading indicator properties of credit has focused on the credit-to-GDP gap, a measure of excess credit that compares the current credit-to-GDP ratio with an estimate of its longer-term trend (Borio and Lowe (2002)). We exclude this measure from our list for two reasons. First, in our experience, following a period of rapid unsustainable credit growth, the statistical filter employed to calculate the trend can provide a misleading impression of the risk level. This is a particular concern in the post-crisis environment in the United Kingdom, where the credit gap is currently extremely negative (-16.5% in 2017Q2), suggesting highly subdued risks, despite credit

[12] At face value, the inclusion of nonstationary variables in this framework may seem to present a problem in terms of linking these indicators to some concept of the 'financial cycle'. However, we prefer to work with the variables in Table 1 directly, rather than to work with differenced data, because the trends are likely to carry important economic information. For example, the fact that the level of household debt as a share of income has generally been growing in the UK over the past thirty years is likely to imply that risks have been rising, and so it is not obvious we would want to transform the data further to make such levels series stationary.

Figure 3: *Heat map of the 29 individual risk indicators, 1987-present*



Notes: The figure presents a heat map of the 29 individual indicators presented in Table 1. Each series has been rescaled into probability-space by using the kernel density estimate of its cumulative distribution function. Reds indicate readings in excess of the 80th percentile; blues indicate readings below the 20th percentile.

having grown broadly in line with nominal GDP for 3 years. Moreover, in the episode prior to the crisis, where the credit gap performs strongly as an early warning indicator, it is also highly correlated with the rolling three-year growth rates of credit by sector that we employ here (its correlation with mortgage credit growth is 0.88; with consumer credit growth is 0.58; with CRE credit growth is 0.80; and with non-CRE corporate credit growth is 0.74). We discuss how the risk indices introduced in this paper compare to the credit-to-GDP gap in more detail in section 5.

We also employ multiple measures of valuation pressures in financial and property markets. These capture the asset's *current yield*, *i.e.* current earnings relative to its price (in equity markets, this is given by the earnings-price ratio), its implied *risk premium*, defined as the discount rate in excess of risk-free returns required to explain current prices given expected future earnings (in equity markets, this is given by the equity risk premium), and the *growth rate* of prices. We do so for robustness reasons. While these measures will often be highly correlated, that is not always the case. For example, at present yields on various assets are low by historical standards, implying elevated risks of a correction. But risk premia are high, implying subdued risks. Which perspective is correct is likely to depend, amongst other factors, on the reason why long-term risk-free interest rates are so low at present. For instance, if low long-term interest rates reflects pessimistic expectations of growth over this horizon, then risks from asset valuations may be greater than implied by risk premium-based measures alone.[13]

5. Aggregation of indicators into composite measures

While the information contained in the 29 indicators presented above is valuable in its own right for risk monitoring purposes, composite measures provide a sharper signal of emerging trends across the indicators. They are useful too for establishing some basic statistical characteristics of the data, such as the leading indicator properties of some categories of indicator. In what follows we therefore aggregate the information into three summary composite measures capturing (a) nonfinancial private sector indebtedness, (b) asset valuations, and (c) terms and conditions on new credit. In this section, we first explain how we standardise and aggregate the indicators into these three categories. As a cross-check on the plausibility of the diagnosis provided by these measures,

[13] Kiley (2004) emphasises the need to view earnings forecasts, discount rates and asset valuations as being jointly determined in general equilibrium.

we then use them as a lens through which to view developments in U.K. macro-financial history over the past three decades.

5.1. Normalisation and aggregation

Our aggregation procedure follows Aikman *et al.* (2017) and involves the following steps. First, we calculate a Z-score for each individual series by subtracting the sample mean of the indicator and dividing by its sample standard deviation, where both mean and standard deviation are calculated over the full sample. We consider a pseudo real time measure at the end of this section, which normalises the period t value of an indicator using only data up to and including period t.

Second, we combine small sets of related indicators by each sector and market by taking their unweighted arithmetic average. This generates a set of 8 *component indices*, capturing: (1) household indebtedness, (2) corporate indebtedness, (3) external leverage, (4) housing valuations, (5) commercial real estate valuations, (6) financial market valuations and risk appetite (capturing government bonds, corporate bonds, equities and volatility), (7) household credit terms, and (8) corporate credit terms. For example, the household indebtedness component is obtained by combining the normalised household debt to income ratio, income gearing, mortgage and consumer credit growth series. Table 1 describes the mapping from indicators to component indices. The weight of each indicator is inversely proportional to the number of indicators included in each component.

Third, we combine the components into 3 *composite measures* capturing: private nonfinancial sector indebtedness (given by the unweighted average of the household and corporate indebtedness and external imbalance components); asset valuations (which combines the housing, commercial real estate and financial market valuation components); and terms and conditions (which combines the household and corporate credit terms components).

We do not make any exaggerated claims about this weighting scheme. It is deliberately simple and transparent – characteristics, we argue, which are desirable both from communications and robustness perspectives.[14] But we demonstrate in what follows that the resulting aggregates have

[14] We argue that our approach of weighting indicators equally is well suited to the problem at hand: uncertainty, in the Knightian sense, is high, and we have few crisis observations available with which to estimate predictive weights. The risk of overfitting (e.g. attaching excessive weight to indicators that happened to predict the last crisis) is therefore high. There is evidence from the forecasting literature that simple average weights can outperform Bayesian model averaging in environments such as these (Graefe *et al.* (2014) and Green and Armstrong (2015).

appealing features both in their narrative and their empirical properties. We explore alternative weighting schemes in Appendix C, including those based on the correlation structure of the indicators (principal components analysis (PCA)). In headline terms, we find that these approaches deliver similar results to the simpler weighting scheme described in this section.

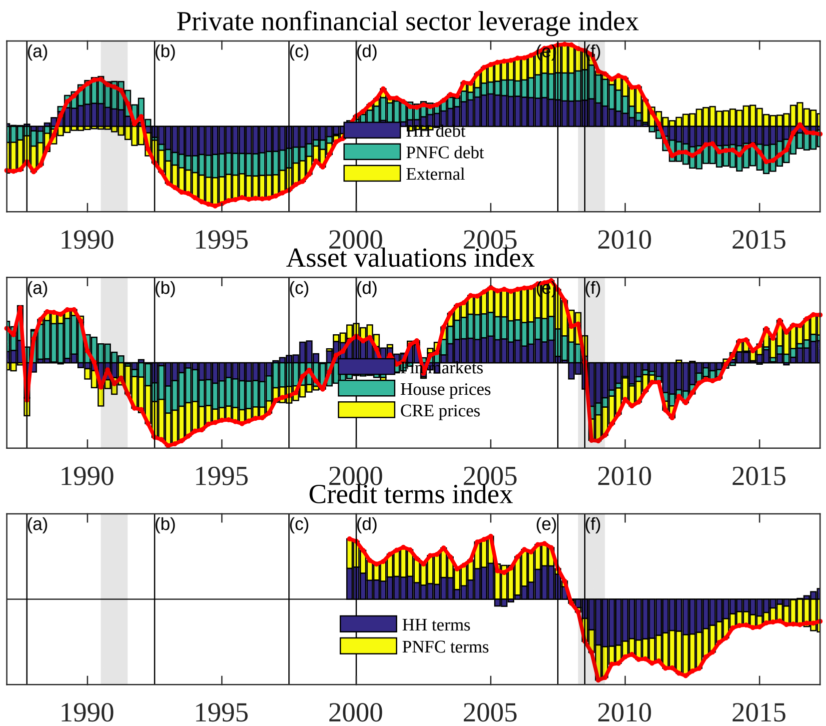
5.2. Historical evolution of risk measures

Figure 4 presents the three composite measures for private nonfinancial sector indebtedness, asset valuations and terms of credit over the period 1987 Q1 to 2016 Q4, alongside the contributions of the component indices underlying each category. What insights do these risk measures provide as to the ebbs and flows of the U.K.'s financial cycle over the past 30 years?

The episode begins with the 'Lawson boom' of the late 1980s. Notwithstanding the sharp fall in equity markets in October 1987 ('Black Monday'), this was a period during which credit and asset prices grew rapidly, presaging the recession of the early 1990s. Our framework captures the building financial vulnerabilities during this period via sharp increases in the private non-financial sector indebtedness index and elevated levels of the asset valuations index. The build-up in risks was broad-based across sectors and markets, with household indebtedness and housing valuations appearing particularly stretched. Both indices then fell sharply in the ensuing recession, with asset valuations reacting more quickly. The asset valuations index declined from its 82nd percentile in the first quarter of 1990 to its 19th percentile by the fourth quarter of that year, reaching a trough at its 3rd percentile in 1993Q2. Private nonfinancial sector indebtedness followed a similar profile, albeit with a 6 quarter lag, declining from its 75th percentile in the second quarter of 1991 to a trough at its 9th percentile in the fourth quarter of 1994. Thereafter, the level of vulnerability implied by both indices remained subdued through most of the 1990s.

Risks began to build once more in the late 1990s, however, with synchronised increases in private nonfinancial sector indebtedness and asset valuations. Both indices reached their 50th percentiles in 1999H1 and continued to increase thereafter. The increase in risks was particularly pronounced in the equity market; equity valuations had been recognised as being elevated since 1996 at least, when then-Chairman of the Federal Reserve Board, Alan Greenspan, delivered his infamous 'irrational exuberance' speech, four years ahead of the eventual correction. Some commentators at the time

Figure 4: *Private nonfinancial sector leverage, asset valuation and terms of credit indices, 1987-present*



Notes: This figure presents the three composite indices for private nonfinancial sector leverage, asset valuations and credit terms, as defined in section 5 (red lines), alongside the contributions of the underlying composite measures, as labelled in the chart. We have mapped each index into the $(0, 1)$ interval based on its percentile in its historical distribution, which in turn is estimated via a non-parametric kernel estimator. Contributions are defined in terms of deviations from the median value of the series. The vertical lines are: (a) the October 1987 'Black Monday' crash in global stock markets; (b) the September 1992 withdrawal of the £sterling from the Exchange Rate Mechanism; (c) the July 1997 beginning of the Asian financial crisis; (d) the March 2000 peak of the NASDAQ index during the 'dotcom' bubble; (e) the September 2007 run on Northern Rock; and (f) the October 2008 collapse of Lehman Brothers. Shaded areas designate recessions in the United Kingdom.

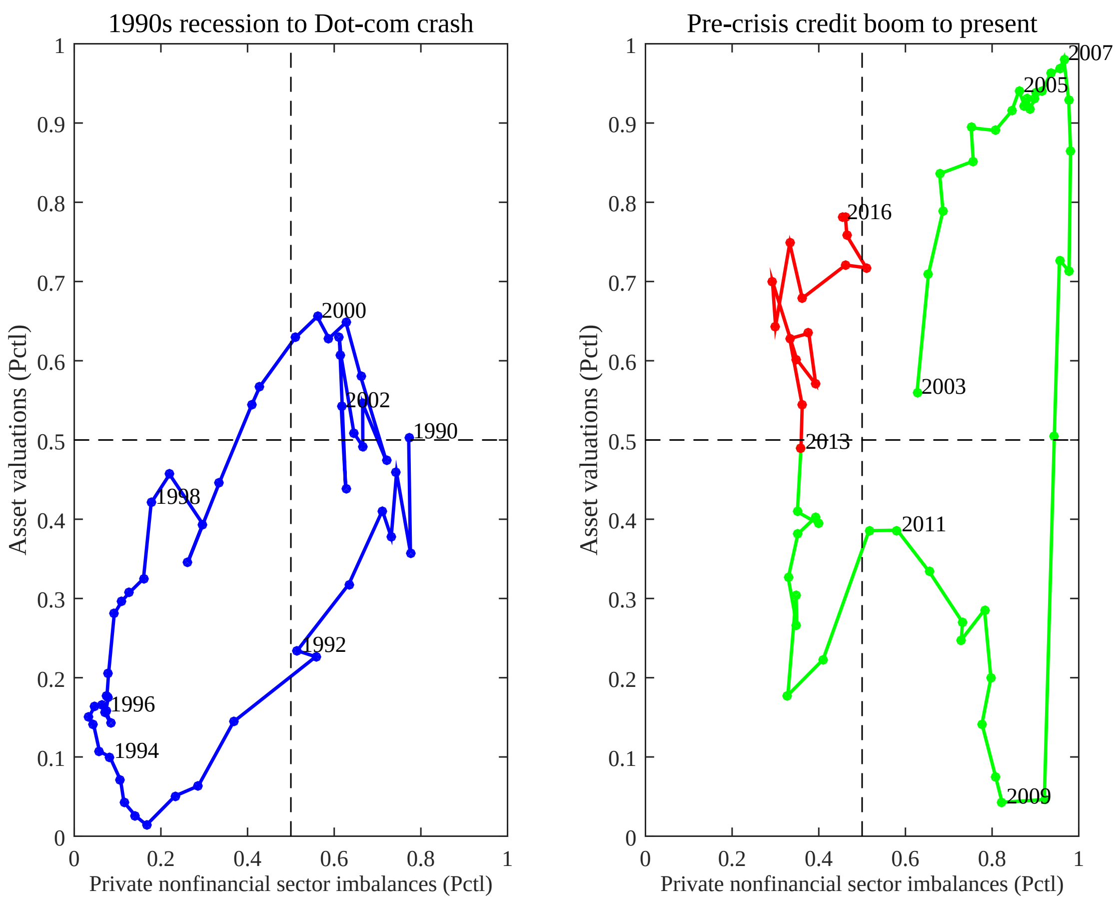
emphasised the role of falling real interest rates in driving equity valuations (see, *e.g.*, Brealey and Vila (1998)), alongside rising earnings expectations. Others were concerned that the market was being driven by unsustainable bubble dynamics, reflecting unrealistic earnings expectations and herd behaviour (Shiller (2000)). Asset valuations peaked locally in the first quarter of 2000, at the height of the 'dot-com' bubble in equity markets; equity markets in the United Kingdom then fell persistently. The decline was relatively contained, however, and led to only a modest decline in our valuations index. Thereafter, valuations increased strongly, pushing our index to unprecedented levels, driven by real estate markets.

This provides the initial backdrop to the further build-up in financial system vulnerabilities that characterised the early-to-mid 2000s. Our framework identifies the rapid expansion in household and corporate sector balance sheets that occurred over this period, alongside increasingly stretched asset valuations, particularly in housing and commercial real estate markets, which in turn facilitated further increases in credit via higher collateral values. And the growth of the imbalances between the United Kingdom and the rest of the world. By early 2005, 10 quarters before the first signs of the crisis became apparent, all three component measures were in the upper quartile of their historical distributions, implying a broad-based build-up in risk. Evidently, the 'Great Moderation' observed in macroeconomic variables during the period (Stock and Watson (2003)) coincided with an unprecedented build-up of financial risks, culminating in the 2007-9 global financial crisis.

When the crisis struck, asset valuations plummeted, reflecting large declines in financial market and property prices, with the valuations index falling from its 97th percentile peak in the second quarter of 2007 to a low of its 4th percentile by the first quarter of 2009. Terms and conditions on credit extended to households and corporates tightened simultaneously, with that index falling from its 82nd percentile to its 2nd percentile over the same period. Private nonfinancial sector leverage took much longer to correct, however. This index continued to increase as the crisis deepened, peaking at its 97th percentile in the first quarter of 2008. Thereafter, it declined gradually, reaching its post-crisis low (37th percentile) only in the fourth quarter of 2011.

Figure 5 presents this information a different way, by plotting the private nonfinancial sector leverage index alongside the asset valuations index over the two distinct financial cycles over this period: the cycle that began during the recession of the early 1990s, ending with the collapse of the dot-com bubble a decade later; and the cycle that encompassed the pre-crisis credit boom, the global financial crisis, and the post-crisis repair phase running to the present. Both cycles have similar time signatures. The boom phase is characterised by observations in the north-east quadrant of the figure, where asset valuations and private sector leverage are both elevated. When the crash comes, observations tend to be concentrated in the south-east quadrant, reflecting the fact that asset prices adjust more rapidly than private sector leverage. The financial system then enters a balance sheet repair phase, with observations concentrated in the south-west quadrant, as households and companies reduce their leverage and asset prices remain depressed.

Figure 5: *Dynamics of private nonfinancial sector leverage and asset valuations, 1990-present*



Notes: This figure tracks the evolution of the private nonfinancial sector leverage index alongside the asset valuations index through two financial cycles. The chart on the left covers the period 1990Q3 to 2002Q4; the chart on the right the period 2003Q1 to 2016Q4.

The observations marked in red show the most recent values of our risk metrics. The combination of increasingly elevated asset valuations with no apparent broad-based pick-up in private nonfinancial sector leverage at this stage of the cycle is historically unusual. This could reflect the impact of low long-term interest rates in boosting asset prices, combined with the scale of deleveraging required in the private nonfinancial sector following the crisis.

5.3. Comparison of risk indices with the credit-to-GDP gap and real GDP growth

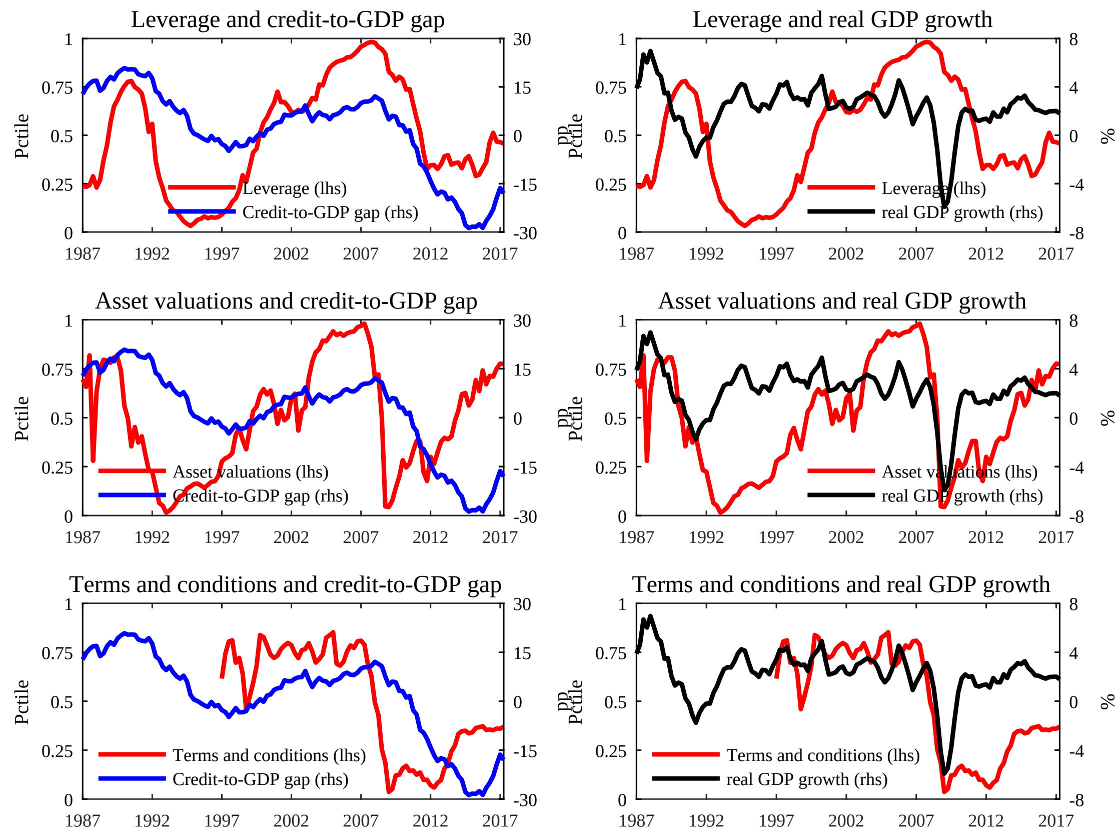
The risk indices introduced in this section have time series properties that are quite distinct from the credit-to-GDP gap, an indicator that has received prominence in the early warning literature and which has a role in Basel III as a guide variable for the CCyB (Basel Committee (2010)). As Figure 6 shows, the U.K.'s credit-to-GDP gap was highly elevated in the late 1980s and early 1990s. Following a prolonged decline, it then became moderately elevated again in the run-up to the global financial crisis. Since the crisis, it has been highly subdued. By contrast, our risk indices reach their peaks in advance of the global financial crisis and have recovered to a larger extent since the crisis than has the credit-to-GDP gap. Over the full sample, the correlations between the credit-to-GDP gap and our risk indices are moderate to low (private nonfinancial leverage: 0.34; asset valuations: 0.03; terms and conditions: 0.53).

The right-hand panels of Figure 6 compare our risk indices with the time series of real GDP growth. The dynamics of the financial cycle differ markedly from those of the business cycle. The fluctuations in GDP growth from the mid-1990s onwards were small, with the economy growing at trend until the financial crisis struck. Evidently, stable growth was not a sufficient condition for a stable financial system. The correlations between growth and our risk indices is also moderate to low (private nonfinancial sector leverage -0.38; asset valuations 0.35; terms and conditions: 0.58).

5.4. Assessing the performance of our framework in pseudo-real time

Finally, in this section we compare the *ex post* estimates of the risk indices defined in section 5 with pseudo real time estimates, *i.e.* whereby the period t estimate uses data up to and including period

Figure 6: *Risk indices, the credit-to-GDP gap and real GDP growth*



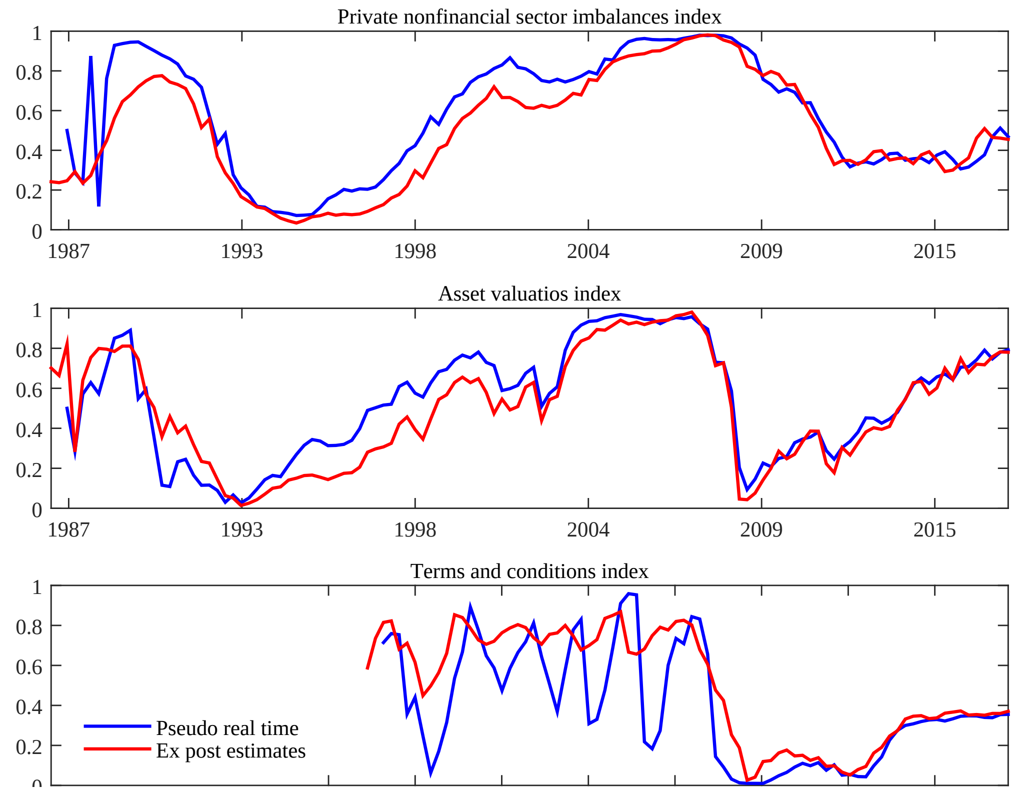
Notes: This figure presents time series of the three risk indices described in this section alongside the credit-to-GDP gap (left-hand panels, blue lines) and annual growth in real GDP (right-hand panels, black lines). The credit-to-GDP gap is the defined as the ratio of total credit extended to the household and non-financial corporate sectors divided by nominal GDP minus its long-term trend. The trend is calculated by applying a two-sided Hodrick-Prescott filter with a smoothing parameter of 400,000.

t, but no subsequent periods.[15] Figure 7 compares pseudo real times estimates with the ex post estimates for the 2017Q2 vintage of data.

For the private nonfinancial sector leverage and asset valuations indices, the real time estimates are more elevated than the *ex post* estimates throughout the mid-1990s through the early 2000s. This reflects the fact that the extreme values of our indices before and after the crisis are not, at this point, part of the sample used to calculate reference points. Thereafter both estimates are extremely

[15] These indices aim to replicate what policy makers might have looked at in real time, taking into account publication lags in the individual series. The label 'pseudo' recognises that these series were not generated at the time, and that some of the underlying data have only been created more recently, *e.g.* granular information on loan to income ratios on new mortgages.

Figure 7: *Pseudo real time estimates of category indices*



Notes: This figure presents the locus of pseudo real time estimates of the three category indices (blue lines) alongside the 2017Q2 vintage of ex post estimates (red lines).

close to one another. The real time estimates for the terms and conditions index are highly volatile throughout the late 1990s and early-to-mid 2000s, reflecting the limited back-run of data available in this category. Overall, these estimates suggest that, even in real time, the monitoring framework introduced in this paper would have signalled heightened vulnerabilities in the period leading up to the global financial crisis.

6. Relationship between composite risk indicators and GDP tail risk

In this section, we assess formally the relationship between our composite indicators and projected tails of the distribution of U.K. GDP growth. As discussed in section 2, in this paper we take this 'GDP at risk' measure as a summary statistic of the ultimate objective of macroprudential policy. Our aim is to understand the extent to which increasing risks to financial stability, as measured via our composite measures, create downside risks for U.K. growth, and if so, over what horizon. To do so, we estimate the conditional distribution of U.K. GDP growth using quantile regressions (Koenker and Bassett (1978)).

In our specification, we model the quantiles of GDP growth, Y, as a linear function of lagged values of a vector of regressors, X, for which we use the private nonfinancial leverage and asset valuations indices defined in section 5, plus a constant. We do not include the terms and conditions index in this analysis given its limited back-run, which would restrict the sample significantly.

The quantile regression estimates a slope coefficient, β, for each specified quantile, τ , according to:

$$\beta_\tau = arg_b minE[\rho_\tau(Y_i - X_i'b)] \tag{3}$$

where $\rho_\tau(u) = 1(u > 0)\tau|u| + 1(u <= 0)(1-\tau)|u|$ is a function that weights positive and negative error terms. For $\tau = 0.5$, this function estimates the conditional median; otherwise it weights positive and negative terms asymmetrically to estimate other quantiles. We estimate this equation over the sample 1987-2016Q2, with lags of 1 quarter, 4 quarters, and 12 quarters in the regressors. We perform the estimation for each decile of the GDP growth distribution ($\tau = 0.1, 0.2$ etc.), but also for quantiles $\tau = 0.05$ and $\tau = 0.95$ in order to obtain more detailed information about the tails of the distribution.

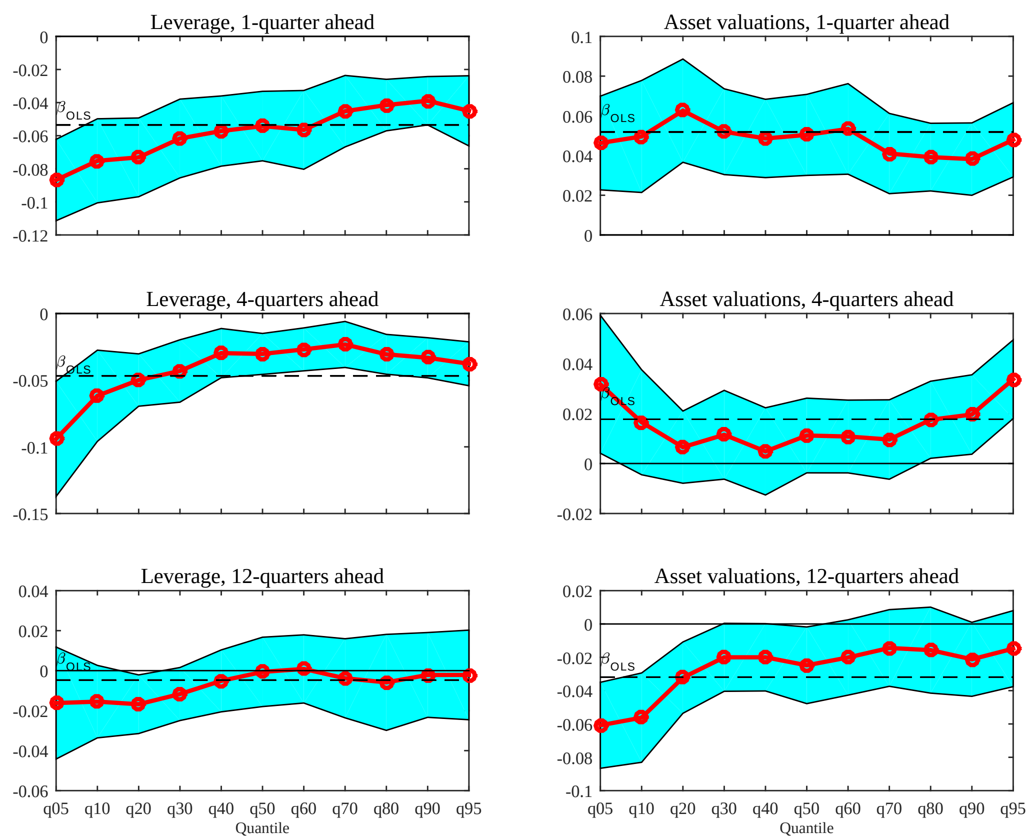
This approach follows the methodology proposed by Adrian *et al.* (2016), who estimate the distribution of real GDP growth for the United States conditional on indices of financial conditions – namely the Federal Reserve Bank of Chicago's National Activity Index (CFNAI) and National Financial Conditions Index (NFCI).[16] These measures are typically thought to contain information about near-term growth prospects. By contrast, we estimate the distribution of U.K. growth conditional on our composite risk measures, which we anticipate contain information about growth prospects at more distant horizons.

Figure 8 presents coefficient estimates at horizons of 1, 4 and 12 quarters, alongside 95% confidence intervals. The figure also shows the equivalent OLS estimate in each case for comparison

[16] The International Monetary Fund's October 2017 *Global Financial Stability Report* presents a similar analysis, conditioning projections of future growth in various countries on a financial conditions index. Cecchetti (2006) and Cecchetti and Li (2008) present earlier applications of this approach.

(dashed line). Overall, the quantile regression estimates are significantly different from zero in many cases. Moreover, they vary in many cases from the OLS estimate, providing justification for the quantile regression approach. As a robustness check, we also ran the same set of regressions with lags of GDP growth included as additional regressors and obtained similar results. The results are available upon request.

Figure 8: *Estimated impact of risk metrics on quantiles of GDP growth*



Notes: This figure presents GDP quantile coefficient estimates obtained from regressions of GDP growth on the leverage and asset valuations indices at horizons of 1, 4, and 12 quarters (red lines). The teal shaded area represents a 95% confidence interval. The dashed horizontal line on each chart is the corresponding OLS coefficient estimate.

Beginning with private nonfinancial sector leverage, we find that, at a horizon of one to four quarters, this index perhaps surprisingly has a negative impact on GDP growth at all quantiles.[17] But

[17]This is surprising because one might expect a relaxation in credit constraints to boost growth in the near term, creating risks further out. We plan to explore this result further in future work, assessing whether it is present in richer cross-country panel datasets and whether it is driven by the inclusion of specific indicators. In particular, some indicators within our leverage composite are 'levels' measures (e.g. the household debt-to-income ratio), which are less likely to boost growth, even in the near term.

this negative impact is largest in the lower tail of the growth distribution. This suggests that, when our private nonfinancial sector leverage measure is elevated, not only is GDP growth more likely to be weak over this horizon, but the distribution is also likely to have a significant negative skew. In quantitative terms, the estimates imply that a 10 percentile change in the private nonfinancial sector leverage index would be expected to lower the median value of annual GDP growth four quarters ahead by around 0.2 percentage points. At the fifth percentile, however, the effect is much larger: projected growth declines by 0.7 percentage points. As documented in section 5, the leverage index increased by more than 40 percentiles between 1999 and 2007. Our quantile regression estimates imply that median projected growth would have declined by around 0.8 percentage points. But the fifth percentile would have declined by around 2.6 percentage points, signifying the emergence of significant downside risks to growth over this period.

When we extend the horizon to twelve quarters, however, the private nonfinancial sector leverage metric is insignificant at almost all quantiles. This is again surprising given other results in the literature, which find that credit booms tend to depress growth in the medium-run (Cecchetti and Kharoubi (2012), Mian, Sufi and Werner (2017); for a contrasting view, see Ranciere *et al.* (2008)). So while this index does appear to provide an informative signal of downside growth risks, by the time it has become elevated it may be too late to take effective policy action. There could be value, therefore, in conditioning policy responses on *forecasts* of this index – we discuss the ability to forecast this index in section 7.

By contrast, we find that the asset valuations index generally has a positive impact on GDP growth at short horizons, with fairly limited variation across the quantiles. At a horizon of 4 quarters out, however, the relationship becomes insignificant at most quantiles, except at the extreme tails of the growth distribution, where the effect is positive. At a horizon of 12 quarters, however, elevated valuations have a significantly negative effect on the lower left-hand tail of the growth distribution - controlling for the effects of private nonfinancial sector leverage. This result provides a rigorous rationale for incorporating information on asset valuation pressures in a risk monitoring framework. We explore this idea further in section 7, where we analyse whether information on asset prices can also help forecast future developments in the broader financial cycle.

With estimates of equation (3) in hand, we next present estimates of the evolution in the distribution of GDP growth over time using the expression:

$$Q_{Y|X}(\tau) = X_t \beta_\tau \tag{4}$$

where X_t is a vector that includes the level of our private nonfinancial sector leverage and asset valuation indices plus a constant, and β_τ is the vector of estimated slope coefficients from Figure 8. For each quarter, this produces 11 discrete estimates of predicted GDP growth, one for each quantile. Figure 9 presents the time series of one-year-ahead conditional distributions of U.K. GDP growth obtained via this procedure. The smooth distribution each quarter is estimated by fitting a kernel density function around the set of discrete quantile estimates generated.

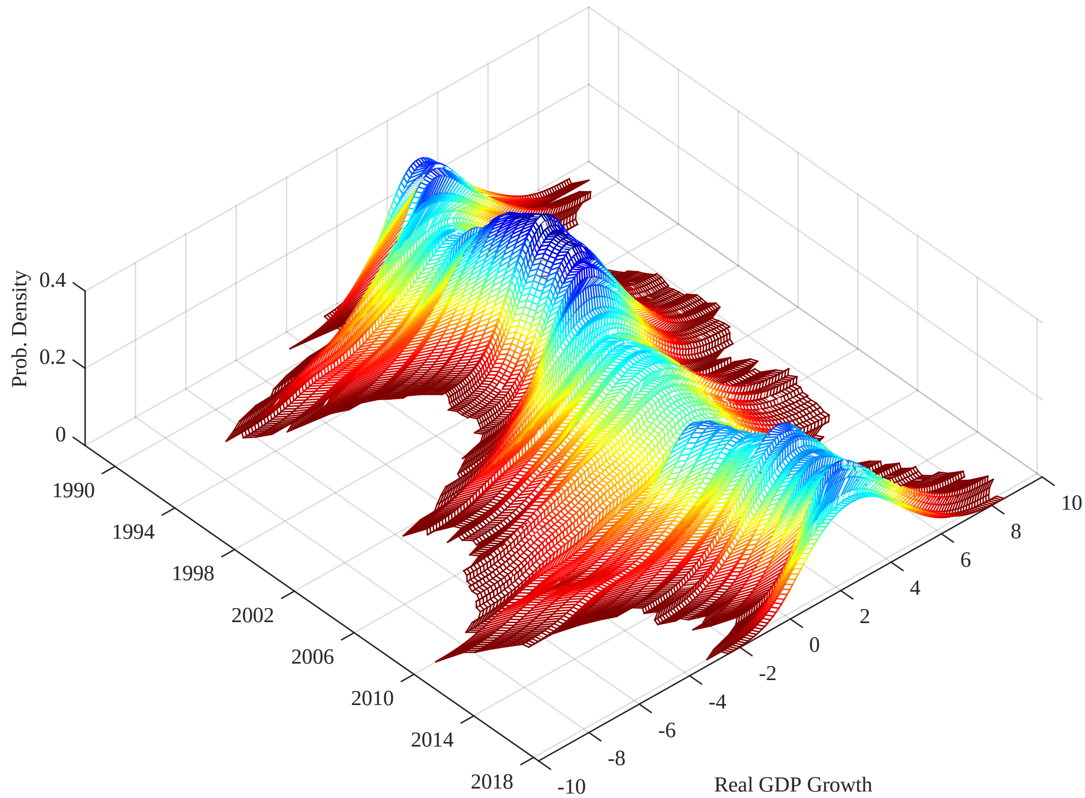
Shifts in the projected growth distribution over time reflect the estimated impact of our private nonfinancial sector leverage and asset valuations indices. The striking feature of the chart is the large variation in the bottom-tail of the growth distribution through time. By contrast, the modal growth projection remains relatively unchanged. The left-hand tail of the projected growth distribution increased significantly over the 2000s up, such that by 2007, on the eve of the global financial crisis, it had reached extremely large proportions. Once the crisis broke out, and these tail risks crystallised, the probability of further large falls in GDP declined significantly. In recent years, the left-hand tail of the growth distribution has widened once more, but remains significantly smaller than in the mid-2000s.

Figure 10 presents 3-year ahead conditional projections out as of 2002Q4, 2004Q4 and 2006Q4, allowing us to examine how the shape of the projected growth distribution changes beyond one-year given the waning influence of leverage at longer horizons, but the increasing impact of asset valuations. Projected growth 8 (12) quarters ahead at the fifth quantile was -1.3% (-1.4%) in 2002Q4. Given the build-up in leverage and asset prices that occurred, by 2004 it had reached -3.7% (-4.1%). And by end-2006, it was -4.1% (-4.5%).

7. Forecasting composite measures to assess future crisis risk

In the previous section, we found that the private nonfinancial sector leverage index has information content for downside growth risks four quarters ahead. But, taking our estimates at face value,

Figure 9: *Time series of one-year ahead projections of the distribution of U.K. GDP growth*

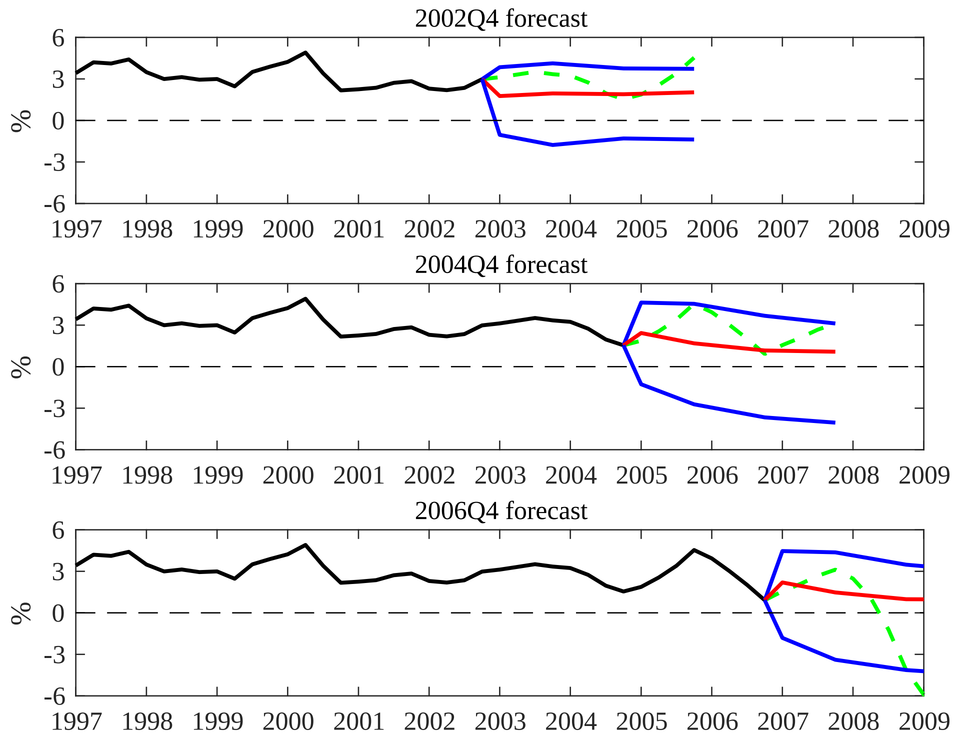


Notes: This figure presents successive one-year ahead forecasts for the density of U.K. GDP growth conditional on the leverage and asset valuations indices. Each date presents the predicted distribution for that quarter, given the information available four quarters previously. The density is produced by first projecting 11 discrete quantiles using the estimates reported in Figure 8, and second, using a kernel density function to estimate the underlying distribution.

it does not have reliable information content for growth risks beyond this horizon. This finding casts some doubt on the utility of such measures for a macroprudential policy maker, given the need for an early warning about building vulnerabilities to inform effective policy responses given implementation and transmission lags. In this section, we examine whether movements in our private nonfinancial sector leverage index can be reliably forecasted to inform the calibration of policy responses.

We begin by examining the lead-lag correlation structure between our composite measures. Figure 11 presents the cross-correlation between the private nonfinancial sector leverage index with leads and lags of asset valuations (blue bars) and credit terms (red bars). The peak of the

Figure 10: *3-year ahead conditional distribution forecasts for U.K. GDP growth*



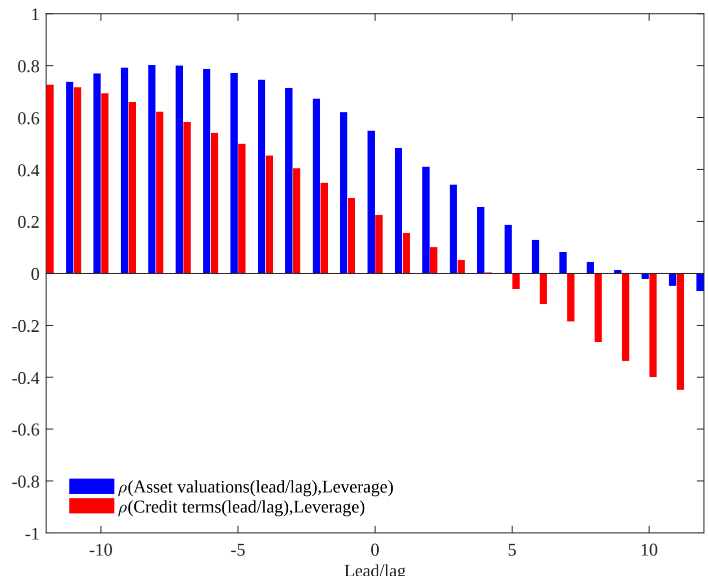
Notes: This figure presents 3-year ahead projections for the distribution of U.K. GDP growth, conditional on the quantile coefficient estimates obtained from estimating equation (3). The black solid line in each panel shows historical data; the red and blue lines show median and 5th and 9th quantile projections respectively; the green dashed line shows subsequent growth outcomes over the forecast period.

cross-correlation function between private nonfinancial sector leverage and asset valuations occurs at a lead of around 8 quarters. Similarly, the peak cross-correlation between credit terms and leverage occurs at a lead of 8-12 quarters. These findings suggest that asset valuations and credit terms have the potential to provide an early steer on future movements in private nonfinancial sector leverage to inform policymaking.

To examine the dynamic relationships between these measures more formally, we estimate a bivariate Bayesian vector autoregression (BVAR) containing the leverage and asset price composite indices described in Section 5.[18] We exclude the credit terms index because of its short historical

[18] We estimate this model using the BEAR 4.0 toolbox developed by Alistair Dieppe, Bjorn van Roye, and Romain Legrand of the European Central Bank.

Figure 11: *Cross-correlation between asset valuations, terms of credit and private nonfinancial sector leverage*



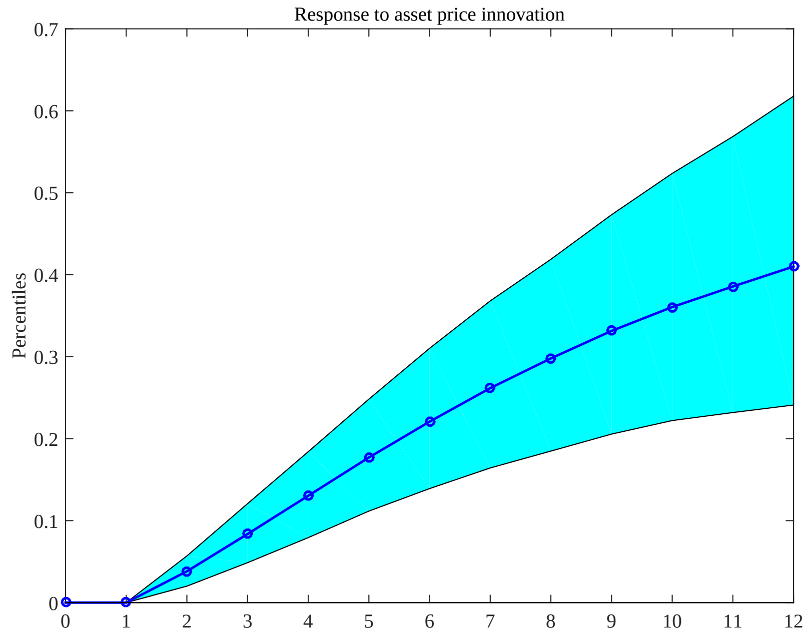
Notes: This figure presents the cross-correlogram between our private nonfinancial sector leverage index with leads and lags of asset valuations (blue bars) and terms of credit (red bars).

time series. We use this model to examine first the response of the nonfinancial sector leverage to innovations in the asset valuations index via generalised impulse responses (Koop *et al.* (1996)). As Figure 12 shows, the leverage index forecast is revised up persistently following data news in asset prices.[19]

Next, we use the BVAR to produce forecasts of the private nonfinancial sector leverage index. To do this, we estimate the model on a dataset up until 2002Q1 and project the system forward 12 quarters. We then expand the estimation window one quarter at a time in real time, generating a set of successive forecast vintages. We calculate the root mean square error (RMSE) of these forecasts at different horizons, and compare the results with those obtained from a random walk model (without drift). Figure 13 presents results from this exercise. Overall, the BVAR has smaller RMSEs - the difference is statistically significant at 4 quarters ahead and 12 quarters ahead horizons using the Diebold-Mariano test.

[19] We have also performed Granger-causality tests on these series. These show that asset valuations Granger-cause private nonfinancial sector leverage - that is, they contain incremental information for forecasting developments in leverage. But not *vice versa*.

Figure 12: *Response of private nonfinancial sector leverage following shocks to asset valuations and terms of credit*



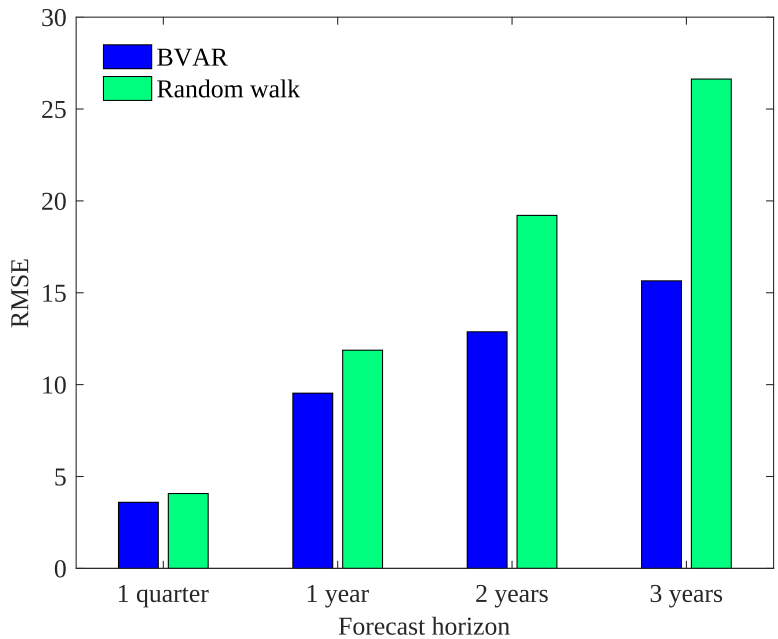
Notes: This figure presents the generalised impulse response of the private nonfinancial leverage index in response to a 1 percentile innovation in asset valuations. The horizontal axis measures quarters. The teal-coloured swathe plots a 95% confidence interval.

8. Conclusion

This paper presents an initial attempt at providing a framework for measuring risks to financial stability, applied to the United Kingdom. We identify 29 indicators of financial stability risk. The choice of indicators draws heavily from the literature on early warning indicators of banking crises, while recognising the need for sectoral diversity, as well as practical considerations such as the need for established time series information.[20] We normalise and aggregate these indicators to produce three composite measures capturing: leverage in the private nonfinancial sector, including the level and growth of household and corporate debt, as well as the external debt of the United Kingdom; asset valuations in residential and commercial property markets, and in government and corporate bond and equity markets, and credit terms facing household and corporate borrowers. We define threshold values for these composite measures based on their historical distributions.

[20]These indicators are presented for the purposes of this paper only. They are not an established set agreed upon by the Financial Policy Committee of the Bank of England.

Figure 13: *Root mean squared errors: leverage forecasts*



Notes: This figure compares the root mean square errors (RMSE) from BVAR-generated forecasts of the private nonfinancial sector leverage index with those obtained from a random walk model. The sample size for estimating the BVAR model is initially set at 1987Q1-2004Q1. We produce a forecast out to 12 quarters ahead, and then incrementally add additional quarters, expanding the estimation window each time. A lower RMSE suggests greater forecast accuracy.

And we present preliminary evidence for how these composite measures influence downside risks to economic growth at different horizons. The measures could lend themselves to simple communication, both with macroprudential policymakers and the wider public.

The analysis could be extended in several directions. First, a natural next step would be to develop a comparable framework for assessing the resilience of the financial system, drawing on indicators of bank leverage, maturity transformation, interconnectedness, and the propensity for fire sales outside the banking system. Second, it would be valuable to explore whether the GDP-at-risk analysis presented is robust to using alternative methodologies and datasets, including a cross-country panel analysis. Third, to inform policy setting, a clear priority is to develop estimates of the impact of alternative policy interventions on GDP-at-Risk. For instance, what impact do higher bank capital requirements have on near-term versus medium term risks, and how does this differ to the impact of tighter monetary policy? Such an analysis would allow for a rigorous cost-benefit assessment of alternative interventions.

Appendix A: Testing the predictive power of a set of macroprudential indicators

In this Appendix, we present an analysis of the signal-to-noise ratios of some commonly-cited macroprudential indicators. We do this for a sub-sample of 16 of the 29 indicators used in this paper for the United Kingdom. This is because a consistent cross-country panel of data for each metric is required to assess predictive power. The analysis is based on a sample of 27 advanced economies, extending from 1960-2015 for most countries. Table A.1 lists the countries in the sample, as well as their credit-to-GDP ratios as a summary metric of their financial development.[21]

Table A.1: *Cross country sample summary statistics*

Country	OECD	Sample length	Number of crises		Credit-to-GDP ratio (%)
			Reinhart-Rogoff (2010)	Laven-Valencia (2013)	
Australia	Yes	1960q1 - 2015q4	1	0	109.5
Austria	Yes	1970q1 - 2015q4	1	1	91.6
Belgium	Yes	1960q1 - 2015q4	1	1	119.3
Canada	Yes	1961q1 - 2015q4	1	0	130.9
Czech Republic	Yes	1995q1 - 2015q4	1	1	77.3
Denmark	Yes	1966q1 - 2015q4	2	1	161.7
Finland	Yes	1960q1 - 2015q4	1	1	120.0
France	Yes	1960q1 - 2015q4	2	1	125.1
Germany	Yes	1991q1 - 2015q4	2	1	100.0
Greece	Yes	1995q1 - 2015q4	2	1	61.6
Hong Kong	No	1973q1 - 2015q4	1	0	162.9
Ireland	Yes	1990q1 - 2015q4	1	1	132.6
Israel	Yes	1970q1 - 2015q4	2	1	110.3
Italy	Yes	1960q1 - 2015q4	1	1	78.8
Japan	Yes	1960q1 - 2015q4	1	1	168.8
Korea	Yes	1970q1 - 2015q4	3	1	107.5
Luxembourg	Yes	1960q1 - 2015q4	1	1	356.8
Netherlands	Yes	1960q1 - 2015q4	1	1	139.5
Norway	Yes	1960q1 - 2015q4	1	1	153.2
New Zealand	Yes	1960q1 - 2015q4	1	0	93.4
Portugal	Yes	1960q1 - 2015q4	1	1	123.6
Singapore	No	1975q1 - 2015q4	1	0	98.1
Spain	Yes	1960q1 - 2015q4	2	1	122.3
Sweden	Yes	1960q1 - 2015q4	1	1	136.9
Switzerland	Yes	1965q1 - 2015q4	1	1	155.7
UK	Yes	1960q1 - 2015q4	5	1	109.4
USA	Yes	1960q1 - 2015q4	2	2	116.6
Total			40	23	

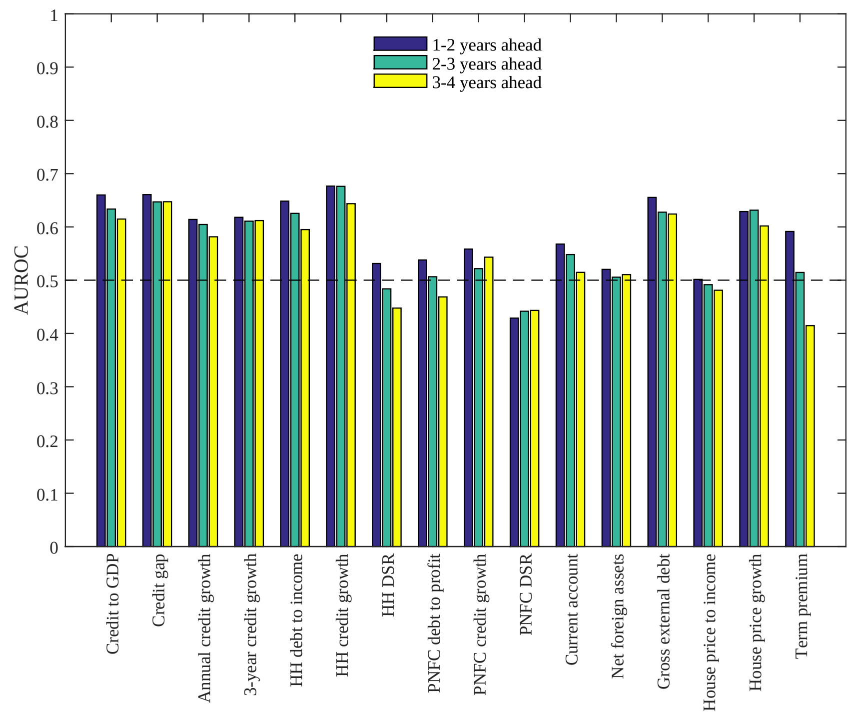
Notes. Credit-to-GDP ratio data are from BIS website for 2015 Q3. The credit series is defined as credit to the non-financial sector.

We measure each indicator's performance using the so-called Receiver Operating Characteristic

[21] There is substantial heterogeneity across these countries in terms of financial liberalisation. For example, Luxembourg's credit-to-GDP ratio reached a maximum of 462%, while the maximum in the Czech Republic was just 92%. While a large dataset such as this is necessary for performing econometric analysis of indicators' signalling properties, the large variation across countries should be taken into account when interpreting the results.

(ROC) curve. This curve plots the proportion of correctly-predicted crises against the proportion of false alarms obtained as we vary the prediction threshold for each indicator. It is common to summarise the predictive power of an indicator by the area under this ROC curve: the so-called AUROC statistic.[22] The AUROC statistic is bound between 0 and 1. A perfect indicator, which calls all crises correctly and produces no falls alarms, would have an AUROC statistic of 1; an indicator whose signals are no better than a coin toss would have an AUROC statistic of 0.5.

Figure A.1: *AUROC statistics for a range of macroprudential indicators*



Notes: This chart presents Area Under the Receiver Operating Characteristic (AUROC) statistics for a range of macroprudential indicators at horizons of 1-2 years, 2-3 years and 3-4 years. We report the average AUROC statistic achieved across both the Reinhart and Rogoff (2010) and Laeven and Valencia (2013) crisis databases. AUROC statistics at or below 0.5 suggest the indicator has prediction value no better than a coin toss.

We consider the ability of indicators to correctly call crises at three different horizons: 1-2 years ahead, 2-3 years ahead, and 3-4 years ahead. We consider crises defined by two studies: Reinhart and Rogoff (2010) and Laeven and Valencia (2013). As Table A.1 shows, the coding of crisis dates

[22]See Drehmann (2013) for a full discussion of the test.

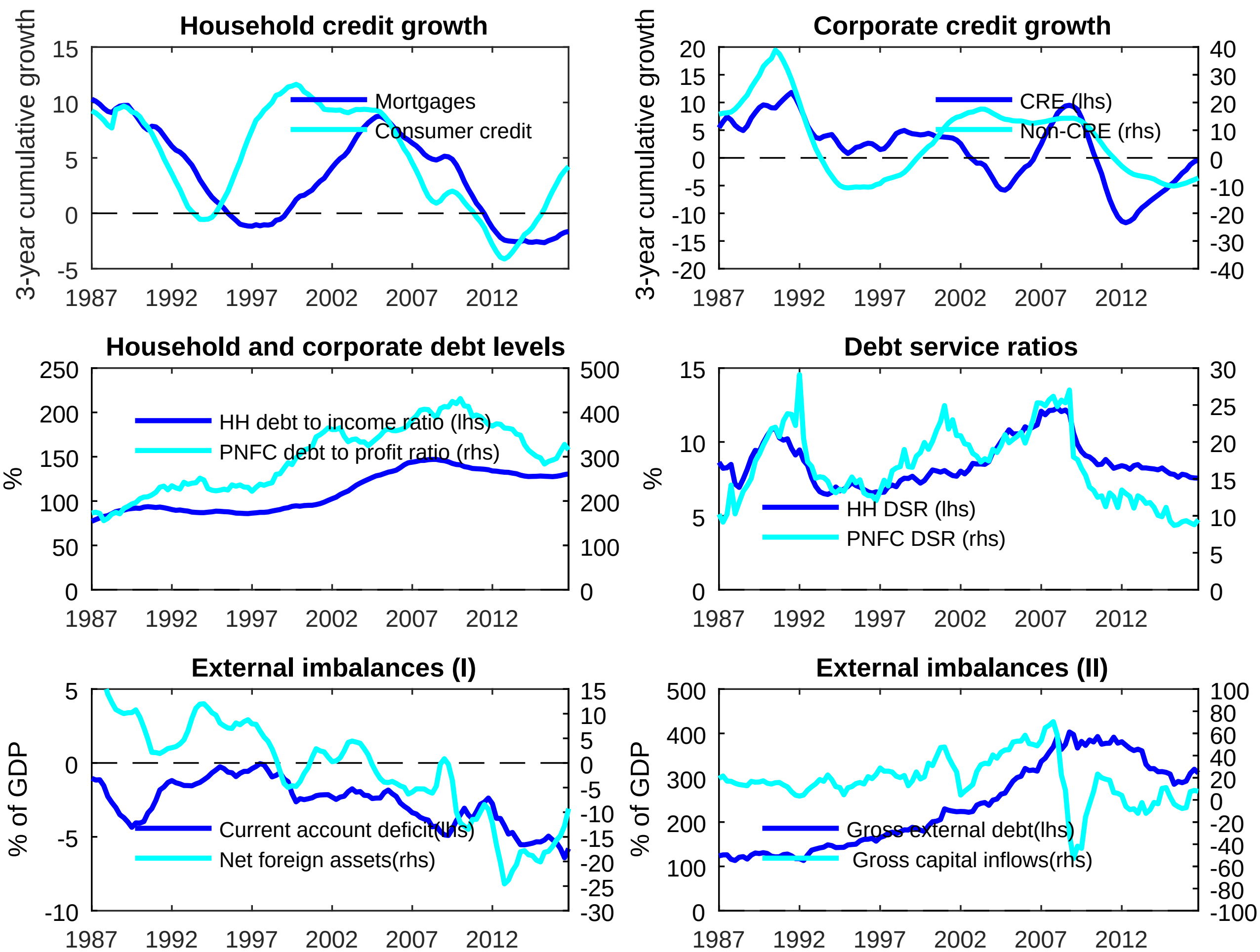
varies between these studies: for example, the Reinhart-Rogoff database identifies 40 banking crises over our sample while the Laeven-Valencia database identifies 23. For robustness, we take the average performance of each indicator across these alternative crisis dates.

Figure A.1 reports AUROC statistics for a range of indicators at the three horizons defined above. In absolute terms, the results suggest that, while these indicators have some information-content as crisis predictors, the signals they provide are quite noisy and come with a large number of false positives. Credit-based variables are the best-performing indicators at all horizons. The level of private non-financial sector credit-to-GDP ratio, its growth rate and the credit gap all score above 0.6. Household credit growth, along with household debt to income and house price growth metrics also perform score relatively highly. Intuitively, the signals of almost all indicators progressively deteriorate as the horizon increases. This is particularly so for the term premium, which has good predictive power about crisis likelihood 1-2 years ahead, but contains no value about risks at more distant horizons. While these results generally accord with those found elsewhere in the literature, our findings on the information-content of the debt-service ratio differ from Drehmann and Juselius (2014), who obtain stronger results for this series.

Appendix B: Time series plots of the 29 indicators

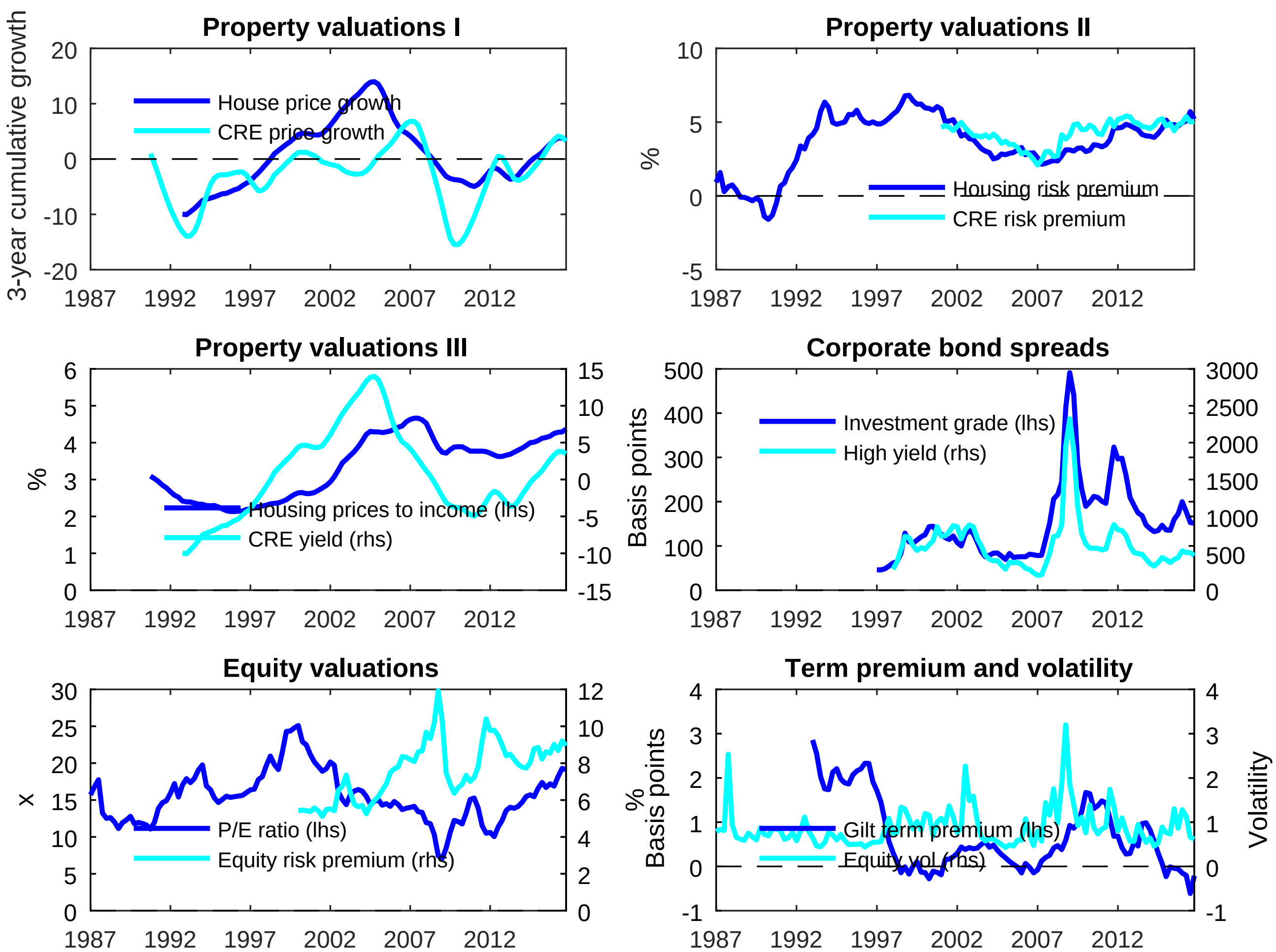
This appendix presents time series plots of the 29 indicators used in the analysis in the main text.

Figure B.1: *Time series plots of private nonfinancial sector leverage indicators*



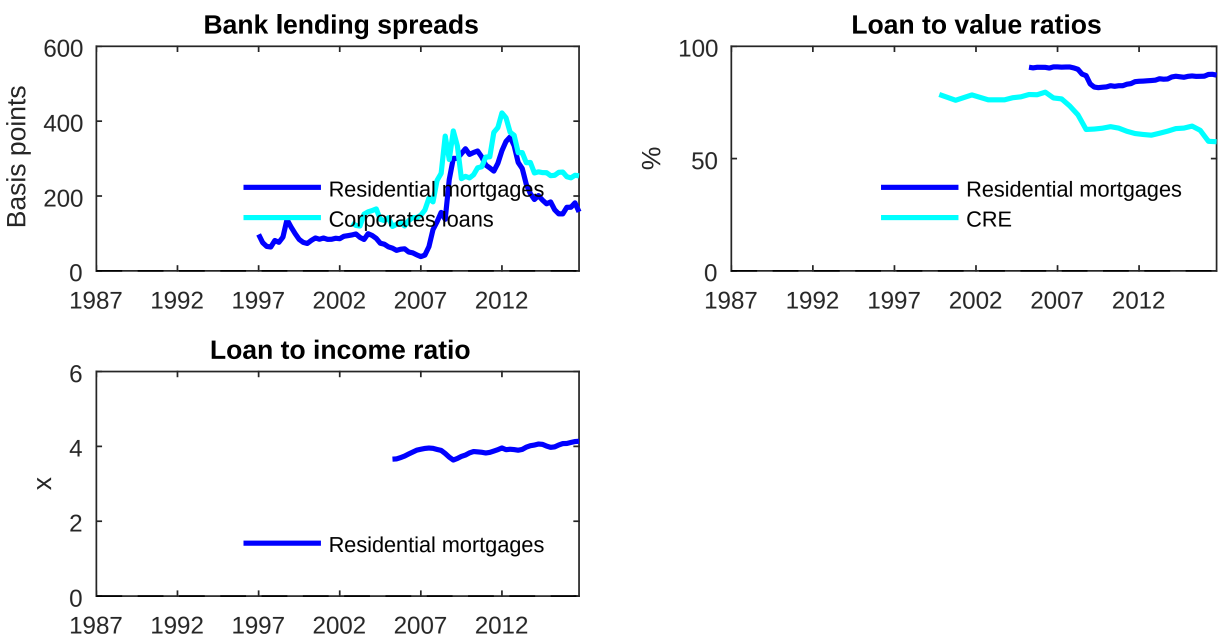
Notes: This figure presents the time series plots of the indicators used in this paper. Sources: Household credit growth: Bank of England and CML; Household and corporate debt levels: ONS and Bank calculations; External imbalances (I): ONS and Bank calculations; External imbalances (II): ONS and Bank calculations .

Figure B.2: *Time series plots of asset valuations indicators*



Notes: This figure presents the time series plots of the indicators used in this paper. Sources: Property valuations I: IHS Markit, Halifax, Nationwide, MSCI IPD and Bank calculations; Property valuations III: ONS/Halifax/Nationwide, MSCI IPD and Bank calculations; Corporate bond spreads: Thomson Reuters Datastream and Bank calculations; Equity valuations: Bloomberg Finance L.P, IMF WEO, Thomson Reuters Datastream, Tradeweb and Bank calculations; Term premium and volatility: Bloomberg Finance L.P, HMT, Thomson Reuters Datastream, TradeWeb and Bank calculations: Daily term premium estimate is an average from four model outputs: benchmark and survey models from Malik, S. and Meldrum, A. (2016), Evaluating the robustness of UK term structure decompositions using linear regression methods. Journal of Banking Finance, Volume 67, June 2016, Pages 85-102; Guimares, R. and G. Vlieghe (2016), Monetary Policy Expectations and Long Term Interest Rates, unpublished working paper; and Andreasen, M. and Meldrum, A. (2015), Market beliefs about the UK monetary policy lift-off horizon: a no-arbitrage shadow rate term structure model approach, Bank of England Staff Working Paper No. 541.. and database right Investment Property Databank Limited and its licensors 2018. All rights reserved. IPD has no liability to any person for any losses, damages, costs or expenses suffered as a result of any use of or reliance on any of the information which may be attributed to it.

Figure B.3: *Time series plots of terms of credit indicators*



Notes: This figure presents the time series plots of the indicators used in this paper. Sources: Bank lending spreads: Bank of England and CML; Loan to income ratio: FCA Product Sales Database and Bank calculations.

Appendix C: Alternative aggregation approaches

In this section, we present results from two alternative approaches to weighting the indicators: (i) a principal-components based approach, where indicators are weighted using the variance-covariance matrix, and (ii) an 'intensity score' measure, following a variant of the approach proposed in Kaminsky (1999).

PCA-based weights

Given that the 29 indicators form an unbalanced panel, we use Stock and Watson's (2002) approach to generate PCA weights.[23] We apply this algorithm to the set of indicators in each of the three categories discussed in the main text (private nonfinancial sector leverage, asset valuations, and terms and conditions), one by one. We then compare the resulting factors obtained with the composite indices defined in section 5. The results are described in Figure C.1.

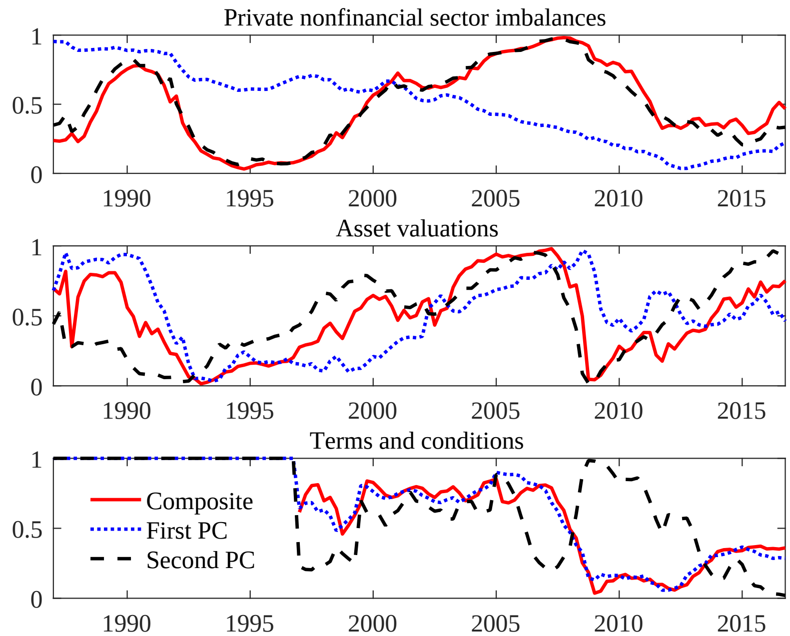
The top panel compares the first and second factors with the private nonfinancial sector leverage index. These factors explain 43% and 28% of the variation in the indicators. It is striking that the second factor corresponds almost exactly with the leverage index; the correlation between the series is 0.96. By contrast, the first factor appears to pick up the trend across indicators in this set, and has a correlation of almost zero with the leverage index. The central panel presents the same exercise for the asset valuations index. The first and second factors explain 30% and 27% of the variation in asset price indicators, respectively. This index is positively correlated with both the first and second factor, but is most strongly correlated with the latter. Finally, the bottom panel presents the analysis for the terms of credit indicators. The factors explain 70% and 23% of the variation in these indicators. The first factor has a correlation of 0.97 with the composite index presented in the paper.

Table C.1: *Correlation between PCA factors and composite indices*

	1st factor:	2nd factor:
Private nonfinancial sector leverage	-0.13	0.96
Asset valuations	0.58	0.67
Terms of credit	0.97	-0.10

[23] The limited sample of data available for some of our indicators makes implementing PCA, which requires a balanced panel of data, problematic. Stock and Watson's (2002) expectation-maximisation (EM) algorithm works by initially applying PCA to a balanced subset of the data, before regressing individual indicators on the resulting index to predict missing values. See Brave and Butters (2011) for a more detailed description and application of this approach.

Figure C.1: *Comparison of PCA-based weights with composite indices from main text*



Notes: This figure presents a comparison between the composite indices introduced in section 5 alongside the first and second principal components obtained by applying Stock and Watson's (2002) Expectation Maximization algorithm to the (standardised) underlying indicators within each cateogory. We map each factor into the (0,1) interval based on its percentile in its historical distribution, which is estimated via a kernel estimator.

8.1. Intensity-score based weights

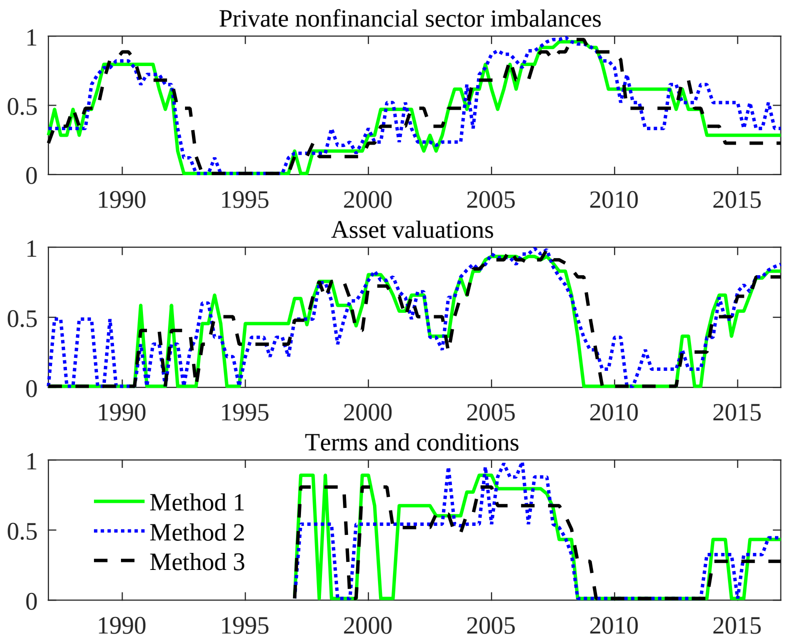
We next explore a set of alternative weighting schemes that are inspired by the approach advocated by Kaminsky (1999). Under this simple approach, we score an indicator 1 if it exceeds a threshold and 0 otherwise. Intuitively, this allows for less substitution between indicators than is inherent averaging, where 'hot' and 'cold' indicators offset.

We use this approach to create three different composite indicators: the first is a simple tally of the number of indicators above their 75th percentile; the second is a tally of the number of indicators between their 65th and 85th percentiles plus two times the number of indicators above their 85th percentile; the third is a variant of the first, but where we tally the number of indicators exceeding their 75th percentile at any point over the past 4 quarters. We express each index as a proportion of its maximum possible score.[24]

Figure C.2 presents the results of this exercise. All three methods present a similar picture to that of the more complicated weighting schemes presented in the paper. Asset valuations appear as the most elevated category at present, whereas private nonfinancial sector leverage appears subdued.

[24]The maximum value of the first and third indices is N, the number of indicators in our sample. The maximum value of the second index is 2N. We report each index as a percentage of its maximum value in that quarter to take account of the fact that the number of time series increases over time.

Figure C.2: *Kaminsky-style 'intensity score' indices*



Notes: This figure presents time series plots of the three intensity score metrics introduced in this appendix. Method 1 tallies the proportion of indicators exceeding their 75th percentiles each quarter; Method 2 is the sum of the proportion of indicators between their 65th and 85th percentiles and two times the proportion exceeding their 85th percentiles; Method 3 tallies the proportion of indicators that have exceeded their 75th percentiles at any point over the past 4 quarters. The y-axis is the proportion of indicators meeting the criteria as a proportion of the total number available each quarter.

References

ADRIAN, T., AND H. SHIN (2014): 'Procyclical Leverage and Value-at-Risk', Review of Financial Studies, Society for Financial Studies, 27(2), 373-403.

ADRIAN, T., A. ASHCRAFT, AND N. CETORELLI (2013): 'Shadow Bank Monitoring', Federal Reserve Bank of New York Staff Report No. 638.

ADRIAN, T., N. BOYARCHENKO, AND D. GIANNONE (2016): 'Vulnerable Growth', Federal Reserve Bank of New York Staff Report No. 794.

AIKMAN, D., A. LEHNERT, N. LIANG, AND M. MODUNGO (2016): 'Financial Vulnerabilities, Macroeconomic Dynamics, and Monetary Policy', Finance and Economics Discussion Series 2016-055, Washington: Board of Governors of the Federal Reserve System.

AIKMAN, D., M. KILEY, S.J. LEE, M. PALUMBO, AND M. WARUSAWITHARANA (2017): 'Mapping Heat in the U.S. Financial System', Journal of Banking and Finance, Vol. 81(C), pages 36-64.

AL-SAFFER, Y., W. RIDINGHER, AND S. WHITAKER (2013): 'The Role of External Balance Sheets in the Financial Crisis', Financial Stability Paper No.24.

ALESSI, L., AND C. DETKEN (2011): 'Quasi Real Time Early Warning Indicators for Costly Asset price Boom/Bust Cycles: A Role for Global Liquidity', European Journal of Political Economy, 27(3), 520533.

BARON, M., AND W. XIONG (2017): 'Credit Expansion and Neglected Crash Risk', The Quarterly Journal of Economics, 132 (2), 713764.

BELL, J., AND D. PAIN (2000): 'Leading Indicator Models of Banking Crises A Critical Review', Financial Stability Review, Bank of England, Issue 9, Article 3,11329.

BIANCHI, J., AND E. MENDOZA (2016): 'Optimal Time-Consistent Macroprudential Policy', National Bureau of Economic Research Working Paper No. 19704.

BILSON, J. (1979): 'Leading Indicators of Currency Devaluations', Columbia Journal of World Business 14, 6276.

BERG, A., AND C. PATTILLO (1999): 'Are Currency Crises Predictable? A Test' International Monetary Fund Staff Paper, Vol. 46, No. 2.

BORDO, M., B. EICHENGREEN, D. KLINGEBIEL, AND M. MARTINEZ-PERIA (2001): 'Is the Crisis Problem Growing More Severe?' Economic Policy: A European Forum 32, 51-75

BORIO, C., AND M. DREHMANN (2009): 'Assessing the Risk of Banking Crises - Revisited', Bank for International Settlements Quarterly Review, March 2009.

BORIO, C., AND P. LOWE (2002a): 'Asset Prices, Financial and Monetary Stability: Exploring the Nexus', Bank for International Settlements Working Paper No. 114.

BORIO, C., AND P. LOWE (2002b): 'Assessing the Risk of Banking Crises', Bank for International Settlements Quarterly Review, December, 4354.

BORIO, C., AND P. LOWE (2004): 'Securing Sustainable Price Stability: Should Credit Come Back from the Wilderness?', Bank for International Settlements Working Paper No. 157.

BOYD, J., P. GOMIS, S. KWAK, AND B. SMITH (2000): 'A Users Guide to Banking Crisis', University of Minnesota Working Paper.

BRAVE, S., AND R.A. BUTTERS, (2011), 'Monitoring financial stability: A financial conditions index approach', Federal Reserve Bank of Chicago Economic Perspectives, Vol. 35, No. 1.

BRIDGES, J., C. JACKSON, AND D. MCGREGOR (2017): 'Down in the Slumps: The Role of Credit in Five Decades of Recessions', Bank of England Working Paper No. 659.

BREALEY, R., AND A. VILA (1998): 'Equity Prices and Financial Stability', Bank of England Financial Stability Review Autumn 1998, 10-18.

BRUNNERMEIER, M., AND Y. SANNIKOV (2014): 'A Macroeconomic Model with a Financial Sector', American Economic Review, Vol. 104, No. 2, 379-421.

CATO, L., AND G. MILESI-FERRETTI (2013): 'External Liabilities and Crises', International Monetary Fund Working Paper 13/113.

CECCHETTI, S. (2006): 'Measuring the Macroeconomic Risks Posed by Asset Price Booms', Asset Prices and Monetary Policy, University of Chicago.

CECCHETTI, S. (2008): 'Monetary Policy and the Financial Crisis of 200708', Centre for Economic Policy Research Policy Insight No. 21.

CECCHETTI, S. AND L. LI (2008): 'Do Capital Adequacy Requirements Matter for Monetary Policy', Economic Inquiry, 46(4), 643-659.

CHAMON, M., AND C. CROWE (2012): 'Evidence on Financial Globalization and Crisis: 'Predictive' Indicators of Crises - Macroprudential Indicators, Institutional Environment, Micro', International Monetary Fund Working Paper.

DANIELSSON, J., K. JAMES, M. VALENZUELA AND I. ZER (2016): 'Model Risk of Risk Models', Journal of Financial Stability, 23, 79-91.

DATTELS, P., R. MCCAUGHRIN, K. MIYAJIMA, AND J. PUIG (2010): 'Can You Map Global Financial Stability?', International Monetary Fund Working Paper 10/145.

DAVIS, E. (1995): 'Debt, Financial Fragility and Systemic Risk', Oxford University Press.

DAVIS, E., AND D. KARIM (2008a): 'Comparing Early Warning Systems for Banking Crises', Journal of Financial Stability, Vol. 4, 89120.

DAVIS, J., A. MACK, W. PHOA, AND A. VANDENABEELE (2014): 'Credit Booms, Banking Crises and the Current Account', Federal Reserve Bank of Dallas Working Paper No. 178.

DELL'ARICCIA, G., D. IGAN, L. LAEVEN, H. TONG, B. BAKKER, AND J. VANDENBUSSCHE (2012): 'Policies for Macrofinancial Stability: How to Deal with Credit Booms', International Monetary Fund Staff Discussion Note No. 12/06.

DEMIRGUC-KUNT, A., AND E. DETRAGACHE (1998): 'The Determinants of Banking Crises in Developing and Developed Countries', International Monetary Fund Staff Paper, Vol. 45, No. 1.

DEMIRGUC-KUNT, A., AND E. DETRAGACHE (2000): 'Monitoring Banking Sector Fragility: A Multivariate Logit Approach', World Bank Economic Review, 14(2), 287-307.

DEMIRGUC-KUNT, A., AND E. DETRAGACHE (2005): 'Cross-Country Empirical Studies of Systemic Bank Distress: A Survey', International Monetary Fund Working Paper No. 05/96.

DREHMANN, M. (2013): 'Total Credit as an Early Warning Indicator for Systemic Banking Crises', Bank of International Settlements Quarterly Review, June 2013.

DREHMANN, M., AND M. JUSELIUS (2012): 'Do Debt Service Costs Affect Macroeconomic and Financial Stability?', Bank for International Settlements Quarterly Review, September 2012.

DREHMANN, M., C. BORIO, AND K. TSATSARONIS (2011): 'Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates', Bank for International Settlements Working Paper No. 355.

EICHENGREEN, B., AND C. ARTETA (2000): 'Banking Crises in Emerging Markets: Presumptions and Evidence', Centre for International and Development Economics Research Working Papers from University of California at Berkeley.

EICHENGREEN, B., A. ROSE, AND C. WYPLOSZ (1995): 'Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks', Economic Policy, Vol. 21, 249-312.

EICHNER, M., D. KOHN, AND M. PALUMBO (2010): 'Financial Statistics for the United States and the Crisis: What Did They Get Right, What Did They Miss, and How Should They Change?' The Federal Reserve Finance and Economics Discussion Series.

FRANKEL, J., AND A. ROSE (1996): 'Currency Crashes in Emerging Markets: An Empirical Treatment', Journal of International Economics, Vol. 41, 351-366.

FRANKEL, J., AND G. SARAVELOS (2012): 'Can Leading Indicators Assess Country Vulnerability? Evidence from the 2008-09 Global Financial Crisis', Journal of International Economics, Vol. 87, 21631.

GAVIN, M., AND R. HAUSMANN (1995): 'Macroeconomic Volatility and Economic Development', Inter-American Development Bank.

GENNAIOLI, N., A. SHLEIFER, AND R. VISHNY (2015): 'Neglected Risks: The Psychology of Financial Crises', American Economic Review, Papers and Proceedings 105, 310314.

GIESE, J., H. ANDERSEN, O. BUSH, C. CASTRO, M. FARAG, AND S. KAPADIA (2014): 'The Credit-To-GDP Gap and Complementary Indicators for Macroprudential Policy: Evidence From The UK', International Journal of Finance and Economics, 19(1), 2547.

GILCHRIST, S., AND E. ZAKRAJSEK (2012): 'Credit Spreads and Business Cycle Fluctuations', American Economic Review, Vol. 102, No. 4, 1692-1720.

GOURINCHAS, P., AND M. OBSTFELD (2012): 'Stories of the Twentieth Century for the Twenty-First', American Economic Journal: Macroeconomics, 22665.

GOURINCHAS, P., R. VALDES, AND O. LANDERRETCHE (2001): 'Lending Booms: Latin America and the World', Economia Journal, The Latin American and Caribbean Economic Association - LACEA, vol. 0 (Spring 20), 47-100.

GRAEFE, A., H. KUCHENHOFF, V. STIERLE, AND B. RIEDL (2014), 'Limitations of Ensemble Bayesian Model Averaging for Forecasting Social Science Problems', International Journal of Forecasting.

GREEN, K. AND J. S. ARMSTRONG (2015), 'Simple Versus Complex Forecasting: The Evidence', Journal of Business Research, 68, 1678-1685.

HATZIUS, J., P. HOOPER, F. MISHKIN, K. SCHOENHOLTZ AND M. WATSON (2010): 'Financial Conditions Indexes: A Fresh Look after the Financial Crisis', National Bureau of Economic Research Cambridge, M.A Working Paper No. 16150.

HONOHAN, P. (1997): 'Banking System Failures in Developing and Transition Countries: Diagnosis and Prediction', Bank of International Settlements Working Paper No. 39.

JORDÁ, O., M. SCHULARICK, AND A. TAYLOR (2013): 'When Credit Bites Back', Journal of Money, Credit and Banking, 45(2).

JORDÁ, O., M. SCHULARICK, AND A. TAYLOR (2015): 'Leveraged Bubbles', Journal of Monetary Economics, 76, 1-20.

JORION, P. (2001): 'Value at Risk', 2nd edition, McGraw-Hill.

KAMINSKY, G. (1999): 'Currency and Banking Crises: The Early Warnings of Distress', International Monetary Fund Working Paper No. 99/178.

KAMINSKY, G., AND C. REINHART (1999): 'The Twin Crises: The Causes of Banking and Balance of Payments Problems', The American Economic Review, Vol. 89.

KAMINSKY, G., S. LIZONDO, AND C. M. REINHART (1998): 'Leading Indicators of Currency Crises', International Monetary Fund Staff Paper, 45, 1-48.

KILEY, M. (2004): 'Stock Prices and Fundamentals: A Macroeconomic Perspective', The Journal of Business, Vol. 77, No. 4, 909-936.

KINDLEBERGER, C. (1978): 'Manias, Panics and Crashes: A History of Financial Crises', New York: Basic Books, 1st edition.

KOENKER, R., AND G. BASSETT (1978): 'Regression Quantiles', Econometrica, Vol. 46, No. 1, 33-50.

KOOP, G., M. PESARAN, AND S. Potter (1996): 'Impulse Response Analysis in Nonlinear Multivariate Models', Journal of Econometrics, 74, 119-147.

KORINEK, A., AND A. SIMSEK (2016): 'Liquidity Trap and Excessive Leverage'. American Economic Review, 106(3), 699-738.

KRISHNAMURTHY, A., AND T. MUIR (2017): 'How Credit Cycles across a Financial Crisis', National Bureau of Economic Research Working Paper No. 23850.

LAEVEN, L. AND F. VALENCIA (2013): 'Systemic Banking Crises Database: An Update', International Monetary Fund Economic Review, Vol. 61, No. 2, 225-270.

LOPEZ-SALIDO, D., J. STEIN, AND E. ZAKRAJSEK (2015): 'Credit-market Sentiment and the Business Cycle', Finance and Economics Discussion Series Staff Working Paper No. 2015-028.

NOLAN, L. (2016): 'Economic Statistics Transformation Programme: Developing the Enhanced Financial Accounts (UK Flow of Funds)', Office for National Statistics.

MCKNINNON, R., AND H. PHIL (1996): 'Credible Liberalizations and International Capital Flows: The Over Borrowing Syndrome', Financial Deregulation and Integration in East Asia, 7-42.

MENDOZA, E., AND M. TERRONES (2011): 'An Anatomy of Credit Booms and Their Demise', Fifteenth Annual Conference of the Central Bank of Chile.

MIAN, A., AND A. SUFI (2010): 'Household Leverage and the Recession of 2007 - 2009', International Monetary Fund Economic Review, 58(1), 74117.

MIAN, A., A. SUFI, AND E. VERNER (2017): 'How Do Credit Supply Shocks Affect the Real Economy? Evidence from the United States in the 1980s', National Bureau of Economic Research Working Paper.

MINSKY, H. (1977): 'The Financial Instability Hypothesis: An Interpretation of Keynes Alternative to 'Standard' Theory', Nebraska Journal of Economics and Business 16(1), 5-16.

MISHKIN, F. (1990): 'Asymmetric Information and Financial Crises: a Historical Perspective', National Bureau of Economic Research Working Paper No. w3400.

REINHART, C., AND K. ROGOFF (2010): 'Growth in a Time of Debt', National Bureau of Economic Research Working Paper No. 15639.

SCHULARICK, M., AND A. M. TAYLOR (2012): 'Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008', American Economic Review, 102(2), 102961.

SHILLER, R. (2000): 'Measuring Bubble Expectations and Investor Confidence', Journal of Psychology and Financial Markets, 1(1), 49-60.

STEIN, J. (2013): 'Overheating in credit markets: origins, measurement and policy responses', speech given at the Restoring household financial stability after the great recession: why household balance sheets matter research symposium sponsored by the Federal Reserve Bank of St. Louis, St. Louis, Missouri, February 7.

STOCK, J., AND M. WATSON (2002): 'Forecasting Using Principal Components from a Large Number of Predictors', Journal of the American Statistical Association December 2002, Vol. 97, No. 460.

STOCK, J., AND M. WATSON (2003): 'Has the Business Cycle Changed? Evidence and Explanations', Monetary Policy and Uncertainty: Adapting to a Changing Economy, 9-56.