Financial Stability Focus

SOUTH AFRICAN RESERVE BANK

2022 | July 2022

The Financial Stability Focus is a compilation of South African Reserve Bank Topical Briefings

1. **Topical Briefing: The Financial Stability considerations associated with Central Bank Digital Currency (CBDC)**
   Mulalo Mamburu, Alex Smith

2. **Topical Briefing: Part 1: The application of the International Association of Insurance Supervisors (IAIS) Holistic Framework to South Africa**
   Videshree Rooplall and Christiaan Henning

3. **Topical Briefing: Part 2: Vulnerabilities assessment of the South African insurance sector using elements of the International Association of Insurance Supervisors (IAIS) Holistic Framework and general indicators**
   Aadila Hoosain and Katlego Lesejane

4. **Topical Briefing: Financial conditions and risks to financial stability in South Africa**
   Greg Farrell

5. **Topical Briefing: New Econometric Models for South African House Prices, Mortgage Debt and Residential Investment**
   Janine Aron, John Muellbauer

6. **Topical Briefing: Enhancing Financial Stability and Monetary Analysis in the Core Model of the SARB**
   Janine Aron, John Muellbauer

7. **Topical Briefing: Non-Performing Loans in South Africa: a Scoping paper for Future Model Development**
   John Muellbauer, Janine Aron

Authorised for distribution by:

Dr Nicola Brink

Nicola.Brink@resbank.co.za

TOPICAL BRIEFING: No. 1

SOUTH AFRICAN RESERVE BANK

Financial Stability Department

The Financial stability considerations associated with Central Digital Currency (CBDC)

Abstract

This paper provides an overview of the potential financial stability risks associated with CBDC and makes suggestions on how to address these risks. CBDC will be disruptive to the financial system as it will necessarily involve a movement of funds off commercial bank balance sheets on to the balance sheet of the SARB (although the quantum is unclear). This would be an intentional effect given that the CBDC will be designed as a means of payment superior to those currently in existence in South Africa. We distinguish between two different types of liability migration from bank balance sheets resulting from CBDC: structural and idiosyncratic. Structural migration refers to the re-organisation of deposits in a longer-term equilibrium setting, while idiosyncratic migration refers to potential deposit flight into CBDC during crisis-type episodes. The latter being underpinned by the risk-free nature of CBDC. Structural liability migration would be a feature (rather than a bug) of CBDC, which we argue should be catered for through increased accommodation from the SARB to the commercial banking sector. Depending on the size of the structural migration, this could call for a re-evaluation of the SARB's collateral and liquidity provision frameworks. Idiosyncratic migrations are potentially more concerning from a financial stability perspective. Here we argue that there are trade-offs that will need to be made between payments system efficiency and the degree of bank run risk that the financial system is exposed to. This trade-off will be affected by the degree to which the SARB is willing to supply large amounts of emergency liquidity assistance to the banking sector on demand.

Introduction

The SARB is currently in the process of studying CBDC and exploring whether the introduction of a central bank issued digital currency could confer benefits on South Africa's financial system and broader economy. This analysis is not complete and no decisions regarding the introduction of CBDC have yet been made. To contribute to the SARB's thinking on this subject, this paper offers some views on the potential implications of CBDC for financial stability.

As CBDC can have a variety of different features, our starting point for this discussion is to focus the SARB's previously expressed preferences for the design of a domestically issued CBDC[1]. These are: that the CBDC should be for general purpose retail use and that it should be complementary to cash (i.e., it is not a replacement for cash). We also assume that CBDC will, as the name implies, be issued by the central bank and will therefore be a central bank liability.

While this final point may appear to be an obvious assumption, there is a growing debate in the literature around so-called synthetic CBDCs. These are funds deposited by the public at private financial service providers, which are in turn fully matched by an amount held at the central bank. By ensuring full coverage of any deposits held by these private entities through a simultaneous deposit at the central bank these synthetic CBDCs purport to offer a similar feature to a true CBDC, that of certainty in the stable value of the funds held in the account by the public. However, synthetic CBDCs are liabilities of the private sector and not of the central bank, which means that their implications for the broader financial system are materially different to that of a true CBDC. For this reason, we do not examine synthetic CBDCs in this paper.

The remainder of this paper is laid out over four sections. Section one discusses what CBDC is and examines the key reasons why many central banks are considering issuing CBDC. Section two provides an overview of the financial stability risks and benefits associated with CBDC according to the burgeoning literature on this subject. Section three outlines the policy options available to address the financial stability risks

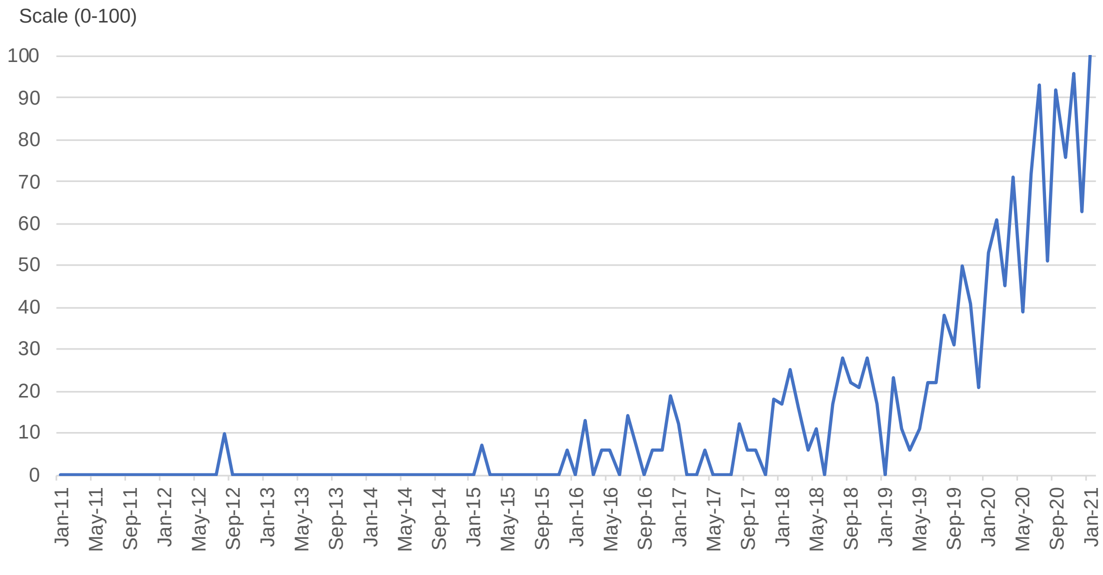
[1] See the following press release from the SARB: https://www.resbank.co.za/content/dam/sarb/publications/media-releases/2021/cbdc-/Feasibility%20study%20for%20a%20general-purpose%20retail%20central%20bank%20digital%20currency.pdf

that could emerge from the introduction of a CBDC. Section four concludes and discusses the way forward.

1. CBDC: Design, unique features and reasons for issuance

CBDC is a term that has grown in prominence since 2019 as many central banks have begun research into their own versions of CBDC (see figure 1). Despite its increased use, the precise meaning of the term can differ depending on what each central bank has in mind. Moreover, some slightly confusing analogies have been put forward to describe CBDC, which may be as misleading as they are illuminating. In this section we will explain how the SARB sees CBDC[2] and we will describe some of its key features. As we will discuss in sections 2 and 3 of this paper, the design features of a CBDC are integral to any discussion of its policy implications and potential economic impact.

Figure 1: Popularity of the search term "central bank digital currency" on Google



Source: Google trends

2 As noted in the introduction, the SARB continues to study CBDC and its views remain subject to change. Thus, the current discussion is based on a point-in-time view.

The scale on the above graph refers to the popularity of a search term, where 100 is when it is most popular and a score of 50 represents a time when the term was half as popular as when it scored 100.

What it is

CBDC is a form of money issued by the central bank. It is a liability of the central bank in the same way that cash is. It is unique because the general public have not historically had access to a digital central bank liability. In the past, digital exchanges of funds between members of the general public have involved transfers from one commercial bank account to another. This represents the transfer of commercial bank liabilities. However, the banking sector has had access to a digital form of central bank money for many years. At the end of each day commercial banks settle their balances with each other by transacting in their deposits held with the SARB. Therefore, in a strict sense, a digital representation of central bank liabilities does already exist, but it is currently restricted to the balances that commercial banks hold with the SARB.

It is currently envisioned that CBDC would operate like cash, in that it would be a widely available central bank liability for use as a payment instrument. However, comparisons with cash are imperfect, because digital transactions face a different set of protocols to those of cash. The most obvious example is that transacting in cash requires a physical exchange between two parties in a particular location, whereas digital transactions can occur remotely, but require sophisticated underlying infrastructure to ensure the transaction is executed securely. It is also important to note that CBDC is not expected to replace cash, but to co-exist with it. While CBDC may substitutive cash to a limited extent, its digital nature could make it more likely to displace digital payments channels (which do not operate with credit) such as debit cards or fintech payments solutions such as Zapper.

CBDC is also expected to allow individuals to transact at any time of the day, even if one, or both, of the parties are not connected to the internet. Sending CBDC from one person to another would be instantaneous (or at least occur in a matter of seconds). In this way, CBDC would differ from a commercial bank deposit as it currently takes a day or more for payments to flow from the deposit account of one bank to that of a different bank. The mechanics of affecting a CBDC payment would also be fundamentally different. While interbank payments involve the transfer of a bank liability (in other words a bank deposit) from one person to another, CBDC involves the transfer of a SARB liability. When individuals hold bank deposits and use them to transact, the inflow of these deposits allows banks to acquire assets (generally loans) on the other side of their balance sheet. Conversely, a switch to holding CBDC as a payments instrument (which is a SARB liability) will instead create the need for SARB to acquire additional assets on its balance sheet to match the rise in liabilities. Thus, if banks hold CBDC on behalf of their clients, this would be fundamentally different to holding a deposit on behalf of a client. In this sense, CBDC safekeeping by a bank is analogous to a client holding cash in a safe deposit box at a bank. The CBDC remains a liability of the SARB and an asset of the owner. Therefore, CBDC cannot provide scope for lending by a bank in the same way that a deposit can, because the CBDC does not reside on the balance sheet of the commercial bank.

CBDC would also likely be accessible to those without a bank account, providing an alternative means of payment to the unbanked. Therefore, CBDC would be a new kind of instrument which is similar to, but also different from other forms of money that currently exist in South Africa. It is this uniqueness which makes CBDC a potentially important financial innovation, but also something that could create risks to the financial system.

Why would the SARB issue a CBDC?

The primary justifications for issuing a CBDC include the enhancement of financial inclusion, a reduction of the cost and time delays associated with executing retail payments (both domestically and across borders), promoting new innovations and competition within the payments space, offering a central bank issued alternative to new types of privately created digital money (and crypto assets) and providing a reliable alternative to both cash and commercial bank deposits (Bank of England, 2020; DNB, 2020). CBDC could benefit consumers by providing access to a secure, fast, low cost, easy to use means of payment. It could also benefit businesses, particularly those in the informal sector by making it easier to engage in digital transactions. New types of financial services may emerge as CBDC could become a fully interoperable payment instrument upon which value-added services may be built. Finally, CBDC could make transactions safer by reducing reliance on cash, which is susceptible to being stolen or damaged.

What remains to be decided

There are various aspects of the potential CBDC design which remain undefined. We will discuss three particularly important features which could have implications for the way in way CBDC is used and how it will affect the financial system.

The first is whether or not CBDC would be interest bearing. Commercial bank deposits with the central bank do generally pay interest. Various academics have argued that a CBDC which pays interest at the central bank policy rate could enhance the transmission of monetary policy (Bank of England, 2020). Moreover, this could provide the public with a savings vehicle and could then encourage commercial banks to pay a similar rate of interest on the deposits that the public holds with commercial banks. The risk associated with a CBDC that pays interest is that it could make commercial bank deposits less attractive and could lead to a migration of funds off commercial bank balance sheets and onto that of the SARB (an issue that is discussed further in section 2).

A second issue is the extent to which the public will have unfettered access to the CBDC. This question captures a number of different issues, including whether foreign citizens could hold domestically issued CBDC, whether there would be any limits on the amount of CBDC that a person (or business) can hold, whether there would be any frictions introduced in using CBDC (such as transaction/holding fees or limits to the size of transactions) and whether banks or other institutions could refuse the transferal of funds into CBDC.

The third issue relates to the technology used in the design of CBDC and the degree of anonymity associated with its use. Two kinds of technology are possible for the infrastructure used in exchange: account or token based. In an account-based system, the emphasis is on verification of the account holder's identification to ensure that the payer has sufficient funds in the account and has initiated the transaction (this is analogous to a bank account and the need to verify your credit card transactions through an OTP). Meanwhile, in a token-based system the emphasis is on the verification of the payment instrument. For example, cash payments are token based and only require that the cash itself is not counterfeit, information about the payer is generally not required to conclude the cash transaction (Committee on Payments and Market Infrastructures, 2018). As pointed out by Risberg and Segal (2020): “while a decentralized digital token would allow for greater anonymity and arguably be more resilient to infrastructure outages and cyberattacks, a centralized ledger could promote greater transparency and facilitate compliance with anti-money laundering and countering the financing of terrorism (AML-CFT) and know-your-customer (KYC) frameworks”. The degree of anonymity could also constrain regulators from introducing frictions such as holding limits, because an anonymous token is one about which the regulator has limited information. A range of options for anonymity are possible however and would depend on the SARB’s preferences.

There are various other design choices that policy makers will have to face, but we don’t dwell on them in this paper given their limited implications for financial stability.

To summarise, a domestically issued CBDC will likely be accessible to the general public (although perhaps not foreign citizens), but may have some limitations (or added costs) imposed on the size of allowed holdings or the ease with which funds can be transferred between CBDC and commercial bank accounts. CBDC could be account or token based and could have varying degrees of anonymity associated with holding and transacting in it. CBDC may or may not pay interest. It is intended primarily as a secure, cheap and fast means of payment, but it could also be used as a risk-free (albeit low or zero return) option for storing savings. Apart from the efficiency gains associated with introducing an alternative (better) payments mechanism, CBDC is also potentially a central bank response to privately issued crypto-assets (such as Bitcoin). Crypto assets, if they become sufficiently ubiquitous, could migrate domestic payments, saving, borrowing away from Rands. This, in turn, could reduce the efficacy of monetary policy, limit the SARB’s effectiveness in combating illicit financial flows and create financial stability risks associated with currency mismatches on firm balance sheets. It could also constrain the SARB’s ability to address these risks as policies such as lender of last resort would be less effective in a setting where a significant share of bank liabilities is denominated in currencies (or crypto-assets) other than the Rand.

Finally, CBDC could also create an interoperable base upon which private sector entitles could build value added products and services. In this sense, CBDC could unleash a wave of new financial innovations and could spur greater competition in the financial sector (including from traditionally non-financial entities such as telecoms and retail firms).

2. The financial stability risks and benefits associated with CBDC

The potential financial stability implications of CBDC issuance are largely contingent on the design of the CBDC as well as the governance and infrastructure that underpins it. While the literature on the policy impacts of CBDC issuance is still new, authors have converged on two major issues regarding CBDC and financial stability: the potential disintermediation of commercial banks and the risk of runs on commercial bank deposits during crisis-like episodes. We refer to these as structural and idiosyncratic liability migrations, respectively. To a lesser extent, the literature also discusses operational and security risks associated with CBDC which could have implications for financial stability. However, these are not specific to CBDC as current payment methods face similar risks, hence they are not discussed in this section.

Disintermediation of banks

Depending on the design of the framework, CBDC issuance could result in a large structural migration of bank deposits into CBDC. This, in turn, could lead to the (partial) disintermediation of commercial banks, threatening the viability of their current business models (Meaning *et al.*, 2018). While commercial banks could continue to offer value added services such as loans, financial advice and cash withdrawals, CBDC could become an attractive alternative to a purely transactional bank account. There are three key features that could make CBDC more attractive: 1) if CBDC holding and transactional fees are considerably lower than transactional bank account fees, 2) if CBDC provides a more efficient means of payment than interbank transfers currently are, and 3) CBDC would be free of any default and liquidity risks associated with holding a commercial bank deposit.

Commercial bank deposits carry credit risk (i.e. the risk that the bank fails), which is mitigated to some extent by a deposit insurance scheme. South Africa is in the process of implementing such a scheme. Deposit insurance guarantees the value of a bank deposit up to a certain size (in South Africa this is expected to be set at R100 000).

BDC deposits, on the other hand, would be risk-free[3] as they are liabilities of the entity that has the monopoly right to the creation of domestic currency. Bindseil (2020) argues that depositors could shift low-remuneration deposits from commercial banks to riskless CBDC balances. This could particularly be the case for large corporate deposits which are more likely than household deposits to exceed the insurance limit.

If bank deposits do shrink on a structural basis, banks may either have to pay higher interest rates to attract new forms of funding (thereby incurring increased funding costs) or will experience a commensurate reduction in their assets (most of which are loans). According to data from the BIS, more than 80% of credit provided to the private sector in South Africa is from banks. Hence, credit could become more expensive or difficult to acquire for many borrowers, should a CBDC become a popular alternative to a bank deposit. The magnitude of this effect would depend on how the SARB manages its own balance sheet as discussed below.

Brunnermeier and Niepelt (2019) develop a theoretical model which indicates that the financial stability impact of increased CBDC holdings at the expense of commercial bank deposits would depend on the response of the central bank's open market operations. Specifically, the authors argue that the adoption of CBDC would not undermine financial stability if the central bank committed to automatically replacing deposit funding with central bank funding to banks, thereby altering the composition, but not the volume of bank funding.

However, if this proposal were to be implemented, increased commercial bank borrowing from the central bank would create the need for additional collateral against which the central bank extends credit[4]. This could raise challenges for the current framework used in South Africa. At present only government securities are accepted as collateral, hence banks may need to acquire additional government securities or the SARB would need to consider whether it would accept a broader the pool of assets as

[3] Risk-free in this context means that the CBDC will retain its value in rand terms. However, CBDC, like other rand assets would remain subject to inflation risk and may change in value relative to other currencies. Thus, the term risk-free in this context applies to credit risk.

[4] The central bank does not provide uncollateralized loans to commercial banks.

collateral. Increasing commercial bank holdings of government debt as collateral would in turn increase the sector's exposure to sovereign risk, while a broadening of assets eligible for posting as collateral could increase the credit risk exposure of the central bank (Bindseil, 2020).

Whether commercial banks raise their additional funding through wholesale issuance or the central bank, this would likely have an adverse impact on their profitability. This is because these alternative forms of funding are typically more expensive than deposits. In response, banks would likely pass along the cost to customers, which could in turn result in a reduction in credit demand through higher bank lending rates. This effect is the potential source of bank disintermediation. In particular, non-bank financial institutions may be more competitive in an environment of higher bank funding costs, creating the scope for more credit provision to shift away from the banking sector. If this effect occurs rapidly or in a large magnitude, it could create financial stability risks, including by adversely impacting on bank liquidity and profitability, forcing a contraction in bank lending and indirectly slowing economic activity by lifting the cost of credit.

Contrary to the view of bank disintermediation as a financial stability risk, some authors have noted the potential for CBDC to enhance financial resilience. The Bank of England (2020) notes that central bank money "plays a fundamental role in supporting financial stability by acting as a risk-free from of money that provides the ultimate means of settlement" for all payments in the financial system. Introducing another form of central bank money in the form of CBDC could thus enhance the stability of the payments system. Furthermore, Dyson and Hodgson (2016) suggest that CBDC could enhance financial stability by addressing the moral hazard created by deposit insurance schemes to the extent that insurable deposits could be shifted from commercial banks to CBDC balances, mitigating the credit risk inherent in a financial system based on bank deposits. This mechanism could introduce additional discipline on banks, who may otherwise be inclined to take on excessive risks with depositor funds.

Other potential benefits to bank disintermediation are related to the payments settlement system. At present, the settlement of payments takes place between banks through their central bank deposits. CBDC could enhance financial stability by allowing more firms direct access to the central bank payment system, reducing credit and liquidity risk within the system or the systemic impact of an outage or collapse of a single bank (Dyson and Hodgson, 2016). The diversification of payment systems through the introduction of CBDC would also help to improve resilience since it is unlikely that both traditional payments networks and a CBDC network would suffer outages simultaneously.

The risk of runs on commercial bank deposits

Most discussions in the literature on CBDC assume a framework in which these central bank deposits would be close substitutes for commercial bank deposits with perfect convertibility between CBDC and other forms of central bank money (e.g. Meaning *et al.* (2018), Bindseil (2020)). Interoperability between bank deposits and CBDC could support a rapid, widespread shift from bank deposits to risk-free CBDC during periods of heightened real or perceived stress in the banking system, which could exacerbate a systemic banking crisis (Mersch, 2018). Although CBDC and cash are both central bank liabilities, it is practically more challenging to withdraw large amounts of cash from a bank due to the costs, security risks and the physical constraints on transporting it. Meanwhile, CBDC as a 'digital form of cash' would potentially not be subject to these challenges.

Dyson and Hodgson (2016) note that CBDC can aggravate what could start out as a minor panic in the banking system, as account holders could decide to temporarily shift to CBDC until the issue is resolved, making the bank's liquidity issues worse. The increased risk of runs by depositors may also incentivise banks to adopt actions to protect themselves during such periods, such as hoarding reserves, which could affect the functioning of money markets (Bank of England, 2020).

Given the lack of empirical evidence, there is still significant uncertainty regarding the extent to which the introduction of CBDC could increase the risk of bank runs. Meaning *et al.* (2018) note that while the existence of CBDC may indeed make it easier to run in the event of increased credit risk in the banking system, it is possible that risk-averse depositors most likely to run on a bank are also most likely to shift from bank deposits to CBDC upon its introduction, limiting the scale of a run during a stress event. Empirical evidence also suggests that widespread runs on the banking system are unlikely. For instance, Bindseil (2020) uses Eurosystem banking data for 2008 and 2011 to show that historically runs from weak banks to strong banks were a more significant component of bank runs than systemic runs on banking deposits. Furthermore, run-risk can potentially be mitigated through the design of the CBDC system, with features such as a tiered-remuneration system (Bindseil, 2020) and notice periods for large withdrawals into CBDC accounts (Meaning *et al.*, 2018). Other design suggestions, such as Kumhof and Noone's (2018) approach of limiting the convertibility between bank deposits and CBDC, are more controversial in the literature as they challenge fundamental principles of central banking (Bindseil, 2020).

While CBDC could increase the risk of bank runs, it may also offer the SARB the benefit of an early-warning indicator of bank stress, depending on how CBDC data is reported and aggregated. For example, the SARB may be able to monitor flows of funds into and out of CBDC from the accounts of each commercial bank in real time. If large outflows from a particular bank are recorded, the SARB may be able to act quickly to address the incipient risk before it becomes systemic in nature.

In summary, it appears as if CBDC could pose a trade-off between financial system efficiency and stability. The easier it is to move into and out of CBDC, the greater the potential risk of disintermediation and bank runs. Policy makers may also have to face an intertemporal trade-off because a CBDC could create short term instability in the financial system as adoption increases and financial institution business models adapt to this new product. However, the benefits of a more efficient and competitive financial system will likely materialize over a longer-term horizon. The following section of this paper discusses the options available to policy makers to mitigate these financial stability risks associated with CBDC, whilst leveraging its potential benefits.

How to address the Financial Stability risks associated with CBDC

To address the risks associated with structural and idiosyncratic liability migration from the banking sector, the central bank has various options which fall into two broad categories. The first category includes frictions or limitations in the use or holding of CBDC. The second category relates to the commitment by the central bank to refinance any migration of liabilities from commercial bank balance sheets. These two approaches are potentially complementary and neither one is without its drawbacks.

In particular, as will be elucidated below, the extent to which these policy interventions are adopted may pose trade-offs with other goals associated with CBDC and with the central bank's broader mandate

Frictions and holding limits on CBDC

The literature points to a range of frictions that can be introduced into the CBDC architecture, which in turn aim to constrain the extent to which either structural or idiosyncratic bank liability flight can occur. For example, Bindseil argues for a tiered approach to the remuneration of CBDC. Under this approach a set holding value per individual or household is remunerated at a higher interest rate (for example, the prevailing policy rate), while holdings above this level are remunerated at a punitive interest rate which may fall during times of high demand for CBDC to disincentivize CBDC holdings (this rate could become negative). Other approaches such as hard holding limits, taxes or fees on CBDC transfers and a threshold above which transfers are non-instantaneous have also been suggested.

The downside of these approaches is that they aim to constrain the degree to which CBDC is widely used and held. Assuming that CBDC is a more efficient and/or cost-effective means of payment than the current payments system (which is its raison d'etre), such frictions will constrain the efficiency gains from CBDC. Thus, there emerges a clear trade-off between stability and efficiency in the financial system if an approach of this nature is followed. This is particularly severe as large corporations, which are key drivers of the economy would likely be excluded from using CBDC (or would face punitive costs), while low balance depositors would benefit from the most efficient payments instrument.

A more extreme solution proposed by Kumhof and Noone (2018) to the issue of liability migration from the banking sector is that commercial banks should be allowed to refuse conversation of deposits into CBDC. These authors argue that optional non-convertibility is the only way to guarantee that bank runs don't deplete bank liquidity and force the central bank to provide large scale credit back to the banking sector to ensure that CBDC convertibility can occur (a situation which can lead to collateral shortages as discussed further below). However, this solution may result in a parallel market for CBDC in which it trades at a higher value than commercial bank deposits even though both are denominated in the same currency. To address this risk, Kumhof and Noone (2018) propose that interest rates on CBDC be allowed to fluctuate in order to balance demand and supply. This could however result in large discrepancies between the policy rate and interest rates on CBDC, which if persistent may pose challenges to monetary policy transmission. More importantly, if South Africa opted to introduce a non-interest bearing CBDC, there would be no mechanism to ensure that CBDC and commercial bank deposits remain at parity under the condition that banks are not obliged to offer convertibility between these two assets.

Central Bank funding commitments

A second approach (alternative or complementary) to contain the financial stability risks associated with bank liability flight caused by CBDC is for the central bank to stand ready to refinance these liabilities. Through this approach banks simply swap one liability (a household or corporate deposit) for another (a central bank loan). As pointed out by Brunnermeier and Niepelt (2019), "the issuance of CBDC would simply render the central bank's implicit lender-of-last-resort guarantee explicit. By construction, a swap of CBDC for deposits thus would not reduce bank funding; it would only change the composition of bank funding."

The implementation of such an approach could remove the need for frictions and limits to the holding of CBDC and could thereby harness the full efficiency gains associated with CBDC. However, this would require a material change to the SARB's lender of last resort policy. Given that large scale runs could occur rapidly, SARB would need to stand ready to offer unlimited amounts of liquidity to the banking system. Furthermore, banks may not have sufficient government securities as collateral to pledge, so additional forms of collateral would need to be included to ensure that banks are able to access a sufficient amount of SARB funding.

Such an approach would imply a large philosophical shift from the current Basel III approach to mitigating liquidity risk. Basel III encourages banks to self-insure through the adequate holdings of high-quality liquid assets to mitigate deposit runs. However, under the framework discussed above, SARB would need to stand ready to insure banks against liquidity risk. Such insurance may never need to be called upon. Indeed, simply by stating its intention to insure banks against such a risk, the risk may never materialize. But, this would only be possible if the SARB's commitment is considered credible, with the credibility requiring a clear shift in policy to operationalize the commitment.

Such a commitment would also remove the discipline that banks face from a potential depositor run. By removing a key source of bank failure, the SARB could in fact spur excessive risk taking by banks. Thus, alternative oversight mechanisms may need to be put in place to ensure that banks manage risks appropriately and that there isn't an overreliance on SARB funding. One option in this regard is to charge a punitive interest rate on SARB lending above a certain volume.

3. Policy Discussion

We suggest that SARB should carefully consider the policy outcomes that it hopes to achieve with CBDC. We suggest two potential aims: 1) targeting primarily financial inclusion of currently non-banked, relatively low-income individuals; or 2) widespread adoption of CBDC across the entire economy, achieving not only financial inclusion but significant efficiency gains associated with a faster, cheaper, and more secure payments infrastructure. These two options are a simplification for illustrative purposes. Each has materially different implications for the design of the CBDC and the impact on the broader financial system.

To achieve the benefits associated with the first option, while mitigating the impact on the broader financial system, CBDC could be introduced with a specified holding limit or a steeply increasing fee structure that would strongly disincentivize holding large sums of CBDC. In this way, bank deposits would not be diverted into CBDC *en masse*. However, lower income consumers or micro enterprises with small deposits could benefit from a low cost, safe payments instrument.

Alternatively, under option 2, wherein large corporations and a full range of households could adopt CBDC for use in daily transactions and for use as a store of value, the potential impact on bank deposits would be much larger. Under this scenario, the most practical approach would likely be for SARB to refinance bank deposits that flow into CBDC. While this may seem a trivial issue, it does have material implications for the current refinancing operations of the SARB. In particular, it would call for a significant re-think of the SARB's collateral framework. Under the current framework, banks are only allowed to offer government bonds to the SARB as collateral in the main repo auction. If bank funding requirements were to increase substantially, this would call for a much larger holding of government bonds by banks. Such an outcome would be undesirable for two reasons. First, it would result in banks channeling productive credit away from the private sector, thereby distorting capital allocation decisions by banks. Second, it would further concentrate the exposure of banks to the sovereign, exacerbating the financial stability risks associated with the bank sovereign nexus.

Therefore, under this approach the SARB may have to consider accepting alternative forms of collateral under its main repo. For example, banks could offer securitised mortgage or vehicle loans. Nevertheless, such loans would be subject to an appropriate haircut based on their riskiness given that the SARB aims not to expose itself to credit risk in the provision of financing to the banking sector. If the deposit outflows from the banking sector are sufficiently large, then banks may not have enough collateral to post in order to refinance the total deposit outflows and maintain their balance sheet size. In this case, banks would be faced with the option to reduce lending or access further funding from capital markets. Market based funding typically comes at a higher cost than that of deposit funding, so this approach would increase bank funding costs. However, funding from the SARB would likely also raise funding costs relative to those paid on a transactional deposit.

The consequences of the above discussion highlight that a widely adopted CBDC could increase commercial bank reliance on the SARB for funding and could drive up bank funding costs. This in turn could raise the cost of borrowing. During times of crisis, deposit flight out of banks and into CBDC could be large and rapid, which would call for SARB to stand ready to refinance these deposits to avoid a situation in which the aggregate banking system faces an illiquidity spiral. Recall that at present deposit flight generally can only occur between banks, but not out of the banking system at large (except in the case of a run into cash, but there are practical limits on the ability of businesses and households to do this). However, if the run on the banking system is large enough, collateral haircuts (and potentially a deteriorating quality of collateral during a crisis episode) may limit the SARB's ability to fully refinance deposit flight from banks. In this case banks would be reliant on market-based funding, which itself tends to be costly and unreliable during times of crisis.

Therefore, as a backstop measure, we suggest that limitations on CBDC holdings and a tiered fee structure be incorporated into the design of a CBDC. Even if these measures aren't a permanent feature of CBDC, we believe that they should be available for authorities to use as a macro prudential tool in times of crisis. The degree to which such frictions in the use and holdings of CBDC are implemented will in turn influence the extent to which CBDC can be considered a true substitute for the traditional payments system and in turn will affect the efficiency gains that can be extracted from CBDC. Recall also that a financial stability argument in favour of CBDC is that it provides an alternative payments system that can be used when the traditional payments system faces outages. However, this argument only holds if CBDC can be widely used with no significant limitations. Thus, the authorities must confront the trade-offs inherent within CBDC in order to determine how best to use it, if at all. One piece of good news is that CBDC may generate financial inclusion benefits even if its design makes it unusable or unattractive to individuals or corporation with large transaction requirements. However, under this scenario, the poorest in society would have access to the most efficient and lowest cost payments system. While large corporates would continue to use a relatively inefficient and costly option. Under this 'parallel payments system' option SARB would need to consider introducing many of the key CBDC innovations into the current payment system to ensure that entities performing large transactions can do so cheaply and efficiently.

4. Conclusion

CBDC offers many potential benefits. One key objective of its introduction could be to create a more efficient and/or cheaper payments instrument than is currently available in the market. In this way financial inclusion and the efficiency of financial transactions can be enhanced. However, if CBDC is significantly better than the alternative digital payments option - that of transacting in commercial bank deposits - it could have a material impact on the supply of deposits to the domestic banking system. While the magnitude of this effect is difficult to estimate with any certainty, the potential for the disintermediation of banks raises the need for consideration around how this effect could be managed to avoid systemic risk. The paper outlines the nature of this risk and discusses some potential options available to manage it. We argue that these options should be carefully thought through before CBDC is introduced as the design features of the CBDC must incorporate the appropriate mechanisms to manage systemic risk.

We suggest that if the SARB is to introduce a CBDC, it should be clear what policy goals it hopes to achieve. SARB should decide whether its aims are primarily around financial inclusion, enhanced efficiency of the payments system or something else. Once that decision has been made, the CBDC design features required to meet that objective and maintain financial stability will be clearer. Furthermore, it will become easier to compare the costs and benefits of CBDC against alternative policy approaches to achieve the same goal.

References

Bank of England (2020). Central bank digital currency: opportunities, challenges and design. Discussion Paper.

Bindseil, U. (2020). Tiered CBDC and the financial system. ECB Working Paper 2351.

Brunnermeier, M. K., & Niepelt, D. (2019). On the equivalence of private and public money. *Journal of Monetary Economics*, *106*, 27-41.

Committee on Payments and Market Infrastructures, Bank for International Settlements. (2018). Central bank digital currencies. Available from: https://www.bis.org/cpmi/publ/d174.pdf

DNB. (2020). Central Bank Digital Currency. Occasional Studies. Volume 20-01. Available from: https://www.dnb.nl/media/espadbvb/central-bank-digital-currency.pdf

Dyson, B., & Hodgson, G. (2016). Digital cash: Why central banks should start issuing electronic money. *Positive money*.

Kumhof, M., & Noone, C. (2018). BoE Staff Working Paper No. 725 Central bank digital currencies - design principles and balance sheet implications.

Meaning, J., Dyson, B., Barker, J., & Clayton, E. (2018). Staff Working Paper No. 724 Broadening narrow money: monetary policy with a central bank digital currency.

Mersch, Y. (2018). Virtual or virtueless? The evolution of money in the digital age. In *Lecture at the Official Monetary and Financial Institutions Forum, London* (Vol. 8).

Risberg, P., and Segal, S. (2020). Central Bank Digital Currency, Design Choices, and Impacts on Currency Internationalization. CSIS policy brief. Available from: https://www.csis.org/analysis/central-bank-digital-currency-design-choices-and-impacts-currency-internationalization

TOPICAL BRIEFING: No. 2

South African Reserve Bank

Financial Stability Department

Part 1: The application of the International Association of Insurance Supervisors (IAIS) Holistic Framework to South Africa[5]

1. Introduction

During 2007 and 2008, the world experienced one of the most severe financial crises since the Great Depression of the 1930s. Due to the complexity and interconnectedness of the financial system and the fact that risks were not fully addressed on a system-wide basis, many institutions faced serious challenges, with some institutions even failing. The global financial crisis (GFC), in fact, revealed a lack of adequate tools and models to monitor risks emanating from the interconnected global financial system. Subsequently, a comprehensive programme of financial reforms was launched by the Group-of-Twenty (G20) countries to increase the resilience of the financial system. As part of this programme and specific to the insurance sector, the International Association of Insurance Supervisors (IAIS) also contributed through an initiative to identify global systemically important insurers (G-SII). These insurers are typically very large insurers in the insurance sector whose distress or disorderly failure can result in significant shocks to the financial system.

An assessment methodology was adopted by the IAIS in 2013 to support recommendations on the identification of G-SIIs which also included policy measures that could be applied to these entities. Since then, the IAIS approach to addressing systemic risk has evolved and the recognition that systemic risk does not only originate

5 The PA and SARB have member representation on IAIS committees and working groups that include the IAIS Executive Committee, Policy Development Committee, Macroprudential Committee, Macroprudential Monitoring Working Group, Macroprudential Supervision Working Group, Climate Risk Steering Group, among others.

from the disorderly failure or distress of G-SIIs but also from the collective exposures of insurers at a sector-wide level, gained more prominence. This recognition led to the design of the Holistic Framework, an integrated set of key elements aimed at assessing and mitigating both potential sources of systemic risk. The IAIS adopted this framework in November 2019 which took effect from 2020.

A key element of the Holistic Framework is macroprudential insurance supervision which is aimed at identifying and addressing both vulnerabilities and the build-up of systemic risk at the individual insurer level and the insurance sector as a whole. An important consideration is that vulnerabilities building-up in certain jurisdictions may have cross- jurisdictional implications that merit monitoring at the global level. Correspondingly, supervisors may gain a better understanding of underlying trends at the jurisdictional level when such trends are assessed at a global level. Hence, a jurisdictions macroprudential supervision processes and procedures should be proportionate to the nature, scale and complexity of its insurance sector's exposures and activities.

This note is *Part 1* (of a two-part series) that focusses on the key elements of the IAIS Holistic Framework and how South African can implement it to promote financial stability by enhancing frameworks to mitigate the potential build-up of systemic risk in the insurance sector. *Part 2* will provide a vulnerabilities assessment of the South African insurance sector. The risk assessment will take into consideration additional monitoring indicators identified by the IAIS as part of the Holistic Framework. The Financial Sector Regulation Act (FSRA) has key synergies with the IAIS Holistic Framework, explicitly stating the need for increased focus on building indicators to detect systemic risk in the insurance sector on both an entity and industry-wide level.

The note is structured as follows: Section 2 provides an overview of the key elements of the IAIS Holistic Framework and the IAIS committees involved in the Holistic Framework; Section 3 discusses the application of the Holistic Framework to South Africa and monitoring indicators available to South Africa based on its regulatory reporting requirements. Section 4 addresses other areas where the Holistic Framework could support the systemic importance of macroprudential insurance supervision by the SARB and PA and future considerations for South Africa. Section 5 provides a summary and conclusion.

2. Key elements of the IAIS holistic framework and committees

The Holistic Framework for the assessment and mitigation of systemic risk in the global insurance sector supports the IAIS mission of promoting effective and globally consistent supervision to protect policyholders and contribute to global financial stability. The Holistic Framework recognises that systemic risk may arise not only from the distress or disorderly failure of an individual insurer, but also from the collective exposures and activities of insurers at a sector-wide level.

2.1 Key elements of the Holistic Framework

The key elements of the Holistic Framework are: (i) Supervisory Material; (ii) Global Monitoring Exercise (GME); and (iii) Implementation Assessment. While each key element represents an essential building block in itself, the overall effectiveness of the Holistic Framework depends on the elements working together.

Source: IAIS

2.1.1 Supervisory material

Supervisory material is an enabler of the framework and consists of an enhanced set of supervisory policy measures for macroprudential purposes, designed to increase the overall resilience of the insurance sector and help prevent insurance sector vulnerabilities and exposures from developing into systemic risk. When a potential systemic risk is detected, supervisory powers of intervention enable a prompt and appropriate response.

The Insurance Core Principles (ICPs), a globally accepted framework for the supervision of the insurance sector, form the base of the IAIS supervisory material. In addition, the Common Framework for the Supervision of Internationally Active Insurance Groups (ComFrame) establishes supervisory standards and guidance focussing on the effective group-wide supervision of Internationally Active Insurance Groups (IAIGs).

The ICPs and ComFrame aims to protect policyholders and contribute to global financial stability through the maintenance of consistently high supervisory standards in IAIS Member jurisdictions. As part of the Holistic Framework, the IAIS revised certain ICPs and ComFrame materials by enhancing or adding supervisory policy measures specifically designed to assess and mitigate potential systemic risk building up in the insurance sector.

The Holistic Framework moves away from the previous binary approach, in which a set of pre-determined policy measures applied only to a small group of identified global systemically important insurers (G-SIIs). Instead, it promotes a proportionate application of an enhanced set of supervisory policy measures and powers of intervention for macroprudential purposes to a broader portion of the insurance sector through the ICPs and ComFrame.

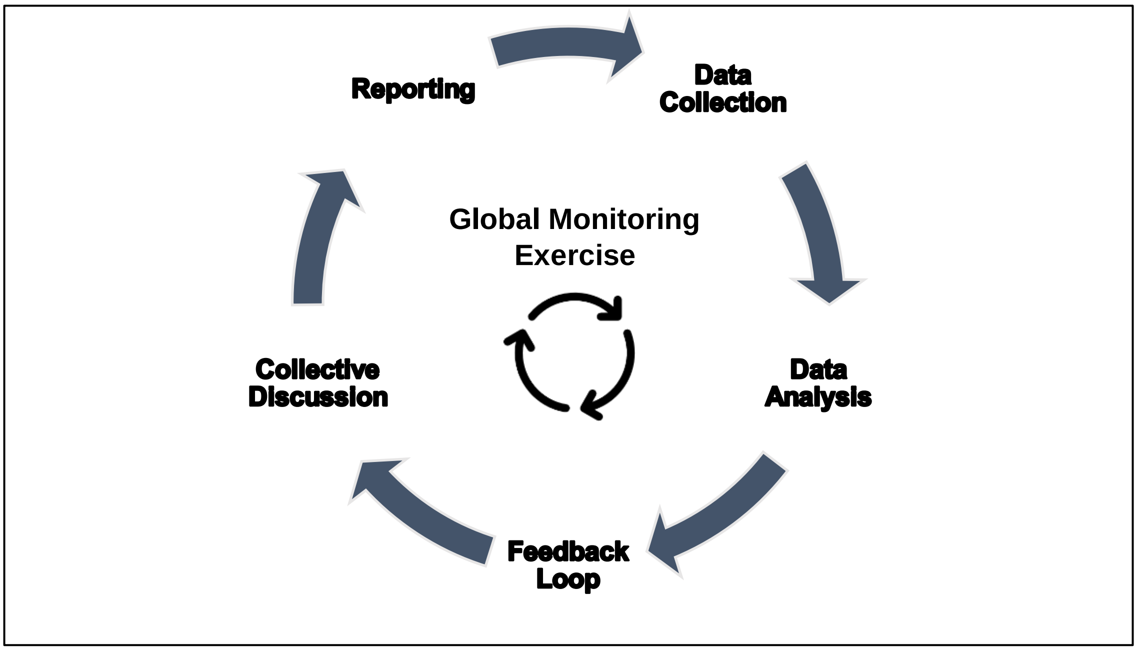
The policy measures include, but are not limited to:

(a) *On-going supervisory requirements* applied to insurers, targeted at key potential systemic exposures: liquidity risk, macroeconomic exposure and counterparty exposure;
(b) *Macroprudential supervision*, aimed at identifying vulnerabilities and addressing the build-up of systemic risk at the individual insurer and sector-wide levels;
(c) *Crisis management and planning*, which includes requirements on recovery and resolution planning, as well as the establishment of crisis management groups; and
(d) *Powers of intervention*, that require supervisors to have a sufficiently broad set of preventive and corrective measures in place to enable a prompt and appropriate response when a potential systemic risk is detected.

2.1.2 Global monitoring exercise (GME)

An IAIS global monitoring exercise is designed to assess global insurance market trends and developments and detect the possible build-up of systemic risk in the global insurance sector. This includes, at an individual insurer and sector-wide level, a collective discussion at the IAIS on the assessment of potential systemic risks and appropriate supervisory responses and reporting to the Financial Stability Board (FSB) on the outcomes of the global monitoring exercise (Refer to Annexure A for a schematic of the GME).

Process of the global monitoring exercise



Source: IAIS

The IAIS undertakes annual GME to assess insurance market trends and developments and determine any potential build-up of systemic risk in the global insurance sector. This includes an assessment of potential systemic risk arising from sector-wide trends regarding specific activities and exposures, but also the possible concentration of systemic risks at an individual insurer level (using an updated assessment methodology) arising from these activities and exposures.

(a) Individual Insurer Monitoring (IIM)

The IIM is aimed at assessing systemic risk stemming from an individual insurer's distress or disorderly failure, recognising that potentially systemic activities or exposures may become concentrated in an individual insurer, such that its distress or disorderly failure would pose a serious threat to global financial stability.

(b) Sector-Wide Monitoring (SWM)

The SWM is aimed at assessing sector-wide trends regarding specific activities and exposures and consists of both a qualitative and quantitative part. It is a complement to the IIM, and both their outcomes will feed into the IAIS's assessment of systemic risk as well as in the IAIS collective discussion. The SWM brings together existing IAIS efforts related to macroprudential surveillance and broader market surveillance, including the:

- IAIS Key Insurance Risk and Trends (KIRT) Survey: a voluntary, annual survey amongst IAIS Members about their qualitative assessment of risk;
- IAIS Global Reinsurance Market Survey (GRMS): a data collection amongst relevant IAIS Members, the results of which are annually reported to the public within the Global Insurance Market Report (GIMAR); and
- IAIS GIMAR: which provides an overview of trends and developments in global insurance markets along with a series of topical chapters which allow to develop a global view on relevant issues from the perspective of insurance supervisors.

The GME also includes a collective discussion by the IAIS of the assessment of potential systemic risk in the global insurance sector, at both a sector-wide and individual insurer level, and appropriate supervisory responses to systemic risk if it arises. The discussion of appropriate supervisory responses will include the consideration of enhanced supervisory policy measures and/or powers of intervention, taking into account the IAIS’ assessment of those supervisory policy measures and/or powers of intervention that have already been implemented.

The IAIS will share the outcomes of the GME each year with participants in the global monitoring exercise (participating insurers as well as participating IAIS Members), other IAIS Members, the FSB and the general public.

2.1.3 Implementation assessment

A key element of the Holistic Framework is the IAIS's assessment of the consistent implementation of the Holistic Framework supervisory material. It aims to promote globally consistent and effective implementation of the relevant supervisory material among its members. This is critical for supporting financial stability as the potential build-up of systemic risk may be global in nature, so should there be a globally consistent and effective application of policy measures aimed at assessing and mitigating these risks.

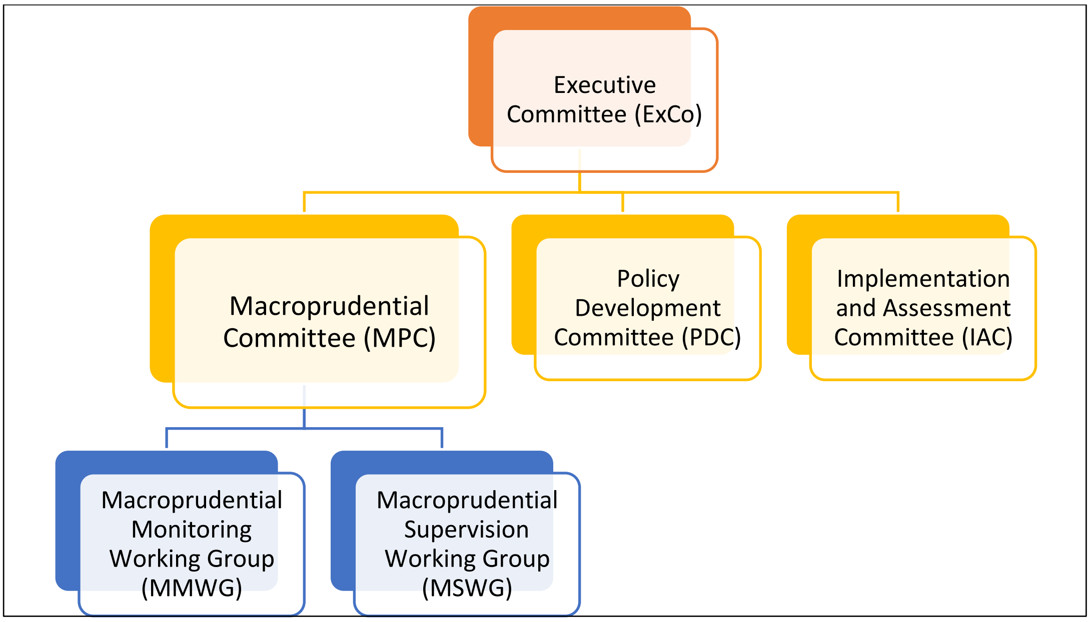
The IAIS's implementation assessment approach builds on existing methodology for assessing implementation of ICPs and ComFrame, while taking into account the specific nature of the Holistic Framework as a subset of ICP and ComFrame material that is relevant to the assessment and mitigation of systemic risk. Assessments are conducted in phases, that started with a baseline assessment in 2020 and moved towards more intensive jurisdictional assessments in 2021, which included targeted in-depth verification of supervisory practices.

The IAIS typically shares the outcomes of the Holistic Framework implementation assessments with the FSB and the general public.

2.2 Overview of the IAIS committees involved in the Holistic Framework

The IAIS organisational structure consists of various committees and working groups. The chart below highlights some of the committees that forms a big part of the Holistic Framework. Although there will be elements of the Holistic Framework discussed at all of these committees and at others not highlighted by this extract, most of the groundwork happens within the IAIS Macroprudential Monitoring Working Group (MMWG) and the Macroprudential Supervision Working Group (MSWG), highlighted in blue in the chart below.

IAIS Committees relating to the work of the Holistic Framework



Source: IAIS

The ***MMWG*** is responsible for:

- The coordination of the global monitoring exercise (including the IIM and SWM); and;
- Macroprudential assessment of trends, developments and risks to the financial stability of the global insurance sector.

The ***MSWG*** is responsible for matters relating to macroprudential supervision:

- Developing and maintaining related supervisory and supporting material related to the holistic framework; and
- Coordinating with other working groups on related matters, including the workstream on implementation assessment of holistic framework related material.

3. The application of the Holistic Framework to South Africa

3.1 South Africa's participation through the MMWG

The Prudential Authority (PA) participates in the GME and provides data through both the IIM and SWM data collection exercises. The PA in South Africa is the Group Wide Supervisor (GWS) for two identified IAIGs, namely Sanlam Limited and Old Mutual Limited. The PA coordinates the data collection from these entities as part of the IIM data collection process while also providing input into the SWM data collection process on an aggregate basis.

An annual update is provided to the Financial Stability Board (FSB with outcomes of the GME. In November 2022, the FSB will review the need to discontinue or re-establish an annual identification of G-SIIs in consultation with the IAIS and national authorities.

3.2 South Africa's participation through the MSWG

The Financial Stability Department (Finstab) participates in the IAIS MSWG in the following way:

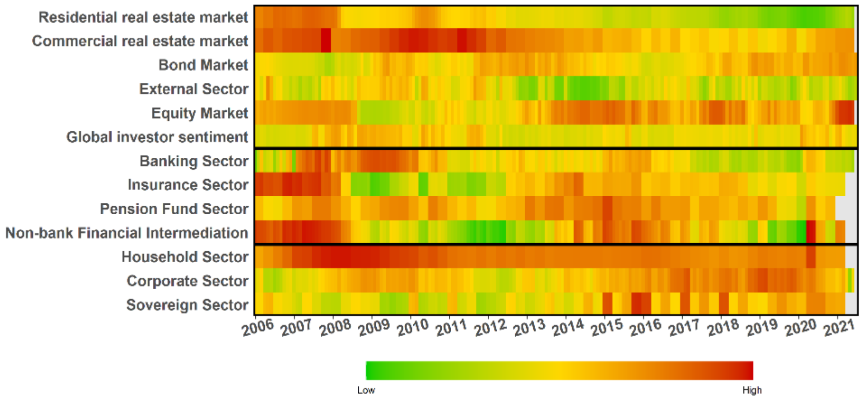
(i) Developing and maintaining supervisory and supporting materials related to the Holistic Framework and macroprudential policy, for example, the Application Paper on Macroprudential Supervision and the Application Paper on Liquidity Risk Management;
(ii) Providing input into other IAIS Committees such as the Macroprudential Policy Committee and the Executive Committee; and
(iii) Coordinating with other international bodies dealing with macroprudential policy of insurers and providing input where relevant.

3.2.1 Application of MSWG initiatives to South Africa

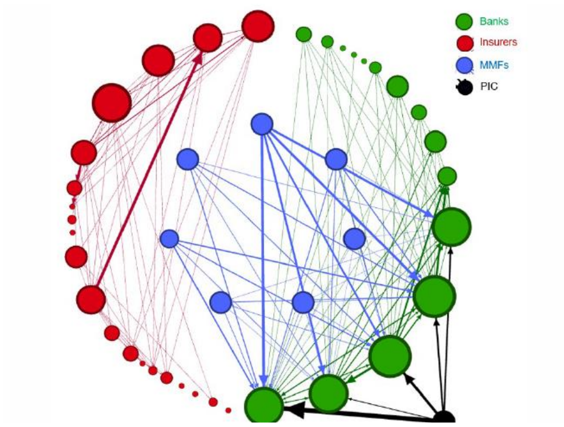
The SARB is in the process of implementing MSWG recommendations to enhance its systemic risk assessment framework for monitoring vulnerabilities across all sectors, including insurance. Such examples include building up insurance indicators in the

SARB’s Financial Stability Heatmap and systemic risk assessment framework as well as near-term considerations on interconnectedness (network analysis) between banks, insurers, money-market funds (MMFs), among others.

SA Financial Stability Heatmap



Financial System Interconnectedness



Source: SARB, IMF

3.3 Assessment of the SA insurance sector using selected Holistic Framework macroprudential indicators

The MSWG through supervisory material such as the ‘IAIS Application Paper on Macroprudential Supervision’ makes recommendations for institutions to monitor vulnerabilities within the insurance sector using various indicators.

The Application Paper includes macroeconomic and microeconomic data, indicators to support the assessment of liquidity risk, macroeconomic exposure, counterparty risk, as well as cross-sectoral indicators. Whilst not an exhaustive list, the IAIS indicators and data elements identified are deemed as adequately sufficient for assessing potential systemic risk in the insurance sector (for both individual insures and on a sector-wide level). Table 1 provides examples of these indicators and data elements.

Table 1: Examples of indicators and data elements identified in the IAIS Application Paper on Macroprudential Supervision

	Indicators	
Macroeconomic data examples to assess the exposure of the insurance sector to economy-wide factors	Solvency	Solvency Capital Ratio; changes in interest rate; changes in GDP growth, financial cycle; changes in inflation; changes in equity valuations; downgrades in credit ratings and outlooks (fixed income portfolios); financial strength ratings; insurance outlook (industry-wide): life insurance and property and casualty; changes in sovereign and major indices credit default swap (CDS) spreads; changes in real estate valuations; changes in equity prices (local and sectoral); duration mismatch; changes in exchange rates (impact on valuations); changes in volatility indices
	Profitability	Changes in financial revenue (impairments, investment losses from higher risk aversion; return on equity; changes in corporate dividends; changes in claims (life and non-life); changes in banking sector profitability; changes in combined ratio (loss ratio plus expense ratio); change in reinvestment rates versus guaranteed rates; changes in new lines of business; changes in premium income of different segments; changes in corporate sector profitability; performance of equity prices and expected profits of the national companies and the belonging area (insurance or non-insurance activities); changes in paid-up rates; changes in credit-to-GDP gap; Asset-to-GDP changes in household debt service ratio; changes in unemployment and corporate solvencies; changes in gross and net premium income; changes in household debt as a percentage of disposable income; growth in household disposable income
Microeconomic data examples to identify variances in insurance trends for more in-depth monitoring	General Data	Market share of the insurance sector; changes in insurance pricing and underwriting performances (individual and industry); changes in expenses; cancellations and policy lapses; changes in equity value (capital and surplus); changes in asset allocation (bonds, equities, cash, deposits, collective investment schemes etc); changes in shareholder and policy dividends; changes in capital requirements; capital contributions to shareholders; changes in interest rate and inflation; changes in morbidity and mortality rates; changes in assets and liabilities; Jaws ratio; changes in underwriting clauses; changes in legal coverages
	Data relating to specific and unforeseen events, such as pandemics, natural	Changes in the frequency and severity of events; changes in the solvency position; changes in the liquidity position; changes in profitability (realised gains or losses); changes in assets; changes in asset allocation; collateral requirements as a result of changing market conditions; changes in liabilities;

	Indicators	
	disasters, cyber-attacks: (individual and industry)	switch to marked-to-model valuations following illiquid markets; changes in operations and business continuity
Liquidity risk data examples to detect possible liquidity mismatches between assets and liabilities (individual and sector-wide level)	Assets side	Degree of liquidity of assets (insurance liquidity ratio); ratio of bank loan funds in asset portfolio; changes in sovereign bond investments; changes in equity investments; changes in investment funds; changes in asset composition (equities, debt, cash); decrease in corporate debt investments; changes in leveraged loans; changes in collateralised loan obligations; changes in financial guarantees; revaluations (real estate and equities); deterioration in credit quality of assets (due to credit rating downgrades); changes in derivatives holdings; changes in securitised assets; average duration of assets; changes in level 1, level 2 and level 3 assets (Fair Value Hierarchy); changes in sovereign bond yields and spreads; changes in interest rates; higher market volatility (VIX), higher margin calls on options or derivatives
	Liability side	Changes in claims (life and non-life), business interruption insurance, pandemic insurance; changes in claims (due to lower economic activity in motor, aviation, marine insurance); changes in net and gross incurred claims; insurance claim triangles; changes in net and gross written premiums (life and non-life); changes in direct premiums written for LOB; changes in surrenders and lapses; total borrowing by insurance companies (short term and long term); average duration of liabilities; changes in maturity or redemption structure of non-insurance liabilities; short term debt issued by insurers; financial guarantees on life insurance; line of credit or letter of credit drawdowns; litigation and reputational risk
Counterparty risk examples to assess the probability of default	Probability of default	Capital adequacy ratio of insurers; concentration: insurance sectors assets and liabilities holdings by life and non-life insurers; measuring credit quality of insurers (non-performing loans, share prices, implied CDS and CDS spreads, market capitalisation, changes in short term and long term debt, loans loss reserve); expected default frequencies to measure credit risk; profitability of insurers; derivatives holdings(local and foreign holdings); specific sectors holdings (e.g. financial or real estate), and geographical areas; market concentration risk (holding of equities and debt – local and foreign); reinsurance coverage (with local and foreign reinsurers); cross-sectoral holdings of other insurers, banks (and money-market funds), other financial institutions and corporates

		Indicators
Macroeconomic exposure data examples to monitor to assess the insurance sector's vulnerability to macroeconomic shocks (life and non-life)	Macroeconomic exposure	Changes in GDP growth; unemployment levels; inflation rate; interest rate; savings rate; changes in equity prices; changes in bond yields

Source: IAIS Application Paper on Macroprudential Supervision (2021)

The SARB is currently in the process of building up indicators to monitor vulnerabilities and systemic risk in the South African insurance industry. Table 2 captures the indicators specifically applied to South Africa for macroprudential analysis through its supervisory reporting data channel[6]. Additionally, the Prudential Authority uses a risk-based approach in the analysis of information, implying that not every single line item in the returns are monitored even though the data might be collected. A risk-specific and entity-specific approach is used in deciding on which indicator or element to focus on.

Table 2: Detailed list of monitoring indicators applied in the analysis of the South African insurance sector

Solvency	Solvency Capital Ratio;
Profitability	Changes in financial revenue (impairments, investment losses from higher risk aversion; changes in claims (life and non-life); Changes in combined ratio (loss ratio plus expense ratio); Changes in new lines of business (LOB); Changes in premium income of different segments; Changes in lapse rates
General data	Market share of the insurance sector; Changes in underwriting performances (individual and industry changes in expenses; cancellations and policy lapses; changes in equity value (capital and surplus);

6 South African insurance companies are obliged to submit quarterly returns in line with regulation in addition to audited annual returns.

	changes in asset allocation (bonds, equities, cash, deposits, collective investment schemes etc); changes in shareholder and policy dividends; changes in capital requirements; capital contributions to shareholders; changes in morbidity and mortality rates; changes in assets and liabilities; Jaws ratio;
Data relating to specific and unforeseen events, such as pandemics, natural disasters, cyber-attacks: (individual and industry)	Changes in the frequency and severity of events; changes in the solvency position; changes in the liquidity position; changes in profitability (realised gains or losses); changes in assets; changes in asset allocation; collateral requirements as a result of changing market conditions; changes in liabilities; switch to marked-to-model valuations following illiquid markets; changes in operations and business continuity
Assets	Degree of liquidity of assets (insurance liquidity ratio); ratio of bank loan funds in asset portfolio; changes in asset composition (equities, bonds, cash); changes in leveraged loans; changes in collateralised loan obligations; changes in financial guarantees; deterioration in credit quality of assets (due to credit rating downgrades); average duration of assets; changes in level 1, level 2 and level 3 assets (Fair Value Hierarchy)
Liabilities	Changes in claims (life and non-life), changes in claims (due to lower economic activity in motor, aviation, marine insurance); changes in net and gross incurred claims; insurance claim triangles; changes in net and gross written premiums (life and non-life); changes in direct premiums written for lines of business; total borrowing by insurance companies (short term and long term);

	average duration of liabilities; changes in maturity or redemption structure of non-insurance liabilities; short term debt issued by insurers; financial guarantees on life insurance; line of credit or letter of credit drawdowns; litigation and reputational risk
Probability of default	Capital adequacy ratio of insurers; concentration: insurance sectors assets and liabilities holdings by life and non-life insurers; expected default frequencies to measure credit risk; profitability of insurers; reinsurance coverage (with local and foreign reinsurers); cross-sectoral holdings of other insurers, banks (and money-market funds), other financial institutions and corporates

Source: PA, SARB

Looking ahead, a key priority of the SARB is to build more indicators into its systemic risk monitoring framework - to align itself with the work being done by IAIS member jurisdictions and other best practices to provide a more comprehensive view of potential systemic risk in the domestic insurance sector.

3.4 Holistic Framework baseline assessment that was performed on South Africa and the outcomes thereof

During the third quarter of 2020, the IAIS performed a Baseline Assessment (BLA) as the first phase in assessing the implementation of the Holistic Framework by supervisors. 26 Jurisdictions, including South Africa, took part in the BLA and is included in the aggregate report that was published by the IAIS in June 2021[7]. For South Africa, 20% of the Standards was assessed as being not observed but initiatives were already in motion to address these gaps. 50% Of the Standards was observed while 10% and 20% were largely observed and partially observed, respectively.

7 IAIS Aggregate Report on Holistic Framework Baseline Assessment Results

4. Other areas where the Holistic Framework could support the systemic importance of macroprudential insurance supervision by the SARB and the PA

(i) Stress testing of the South African insurance sector

As part of assessing systemic risks and vulnerabilities in the financial sector, the SARB recently extended its stress testing framework to cover the insurance industry by conducting an exploratory bottom-up (BU) sensitivity stress test of the South African insurance industry. Eleven non-life and eight life insurers participated in the exercise with coverage of 64% of the non-life sector as measured by total gross premiums and 69% of the life sector in terms of total assets. The risk types covered in the exercise were market risk and underwriting risk.

Overall, the insurance industry was found to be largely resilient to the identified shocks. For market risk, counterparty defaults had a material impact on the solvency positions of both life and non-life insurers. For underwriting risk, the increase in mortality stress parameter had the largest impact on life insurers while the large claims stress parameter impacted non-life insurers severely. Looking ahead, the Finstab will continue engaging with the PA and the insurance industry while taking into consideration IAIS work on the Holistic Framework and stress testing, in order to develop a more complex stress test exercise in the near or medium term. Future macroprudential stress test exercises of the insurance industry are envisaged to be based on forward-looking scenarios, incorporate elements of climate change risks, and focus more specifically on D-SIIs once these have been designated.

(ii) Climate risk

Similar to most global jurisdictions, the SARB and the PA are in the process of integrating climate risk into their general macroprudential and microprudential supervisory practices and their stress testing capability. There is increased focus enhancing its understanding of and translating the impact of the climate scenarios developed in the following areas: micro analysis; the identification of physical and transition risks; conducting balance sheet analysis/surveys to identify direct and indirect exposure of financial institutions and the possible credit, operational, market, underwriting and liquidity risks; understanding feedback loops between the insurance sector and other financial institutions; designing suitable scenarios and stress testing approaches following the approach outlined by NGFS; incorporating climate considerations into the SARB investment management framework and supporting markets for green financial instruments; among others.

There will continue to be increased focus on the implementation of the Task Force on Climate-Related Financial Disclosures (TFCD) recommendations, as guidance from international organisations such as the IAIS, FSB, IMF as well as continuous engagements with international central banks and experts on this matter.

(iii) South African domestically systemic important insurers (D-SII) methodology

The IAIS has replaced the methodology to determine global systemically important insurers (G-SIIs) with the Holistic framework, moving away from identifying G-SIIs. South Africa is in the process of finalising its proposed methodology to identify D-SIIs. A discussion paper was published for public consultation in 2020, which discusses the relevant indicators and sub-indicators used in the methodology. The indicators used in the methodology are size, interconnectedness, substitutability, and complexity. Each indicator is weighted and consists of a set of sub-indicators unique to the insurance industry and is in alignment with the Holistic Framework and FSB initiatives.

D-SII methodology for South Africa

Category	Indicator	Composite	Life	Non-Life
Size 40%	- Total assets	✓	✓	
	- IFRS Profit before tax	✓		✓
	- Gross written premiums	✓		✓
	- Number of individual policies		✓	✓
	- Number of group schemes	✓	✓	
Interconnectedness 30%	- Derivatives	✓	✓	✓
	- Gross written premiums ceded	✓	✓	✓
	- Intra-financial assets	✓	✓	✓
	- Intra-financial liabilities	✓	✓	✓
	- Reinsurance non-life	✓		✓
	- Reinsurance life	✓	✓	
Substitutability 20%	- Gross written premiums per business line	✓	✓	✓
	- Best Estimate Liabilities per business line	✓	✓	✓

Complexity 10%	Indicator	Comp Reins	Life cell Cap	Life Micro Insurer	Life primary	Life Re-insurer	Non-Life Cap	Non-life Cell Cap	Non-life Lloyd's	Non-life Primary	Non-life Reins
	Number of lines of business for life	✓	✓	✓	✓	✓					
	Number of lines of business for NL	✓					✓	✓	✓	✓	✓
	Number of cells		✓					✓			

Source: SARB

(iv) South Africa's Resolution Framework

The Financial Sector Laws Amendments Bill (FSLAB) that was promulgated in early 2022, sets out proposed amendments to the FSRA to introduce a new resolution framework. The resolution framework will apply to designated institutions which include all banks, non-bank SIFIs and their holding companies. D-SII insurers (excluding non-SIFIs) will fall under the resolution framework as per the FSLAB. The FSLAB designates the SARB as resolution authority responsible for developing resolution plans for and conduct the orderly resolution of a designated institution. The FSLAB also provides for the necessary resolution powers and tools to develop resolution plans and conduct an orderly resolution, including the stabilisation powers set out in the FSB's Key Attributes for effective resolution regimes.

5. Summary and conclusion

Macroprudential insurance supervision is a key element of the IAIS holistic framework in identifying and addressing both vulnerabilities and the build-up of systemic risk at the individual insurer level and the insurance sector as a whole. It is important that a jurisdiction's macroprudential supervision processes and procedures be proportionate to the nature, scale and complexity of its insurance sector's exposures and activities. The IAIS will continue with the execution of the Holistic Framework through the collection of individual data from selected insurers as well as sector-wide data from supervisors to support the identification and assessment of systemic risk.

The IAIS Holistic Framework supports global financial stability by providing sound guidance to enhance frameworks to mitigate the potential build-up of systemic risk in the insurance sector. In South Africa, the FSRA has alignment with the Holistic Framework, explicitly stating the need for increased focus on building indicators to detect systemic risk in the insurance and other sectors (entity and industry level). Domestically, the SARB will continue with the implementation of the key elements of the Holistic Framework and Application Papers on a broader scale to enhance its risk assessment frameworks and close critical data gaps. Additional key focus areas will be interconnectedness between banks, insurers and money market funds and bottom-up stress testing in the domestic insurance sector, including scenarios for climate risk.

TOPICAL BRIEFING: No. 3

Financial Stability Department

SOUTH AFRICAN RESERVE BANK

Part 2: Vulnerabilities assessment of the South African insurance sector using elements of the International Association of Insurance Supervisors (IAIS) Holistic Framework and general indicators

1. Introduction

This note is *Part 2* (of a *two-part series*[8]). *Part 1* focussed on the application of the IAIS Holistic Framework to South Africa in enhancing financial stability by identifying vulnerabilities and mitigating the build-up of systemic risk in the domestic insurance sector. This paper *(Part 2)* provides a vulnerabilities assessment of the South African insurance sector, taking into consideration additional monitoring indicators identified by the IAIS as part of the Holistic Framework.

The note is structured as follows. Section 2 provides an overview of the South African insurance sector; Section 3 discusses a vulnerabilities assessment on the domestic insurance sector using proposed indicators from the IAIS Holistic Framework; Section 4 identifies data gaps for the SARB from the Holistic Framework's list of indicators and elements; and Section 5 concludes the paper.

[8] Part 1: The application of the International Association of Insurance Supervisors (IAIS) Holistic Framework to South Africa: authored by Videshree Rooplall (Financial Stability Department) and Christiaan Henning (The Prudential Authority).

2. An overview of the South African insurance sector

The South African insurance sector is robust and serves an important role within the broader economy. Under the Insurance Act 18 of 2017 (the Act), the Prudential Authority (PA) regulates and supervises this sector with an objective to protect policyholders. Given the size, complexity, and interconnectedness of the insurance sector with other financial and non-financial institutions, it is deemed as systemically important, and thus consequential to financial stability. A new risk-based approach for the prudential supervision of the insurance sector became effective on 1 July 2018. This approach was built on the international insurance best practice and enabled through the Act and the related Prudential Standards. It is also largely based on the European Union risk regime, Solvency II, as well as relevant developments from other international jurisdictions, considering specificities in the South African context.

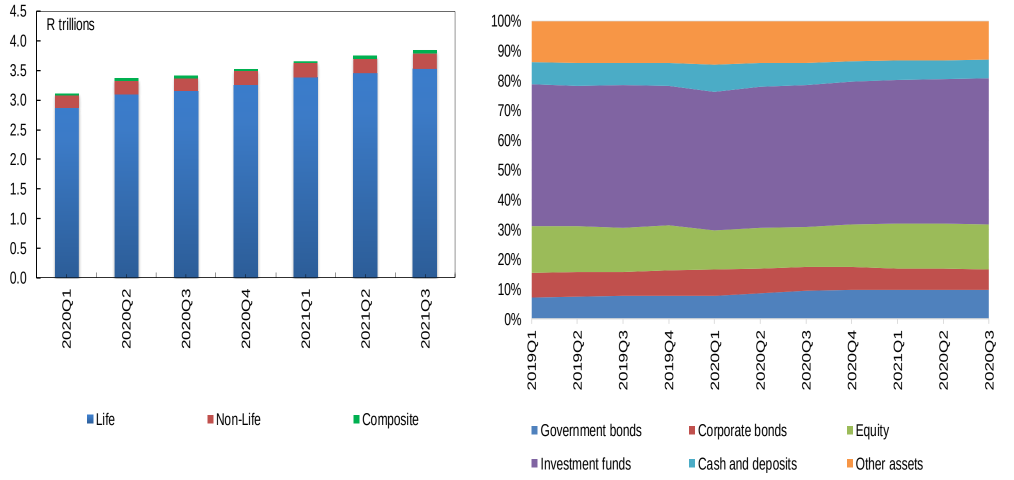
The approach is based on three Pillars, namely quantitative requirements (Pillar I); governance of the insurer and supervisory activity (Pillar II); and supervisory reporting and public disclosure (Pillar III). The framework was also designed to create a macroprudential layer to reduce systemic risk. This aligns with the Group-of-Twenty's (G20) agenda of moving towards a system of financial soundness and stability for financial services institutions and the IAIS Holistic Framework.

The domestic insurance sector forms part of a critical source of funding for both the private and public sector evidenced by the distribution of assets across various asset classes (Figure 1). Total assets as at the end of Q3:2021 stood at R3.8 trillion, compared to banking sector assets of R6.6 trillion. There are currently 155 insurance companies authorised by the Financial Services Board[9] to conduct business in South Africa[10]. Although almost half the size of the banking sector in terms of assets, the South African insurance sector is a potential source of systemic risk. The sector is highly concentrated with over 90% of the assets within the life insurers segment. Since 2015, the top five life insurers have consistently accounted for more than 70% of the life insurance business. Non-life insurance has been less concentrated. The top five non-life insurers account for 46% of gross written premiums (GWP). However, this

9 This includes life and non-life insurance companies.
10 https://www.resbank.co.za/en/home/what-we-do/Prudentialregulation/insurers-list.

segment of insurers is characterised by product concentration, with over 80% of the premiums coming from motor, property, and liability business.

Figure 1: Insurance assets and distribution in South Africa[11]

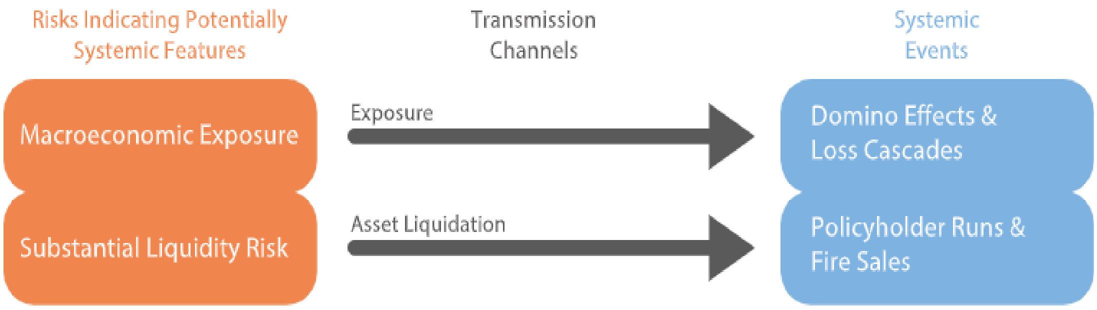


Source: PA

Against this backdrop, an insurer can transmit risk to the financial system and the economy through the insurance products it underwrites, that is, systemic risk insurance product features (SRIPF). Figure 2 displays the possible channels through which systemic risk may be transmitted, namely the macroeconomic exposure channel; and the substantial liquidity channel. For example, in 2007, American International Group Inc. (AIG) had the potential to cause systemic risk through the macroeconomic exposure channel from the sale of credit default swaps (CDS) to financial institutions. This high interconnectedness would have exposed purchasing financial institutions to credit risk had AIG failed.

11 A composite insurer provides both life and non-life insurance products.

Figure 2: Insurance product features can create macroeconomic exposure and substantial liquidity risk



Source: FSI connect

Macroeconomic exposure and the asset liquidation channel are the two transmission channels through which systemic events may occur. Domino effects, loss cascades, policyholder runs, and fire sales represent potential events through which systemic risk may manifest themselves. A domino effect refers to the collapsing of financial institutions, arising from the financial stress in one or more insurer. This effect is amplified in an environment that is highly interconnected. Like the domino effect, loss cascades may be exacerbated through interconnectedness, for example, via shareholding between different institutions. Policyholder runs and fire sales[12] are systemic events that may occur through the liquidation channel. Policyholder runs are characterised by herd behaviour with policyholders surrendering policies simultaneously. For example, Ethias, an insurer in Belgium, experienced a run on its policies in 2008 due to asset-liability mismatches and relatively high exposure (5%) to a Belgian bank.

Evidence in the case studies of both AIG and Ethias have highlighted the need for increased monitoring of the insurance sector for financial stability purposes. This also brings to the forefront the importance of building indicators and data elements to detect the build-up of systemic risk and vulnerabilities in the sector earlier.

[12] Refers to the selling of a security at a price below the market value.

3. The application of the IAIS Holistic Framework to South Africa using selected indicators

As discussed in *Part 1* of this two-part series, the IAIS Application Paper on Macroprudential Supervision, assesses systemic risk across six broad categories, namely size, global activity, interconnectedness, asset liquidation, and substitutability. Each category comprises of a list of indicators and elements.

The indicators and data elements identified by the IAIS are not meant to be an exhaustive list and are institution- and country specific, depending on market structures and financial conditions. For example, during the Covid-19 crisis, business interruption insurance and pandemic insurance data, was specifically requested by the PA, forming part of the stress testing exercise in the Own Risk and Solvency Assessment (ORSA) submissions. To this end, insurers whose business models were materially impacted by the Covid-19 crisis submitted out-of-cycle ORSA's. While those not materially impacted submitted specific COVID-19, business interruption and impact of strikes.

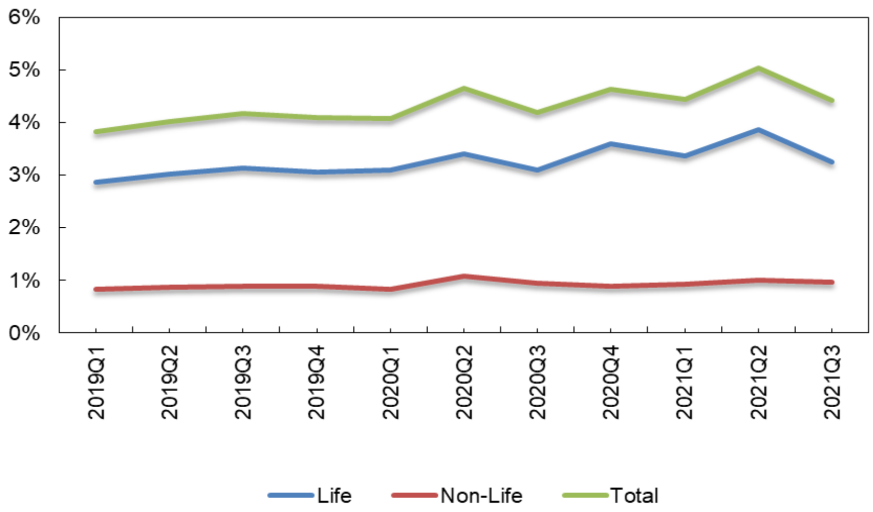
The section below consists of a trend analysis of the South African insurance sector, using data (for the period Q1:2019 to Q3:2021) for select indicators proposed by the IAIS Holistic Framework as well as other general indicators of the SARB and PA. Definitions of each indicator are included in footnotes below.

(i) Insurance Penetration[13]

Total insurance penetration moderated between Q1:2020 and Q3:2021 against a backdrop of weak economic growth, leaving households and corporates vulnerable. Growth in Q2:2021 was due to strong premium growth coupled with subdued economic growth. In Q3:2021, total insurance penetration was 4.4%, in line with the 3-year average and significantly higher than that of the emerging market average (3.3%). Life insurance penetration grew from 2.9% in Q1:2019 to 3.2% in Q3:2021, non-life insurance penetration also increased in that period, albeit slower, from 0.8% to 1.0%, reflecting the growing importance of the insurance sector for financial stability.

[13] Insurance penetration is calculated as the ratio of gross written premium to GDP and provides a measure of market development for the insurance sector. On average, the higher the ratio the more developed the sector.

Figure 3: Penetration Ratio[14]



Source: PA

(ii) Concentration Ratio[15]

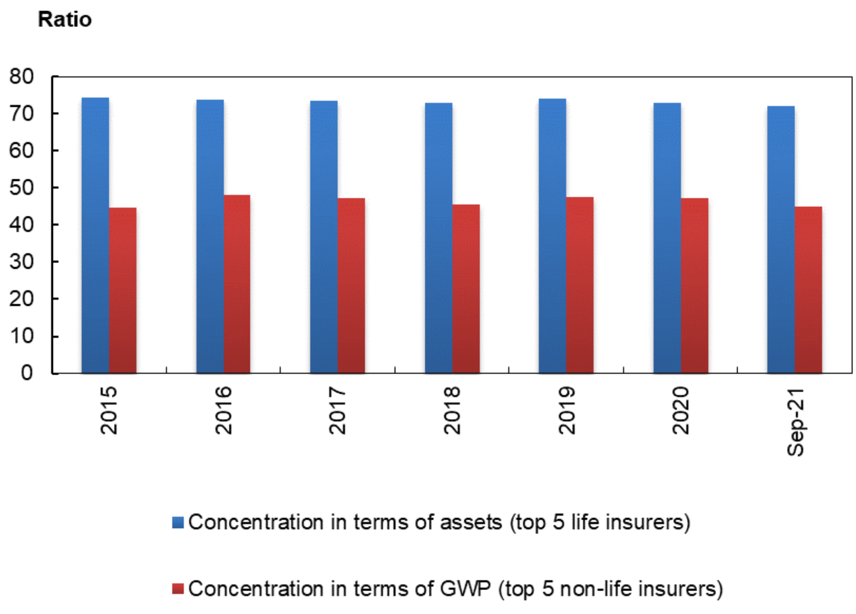
The concentration ratio for life and non-life insurance remained relatively unchanged since 2015 at relatively high levels. The life insurance concentration ratio in Q3:2021 stood at 72% while that of non-life stood at 45%. These high concentration levels are seen as a potential vulnerability for the sector and will continue to be monitored closely for potential spillovers, particularly from the top five life insurers to other insurers, and markets during times of market turmoil. The high concentration levels in the general South African financial system makes an understanding of the channels of interconnectedness even more pertinent. In this regard, at the end of 2021, the SARB initiated a multi-year industry-wide project to develop a monitoring and assessment framework for interconnectedness and concentration in the South African financial system. Additionally, in terms of the SARB's new responsibilities as resolution authority, an understanding of interconnectedness helps to estimate the likely contagion effects of institutional failures (including insurers) as well as the systemic impact of specific resolution actions, such as bail in.

---

14 Calculated on a quarterly basis. Annual figures for 2021, 2019 and 2021 are 16.1% and 17.5%, respectively.

15 The concentration ratio for life insurance is calculated using top five life insurers divided by total life assets, while non-life insurance is calculated using top 5 non-life insurers divided by total non-life gross written premium. This indicator measures the extent of competition in the market. The higher the concentration ratio the less competitive the sector is and the opposite holds.

Figure 4: Concentration Ratio



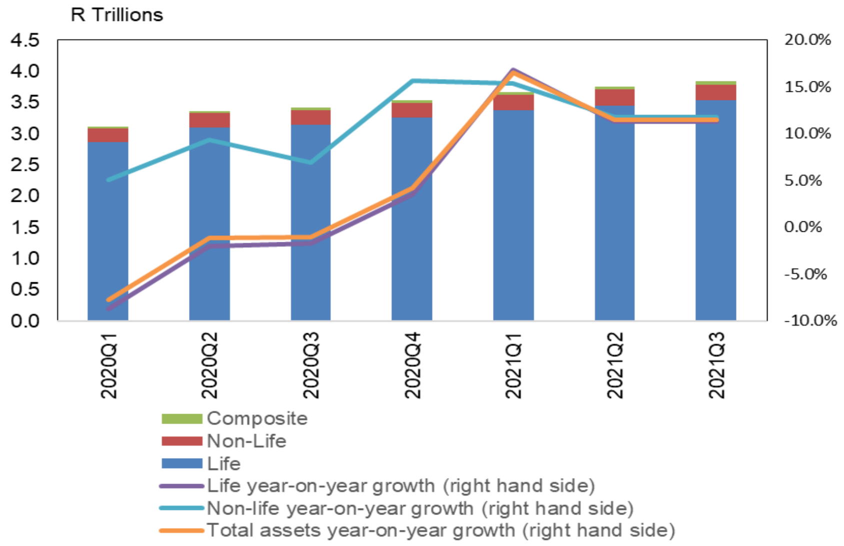
Source: PA

(iii) Changes in assets[16]

Despite being weighed down by low economic growth, lockdown restrictions in 2021 and increasing claims, total assets continued to grow since Q4:2020. Life insurers' assets accounted for more than 90% of total assets. Asset growth in the life segment grew by 11% year on year to R3.8 trillion in Q3:2021 compared to the 9% decline observed in Q1:2020. Changes in assets for life and non-life segments have differed, with life assets being more severely impacted by the pandemic than non-life assets, largely due to the varying distribution of assets between the two segments. Life assets have a significantly higher weighting in investment funds than non-life (since Q2:2020 an average of 50%, compared to non-life average of 10%), and this asset type was substantially impacted by fire sales at the beginning of the pandemic. While the growth in assets may reflect capital asset growth, it is also indicative of the insurance sector's ability to reinvest its premiums in assets, thereby generating additional investment income.

16 The change in assets indicator is calculated as year-on-year growth in total assets and provides a measure to assess the insurers' ability to increase assets through operations and investments. All else equal, high growth in assets is seen as favorable. The insurance sector is a key source of funding for financial and non-financial corporation and growth in assets is important for financial stability.

Figure 5: Change in Assets



Source: PA

(iv) Assets-to-GDP[17]

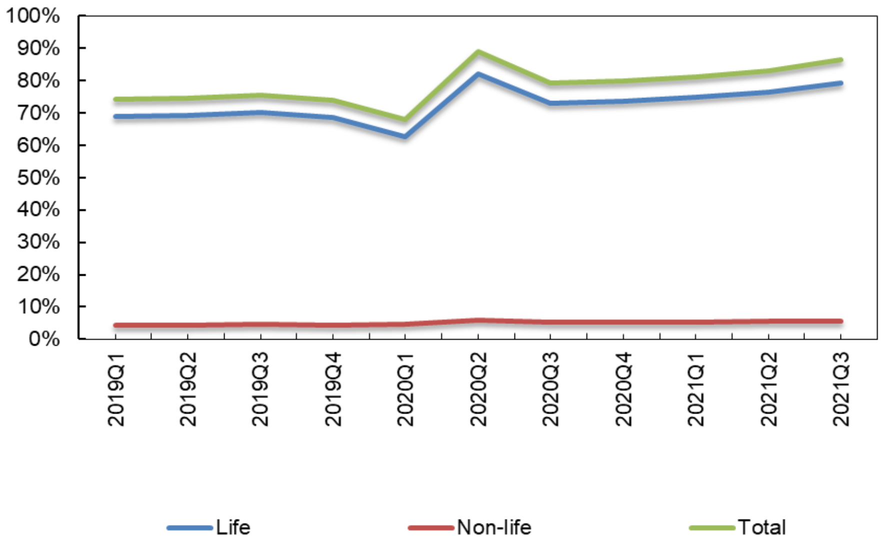
The insurance assets-to-GDP ratio has been on an upward trend for the period since Q3:2020, largely due to strong growth in assets of insurers as well as lackluster economic growth. The total assets-to-GDP ratio for Q3:2021 was 86%, indicating the sectors high contribution to the economy. The ratio, however, remains relatively lower than the bank's assets-to-GDP ratio of 148% for the same period. Non-life assets-to-GDP ratio averaged 5% since the first quarter of 2021. In terms of global rankings, in 2019 South Africa ranked favourably among the likes of Germany, Netherlands, and South Korea[18].

---

17 Assets-to-GDP is calculated as the ratio of total assets to GDP, and it provides the relative importance of insurers to the size of the economy. A high ratio shows relatively high importance of the sector to the economy.

18 https://www.theglobaleconomy.com/rankings/insurance_company_assets/

Figure 6: Assets-to-GDP



Source: PA

(v) Changes in asset allocation[19]

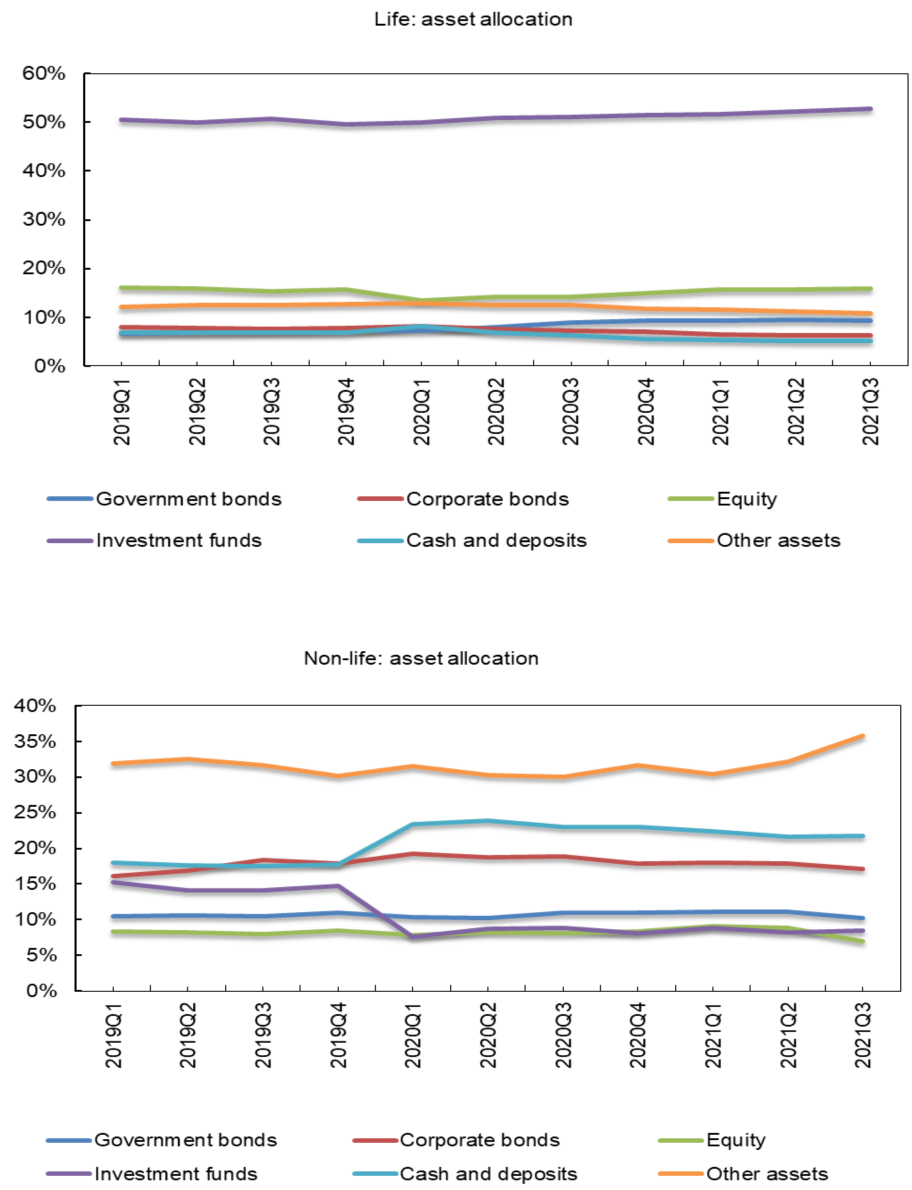
Figure 7 shows a breakdown of the relative asset allocation for insurers. Asset allocation remained stable overall, in line with the long-term average since Q1:2019. Investment funds (supported by a look-through approach)[20] averaged 48% and equity averaged 15% from Q1:2019 to Q3:2021. While asset allocation has remained relatively unchanged, in the current environment of a search for yield, insurers have increased their exposure to government bonds from 7% in Q1:2019 to 10% as at the of end Q3:2021.

---

19 Changes in asset allocation measures the extent to which assets are diversified across different investment type e.g., equity, bonds etc.

20 The look-through approach requires insurers to assess the risks of the assets underlying the investment vehicle and apply capital requirements for the relevant components of market risk to the underlying assets.

Figure 7: Asset allocation

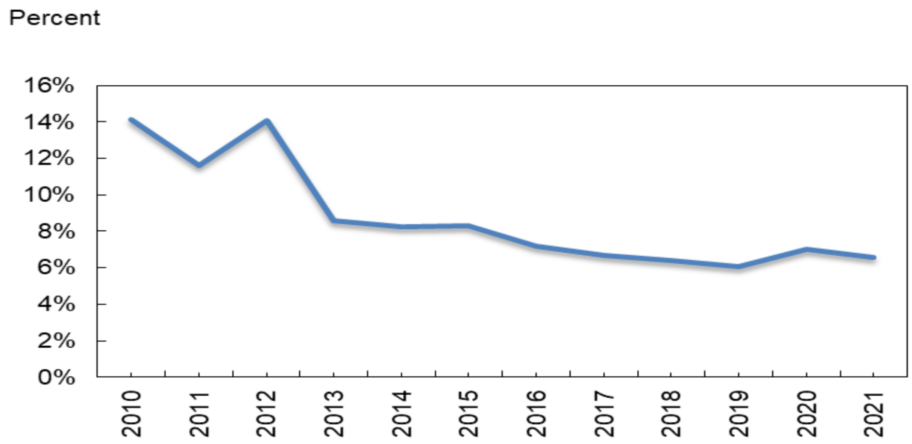


Source: PA

Life insurance assets are mostly allocated to investments funds and equities (Q3:2021: 52% and 16%, respectively) while non-life assets are mostly invested in reinsurance recoverables and cash and deposits (Q3:2021 36% and 22%, respectively). These relatively high weightings of life insurers assets towards investments funds and equities are seen as a moderate vulnerability for insurers, as these types of asset classes tend to exhibit volatility during times of economic uncertainty.

Figure 8 indicates the relative holdings of government bonds by insurers relative to other institutional investors. Since 2010, insurers' holdings of domestic government bonds moderated from 14% in 2010 to 6.6% at the end of 2021. The low interest rate environment and search for yield in 2020 however, resulted in a slight increase in holdings by insurers since 2019 to 7.0% in 2020.

Figure 8: Holdings of Domestic Government Bonds: Insurance[21]



Source: National Treasury

As witnessed in 2020, large scale market related selloffs tend to depress asset prices, affecting trading and funding markets, thereby causing firms with similar asset holdings to incur losses. Although relatively small when compared to the weighting of other assets, both life and non-life insurers have similar exposure to government bonds of around 10%. Should sovereign risk continue to rise, bond spreads may widen, with negative implications for profitability through reduced investment income.

(vi) Changes in gross written premium[22]

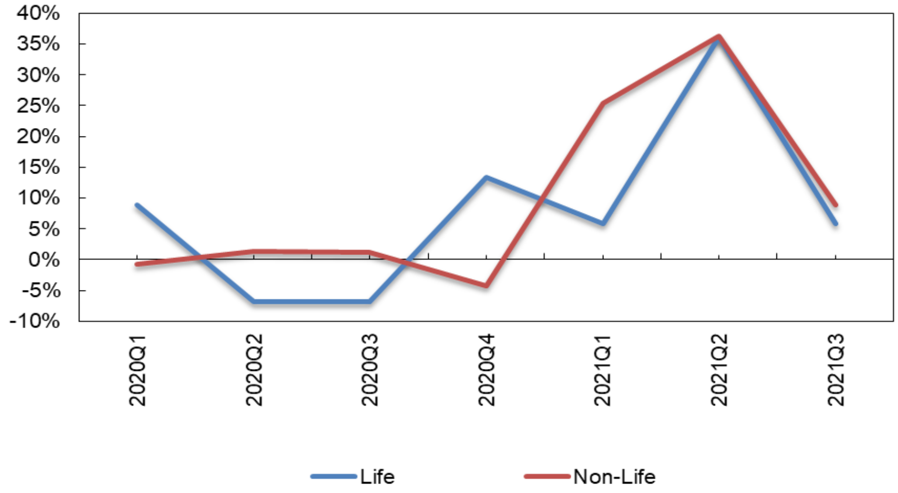
The Covid-19 lockdown restrictions dampened growth for both life and non-life insurers, particularly because of the sector's reliance on face-to-face interactions for sales. Despite this, the sector has been able to innovate showing signs of solid growth

---

21 This data is sourced from the National Treasury while data used in (v) are sourced from the Prudential Authority.
22 Changes in gross written premium (GWP) is calculated as a year-on-year growth in GWP, and it assesses the ability of insurers to grow their business. An increase in GWP indicates a growing insurance business.

in GWP for life and non-life businesses. For instance, one South African insurer was able to implement its digitization initiative three weeks after the pandemic commenced, thus increasing its client base. By the third quarter of 2021, life and non-life GWP increased by 5.86% and 8.94% year-on-year, respectively. The average annual growth rate for the life and non-life segments between 2011 and 2020 was 7.3% and 8.2%, respectively. Despite the growth in GWP, slow economic growth remains a key concern for the sector as it has the potential to negatively impact future demand for insurance products.

Figure 9: Changes in Gross Written Premium (GWP)



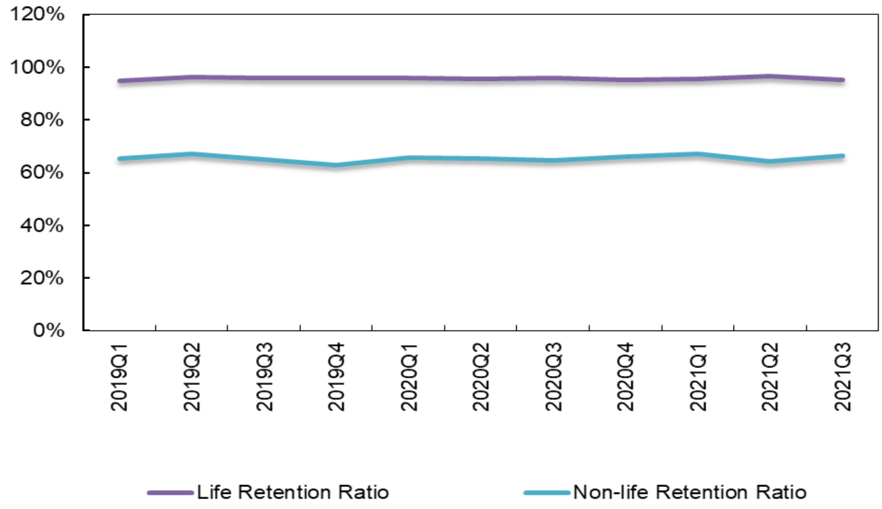
Source: PA

Retention Rate[23]

Retention rates for both life and non-life segments have remained relatively unchanged over time, indicating that overall, there have not been significant movements of insurance obligations to reinsurers. Life insurance retention rate averaged around 96% between Q1:2019 to Q3:2021 and 65% for non-life insurers. Between Q2:2021 and Q3:2021, an observable 2% of non-life insurance obligations were however, transferred to reinsurers.

[23] The ratio is calculated as the ratio of net written premiums to GWP. Net written premiums are calculated as the GWP less reinsurance premiums. It provides a measure of how much of the insurance risk is retained by the insurer and not passed to reinsurers. The thresholds differ significantly between life and non-life segments, life segment typically has higher rates around 90% while non-life segment's rates range between 40% and 80%, depending on the products.

Figure 10: Retention Rate



Source: PA

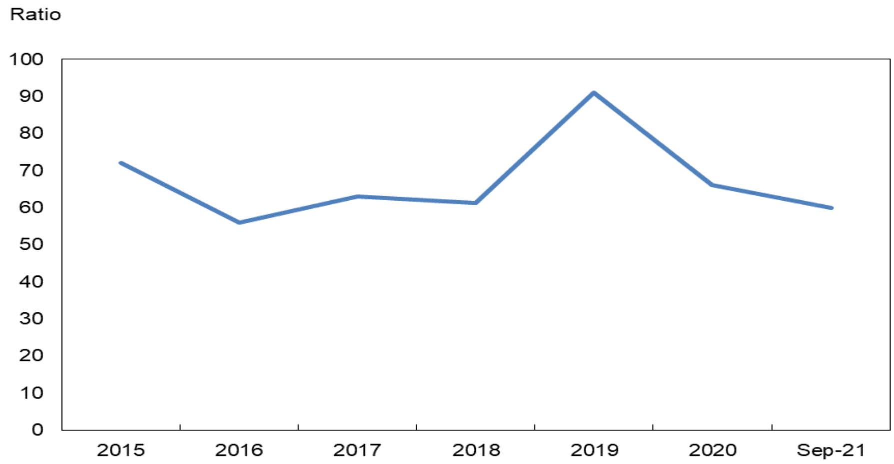
(vii) Life: Individual Lapse Ratio[24]

The life lapse ratio moved downwards from 91% in 2019 to 60% in the third quarter of 2021 as policyholders opted to retain their policies. The improvement in the ratio also reflected a change in consumer behavioural patterns as the realisation of the importance of life policies increased during the pandemic. From the perspective of insurers, even at the height of the pandemic, the lapse ratio remained lower than pre-Covid levels, implying that insurers were able to retain clients under challenging conditions. Currently, this indicator shows limited signs of systemic risk.

---

[24] Lapse ratio is expressed as a percentage of lapsed policies to net written premium during the period. The indicator is a measure of how well the insurer can retain clients and grow the business. High lapse ratio can be an indication of challenging trading conditions and/or poor quality of Insurance underwriting.

Figure 11: Individual Lapse Ratio



Source: PA

(viii) Changes in claims[25]

Life insurance net claims[26] increased by 4.8% year-on-year in September 2021 and declined by 9.5% quarter-on-quarter in Q3:2021, reflecting the impact of a decrease in vaccine apathy among policyholders and increased supply in vaccines. According to the Association for Savings and Investment in South Africa (ASISA)[27], an alarming trend of fraudulent and dishonest claims in the life segment pose a risk to this segment of insurers.

Non-life insurance claims increased sharply by 218.8% year-on-year in September 2021 and 145% quarter-on-quarter in Q3:2021, mainly due to claims arising from the social unrest in July 2021. Excluding claims originating from the state-owned insurer South African Special Risk Insurance Association (SASRIA)[28], non-life insurance claims increased by 50% quarter-on-quarter in Q3:2021, largely owing to claims that arose from the July 2021 social unrest incident. This segment also reported/registered increased incidents of flooding between December 2021 and

---

25 Changes in claims are calculated as year-on-year changes in net claims. The indicator is a measure of profitability. High claims can erode profitability of insurers and changes in this indicator need to be monitored closely. Life insurance claims have increased significantly weighed down by the pandemic. Non-life insurance has also increased because of increases in motor, property and liability insurance which together amount to 80% of the non-life business.

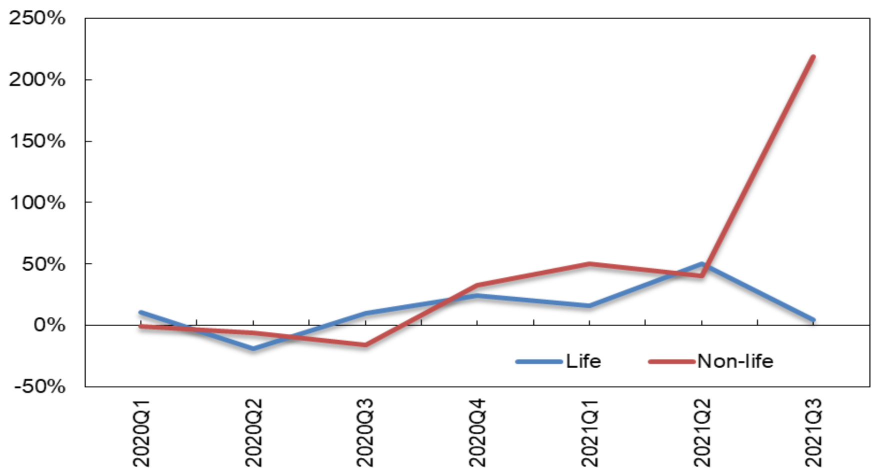
26 Claims and net claims are used interchangeably.

27 https://www.asisa.org.za/media/yr2ntbfn/20210823_life-insurers-report-significant-increases-in-funeral-insurance-fraud-for-2020.pdf

28 SASRIA is the sole provider of insurance cover against public unrest of the nature seen during July 2021.

January 2022 and remains vulnerable to both further flood events and social unrest incidents.

Figure 12: Changes in Claims



Source: PA

(ix) Combined Ratio: Non-life[29]

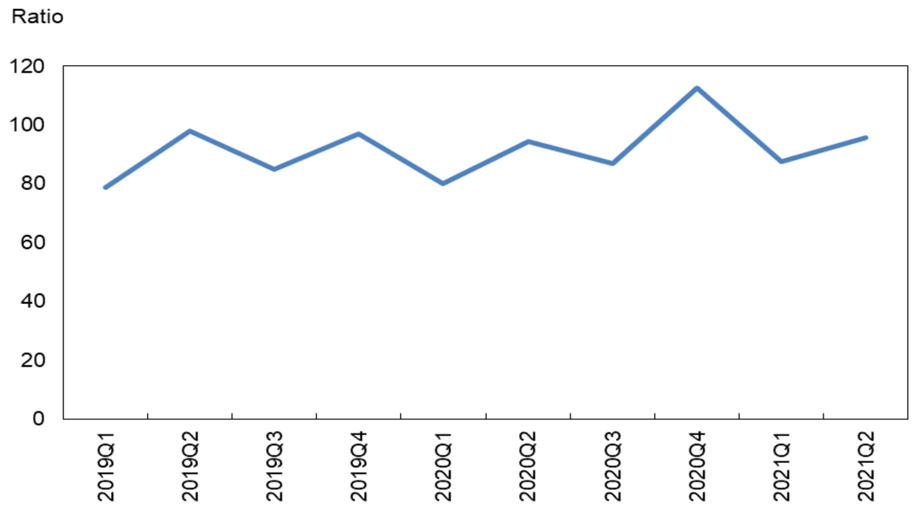
The International Monetary Fund's (IMF) Financial Soundness Indicator (FSI) 2019, provides a recommendation that an acceptable combined ratio should be less than a 100%. However, thresholds can vary considerably, with those between 75% and 95% considered to be somewhat appropriate. A very high ratio can imply that premiums are not able to keep up with expenditures and a very low ratio can imply insufficient expenditure to allow for good business administration[30]. The combined ratio averaged 92% between Q1:2019 and Q3:2021 reaching 113% in Q4:2020, mostly due to higher claims and commission paid as insurers aggressively pursued new business opportunities amid difficult economic conditions. The combined ratio peaked at 197.9% in the third quarter of 2021, reflecting the negative impact of the social unrest in July 2021 on the non-life insurance segment.

29 The combined ratio is calculated as claims plus expenses plus commission divided by net written premiums. It is a measure of how well non-life insurers manage their expenses relative to income.

30 A ratio consistently over 100% is also a sign of mispricing and provides an insurer with an incentive to increase profitability by investing in riskier assets.

The movement of the combined ratio has been mixed over time, trending on the high side of the threshold. This indicator will continue to be monitored closely as reduced profitability is a vulnerability for the sector and has the potential of undermining growth.

Figure 13: Combined Ratio



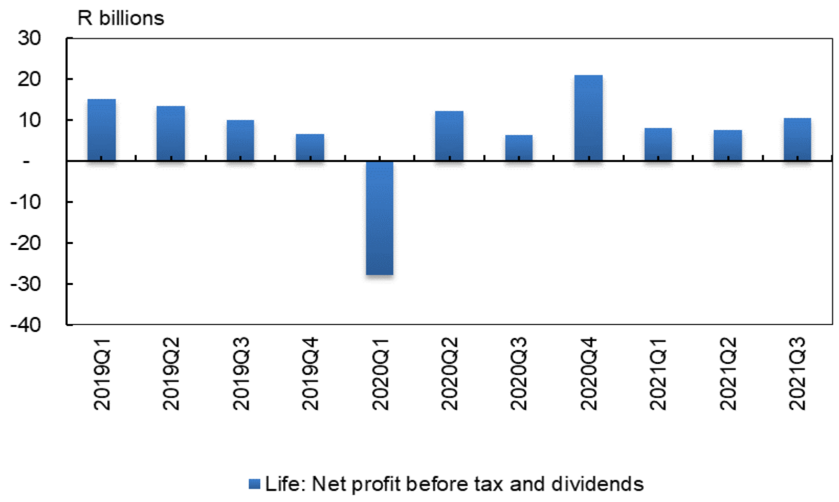
Source: PA

(x) Life Insurance Profitability: Net profit before tax and dividends[31]

Life insurers have experienced challenges in returning to pre-Covid profitability as claims continue to increase and the pandemic lingers. Profits fell sharply at the onset of the pandemic weighed by a decline in investment income. While profitability has not returned to pre-Covid levels, signs of recovery have been evident. Net profit before tax increased to R9.4 billion in Q3:2021 from R6.0 billion in Q2:2021, reflecting a change in policy liabilities and higher investment income.

[31] Net profit before tax and dividends is calculated as total revenue minus total expenses before accounting for tax and dividends. The indicator looks at the profitability of the business, and high numbers show growth and soundness of the company.

Figure 14: Life Profitability: Net profit before tax and dividends



Source: PA

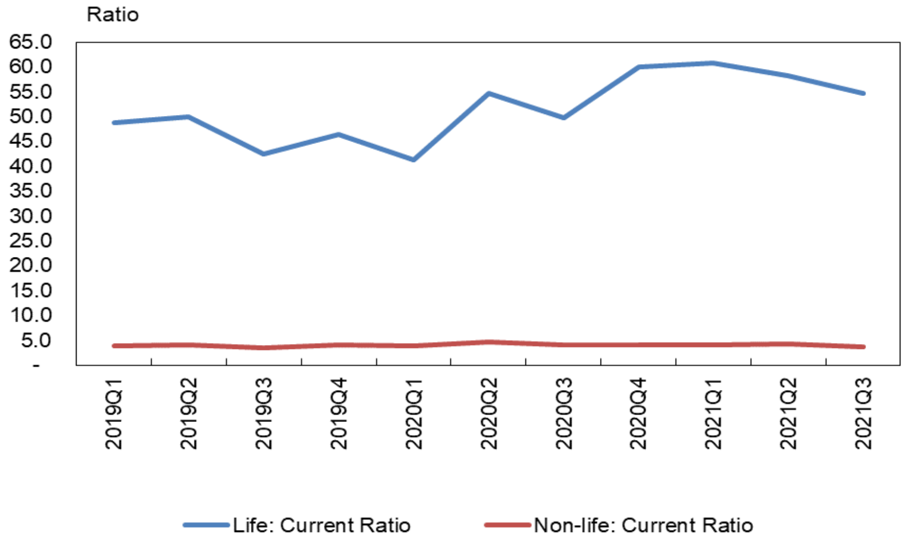
(xi) Current Ratio[32]

The current ratio for the life segment increased from an average of 46 times in Q4:2019 to 54.6[33] times in Q3:2021. While the non-life current ratio remained relatively unchanged over time, it declined slightly in March 2020 to 3.8 times from 4.0 in Q1:2019. There are no serious concerns for financial stability as non-life segments sector's ability to cover its short-term obligations by Q3:2021 was at 3.7 times.

[32] The current ratio is a measure of liquidity and is calculated as a ratio of liquid assets to current liabilities. The indicator measures the insurer's ability to meet short-term obligations. Life insurers' liquidity increased for the period under review driven by growth in investment funds.

[33] Liquid assets include reinsurance, investment funds, cash and cash deposits and current assets.

Figure 15: Current Ratio



Source: PA

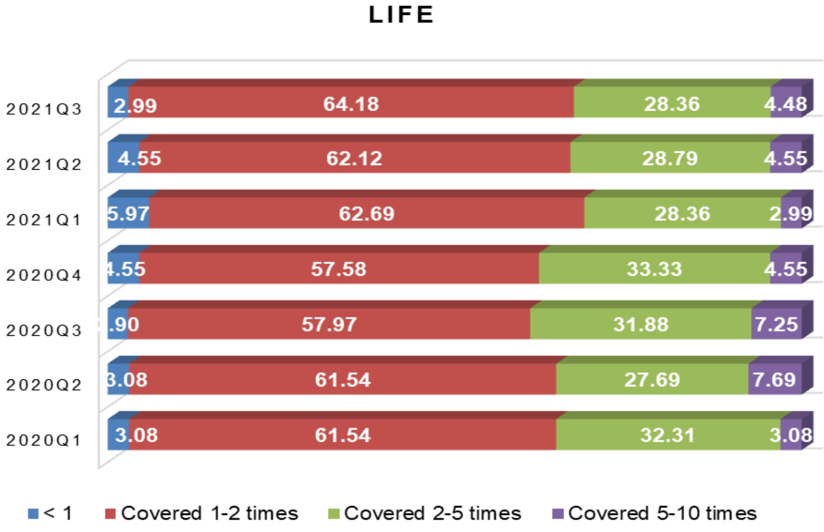
(xii) Solvency Capital Requirement Coverage Ratio (SCR ratio)[34]

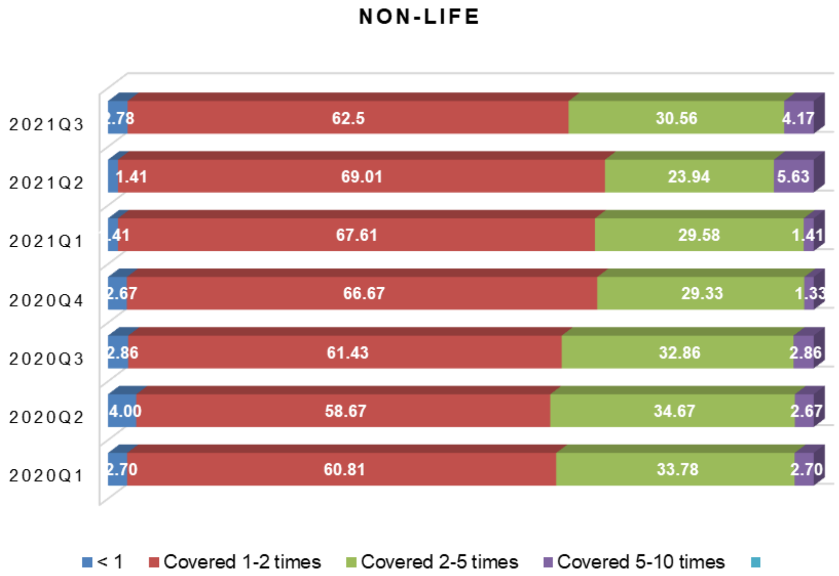
On an average basis, both the life and non-life insurance segments maintained average SCR coverage ratios well above the minimum requirement of 1 times[35]. The ratio deteriorated slightly Q3:2020 because of a reduction in insurance profitability - which has negatively affected the retained earnings of insurers. Overall, however, the insurance sector remains adequately capitalised. Additionally, as at 31 December 2021, SASRIA had a SCR cover of less than 1 times (-2.26). SASRIA is currently in discussions with the National Treasury for the final tranche of R7.1 billion capital injection, which would bring its SCR cover to around 1.8 times.

34 The solvency capital requirement (SCR) is the main regulatory requirement for insurers and reflects the amount of own funds that a company requires to survive a 1-in-200-year loss event.

35 In the third quarter of 2021 one life insurer and two non-life insurers had SCRs below the minimum requirement, however, overall, systemic risk remains limited.

Figure 16: Solvency Capital Ratio





Source: PA

(xiii) SASRIA Claims[36]

SASRIA received claims to the value of R33,8 billion up to the beginning of February 2022, R16,6 billion of which were already paid out. SASRIA entered the social unrest period in a relatively sound financial position, but the social unrest episode has exposed the following vulnerabilities for the firm and its policyholders.

36 SASRIA is the state-owned non-life insurer.

Firstly, the social unrest led to an initial cash injection of R3,9 billion from the National Treasury placing pressure on government finances and secondly, the additional cost burden was transferred to consumers through higher premiums. Effective 1 February 2022, insurers subscribing to the SASRIA option, were required to pay up to 1700% more for premiums related to unrest insurance. In January 2022, additional funds were allocated to SASRIA as part of its larger approved capital injection program. The possibility of additional cash injections required by SASRIA, and the potential of further social unrest incidents owing to slow economic growth and rising inequality pose a vulnerability to this sector. Cash injections place an additional strain on the fiscus through the crowding out of government funds.

Overall, despite some pockets of vulnerabilities in certain segments that will require close monitoring, South Africa's insurance sector is relatively resilient, adequately capitalised and has ample liquidity and does not exhibit major signs of systemic risk.

The insurance sector, however, remains vulnerable to the increasing trend in claims (non-life insurance segment) due to the possibility of further social unrest events; Covid-19 related claims (life insurance segment), a high concentration of assets in the sector, and the risk that SASRIA might not have sufficient funds at its disposal to pay outstanding claims.

4. Other risks and emerging risks in the domestic insurance sector

Other current and emerging risks that pose a threat to stability of the domestic insurance sector include: Cyber-attacks on key financial infrastructures, climate change and, private equity participation.

(i) Cyber-attacks on key financial infrastructure

The growing dependency on Information Technology (IT) for transactions and communications poses a risk to the insurance sector, with financial services firms being among the most attractive targets. The accelerated reliance on digital platforms and IT, while bearing efficacy benefits, have also made insurance firms vulnerable to cyber-attacks. By the very nature of its client-driven business, insurance firms have large amounts of personal client information making it susceptible to cyber-attacks.

Such attacks may lead to business interruption, loss of consumer confidence and reputational damage. As the migration towards digital platforms progresses, insurers are exposed to vulnerabilities in the form of heightened fraud attempts and potential data hacks.

Potential mitigating factors include the SARB's participation in fora to monitor and detect risks such as the Financial Sector Contingency Forum (FSCF) and enhancements of structures to ensure prevention, timely detection, and response and recovery to such attacks; the large information technology security spend and focus from systemically important financial institutions (SIFIs) and the promulgation of the Cybercrimes Act in December 2021, among others.

(ii) Climate change

South Africa is heavily dependent on climate-sensitive sectors such as agriculture and forestry, and rainfall fluctuates around a mean of 464mm compared to a world average of 857mm[37]. Climate-induced changes to rainfall can affect plantations and increase the occurrence of drought. Conversely, climate-induced changes may also result in greater than normal rainfall damaging property through floods. South Africa experienced damage to property because of flooding during December 2021 and January 2022. To this end, South Africa is lagging its peers, in the monitoring of climate change risks. The SARB is however in the process of incorporating climate risk into its stress testing capability. On 2 November 2021, the South African government undertook a partnership with the governments of France, Germany, the United Kingdom, the United States, and the European Union in support of a just transition to a low carbon economy and a climate resilient South Africa.

For South Africa, potential associated vulnerabilities include a relatively high risk of climate related damage to property and a high concentration of carbon intensive activities with potentially large exposures to the financial system. For insurers, this raises both aggregation and concentration risk. Concentration risk increases due to interconnectedness[38], while aggregation risk raises concerns of multiple claims arising

[37] https://www.adaptation-undp.org/explore/africa/south-africa
[38] https://www.mckinsey.com/industries/financial-services/our-insights/climate-change-and-p-and-c-insurance-the-threat-and-opportunity

from one event such as flood and property damage. On the operational side, insurers face the challenge of rising input costs in updating their business models and finding new ways to underwrite risk that will include the effects of climate change.

(iii) Increased private equity participation in the insurance sector

Globally, over the last decade, the insurance sector has witnessed increased private equity participation, particularly in the life/annuity (L/A) insurance segment. Globally, private equity-owned or sponsored insurers' admitted assets[39] grew to US$604.1 billion in 2020 from US$67.4 billion in 2011[40].

Increased participation from private equity dealers bears the risks of passing higher costs to consumers and satisfying the desire for higher returns demanded from investors. These higher costs result in higher premiums paid by the policyholder. Unaffordable or large increases in premiums may therefore culminate in an increase in lapse ratios, increasing the number of uninsured consumers. Although beyond the scope of this paper, private equity participation opens avenues for future research on this topic in South Africa.

Data gaps identified for the SARB and PA in the analysis of the insurance sector

In support of the implementation of the IAIS's Insurance Core Principle (ICP) 24.1, the current Holistic Framework list[41] of indicators and elements as specified in the IAIS Application Paper on Macroprudential Supervision captures a set of indicators that may be used for macroprudential and microprudential analysis and supervision. From the IAIS Holistic Framework, a list of indicators currently available to the SARB through its regulatory reporting was identified and listed in *Paper 1* of this series.

Table 1 provides a list of indicators and data elements that are not reported for regulatory requirements in South Africa but can be sourced through the channels indicated in the table.

---

[39] Assets permitted by an insurance company permitted by state law to be included in the company's financials, usually the balance sheet.

[40] https://www.lifehealth.com/insurance-companies-remain-prime-targets-private-equity/

[41] Detailed list (Annex 1) can be found in the Application Paper on Macroprudential Supervision https://www.iaisweb.org/page/supervisory-material/application-papers/file/98920/application-paper-on-macroprudential-supervision

Table 1: Examples of data currently not reported as regulatory requirements

Profitability: Change in reinvestment rates versus guaranteed rates	This information is specific to life insurers that provide guarantees as part of the products. The quantitative information is not available for individual insurers as part of regulatory reporting, but can be obtained via the report of the Head of the Actuarial control function to the board of directors.
General data: Changes in insurance pricing (individual and industry)[42] Changes in underwriting clauses and changes in legal coverages	The pricing at product level is not reported as part of regulatory reporting, but can be obtained via the report of the Head; Actuarial function to the board of directors. This is more market conduct data that can influence the underwriting results from a prudential perspective but is rather gathered via discussions with insurers or from the FSCA. For instance, it was observed with COVID-19 pandemic cover that is now explicitly excluded in business interruption cover policies.
Data relating to specific and unforeseen events, such as pandemics, natural disasters, cyber-attacks: (individual and industry) Changes in underwriting clauses and changes in legal coverages	This is quite specific reporting requirements and was implemented to track COVID-19 experience as well as part of stress testing in ORSA (Own Risk and Solvency Assessment) submissions.
Liability side Business interruption insurance, pandemic insurance	This is quite specific reporting requirements and was implemented to track COVID-19 experience as well as part of stress testing in ORSA (Own Risk and Solvency Assessment) submissions.
Business interruption insurance, pandemic insurance	This is quite specific reporting requirements and was implemented to track COVID-19 experience as well as part of stress testing in ORSA (Own Risk and Solvency Assessment) submissions.

Source: SARB and PA

[42] However, the analysis of premium income should suggest changes in pricing.

Looking ahead, it will be beneficial for the SARB and PA to fill some of the identified data gaps to enhance their systemic risk assessment monitoring frameworks to conduct deeper assessments and analysis of vulnerabilities in the sector as new risks emerge.

5. Summary and conclusion

This paper provided an analysis of key indicators identified by the IAIS's Holistic Framework and the SARB/PA's monitoring framework. The overall assessment is that there are limited signs of systemic risk evident in the South African insurance sector, amid some pockets of vulnerabilities that will be closely monitored. SASRIA remains a cause for concern as it continuously seeks cash injections from the National Treasury, which adds to the country's fiscal burden. However, overall, the insurance sector remains adequately capitalized with ample liquidity.

The paper also identified data gaps from the IAIS Holistic Framework, which although not submitted through regulatory return requirements, could potentially be obtained through identified channels. As such, the PA uses a risk-based approach to monitor entities which may require the use of more varied indicators over different reporting periods. Looking ahead, the current insurance sector monitoring toolkit could be enhanced with additional available data as identified through the data gap analysis exercise.

TOPICAL BRIEFING: No. 4

SOUTH AFRICAN RESERVE BANK

Financial Stability Department

Financial conditional and risks to financial stability in South Africa

Abstract

In the Growth-at-Risk (GaR) framework of Adrian, Boyarchenko & Giannone (2016, 2019), current financial conditions affect the distribution of future output growth, and the effects are time varying. However, financial conditions are typically characterised by a wide range of indicators, sometimes country-specific, that can be aggregated into Financial Conditions Indices (FCIs) in a number of different ways. These FCIs are usually not constructed with the intention of assessing tail risk in GDP growth in a policy-making environment. In this paper, we provide a framework for constructing a partitioned set of FCIs designed to feed into GaR analysis for South Africa. Indicators are selected based on the literature on early warning indicators of banking crises, and normalised and aggregated into three composite indices or 'partitions', capturing: asset valuations, leverage in the private nonfinancial sector, and external conditions. We argue that using these composite indices of financial conditions as explanatory variables in GaR analysis allows a more intuitive interpretation of the macrofinancial risks to future growth, since each partition captures a different dimension of risk relevant to South Africa. Furthermore, they allow for simpler communication, both with macroprudential policymakers and the wider public.

JEL Classification: E44, E61, G21.

Keywords: Financial stability, financial conditions indices, early-warning indicators.

Contents

1. Introduction

Recent work on Growth-at-Risk (GaR) by Adrian, Boyarchenko & Giannone (2016, 2019) has stimulated interest in modelling how current financial conditions affect the distribution of future output growth. This work requires measures of financial conditions to capture the evolution of risks to financial stability over time. However, financial conditions are typically characterised by a wide range of indicators, sometimes country-specific, that can be aggregated into indices in a number of different ways. These indices are usually not constructed with the intention of assessing tail risk in GDP growth in a policy-making environment.

In this paper, we provide a framework for constructing a partitioned set of Financial Conditions Indices (FCIs) designed to feed into GaR analysis for South Africa. The framework is based on the large and complex literature that deals with the relationship between financial conditions and economic growth, and a preliminary set of indicators is selected based on the literature on early warning indicators of banking crises. These indicators are normalised and aggregated into three composite indices or ‘partitions’, dealing with the categories of risk arising from asset valuations, leverage in the private nonfinancial sector, and external conditions.

Other partitions could of course be chosen and other indicators may be preferred to those presented here. The framework is sufficiently flexible to accommodate this. The motivation for creating the composite indices of financial conditions as explanatory variables in GaR analysis is that they allow a more intuitive interpretation of the macrofinancial risks to future growth. Each partition can be designed to capture a different dimension of risk relevant to South Africa. Furthermore, we argue that they allow for simpler communication, both with macroprudential policymakers and the wider public.

The paper is structured as follows. Section 2 provides the analytical framework that underlies the paper, and Section 3 surveys the literature on financial conditions and risks to real economic activity (including the measurement of financial conditions, and the selection and aggregation of indicators into composite indices). Section 4 discusses existing measures of financial conditions for South Africa, and Section 5 presents the new partitioned composite FCIs. Section 6 examines the dynamic relationships between the composite risk indices, and Section 7 reports quantile regression estimates of GaR for South Africa using the new indices (and compares them to existing work).

2. The analytical framework

The theoretical underpinnings of this analysis are elegantly captured in the framework set out by Aikman *et al.* (2018) to study the tail risks to economic growth. They provide for two types of shocks that impact upon the financial system: endogeous shocks e that are created within the system (e.g. a buildup of risk associated with a credit boom), and exogenous shocks x that are orthogonal to indicators of the financial cycle (cyber threats are mentioned as an example). The set of shocks s is then $s_t = \{e_t, x_t\}$.

A key aspect of the framework is that it also allows for 'vulnerabilities' V that potentially amplify the effect of any given shock, (Adrian *et al.* 2013). V can be assessed in terms of indicators of risk facing the system R and indicators of the resilience of the system K. So $V_t = V(R_t, K_t)$. R_t includes the indicators that are the focus of this paper[43]. It is well known that risks can grow in good times as well as bad (e.g. Brunnermeier and Sannikov, 2014; Danielsson *et al.*, 2018), so the nature of the dynamic relationships between these risk indicators and vulnerabilities, and ultimately therefore with growth, may be complex. K_t could include measures of the extent of banking system capital reserves, the leverage of financial institutions, maturity transformation, and interconnectedness.

Vulnerabilities V_t and shocks s_t are dynamically linked in the Aikman *et al.* (2018) framework by the provision that the likelihood of endogenous shocks, given by the probability $Pr(e_t)$, depends on previously built-up vulnerabilities: $Pr(e_t) = e(V_{t-1})$, with $\frac{\delta e}{\delta V} > 0$.

[43] Aikman *et al.* (2018) list as examples the level and distribution of debt in the private nonfinancial sector; the debt service burden; asset valuation pressures; and measures of the quality of credit extended by the financial system.

Besides s_t and V_t future economic growth and its distribution is also impacted by shocks unrelated to the financial system, termed ξ_t here. Wars and pandemics would be examples. These shocks are viewed as being outside the remit of macroprudential policy, although interactions with vulnerabilities V_t similar to s_t would seem likely.

Using this framework, the GaR (Adrian *et al.*, 2019) or GDP-at-risk (Cecchetti, 2006) of the economy is then

$$V\alpha R_q(GDP_t) = f(s_t, V_t, \xi_t) \quad (1)$$

This models the worst possible decline in GDP over the relevant policy horizon at probability q as a function of s_t, V_t and ξ_t. The derivatives of f are all negative, so it follows that shocks that impact the financial system negatively are exacerbated by vulnerabilities that make the potential fall in GDP even worse.

3. Financial conditions and risks to growth

The relationship between financial conditions and economic growth is complex, and the subject of a large literature. In this section, we highlight the most relevant strands of the literature for this paper and discuss the selection and aggregation of indicators of financial conditions for a GaR assessment.

Following the Global Financial Crisis (GFC), there has been renewed interest in research into the leading indicator properties of financial variables for these episodes. Historically, this has been a common response to crises. The early contributions to this literature include the work on early warning indicators (EWIs) of currency crises (see e.g. Kaminsky *et al.*, 1998 and Frankel and Saravelos, 2010 for surveys) and on forecasting output growth using asset prices and other variables (surveyed by Stock and Watson, 2003). This early work generally found little to encourage researchers. Chamon and Crowe (2012) note that interest stimulated by the crises of the 1990s waned when faced by challenges in predicting the timing of crises and a subsequent period of calm in global financial markets. Stock and Watson (2003) concluded that the evidence that assets prices helped predict output growth outweighed that for inflation, but that even this evidence was disappointing. Some asset prices were useful predictors in some countries, some of the time, but none were reliable predictors all of the time.

Policymakers responded to these early setbacks by focusing less on predicting the timing of crises (predicting the shocks or triggers), and more on identifying vulnerabilities in the financial system (Chamon and Crowe, 2012). In the area of macroprudential policy, this has translated into monitoring frameworks that focus on the systemic vulnerabilities that propagate adverse shocks, rather than the shocks themselves (see Bernanke, 2013 and Adrian *et al.*, 2015, and Farrell & Kemp, 2017 for a South African perspective).

The literature on forecasting output growth using financial variables has also recently begun to focus on features of the GDP growth distribution other than just the mean. The GaR analysis of Adrian *et al.* (2016 / 2019) is a notable example here. They go beyond traditional point forecasts by using measures of financial conditions and macrofinancial vulnerabilities to forecast the entire distribution of GDP growth; specifically, they define GaR to be the 5th percentile of the distribution of future growth, conditional on current economic and financial conditions. Giglio, Kelly & Pruitt (2016) and Adrian *et al.* (2019) suggest that measures of systemic risk are informative about future economic downturns, as they better predict lower quantiles of the conditional distribution of real output growth[44]. Others, such as Plagborg-Møller, Reichlin, Ricco & Hasenzagl (2020) and Brownlees & Souza (2021) are more pessimistic[45].

3.1 Measuring financial conditions

Investigating the relationship between financial conditions and economic growth obviously requires a measure of financial conditions. A large number of potential composite measures have been proposed, although no single consensus measure exists. In this section we assess the wide range of measures of financial conditions that appear in the literature, with the intention of providing a foundation for the

[44] They found that while the upper tail of the distribution of growth was fairly stable over time, the lower tail of the distribution varied with financial conditions. That is, financial conditions do not contain as much information about the upside risks to growth as they do for downside risk to growth.

[45] Plagborg-Møller *et al.* (2020), for example, find that: “moments other than the conditional mean are poorly estimated, and no predictors we consider provide robust and precise advance warnings of tail risks or indeed about any features of the GDP growth distribution other than the mean.” The comments by Gertler and by Liang on the ’forecaster’s perspective’ adopted in the paper provide important context here

modelling choices made in this paper.

A seemingly important distinction is between financial conditions indices (FCIs), financial stress indices (FSIs) and other related measures (vulnerability indices, financial cycles, …). For example, the IMF (GFSR, 2017) argues that FSIs seek primarily to identify episodes of financial stress, i.e. periods where financial intermediation is impaired. By contrast, an FCI is an index of a broad set of financial variables that influence economic behavior and thereby the future of the economy (GFSR, 2017). As Carlson, Lewis & Nelson (2012) argue, FSIs are intended to convey information about the functioning or fragility of financial markets, while FCIs look to map financial conditions onto macroeconomic conditions.

Conceptually, FCIs would appear to be better suited to the requirements of this paper. However, it is easy to link financial stress to vulnerabilities, which Section 2 showed are important for GaR, so the differences between the indices are not necessarily clear cut. A reading of the literature shows that FSIs and FCIs can display broadly similar patterns, use similar variables, be constructed using similar techniques, and claim similar ability to forecast changes in economic conditions[46].

With regard to the last of these points, Hatzius, Hooper, Mishkin, Schoenholtz & Watson (2010) and Koop & Korobilis (2014) argue that FCIs are good predictors of (the conditional mean of) future economic activity. In fact, Adrian *et al.* (2019) and the literature mentioned earlier argue that FCIs are able to flag tail risk associated with future economic contractions. However, many empirical studies report that FSIs also help forecast economic activity. Apostolakis & Papadopoulos (2019), for example, is a recent example that uses a panel VAR approach for 19 advanced economies and finds that a positive shock to financial stress impacts macroeconomic variables negatively, including growth. Kliesen, Owyang & Vermann (2012) find that over a common sample,

[46] Monin (2019, 4) makes similar points: “Financial stress indexes are similar to financial conditions indexes (FCIs). Like FSIs, FCIs combine information from many financial indicators to create a univariate time series that represents conditions in the financial system. The main difference between FSIs and FCIs is in their objectives. The objective of FCIs is to focus on the link between the financial sector and the real economy. Conversely, FSIs are concerned with distress or instability in the financial system without explicit regard for how such distress may manifest in the real economy. Another principal difference between FSIs and FCIs is in the set of indicators used in their construction. FSIs are generally constructed with market price-based measures. FCIs, on the other hand, are constructed with price-based measures and also include other market characteristics such as flows, trading volume, and stock measures. In practice, there is often considerable overlap between FSIs and FCIs in the sets of indicators included and the construction techniques used. Consequently, there is often considerable overlap in the time series properties of FSIs and FCIs.”

both FSIs and FCIs seem to forecast GDP growth equally well.

A related distinction is between early warning indicators (EWIs) and coincident indicators. The definitions provided earlier would suggest that FSIs provide an assessment of the current state of financial markets, whereas FCIs are more naturally perceived as EWIs. As Giordani, Spector & Zhang (2017) put it: “In contrast to financial stress indexes, an EWI is meant to be a leading rather than a coincident indicator, and preferably with a long lead”[47].

Following the oft-cited advice of Drehmann & Juselius (2014), an ideal EWI for banking crises should have a high degree of timing, stability and interpretability. The timing requirement balances the lag that macroprudential policies need before becoming effective, with signals that don’t arrive too early, since policy measures are costly (they aim for EWIs to signal crises 1.5 to 5 years ahead). The stability or persistence of the signal helps reduce uncertainty regarding trends and provides policymakers with more confidence regarding their decisions. Finally, the interpretability requirement relates to the ease of interpretation of the EWI signal. Drehmann & Juselius (2014) argue that the EWI should “make sense” if it is to be adopted by policymakers. Consistent with the motivation for this paper, they also note that ‘if EWIs have sound conceptual underpinnings, they are better suited for clear communication – an important aspect of macroprudential policy making”.

We focus on creating a (partitioned) FCI for South Africa, while accepting that there are close links between FCIs, FSIs and other measures of financial conditions. The analytical framework provided in Section 2 supports this interpretation; GaR is a function of both endogenous shocks and vulnerabilities in the financial system, so measures that capture aspects of both of these will be relevant.

[47] Again, the distinction is less clear in practice. Chatterjee, Chiu, Hacioglu-Hoke & Thibaut (2017) nor regarding their FSI for the UK: “Although our paper is not designed as an early warning signal model, it draws on some of its principles.

3.2 Selecting indicators of financial conditions for a GaR analysis

There are a very large number of indicators that could potentially be included in an FCI[48]. The discussion so far helps to narrow the focus somewhat. For a GaR analysis, we seek indicators that capture the build-up of vulnerabilities and the likelihood of endogenous shocks, and therefore provide information about the tail risk to GDP growth.

This focus differs slightly from that of most existing FCIs. In fact, even in the GaR literature the selection of appropriate indicators can be seen as an open question. The seminal GaR analysis of Adrian *et al.* (2019) used the existing Chicago Fed National Financial Conditions Index (NCFI), a (univariate) FCI that aggregates over one hundred indicators[49]. However, as Plagborg-Møller *et al.* (2020) point out, the intended purpose of such FCIs is generally not to be the best forecaster of tail risk in GDP growth, and alternative FCIs designed explicitly for this purpose may perform better. Also, as Liang (2020) points out, assessing current vulnerabilities is not the same thing as the narrower exercise of forecasting the distribution of GDP. It also seems likely that FCIs that have been adapted to allow for the specific characteristics of an economy will be better suited to a GaR analysis, an issue that we will return to later.

Aikman *et al.* (2018, 13-14) argue that the key requirement for indicators in a risk assessment framework is that they are able to "provide actionable, advance warning for policymakers of the build-up of risk in the financial system". Broadly consistent with the Drehmann & Juselius (2014) advice discussed earlier, they propose that an ideal indicator would have the following characteristics:

1. it signals building vulnerabilities with potential threats to financial stability at least two to three years ahead;

2. it provides reliable signals of building risks, with few erratic movements (high 'signal-to-noise' ratio);

[48] In practice this number varies. Many well-known FCIs use a relatively small set of indicators. For example, the Organization for Economic Co-Operation and Development FCIs for six advanced economies use just seven variables, and the Kansas City Financial Stress Index includes 11 variables. By contrast, the Federal Reserve Bank of Chicago index mentioned here includes more than 100 variables, and FCIs based on factor models even more.

[49] See also Brownlees & Souza for a discussion of the performance of this index in a GaR backtesting exercise.

3. it is available with a long time series.

The wider literature exploring the relationship between financial conditions and real economic activity also provides some broad guidelines on the types of indicators to include in an FCI for GaR analysis. For example, recent evidence suggests that credit growth is an important predictor of financial crises (Schularick and Taylor, 2012; Jordà *et al.*, 2013)[50]. Monitoring frameworks for policymakers have also emphasised credit developments. The Basel Committee on Banking Supervision (BCBS, 2010), for example, suggested in its guidance to national authorities that the credit-to-GDP gap be used as a guide for setting the Basel III countercyclical capital buffer (CCyB)[51].

Some researchers caution against relying too much on credit (see e.g., Kiley, 2021). Indeed, most monitoring frameworks for macroprudential policy opt to consider a range of variables grounded in the literature and the structural characteristics of the economy under consideration. (Aikman *et al.*, 2018; Adrian *et al.*, 2015; Kemp & Farrell, 2017). These frameworks would suggest including in the FCI indicators of developments in the equity, housing, bond, and interbank markets, to capture the various channels through which financial conditions can influence the broader economy. For countries like South Africa, including common global components that impact domestic financial conditions would also seem to be important. Arregui *et al.* (2018) find that a single factor, “global financial conditions,” accounts for 20-40% of the variation in countries domestic FCIs. This factor is found to move together with the US FCI and global risk measures (e.g., the VIX).

Finally, rather than focus exclusively on individual indicators, the approach taken here is to view financial indicators in terms of the specific types of risk they are intended to measure[52]. This approach underlies the use of partioned FCIs, discussed in the next section[53].

---

50 This is a common finding. See e.g. Aikman *et al.* (2018), who provide a meta-analysis of 37 empirical studies on the determinants and impact of banking crises in advanced economies to guide their selection of indicators. They specifically exclude studies on emerging markets, citing structural differences with the UK.

51 See Farrell (2014) for a South African perspective on the CCyB.

52 Indeed, experience and empirical research may show that alternative indicators are better proxies for these types of risks, and these alternatives should then be preferred.

53 Liang (2020), commenting on Plagborg-Møller *et al.* (2020), also takes this view. She cites work by Bernanke (2018) that considers the effects of financial variables on (conditional mean) GDP growth, using four groups of indicators: housing and mortgages, nonmortgage credit availability, short-term funding, and bank solvency. Bernanke finds that the effects of the four groups differ significantly.

3.3 Aggregation of indicators: Partitions and composite indices

We intend to extend existing work on South Africa by grouping the set of macrofinancial indicators into composite component indices or ‘partitions’. This approach reduces dimensionality, although less so than a univariate FCI would, while allowing a more intuitive interpretation of the macrofinancial risks to future growth, since each partition captures a different dimension of risk.

This approach has been advocated in the GaR context by Prasad *et al.* (2019) for IMF country applications[54] and Aikman *et al.* (2018) for the UK. For example, Prasad *et al.* suggest three broad partitions:

1. financial conditions: includes indicators of the price of risks embedded in asset prices, ease of obtaining financing, cost of funding, and degree of financial stress.
2. Vulnerabilities: macrofinancial imbalances and sectoral balance sheet weaknesses (borrowers’ and financial sector).
3. Other: includes external conditions (such as commodity prices and a measure of global risk sentiment).

They provide examples of indicators for each partition, and note that selection should attempt to accommodate country-specific conditions.

Aikman *et al.* (2018) also have three composite measures for their study of the UK: (1) private nonfinancial sector leverage (includes borrowing by households and companies, as well as external leverage); (2) asset valuations in financial and property markets; and credit terms and conditions (captures underwriting standards and credit quality of new lending).

[54] Lafarguette (2019) argues that “partitioning the variables before running the model has useful advantages: first, it allows to reduce dimensionality parameters, as traditionally used in the literature; this is even more important with macro data at the quarterly level, with a limited number of observations. Second, using partitioned data often improves the forecasting estimations by extracting common trends in financial variables, hence filtering information from idiosyncratic noises. Individual financial variables, especially in countries with illiquid markets, often exhibit noisy behavior and erratic volatility. However, co-movement across a sufficiently large number of financial variables often contains valuable information. Finally, it gives the possibility to compute chained-index partitions, which mitigates attrition issues in financial samples.”

Weighting schemes for aggregating the indicators into component indices, composite indices for each partition and for the partitions in an overall index are required when constructing the FCIs. There are many ways to do this. Oet *et al.* (2015) compare methods of combining indicators into indices, noting that the ’delicate issue’ in aggregation is determining how material each indicator and market is to the financial system[55]. Prasad *et al.* (2019) opt to use principal component analysis (PCA) to extract common trends from sets of indicators. Interestingly, Aikman (2018, 20) opt for simple average weights. They argue (2018, 20, footnote 14), with some justification, 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…

4. Existing FCI / FSI measures for South Africa

A number of measures of financial conditions have been published for South Africa. A few examples of these FCIs, FSIs and financial cycle estimates are summarised in Table 1. Most pre-date the recent work on GaR, so only the paper by Kabundi and Mbelu (2020) makes reference to this literature directly, and even then this is not the main focus of the paper.

The indices in Table 1 range from the parsimonious (the 3-variable financial cycle measure of Farrell & Kemp (2020) and the 5-variable FCI of Kasai & Naraidoo (2013)) to the more data intensive (the FCI of Kabundi & Mbelu (2020) includes 39 financial market variables). The latter type indices tend to decompose the financial system into sub-sectors, and then aggregate these. Kabundi & Mbelu (2020) and Kisten (2020) employ 6 sub-indices, with breakdowns similar to Oet *et al.* (2012)[56]. Aggregation methods include simple averages (Kasai & Naraidoo (2013) and Farrell & Kemp (2020)), PCA (Gumata, Klein & Ndou (2012); Thompson, Van Eyden & Gupta (2015); Farrell & Kemp (2020); Kabundi & Mbelu (2020)), PCA and the Kalman filter with

[55] They review four alternative weighting schemes: (1) equal market weights, (2) credit weights, (3) portfolio theoretic weights, and (4) principal component weights.

[56] Oet *et al.* (2012) construct the Cleveland FSI, an index which captures stress in six main financial markets: the funding, FX, credit, equity, real estate and securitization markets.

constant loadings (Gumata *et al.* (2012); Thompson *et al.* (2015)), PCA and the Kalman smoother (Kabundi & Mbelu (2020)) and time-varying cross-correlations (Kisten (2020)). All except the financial cycle measure are assessed in terms of their ability to improve forecasts of (conditional mean) real economic activity and found to be useful in this regard (often with asymmetric effects). Kisten (2020) tests against ad hoc identified benchmark episodes of stress using the 'partial AUROC' metric to make allowance for the well-known paucity of financial crises in South Africa.

The FCI of Kabundi and Mbelu (2020) is currently used as the single measure of financial conditions in the SARB GaR modelling framework (Sing, 2019a; 2019b). This FCI is based on a the time-varying parameter, factor-augmented vector autoregressive (TVP-FAVAR) methodology proposed by Koop and Korobilis (2014). It is calculated in 2 steps:

1. Standard PCA analysis is used to obtain an initial estimate of the FCI, which is passed into a Kalman filter and smoother to calculate time-varying loading factors, and time-varying VAR parameters.

2. These time-varying parameters are then used in a Kalman filter and smoother to extract the FCI.

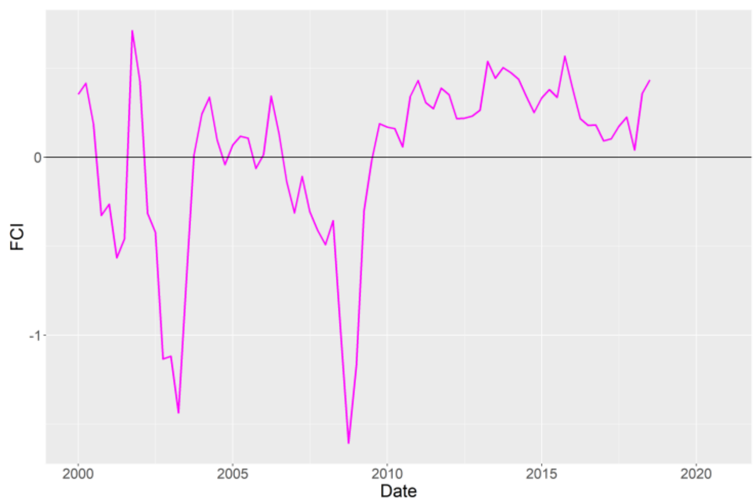
Kabundi and Mbelu (2020) find that this FCI improves forecasts of (conditional mean) inflation and GDP in the medium term. They also use the FCI in a quantile regression to assess the asymmetric relationship between the FCI and GDP growth (they report that tight financial conditions predict the future tail decline in GDP growth well, but the reverse doesn't hold).

Table 1: Measures of financial conditions for South Africa

Study	Index	Methodology	Result: Predicts real economic activity?
Gumata, Klein & Ndou (2012)	FCI	11 nominal variables (quarterly), 1999-2011, PCA, Kalman filtering with constant loadings.	Yes
Kasai & Naraidoo (2013)	FCI	Includes (i) the real effective exchange rate, (ii) real house prices, (iii) real stock price, (iv) credit spread, and (v) the future interest rate spread, 2000M01-2008M12, equal weights.	Yes
Thompson, Van Eyden & Gupta (2015)	FCI	16 monthly financial variables, 1966M02-2012M01, recursive PCA. FCI purged of any potential endogenous feedback effects. Rolling-window estimation techniques	Yes
Farrell & Kemp (2020)	FC	The financial cycle is identified using credit, house prices and equity prices as indicators, and estimated using traditional turning-point analysis, frequency-based filters and an unobserved components model-based approach, 1966–2016.	–
Kabundi & Mbelu (2020)	FCI	Based on Koop and Korobilis (2013). Includes 39 monthly financial market variables and 2 macro variables, Jan 2000 - Apr 2017, time-varying factor modelling (based on PCA, Kalman filtering)	Yes
Kisten (2020)	FSI	Methodology proposed by Chatterjee et al. (2017). SAFSI covers 17 monthly indicators from six market segments including the equity, credit, foreign exchange, money, housing, and commodities markets; market sub-indices are weighted by time-varying cross-correlations, 1995-2017.	Yes

Given that this FCI is the measure currently used by the SARB, we reproduce it here in Figure 1 as a benchmark for comparison with our new FCIs in Section 6. Note that the horizontal axis represents the average of financial conditions over the sample, and that positive (negative) values represent looser (tighter) conditions than average (in Kabundi & Mbelu (2020, Figure 3) this is termed the negative FCI).

Figure 1: Financial Conditions Index: SARB (Kabundi & Mbelu, 2020)



5. Partitioned financial conditions indices for GaR analysis in South Africa

This section presents details regarding the partitions and indicators selected for the new partitioned FCIs for South Africa, and methodology used for aggregation into composite indices. The historical evolution of the new indices relative to key stress periods is assessed, and the indices are compared to one another (with a view to evaluating their ability to capture the specific concepts they are intended to measure) and to the Kabundi & Mbelu (2020) FCI currently used by the SARB.

a. Description of indicators and partitions

We identify 23 indicators of financial stability risk, aggregated to produce three composite measures that each capture a different dimension of risk in the South African financial system:

1. Asset valuations (which combines the housing, commercial real estate and financial market valuation components).

2. Private nonfinancial sector (PNFS) indebtedness (including household and corporate indebtedness and external imbalance components).

3. External conditions (which includes exchange rates, commodity prices and other international indicators).

Other choices are of course possible, if different dimensions of risk are deemed more important. Interest in non-bank financial institutions (NBFIs), e.g., has increased in recent years, and authorities may be interested in including a partition to include risks associated with NBFI activities. Further work on identifying such risks and indicators should therefore guide future development of the FCIs.

Table 2 presents the 23 indicators we include in the framework, together with the aggregation into sub-indices and finally the 3 composite indices described above. Data are quarterly. We also note the start date for each time series, the transformation applied to the series (we average over the quarter where necessary), and the source of the data. We also note whether the inverse of the transformed series is used to ensure increases imply a build-up of risk. Other transformations of the data may also be explored; Aikman *et al.* (2018), e.g., opt to smooth some indicators using moving averages to reduce noise and focus on the build-up of vulnerabilities rather than measuring shorter-term stresses.

More information on data sources is provided in Appendix A.

Table 2: South African FCIs: indicators, components and partitions

Category	Series	Start	Trans	Inv	Source
Asset valuations:					
Real estate	House prices	2002q1	Q ave, %Δ		Haver
Financial	Bond spread: long/short	1960q1	Q ave	Yes	SARB
	Bond spread: long/med	1985q1	Q ave	Yes	SARB
	Sovereign CDS	2000q4	Q ave		BBG
	JSE: Returns	2002q4	Q ave, %Δ		BBG
	JSE: P/E	2002q3	Q ave		BBG
	SRISK	2000q3	Q ave		https://vlab.stern.nyu.edu
Private Nonfinancial Sector Leverage:					
Household	Debt Service	1969q1	Q		SARB
	Debt/disp income	1969q1	Q		SARB
	Debt/GDP	1969q1	Q		SARB
	Credit extension: Mortgage advances	1965q2	Q ave, %Δ		SARB
Corporate	Debt to GDP	1994q1	Q		SARB
	Debt/Net Op. Profit	1994q1	Q		SARB
	ICR	2006q2	Q ave	Yes	SARB
Government	Debt/GDP	1960q1	Q		SARB
	CA/GDP	1960q1	Q ave	Yes	SARB
	Gross K inflows/GDP	1985q4	Cum Q		SARB
External conditions:					
Exchange rate	REER	1970q1	Q ave		SARB
	Volatility	1999q2	Q ave	Yes	BBG
Commodities	Commodity prices	1981q3	Q ave		BBG
	Oil price	1967q3	Q ave		BBG
Other	Vix	1990q1	Q ave	Yes	BBG
	US GDP	1960q1	Q, %Δ		BBG

b. Methodology: Aggregation procedure

The individual indicators are aggregated into component indices and then composite indices using the following 3-step procedure:

1. Calculate a Z-score for each individual indicator 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 (a pseudo real time measure could possibly be used to get a sense of how the indices have performed in real time).

2. Combine small sets of related indicators by taking their unweighted arithmetic average, generating a set of component indices (e.g. in Table 2, the Asset valuations composite index is made up of ’real estate’ and ’financial’ component indices, each calculated as the average of individual series).

3. Combine the component indices into the 3 partitions or composite indices (asset valuations; private nonfinancial sector indebtedness; and external conditions).

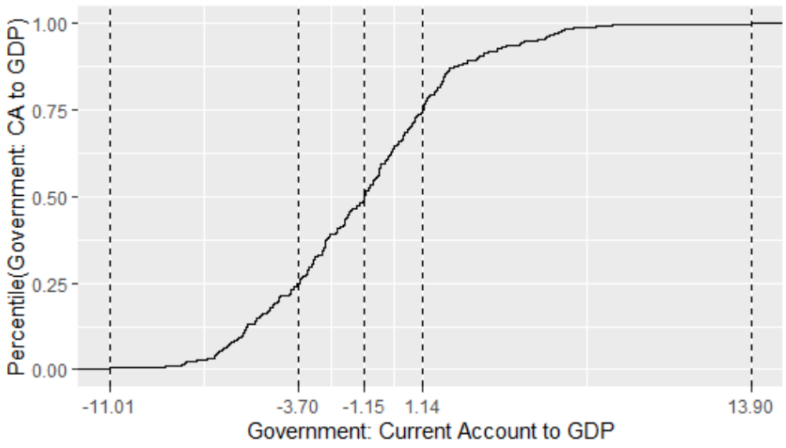
We opt to use simple average weights to aggregate the series, following Aikman *et al.* (2018). It is relatively simple to employ more sophisticated techniques here, although the comparisons presented later in Section 6 suggest this may make little difference and come at the cost of reduced interpretability and ease of communication.

c. Heat map of indicators

Ribbon heat maps are simple data visualisation tools that can be used to provide an overview of the build-up of risks in the financial system over time. Twala & van der Linde (2018), for example, have generated a heatmap for the South African financial system that forms part of the SARB’s financial stability monitoring process.

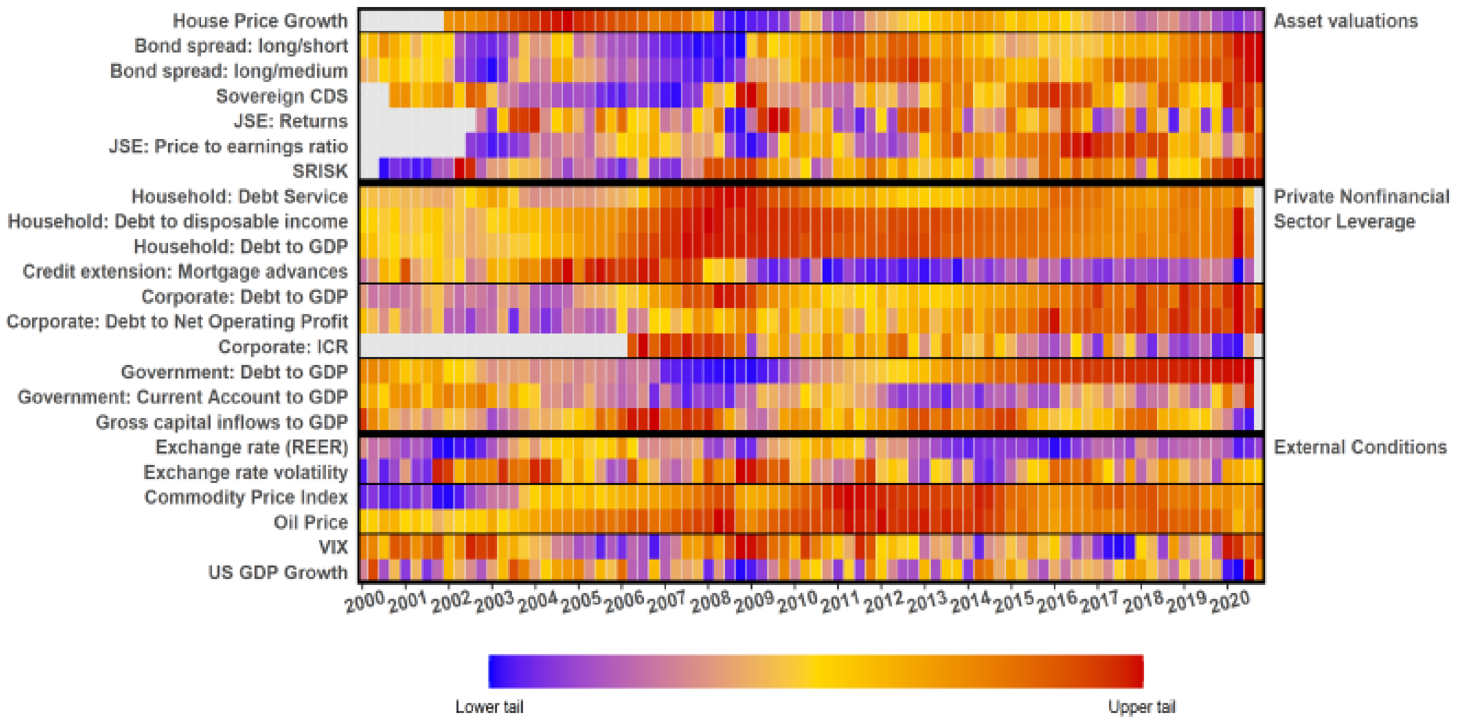
In Figure 3, we present a heat map of the indicators set out in Table 2 for a sample period running from 2000:Q1 to 2020:Q4. It is a simple matter to generate similar heat maps for the component and composite FCIs. Recall that increases in individual series imply a build-up of risk. Here each individual series is normalised using the empirical cumulative distribution function (ecdf) to show the proportion of scores that are less than or equal to each score. Each value is therefore the percentage of observations with that value or less. Figure 2 shows the ecdf for the current account to GDP ratio (which is an indicator in the ’Government’ component index in the PNFS leverage composite FCI). In this case, 50% of current account to GDP values are below the median value of -1.15.

Figure 2: Example: ecdf of the individual risk indicators



In Figure 3, the transformed indicator values for each quarter are represented as colors, where the lowest values are blue and the highest values are red (higher values / 'hotter' colors implying a build-up of risk).

Figure 3: Heat map of the individual risk indicators



Notes:

d. Historical evolution of the new indices

Developments in the composite FCIs for Asset valuations, PNFS leverage, and External conditions are presented in Figure 4(i)-(iii)[57]. The FCIs run from 1960:Q1 to 2020:Q4. Although the focus in this paper is ultimately on FCIs for GaR estimation, an informal analysis of the evolution of the FCIs is facilitated in the graphs by flagging significant stress events in the past 60 years. Vertical lines are included for the 1973 oil price crisis, 1985 South African debt crisis, 1996 Rand exchange rate crisis, 2008 Global financial crisis (GFC) and 2020 Covid pandemic[58].

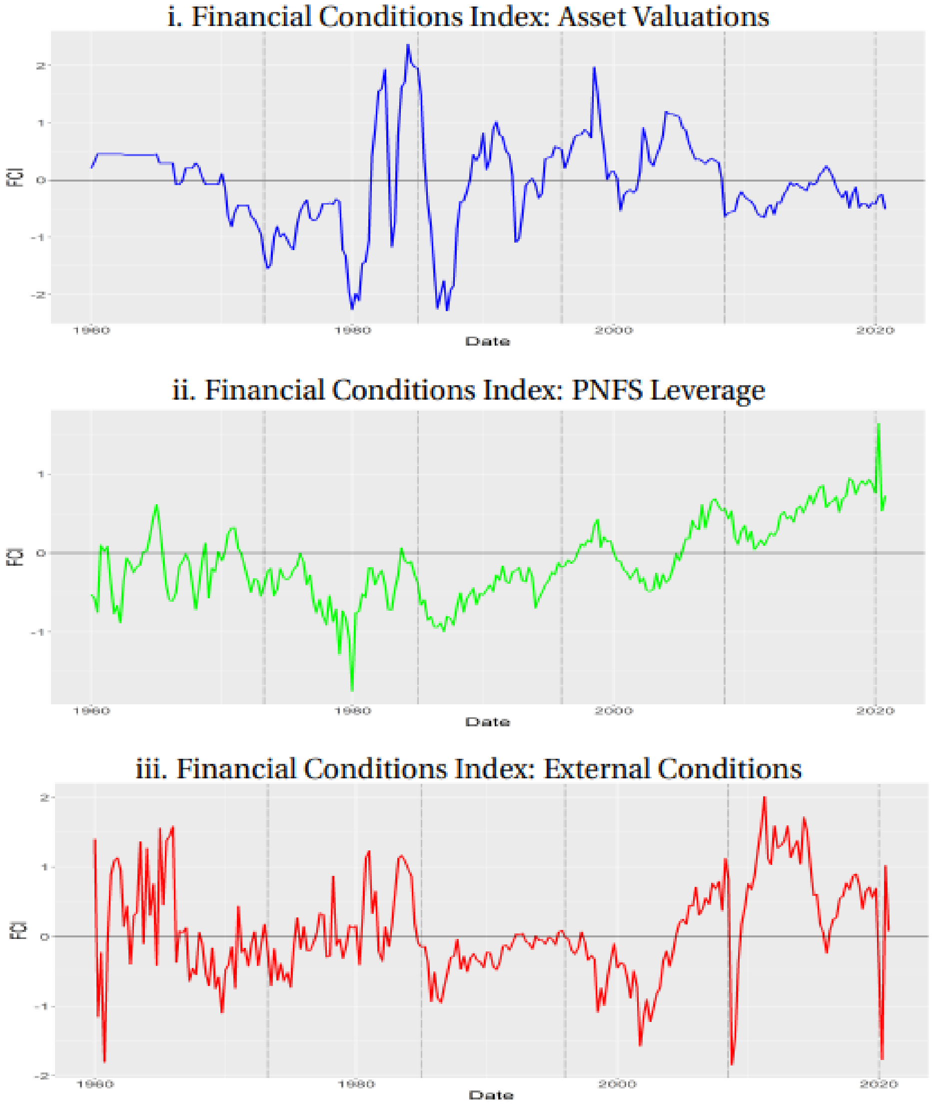
The individual composite FCIs appear to behave differently relative to these stress episodes. Ideally, we would see high, positive values for the FCIs in the period prior to an episode, signifying a build-up of risk in the South African financial system. If we adopt the Aikman *et al.* (2018) criteria, the FCIs would signal building vulnerabilities at least two to three years ahead of time. The Asset valuations and External conditions FCIs seem to meet this requirement in the build-up to the 1985 South African debt crisis, and all 3 FCIs were elevated just prior to the GFC (consistent with the global build-up of vulnerabilities prior to trigger in the US sub-prime market). Perhaps not surprisingly, given the nature of the crises, the FCIs did not capture a build-up of vulnerabilities prior to the 1973 oil price crisis, 1996 Rand exchange rate crisis and 2020 Covid pandemic.

---

[57] It is also possible to calculate the contributions of the various indicators and component indices to the composite indices. See Appendix B for graphs showing these contributions over time.

[58] More formal analysis is possible. Kisten (2020), e.g., uses the 'partial Area Under the Receiver Operating Characteristic Curve' (pAUROC) metric to evaluate financial indicators for her FSI on the basis of their ability to capture ad hoc periods of financial stress in South Africa.

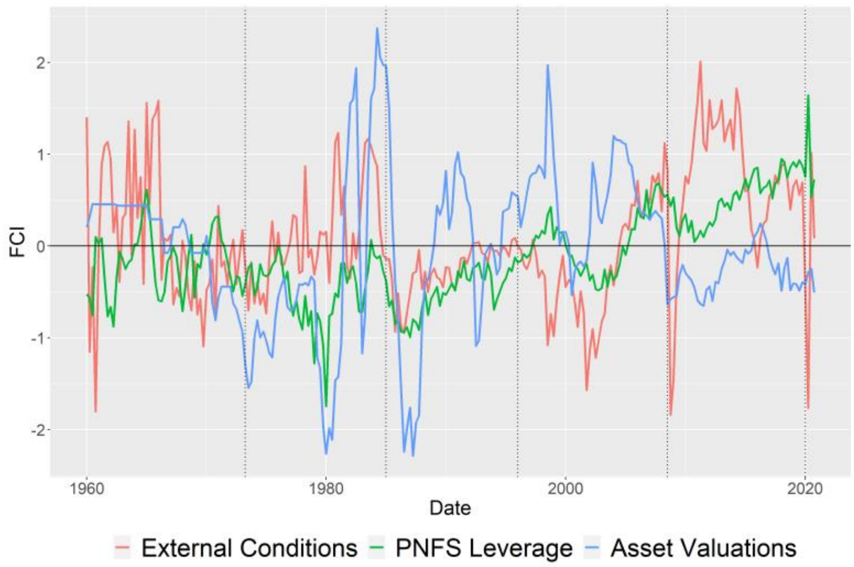
Figure 4: SA Financial Conditions Indices



Note: The dotted vertical lines indicate (from left to right) the 1973 oil price crisis; the 1985 SA debt crisis; the 1996 Rand exchange rate crisis; the 2008 Global financial crisis; and the 2020 Covid pandemic.

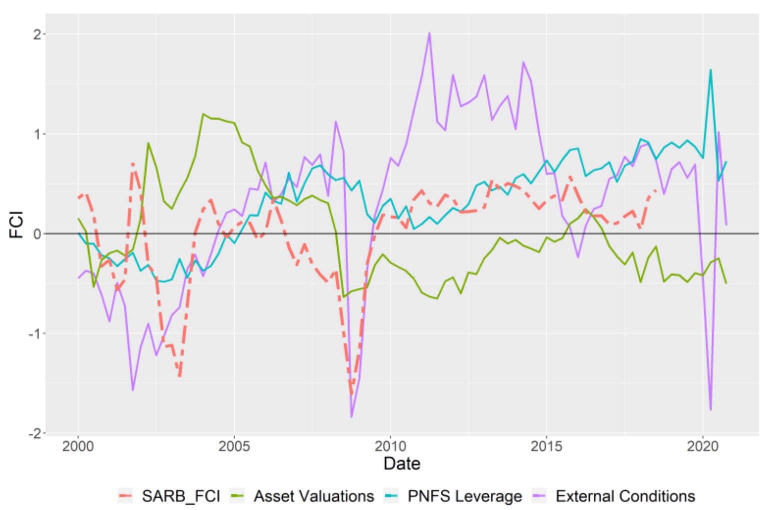
The approach adopted in this paper of grouping indicators into partitions with each partition capturing a different dimension of risk, suggests that our 3 composite indices are likely to differ from one another if the partitions are meaningful. These differences could be in terms of the trends followed, or other time series properties of the FCIs. Figure 5 shows that this is indeed the case. The FCIs sometimes differ in their assessment of building risks (e.g. when one FCI is positive and another negative), and Asset valuations and External conditions FCIs display greater variance than PNFS leverage.

Figure 5: Composite FCIs are not all the same



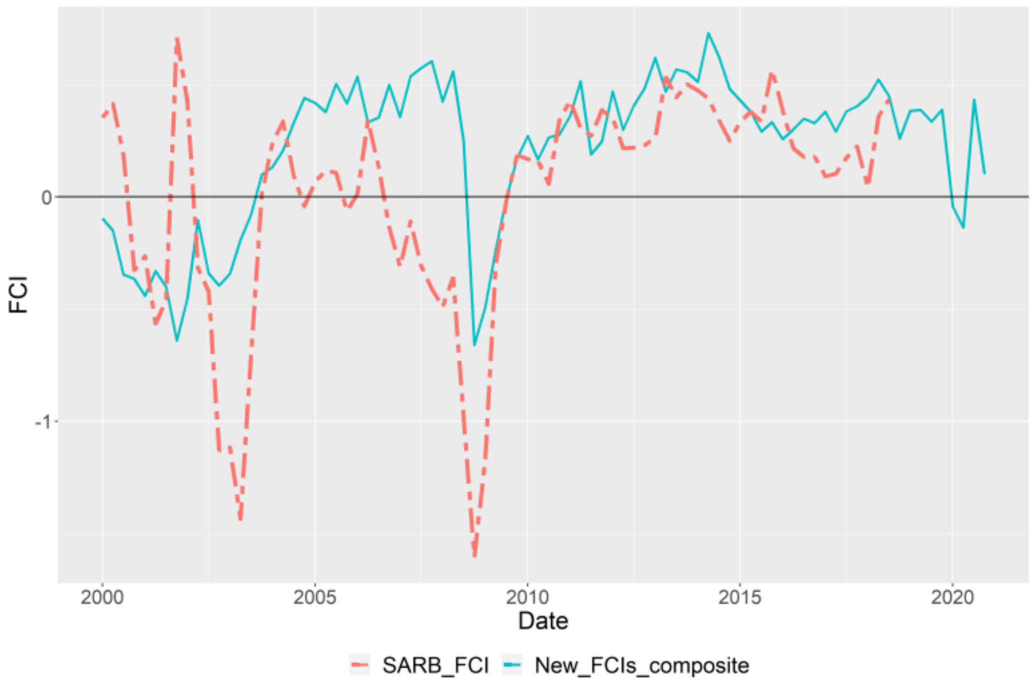
Earlier, in Section 4, we noted that the FCI of Kabundi & Mbelu (2020) is currently used as the (single) measure of financial conditions in the SARB GaR modelling framework (Sing, 2019a; 2019b). In Figure 6 we compare this FCI (denoted SARB FCI) with the new composite FCIs over the period since 2000 when the SARB FCI is first available. Again, we find that our 3 composite indices differ from the Kabundi & Mbelu (2020) FCI. In the early 2000s, e.g., early on in the run-up to the GFC, the SARB FCI, External conditions and PNFS leverage FCIs signaled little evidence of a build-up of risk. Only the Asset valuations FCI signaled this. Only after 2005 did the External conditions and PNFS leverage FCIs turn positive.

Figure 6: Financial Conditions Indices:
SARB (Kabundi & Mbelu, 2020) compared to the new composite FCIs



In Figure 7, we create a univariate index by averaging the 3 new composite FCIs (denoted the New FCIs composite), and compare it to the SARB FCI of Kabundi & Mbelu (2020). Significantly, the New FCIs composite series tracks the SARB FCI for most of the period since 2000. We interpret this as a positive result: it suggests that the approach of grouping indicators into partitions capturing different dimensions of risk is able to provide more information about the build-up of risk than a univariate index made up of the 3 composite indices, while still being consistent with more aggregated approaches.

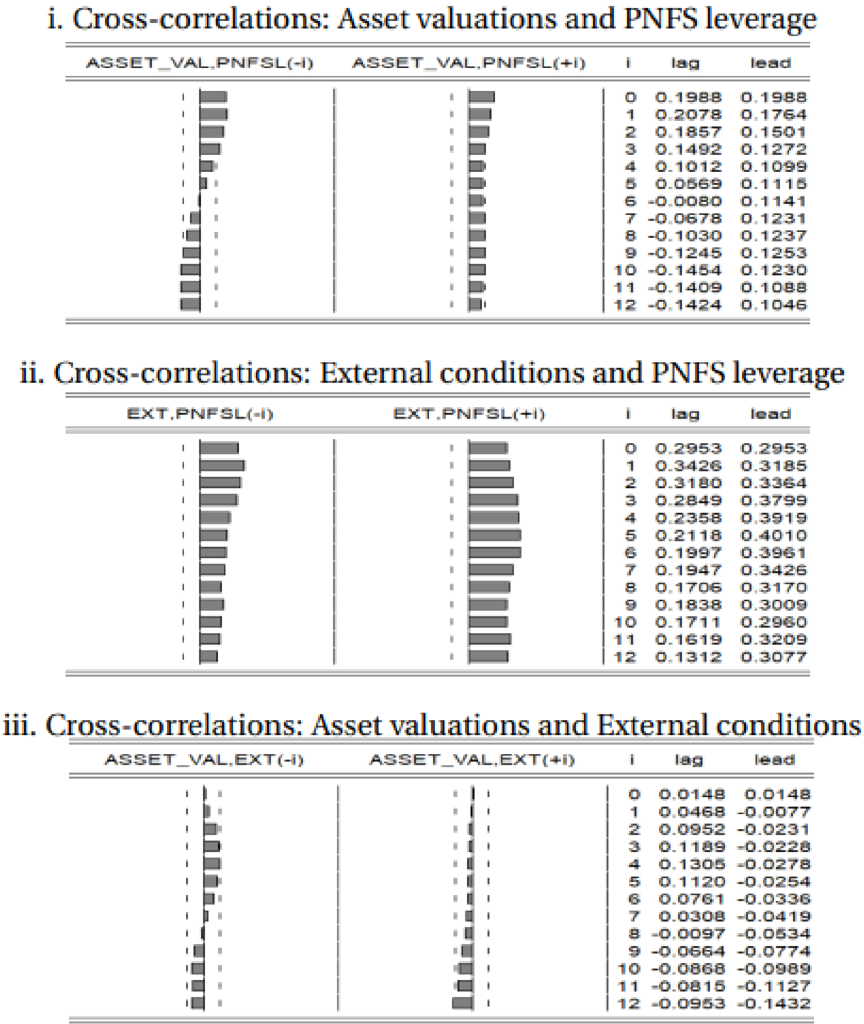
Figure 7: Financial Conditions Indices: SARB_FCI compared to new FCI



6. Relationship between the composite risk indicators: Correlations and Granger-causality

The differences between the composite FCIs noted in Section 6 suggest that a more formal analysis of the dynamic relationships between them may be useful. Following Aikman *et al.* (2018), we consider the lead-lag correlation structure between the composite measures using cross correlograms (Figure 8) and report bivariate Granger-causality tests at various lag lengths (in Table 3). The sample runs from 1960:Q1 to 2020:Q4.

Figure 8: Cross-correlation between the composite FCIs



i. Cross-correlations: Asset valuations and PNFS leverage

ASSET_VAL,PNFSL(-i)	ASSET_VAL,PNFSL(+i)	i	lag	lead
		0	0.1988	0.1988
		1	0.2078	0.1764
		2	0.1857	0.1501
		3	0.1492	0.1272
		4	0.1012	0.1099
		5	0.0569	0.1115
		6	-0.0080	0.1141
		7	-0.0678	0.1231
		8	-0.1030	0.1237
		9	-0.1245	0.1253
		10	-0.1454	0.1230
		11	-0.1409	0.1088
		12	-0.1424	0.1046

ii. Cross-correlations: External conditions and PNFS leverage

EXT,PNFSL(-i)	EXT,PNFSL(+i)	i	lag	lead
		0	0.2953	0.2953
		1	0.3426	0.3185
		2	0.3180	0.3364
		3	0.2849	0.3799
		4	0.2358	0.3919
		5	0.2118	0.4010
		6	0.1997	0.3961
		7	0.1947	0.3426
		8	0.1706	0.3170
		9	0.1838	0.3009
		10	0.1711	0.2960
		11	0.1619	0.3209
		12	0.1312	0.3077

iii. Cross-correlations: Asset valuations and External conditions

ASSET_VAL,EXT(-i)	ASSET_VAL,EXT(+i)	i	lag	lead
		0	0.0148	0.0148
		1	0.0468	-0.0077
		2	0.0952	-0.0231
		3	0.1189	-0.0228
		4	0.1305	-0.0278
		5	0.1120	-0.0254
		6	0.0761	-0.0336
		7	0.0308	-0.0419
		8	-0.0097	-0.0534
		9	-0.0664	-0.0774
		10	-0.0868	-0.0989
		11	-0.0815	-0.1127
		12	-0.0953	-0.1432

Note: Correlations are asymptotically consistent approximations. The dotted lines are the approximate two standard error bounds computed as $\pm 2/\sqrt{T}$. The sample runs from 1960Q1 to 2020Q4 (244 observations).

Here we consider the 12-period lead-lag correlation structure between the composite measures using cross correlograms. The relationships between lags of the second measure and the current-period values of the first measure are presented on the left in Figure 8, and the lead relationships on the right. In Figure 8(i), the largest cross-correlation between current asset valuations and past PNFS leverage occurs at lags of 1-3 quarters. Future PNFS leverage is significantly correlated with current asset valuations for up to 12 quarters. In Figure 8(ii), there are significant cross-correlations between external conditions and future PNFS leverage up to 12 quarters ahead (with the peak cross-correlation at 5 quarters ahead). Lagged PNFS leverage is also significantly correlated with external conditions for up to 12 quarters, although the peak cross-correlation is in the first few quarters. Finally, in Figure 8(iii), as expected, there is little evidence for cross-correlation between domestic asset valuations and future external conditions (disregarding the 12-quarter ahead correlation; at the 5% level, we would expect a probability of Type 1 error of one in twenty). Lagged external conditions show some cross-correlation with domestic asset valuations, peaking at 3-5 quarters.

We also investigate the lead-lag relationships between our three composite measures using Granger causality tests (see Appendix C for details regarding the tests). The results for bivariate tests using 12, 8 and 4 lags are reported in Table 3, and are broadly supportive of those obtained using cross-correlograms. We find that the PNFS leverage FCI Granger-causes Asset valuations for the tests using 12 lags (at the 1% level), 8 lags (at the 1% level) and 4 lags (at the 10% level). The External conditions FCI Granger-causes PNFS leverage using 12 lags (at the 10% level) and 8 lags (at the 5% level). Finally, the PNFS leverage FCI Granger-causes External conditions using 4 lags (at the 10% level).

These findings taken together suggest that PNFS leverage FCI contains information on future movements in asset valuations, and external conditions (in the near term, possibly through the exchange rate). The External conditions FCI may also provide insight into future movements in domestic 30 asset valuations and PNFS leverage.

Table 3: Granger-causality analysis: Composite indices

Null Hypothesis:	Obs	F-Statistic	Prob.
Lags: 12			
External conditions does not Granger Cause *Asset valuations*	232	1.22994	0.2642
Asset valuations does not Granger Cause *External conditions*		0.53379	0.8912
Private NFS leverage does not Granger Cause *Asset valuations*	232	2.56806	0.0034
Asset valuations does not Granger Cause *Private NFS leverage*		0.70935	0.7415
Private NFS leverage does not Granger Cause *External conditions*	232	0.92218	0.5256
External conditions does not Granger Cause *Private NFS leverage*		1.58302	0.0984
Lags: 8			
External conditions does not Granger Cause *Asset valuations*	236	0.86515	0.5467
Asset valuations does not Granger Cause *External conditions*		0.40215	0.9186
Private NFS leverage does not Granger Cause *Asset valuations*	236	2.52869	0.0119
Asset valuations does not Granger Cause *Private NFS leverage*		0.53832	0.8269
Private NFS leverage does not Granger Cause *External conditions*	236	1.20575	0.2967
External conditions does not Granger Cause *Private NFS leverage*		2.03468	0.0436
Lags: 4			
External conditions does not Granger Cause *Asset valuations*	240	0.57668	0.6798
Asset valuations does not Granger Cause *External conditions*		0.31486	0.8679
Private NFS leverage does not Granger Cause *Asset valuations*	240	2.06981	0.0856
Asset valuations does not Granger Cause *Private NFS leverage*		0.38230	0.8212
Private NFS leverage does not Granger Cause *External conditions*	240	2.25631	0.0639
External conditions does not Granger Cause *Private NFS leverage*		1.76175	0.1374

7. Example: Estimating GaR using the South African composite FCI's

In this section we show how the partitioned set of FCIs can be used in a GaR analysis for South Africa using a quantile regression approach (Koenker, 2005), and compare the results to those obtained from existing (single FCI) work. The intention is simply to illustrate the potential usefulness of the approach. A more detailed investigation into the performance of the FCIs in capturing the moments of the conditional distribution of GDP growth in an out-of-sample context is left for future work[59].

a. Quantile Regression

Sing (2019a), following Adrian *et al.* (2019), estimates quantile regressions of the form presented in equation (2) at the 5th, 25th, 50th, 75th, and 95th percentiles. She estimates this for the full sample period (1970:Q1-2018:Q2), and for the period following South Africa's transition into a democracy (1994:Q3-2018:Q2).

$$\text{Growth}_{t+1}^{q} = \beta_0^q + \beta_1^q FCI_t + \beta_2^q Growth_t + \epsilon_{t+i}^q \tag{2}$$

for quantile q, horizon $i = 1,\ldots,12$, and where FCI is the Kabundi & Mbelu (2020) FCI for South Africa.

Using the new partitioned composite FCIs, the quantile regressions are now of the form:

$$\text{Growth}_{t+1}^{q} = \beta_0^q + \beta_F^q X_{A,t} + \beta_V^q X_{L,t} + \beta_O^q X_{E,t} + \beta_G^q Growth_t + \epsilon_{t+i}^q \tag{3}$$

For $i = 1,\ldots,12$, where X_A, $X_L$and X_E represent the composite FCIs for asset valuations, private nonfinancial sector leverage and external conditions, respectively.

b. Quantile Regression Results

To illustrate the GaR analysis, we estimate quantile regressions of equations (2) and (3) for the 5th quantile $(q = 0.05)$ over the sample period from 2000:Q1 to 2018:Q2.

[59] This work could also include an investigation into a potential term structure of GaR in South Africa (Adrian, Grinberg, Liang & Malik, 2018; Sing, 2019a).

We report the results in Tables 4 and 5, respectively.

The results obtained using the single FCI (Table 4) find that higher values of the index have a significantly negative impact on GaR at horizons of 10-12 quarters (the impact is positive in the near term up to 8 quarters ahead, but these effects are not significant at the 5% level). An increase in the financial conditions index (looser conditions than average; see Section 4 and Figure 1) is associated with a deterioration in longer-term tail GDP growth (an increase in GaR over the next 3 years). This appears to be at odds with the finding of Kabundi & Mbelu (2020) that loose financial conditions do not predict future booms, while tight financial conditions predict future tail declines in GDP growth well.

Table 5 reports the results for the new composite FCIs. We discuss each in turn, noting that we are controlling for the effects of the other FCIs. Most strikingly, for the PNFS leverage FCI we find a significantly negative impact on the lower left-hand tail of the growth distribution at horizons from one to twelve quarters (peak effects are at 4 and 5 quarters). Looser credit conditions therefore appear to increase GaR, as expected, and this is so even in the short run. Aikman et al (2018, 29-30) find a similar result for the UK using their PNFS leverage FCI[60].

For the Asset valuations FCI, we find a significantly positive impact on 5th percentile future growth at shorter horizons of 2-4 quarters, and also at 11-12 quarters. The first of these results is expected, the second perhaps less so since we would expect asset booms to signal increased risks to future growth a longer horizons. Finally, the External conditions FCI also has a significantly positive impact on 5th percentile future growth at shorter horizons (1 quarter, and 2-4 quarters). Although this response turns negative at horizons of 9-12 quarters, these coefficients are not significant at the 5% level.

What is evident from this illustration is that the partitioned FCIs allow a more detailed analysis of GaR for South Africa. Over this sample, e.g., we see that PNFS leverage varies negatively with the lower left-hand tail of the growth distribution, while asset valuations and external conditions vary positively with GaR at shorter horizons.

[60] They find a significantly negative relationship, for all quantiles, although only at shorter horizons of one to four quarters. They find these results surprising, given an expectation that loosening credit constraints would not only increase short-term growth but also risks at longer horizons.

Table 4: Quantile regressions of the 5^{th} Percentile (2000Q1-2018Q2): SARB FCI

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Intercept	-3.76^{**}	-1.85	-0.90	-0.11	0.06	0.20	0.23	0.65^{*}	0.70^{***}	1.10^{***}	1.01^{***}	0.80^{***}
	(1.32)	(1.00)	(0.75)	(0.80)	(0.44)	(0.32)	(0.28)	(0.26)	(0.17)	(0.12)	(0.09)	(0.20)
FCI	0.11	2.43	2.43	2.04	1.26	1.34	1.06	0.39	-0.29	-0.37^{*}	-0.48^{**}	-0.60^{*}
	(1.49)	(1.49)	(1.49)	(1.20)	(0.83)	(0.68)	(0.53)	(0.45)	(0.29)	(0.16)	(0.16)	(0.24)
Current Growth	0.95^{*}	0.06	0.06	-0.08	-0.11	-0.01	0.02	-0.05	-0.02	-0.06	0.00	0.08
	(0.41)	(0.32)	(0.28)	(0.30)	(0.21)	(0.13)	(0.10)	(0.08)	(0.05)	(0.05)	(0.03)	(0.05)
Num. obs.	73	72	71	70	69	68	67	66	65	64	63	62
Percentile	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

Each Model(i) has dependent variable (i)-quarter ahead average growth. Standard errors are bootstrapped using 10 000 draws

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5: Quantile regressions of the 5^{th} Percentile (2000Q1-2018Q2): New composite FCIs

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Intercept	0.61	1.36	1.45	2.23^{*}	2.58^{***}	2.61^{***}	2.55^{***}	2.49^{***}	2.51^{***}	2.48^{***}	2.58^{***}	2.56^{***}
	(0.92)	(1.05)	(0.97)	(0.92)	(0.71)	(0.43)	(0.23)	(0.15)	(0.17)	(0.17)	(0.16)	(0.18)
Asset Valuations	2.83	3.42^{*}	3.14^{*}	3.49^{*}	1.75	0.50	0.20	0.43	0.51	0.63	0.86^{*}	0.75^{*}
	(1.45)	(1.37)	(1.45)	(1.45)	(1.17)	(0.70)	(0.47)	(0.39)	(0.36)	(0.38)	(0.36)	(0.31)
PNFS_Leverage	-5.00^{**}	-3.68^{*}	-4.94^{*}	-6.44^{***}	-6.00^{***}	-4.99^{***}	-4.24^{***}	-3.26^{***}	-2.40^{***}	-2.16^{***}	-2.80^{***}	-2.76^{***}
	(1.46)	(1.62)	(2.00)	(1.59)	(1.27)	(0.81)	(0.53)	(0.41)	(0.48)	(0.52)	(0.56)	(0.62)
External_Conditions	1.70^{*}	1.20	1.90^{*}	2.33^{**}	1.39^{*}	0.66	0.32	0.06	-0.24	-0.26	-0.07	-0.12
	(0.84)	(0.74)	(0.79)	(0.81)	(0.60)	(0.35)	(0.23)	(0.22)	(0.26)	(0.21)	(0.23)	(0.28)
Current Growth	-0.09	-0.29	-0.22	-0.55	-0.49	-0.31	-0.15^{*}	-0.09	-0.08	-0.07	-0.07	-0.03
	(0.36)	(0.25)	(0.35)	(0.37)	(0.26)	(0.16)	(0.07)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Num. obs.	73	72	71	70	69	68	67	66	65	64	63	62
Percentile	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

Each Model(i) has dependent variable (i)-quarter ahead average growth. Standard errors are bootstrapped using 10 000 draws

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

8. Conclusion

The paper contributes to the literature that examines the relationship between financial stability risks and the real economy for South Africa. The GaR methodology, popularised by Adrian *et al.* (2019, 2016), has potentially important policy applications. If the financial stability objective can be accurately measured using GaR then comparing the observed and targeted values would identify the extent to which macroprudential policy is “too loose” or “too tight”. Expressing financial stability risks in a common metric would also allow better coordination of macroprudential policy and other policies.

Specifically, we develop a framework for constructing a partitioned set of FCIs designed to feed into GaR analysis for South Africa. Three composite indices or ‘partitions’ are presented, capturing different types of risk. These arise from: asset valuations, leverage in the private nonfinancial sector, and external conditions. We note that other choices are possible, capturing other risks, using the same framework, as are other aggregation methods.

We show that using these composite indices of financial conditions as explanatory variables in GaR analysis allows a more detailed and intuitive interpretation of the macrofinancial risks to future growth. In the example we present, PNFS leverage varies negatively with the lower left-hand tail of the growth distribution over horizons up to 12 quarters, while asset valuations and external conditions vary positively with GaR at shorter horizons.

These types of results invite further research into their robustness, into the ability of the FCIs to capture the moments of the conditional distribution of GDP growth in an out-of-sample context, and into features like the potential term structure of GaR in South Africa.

Appendices

Appendix A: Data sources

1. **Asset valuations**

(ii) House price growth [FNB, Haver Analytics]

(iii) Bond spread: long / short [Government bond yields 10 years plus maturity (SARB KBP2003M); Government bond yields 0-3 years maturity (SARB KBP2000M)]

(iv) Bond spread: long/medium [Government bond yields 10 years plus maturity (SARB KBP2003M); Government bond yields 5-10 years maturity (SARB KBP2002M)]

(v) Sovereign CDS spread (Bloomberg)

(vi) JSE share price returns (Bloomberg)

(vii) JSE share prices: Price to earnings ratio (Bloomberg)

(viii) SRISK [V-Lab; https://vlab.stern.nyu.edu]

2 Private Nonfinancial Sector Leverage

(i) Ratio of debt-service cost to disposable income (SARB KBP6589L)

(ii) Household debt to disposable income (SARB KBP6525L)

(iii) Household debt to GDP (SARB)

(iv) Credit extension to the domestic private sector: Mortgage advances (SARB KBP1364M)

(v) Corporates: Debt to GDP (SARB)

(vi) Corporates: Debt to net operating profit (SARB)

(vii) Corporates interest coverage ratio (SARB) 37

(viii) Government: Debt to GDP (SARB)

(ix) Government: Current account to GDP ratio (SARB)

(x) Gross capital inflows (SARB)

3. External Conditions

(i) Real effective exchange rate (SARB KBP5395M)

(ii) Exchange rate volatility: 1 month implied volatility (Bloomberg)

(iii) Commodity price index (Bloomberg)

(iv) Oil price: Brent crude (SARB KBP5344M)

(v) Vix: Chicago Board Options Exchange Volatility Index (Bloomberg)

(vi) US GDP growth (FRED)

Appendix B: Contributions to FCIs

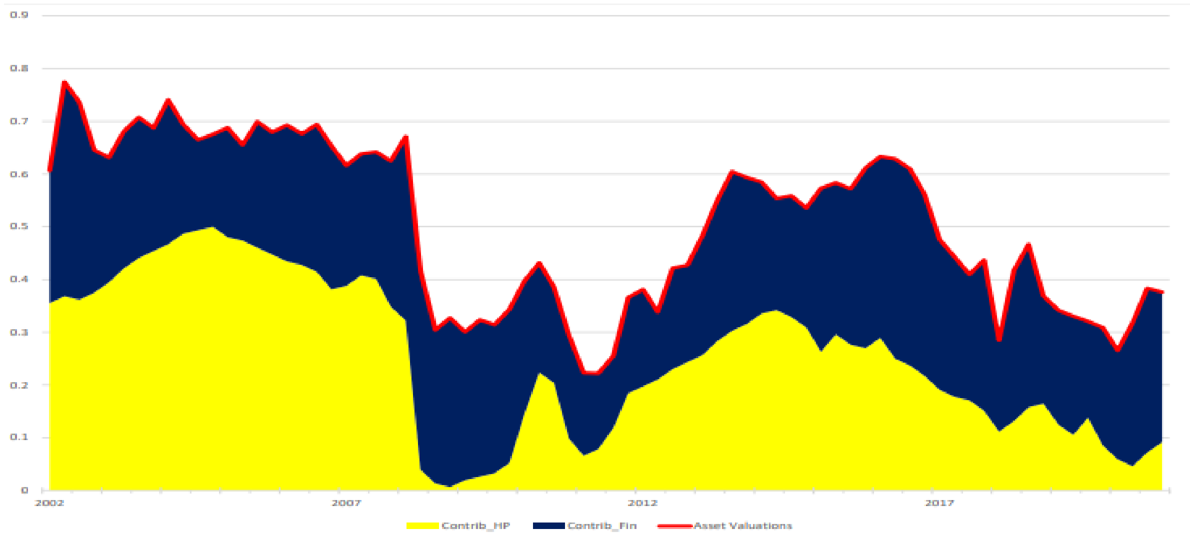


Figure 9: Contributions to Asset Valuations FCI

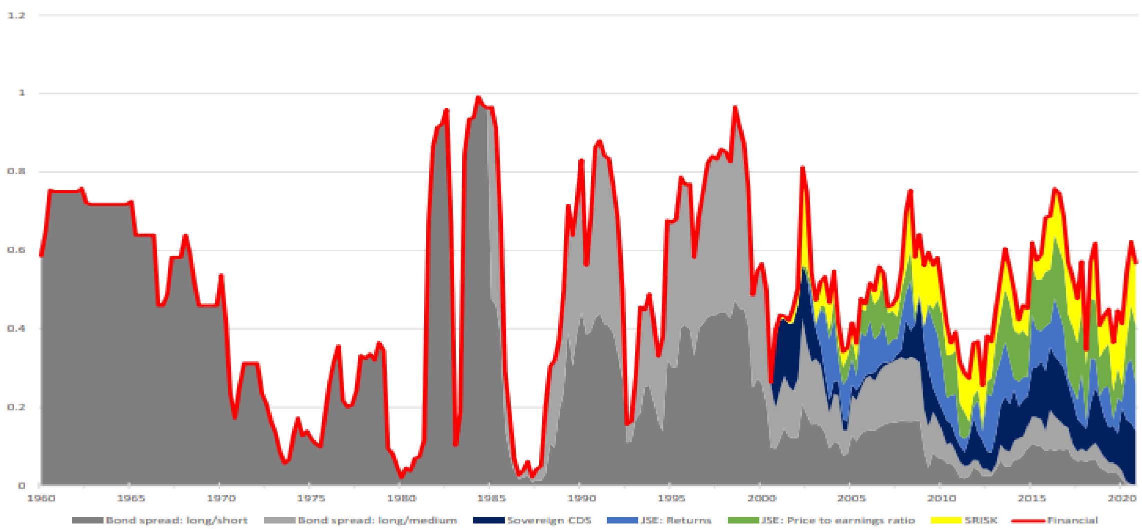


Figure 10: Contributions to Financial component of Asset Valuations FCI

Note: The figures present the composite index (red line), together with the contributions of the underlying components. Each index is mapped into the (0, 1) space based on its percentile in its historical distribution using the ecdf function in R.

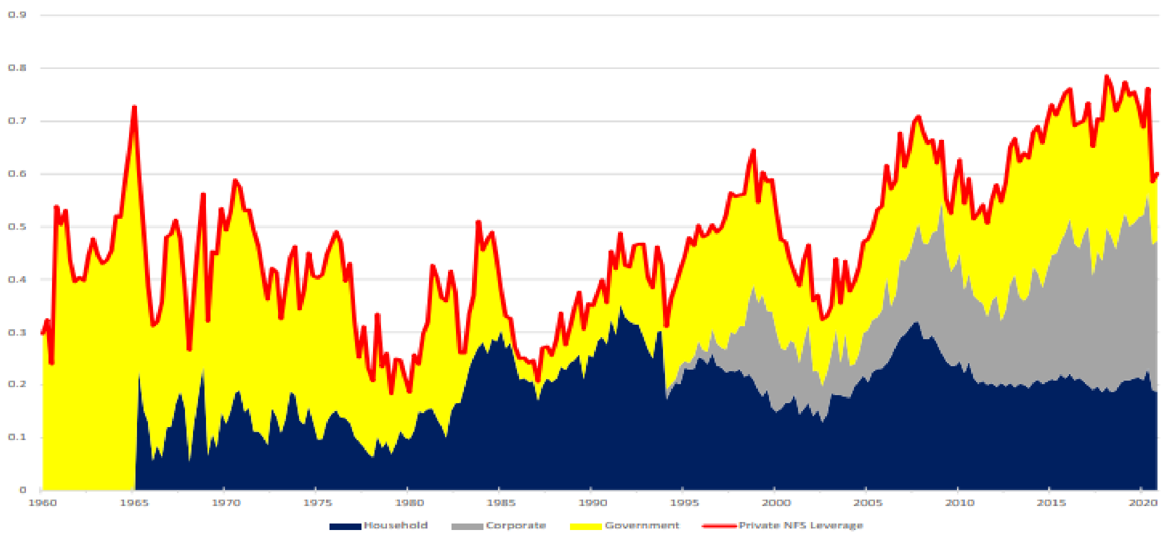


Figure 11: Contributions to PNFS Leverage FCI

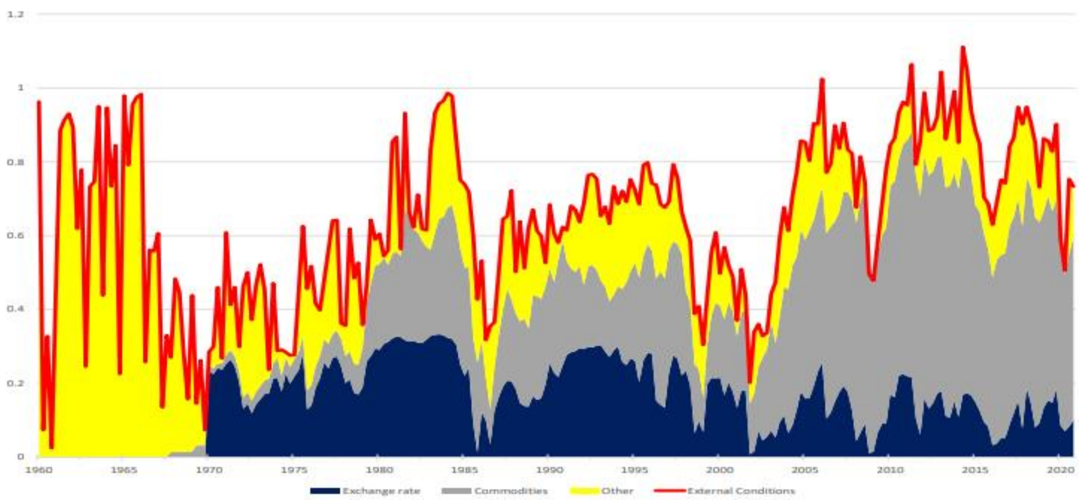


Figure 12: Contributions: External Conditions FCI

Note: The figures present the composite index (red line), together with the contributions of the underlying components. Each index is mapped into the (0, 1) space based on its percentile in its historical distribution using the ecdf function in R.

Appendix C: Granger-causality analysis

Granger (1969) presented a definition of causality that is closely related to predictability. A variable y_1 is Granger-causal for another series y_2 if the information in y_1 helps predict y_2. To formalise this, let $y_{2,t+h|\Omega_t}$ be the optimal h-step-ahead predictor of y_2 at time t based on all available information at this time Ω_t. Then: $y_{1,t}$ is Granger-noncausal for $y_{2,t}$ if and only if

$$y_{2,t+h|\Omega_t} = y_{2,t+h\Omega_t \backslash y_{1,s}|s \leq t} \quad (4)$$

where $\Omega_t \backslash A$ is the set containing all elements of Ω_t which are not in A. Therefore $y_{1,t}$ is Granger-noncausal for $y_{2,t}$ if and only if removing the information contained in past values of $y_{1,t}$ does not change the optimal forecast for $y_{2,t}$ at any horizon h. If including past values of $y_{1,t}$ improves the forecast, then $y_{1,t}$ is Granger-causal for $y_{2,t}$.

To (attempt to) operationalise the definition, consider the bivariate VAR(p):

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \alpha_{11,i} & \alpha_{12,i} \\ \alpha_{21,i} & \alpha_{22,i} \end{bmatrix} \begin{bmatrix} y_{1,t-i} \\ y_{2,t-i} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix}$$

The Granger-noncausality condition in equation (1) is now equivalent to the null hypothesis $\alpha_{21,i} = 0$ for $i = 1,2,\ldots,p$. For stationary processes, the null can be tested using standard χ^2 or F-tests of the Wald type. However, the tests may have nonstandard properties if some of the variables are $I(1)$ (Toda and Phillips, 1993).

References

Adrian T, A Ashcraft & N Centorelli (2013). “Shadow Bank Monitoring”, Federal Reserve Bank of New York Staff Report No. 638.

Adrian T, N Boyarchenko & D Giannone (2019). “Vulnerable Growth”, American Economic Review, 109(4), 1263-89.

—— (2016). “Vulnerable growth””, Staff Report No. 794, Federal Reserve Bank of New York.

Adrian T, D Covitz & N Liang (2015). “Financial Stability Monitoring”. Annual review of Financial Economics, 7, 357–95.

Adrian T, F Grinberg, N Liang & S Malik (2018). “The Term Structure of Growth-at-Risk”, Working Paper No. 42, Hutchins Center.

Aikman D, J Bridges, S Hacioglu Hoke, C O’Neill & A Raja (2019). “How do ˘ financial vulnerabilities and bank resilience affect medium-term macroeconomic tail risk?”. Mimeo, 4 June 2019.

Aikman D, J Bridges, S Burgess, R Galletly, I Levina, C O’Neill & A Varadi (2018). “Measuring risks to UK financial stability”, Bank of England working papers 738, Bank of England.

Apostolakis G & A P Papadopoulos (2019). “Financial Stability, Monetary Stability and Growth: A PVAR Analysis.” Open Economies Review, 30, 157–78.

Arregui N, S Elekdag, G Gelos, R Lafarguette & D Seneviratne (2018). “Can Countries Manage Their Financial Conditions Amid Globalization?”, IMF Working Papers, No 2018/015, International Monetary Fund.

Balcilar M, K Thompson, R Gupta & R Van Eyden (2016). “Testing the asymmetric effects of financial conditions in South Africa: A nonlinear vector autoregression approach”. Journal of International Financial Markets, Institutions and Money, 43, 30-43.

Basel Committee on Banking Supervision. (2010). Guidance for national authorities operating the countercyclical capital buffer. Bank for International Settlements, July.

Bernanke B (2018). “The Real Effects of the Financial Crisis.” Brookings Papers on Economic Activity, Fall, 251–342.

—— (2013). “Monitoring the financial system”. Speech at the 49th Annual Conference on “Bank Structure and Competition”, 10 May.

Borio C, M Drehmann & F Xia (2020). “Forecasting recessions: the importance of the financial cycle”, Journal of Macroeconomics, 66, 1-15.

Brave S & A Butters (2012). “Diagnosing the Financial System: Financial Conditions and Financial Stress.” International Journal of Central Banking, 8(2), 191–239.

Brownlees C & A B M Souza (2021) “Backtesting Global Growth-at-Risk.” Journal of Monetary Economics, 118, 312–330.

Carlson M A, K F Lewis & W R Nelson (2012) “Using Policy Intervention to Identify Financial Stress.” Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Cecchetti S (2006). “Measuring the Macroeconomic Risks Posed by Asset Price Booms”, Asset Prices and Monetary Policy, University of Chicago.

Cecchetti S & L LI (2008). “Do Capital Adequacy Requirements Matter for Monetary Policy”, Economic Inquiry, 46(4), 643-659.

Chamon M & C Crowe (2012). “Predictive Indicators of Crises”, in G Caprio (ed) Handbook in Financial Globalization: The Evidence and Impact of Financial Globalization, London: Elsevier, 499–505.

Chatterjee S, J Chiu, S Hacioglu-Hoke & D Thibaut (2017). “A financial stress index for the United Kingdom,” Bank of England working papers 697, Bank of England.

Danielsson, J, M Valenzuela & I Zer (2018). “Low risk as a predictor of financial crises,” FEDS Notes. Washington: Board of Governors of the Federal Reserve System, May 9, 2018, https://doi.org/10.17016/2380-7172.2169. 43.

Drehmann, M., Borio, C. & Tsatsaronis, K. (2011). Anchoring counter-cyclical capital buffers: The role of credit aggregates. International Journal of Central Banking, 7(4), December, 189–240. Drehmann, M. & Juselius, M. (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. International Journal of Forecasting, 30(3), 759–780.

Farrell G (2016) Countercyclical capital buffers and real-time credit-to-GDP gap estimates: A South African perspective”. Journal of Studies in Economics & Econometrics, 40(1), 1–19.

Farrell G & E Kemp (2019) “Measuring the Financial Cycle in South Africa”, South African Journal of Economics, 88(2), 123–144, June.

—— (2017). “A framework for monitoring financial stability”. SARB Financial Stability Department Topical Briefing No 8 of 2017, November.

Frankel J & G Saravelos (2010). “Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008–09 Global Crisis,” National Bureau of Economic Research Working Paper 16047, June.

Giglio S, B Kelly & S Pruitt, (2016). “Systemic Risk and the Macroeconomy: An Empirical Evaluation”, Journal of Financial Economics, 119, 457- 471.

Giordani P, E Spector & X Zhang (2017). “A new early warning indicator of financial fragility in Sweden”.

Gumata N, N Klein, & E Ndou. 2012. “A financial conditions index for South Africa”. IMF Working Paper, 12/196.

Hatzius J, P Hooper, F Mishkin, K Schoenholtz and M Watson (2010). “Financial Conditions Indexes: A Fresh Look after the Financial Crisis”. Report presented at the Chicago Booth Initiative on Global Markets 2010 US Monetary Policy Forum, New York, 26 February.

IMF (2017a). “Are Countries Losing Control of Domestic Financial Conditions?” In Global Financial Stability Report (April), Chapter 3. Washington, 44 DC: International Monetary Fund.

—— (2017b). “Financial Conditions & Growth at Risk.” In Global Financial Stability Report (October), Chapter 3. Washington, DC: International Monetary Fund.

Jordà O, M Schularick & A M Taylor (2013). “When credit bites back”, ` Journal of Money, Credit & Banking, 45(2), 3–28.

Kaminsky G, S Lizondo & C Reinhart (1998). “Leading indicators of currency crisis”. IMF Staff Papers, 45(1), 1-48.

Kabundi A & A Mbelu (2020). “Estimating a time-varying financial conditions index for South Africa”. Empirical Economics, 60(4), 1817–1844, April.

Kasaï N & R Naraidoo (2013). “The opportunistic approach to monetary policy & financial market conditions.” Applied Economics, 45(18), 2537– 2545.

Kemp E & G Farrell (2017). “A Framework for Monitoring Financial Stability”, Financial Stability Department Topical Briefing 8 of 2017, November.
Kisten T (2020) “Monitoring Financial Stress in South Africa”, Emerging Markets Finance & Trade, 1–18. Published online 20 October 2020.

Kliesen K L, M T Owyang & E K Vermann (2012). “Disentangling Diverse Measures: A Survey of Financial Stress Indexes”. Federal Reserve Bank of St. Louis Review, September/October.

Koenker, R (2005). Quantile Regression. Cambridge University Press: Cambridge, U.K.

Koop G and D Korobilis D (2014). A new index of financial conditions. European Economic Review, 71, 101–116.

Lafarguette R (2019). “GaR Tool: Technical Appendix”, February.
Plagborg-Møller, M, L Reichlin, G Ricco & T Hasenzagl (2020). “When is Growth at Risk?”, Brookings Papers on Economic Activity (Spring), 167-229.

Prasad A, S Elekdag, P Jeasakul, R Lafarguette, A Alter, A Xiaochen Feng and C Wang (2019). “Growth at Risk: Concept & Application in IMF Country 45 Surveillance”, IMF Working Paper WP/19/36, February.

Reichlin L, G Ricco and T Hasenzagl (2020). “Financial Variables as Predictors of Real Growth Vulnerability.” Deutsche Bundesbank Discussion Paper No. 05/2020.

Sing M (2019a) “South Africa’s FCI & Real GDP Growth-at-Risk”. SARB Financial Stability Department Topical Briefing, February.

—— (2019b). “Decomposing & Operationalising the Financial Conditions Index”, mimeo.

South African Reserve Bank (2016). “A new macroprudential policy framework for South Africa”, Discussion Paper, November.

Thompson K, R Van Eyden & R Gupta (2015). “Identifying a financial conditions index for South Africa”. Studies in Economics & Finance, 32(2), 256– 274.

Twala Z & M van der Linde (2018). “A heatmap for the South African financial system”, Topical Briefing, Financial Stability Department.

TOPICAL BRIEFING: No. 5

Financial Stability Department

SOUTH AFRICAN RESERVE BANK

New Econometric Models for South African House Prices, Mortgage Debt and Residential Investment

Abstract

Housing markets and the associated credit markets have important implications for monetary transmission, stabilisation policy and for financial stability. The crucial role of housing wealth as collateral for mortgage borrowing in South Africa, means that changes in housing wealth, mainly due to house price changes, induce large effects on consumer spending. Moreover, residential investment, also driven by house prices, is a volatile component of aggregate demand. It is important therefore to understand the dynamics of house prices, and how they feed through to consumption, mortgage debt and residential investment. This paper develops econometric models for house prices, the mortgage stock and residential investment in South Africa, drawing on insights from theory and the findings of the international literature, and complementing our previous research on consumption (Aron and Muellbauer, 2013). These equations are suitable for incorporation in the SARB's Core Model, improving on the existing house price and mortgage stock equations and clarifying the residential investment channel, currently modelled only through aggregate investment. The estimated equations yield important insights into monetary transmission, with a powerful transmission from interest rates and credit conditions in the mortgage market to house prices, and hence into aggregate demand. There is evidence of a memory of up to four years regarding the expectations of house price appreciation by housing market participants. This implies that a series of positive shocks to housing demand can feed back positively onto housing demand and onto house prices, so extending boom conditions. This can potentially cause house prices to overshoot relative to their fundamentals, with clear implications for risks to financial stability. These new findings

should benefit future development of the SARB's Core policy model of the economy to better inform policy makers.

Acknowledgements: We are extremely grateful to the following persons at the SARB for their advice particularly on data: Wian Boonzaaier, Elriette Botes, Rashad Cassim, Shaun de Jager, Karen Kuhn, Danie Meyer, Caswell Monyela, Mpho Moloto, Lesego Morope, Khumbudzo Muneri, Susana Paulse, Pieter Pienaar and Bart Stemmet. And from the BER: George Kershoff, and Ben Smit. And from the FNB: John Loos and Siphamandla Mkhwanazi. We appreciate the review from Siphiwo Bitterhout.

Content

Executive Summary

Housing markets and the associated credit markets have important implications for monetary transmission, stabilisation policy and for financial stability. The crucial role of housing wealth as collateral for mortgage borrowing in South Africa, means that changes in housing wealth, mainly due to house price changes, induce large effects on consumer spending. Moreover, residential investment, also driven by house prices, is a volatile component of aggregate demand. It is important therefore to understand the *dynamics* of house prices, and how they feed through to consumption, mortgage debt and residential investment.

This paper develops econometric models for house prices, the mortgage stock and residential investment in South Africa, complementing our previous research modelling consumption in South Africa (see Aron and Muellbauer (2013)).

These equations are suitable for incorporation in the SARB's Core Model (see Smal *et al.* (2007), subsequently revised) and the MPRU-Core model which has extended the Core model to addresses macro-prudential issues (De Jager *et al.*, 2021). Our new research improves on the existing house price and mortgage stock equations and clarifies the residential investment channel, currently modelled only through aggregate investment, see our assessment of the SARB's Core models in Aron and Muellbauer (2022a).

The new *house price equation* for South Africa is plausible and well-fitting and yields important insights into monetary transmission. It is based on the 'inverted demand principle', where the price is determined by demand that varies relative to the existing housing stock. There is a powerful transmission from interest rates and credit conditions in the mortgage market to house prices. Both mortgage spreads and loan-to-value ratios (LTVs) appear to be relevant proxies for credit conditions. House price expectations relative to mortgage rates determine 'user cost', which is a key driver of housing demand. There is evidence of a memory of up to four years regarding the expectations of house price appreciation by housing market participants. This implies that a series of positive shocks to housing demand can feed back positively onto housing demand and onto house prices, so extending boom conditions. This can potentially cause house prices to overshoot relative to their fundamentals. Such overshooting has clear implications for risks to financial stability and is relevant when designing stabilisation policy.

From our evidence of a shift in the effect of changes in the exchange rate, there may have been a weakening from 2015 of capital inflows entering the housing market, previously linked with a momentum effect from exchange rate appreciation. Given the linkages from interest rates to the exchange rate, this suggests there may have been a shift in the effect of monetary policy on house prices after 2015 in South Africa

The new *mortgage stock equation* for South Africa, as for the house price equation, finds that interest rates and credit conditions have powerful effects. The relative direct effect of mortgage spreads and LTVs is somewhat different on the mortgage stock than in the house price equation. For mortgages, the direct effect of the level of LTVs is greater than for house prices, while spreads have only temporary effects. However, since a key driver of the mortgage debt to income ratio is the level of house prices to income, there are large *indirect* effects of interest rates and credit conditions on the mortgage stock via the house prices to income ratio. Also, the extrapolative element of expectations of house price appreciation, embedded in house prices, has an *indirect* effect. This implies that mortgage debt, like house prices, can overshoot fundamentals. High levels of mortgage debt relative to income can thus pose risks for financial stability. There may also be risks of sharp downturns in consumer spending if interest rates were to rise.

Estimates for both the house price and the mortgage stock equations are limited by the historical span of data on LTVs and on mortgage spreads, which begin around 2000. In particular, there is only one turning point in the series for mortgage debt to income, from 2001 to just before the arrival of the pandemic in 2020, making robust identification of parameter estimates difficult. Hence, the model for mortgage debt is necessarily provisional. Nevertheless, the model is very consistent with evidence from other countries.

The new residential investment equation for South Africa yielded a stable relationship for data back to 1978, despite the many structural changes and shocks experienced by the South African economy. The key driver of residential investment relative to GDP is the relative price of houses to construction costs. This finding is consistent with international evidence from an important OECD study, Cavalleri *et al.* (2019). The implication for monetary transmission is that the powerful effect of interest rate and credit conditions on house prices in South Africa also transmits to this rather volatile component of aggregate demand, residential investment. While there is no evidence of interest rates effects in the long-run solution – except *indirectly* via house prices, there are powerful short-term effects of changes in the prime rate of interest on residential investment. There was an apparent moderation in residential investment as population growth fell with the AIDS epidemic and from 2017, probably associated with worsening economic and political prospects.

Elevated levels of credit risk indicators like non-performing loans (NPLs) and bank loan loss provisioning are a recurrent characteristic of banking crises. Such crises are typically preceded by poor quality of lending, excessive credit growth and high levels of leverage. As NPLs rise and banks apply tougher lending criteria for firms and households, a credit crunch may follow. Together with the *new model for loan loss provisioning* in South Africa in Aron and Muellbauer (2022b), the three models presented in this paper illuminate this *two-way connection* between credit conditions and credit risk indicators. Liberal credit conditions drive up house prices and mortgage debt, both potentially overshooting and creating financial vulnerability for borrowers and lenders, exacerbated by falling or stagnant GDP.

1. Introduction

Since the global financial crisis, much research has examined the links between credit growth, especially if real-estate linked, the overvaluation of house prices and financial stability. For example, Cerutti *et al.* (2017), analysing an (unbalanced) panel data set of 50 countries for 1970-2012, find that house-price booms are more likely in countries with higher loan-to-value ratios and with mortgage funding arrangements based on securitisation or wholesale sources such as money markets. They find that most house-price booms end in recession, where the downturns are deeper and longer if preceded by booms in both residential mortgages and other private debt. Recession is also linked with reliance on non-retail deposit funding if this generates 'duration mis-match' problems on lenders' balance sheets. Duca *et al.* (2021) consider links between house price overvaluations and financial instability and suggest ways in which empirical evidence can be used to detect episodes of overvaluation. The amplifying, and sometimes stabilising, feedback loops in real estate booms that are accompanied by credit booms are spelled out further by Aron *et al.* (2020).

The latter set of authors consider how such feedbacks vary between countries, and how South Africa compares with other countries. In the short run, there can be strong positive feedbacks from house prices to consumption, when down-payment constraints are loose, access to home equity loans is easy and rates of owner occupation are high. This potentially amplifies risks to financial stability. The UK and the US offer a sharp contrast in this respect to Germany and France, where this amplification does not occur. South Africa is probably closer to the UK than to Germany and France in the consumption channel of transmission. Another transmission from house prices to aggregate demand operates via residential investment, which boosts employment and household income. Aggregate demand in turn feeds back onto house prices. In the UK, where the housing supply elasticity is low, such a feedback would be weaker than in the US, Ireland or Spain. Below we show from our residential investment equation that this transmission channel is strong in South Africa, with a housing supply elasticity that exceeds that for most of Europe.

There can also be pronounced macroeconomic effects from an overshooting of house prices that is induced by a series of strong positive shocks, such as to interest rates, income or credit supply conditions. The effect of these shocks could be amplified by the expectations of house-owners and buyers based on the extrapolation of past capital gains. This type of overshooting is found particularly in economies where high levels of mortgage debt leverage are possible, such as the UK and US (but not Germany or France), since leverage can amplify both returns and risks. In South Africa, the house price equation, that we share below, suggests that expectations are based on the previous four years of house price appreciation, which accords with US findings (Duca *et al.*, 2011; 2016). This implies that this type of overshooting of house prices is of particular significance in South Africa, consistent with the high degrees of household debt leverage available in South Africa.

Understanding what drives house prices, and especially understanding the role of credit conditions, is crucial to evaluating risks from over-valued house prices and over-extended household debtors and mortgage lenders. In this paper, we set out and estimate three equations for house prices, household holdings of mortgage debt and for residential investment, in which variations in credit conditions play an important role. The empirical evidence will be helpful both in improving the understanding of monetary transmission from interest rate policy to aggregate demand, but also in assessing risks to financial stability.

2. Considerations about Credit Condition

Economists have become far more aware of the importance of shifts in credit conditions since the Global Financial Crisis, particularly for mortgage and housing markets, see the literature survey by Duca *et al.* (2021). Mortgage lenders facing endemic asymmetric information use credit scores and information on income (e.g., from payslips) to assess the credit worthiness of potential borrowers. To set credit terms and to ration credit they use loan-to-value ratios and debt-to-income or debt service-to-income ratios as well as risk pricing (i.e., charging higher interest rates on more risky loans). When lenders have access to plentiful capital and feel positive about the economy, with a higher risk appetite, they tend to relax credit conditions, without necessarily cutting mortgage rates. Thus, they will relax credit score requirements and permit borrowing at higher loan-to-value and loan-to-income ratios. A reduction in spreads relative to bank funding costs is often a sign of easier credit conditions. Hence it is important for central banks to track the spreads on new lending, when monitoring mortgage markets. Similarly, they should track loan-to-value and loan-to-income or debt service-to-income ratios, when available, especially for first-time borrowers, the most likely to face credit constraints.

Lacking such information, Chauvin and Muellbauer (2018) used a latent variable approach to estimate a French mortgage credit conditions index, MCCI, from a system of six equations, for house prices, mortgage debt, consumption, non-mortgage debt, liquid assets and permanent income. This is a 'Latent Interactive Variable Equation System (LIVES), see Duca and Muellbauer (2013). Aron and Muellbauer (2013) used a three-equation latent variable model for household debt, consumption and permanent income to estimate a composite South Africa credit conditions index, CCI, covering both mortgage and non-mortgage debt, for 1970 to 2005. However, for models of house prices and of the mortgage stock a *composite* CCI is less useful, as credit conditions for unsecured credit and other non-mortgage borrowing may evolve rather differently from those in the mortgage market.

In future work, it may be desirable to apply the latent variable approach, as in Chauvin and Muellbauer, to a larger equation system for South Africa. In the present paper, a simpler approach is adopted: two indicators closely related to credit conditions in the mortgage market are instead used to define *proxies* for the MCCI. These are used in three separate single equations for house prices, mortgage debt and residential investment. The indicators are loan-to-value ratios and spreads between the actual mortgage rates paid on new advances and the interest rate on prime loans (or the repo rate). Figure 1 illustrates the loan-to-value ratio compiled from Deeds Office data by the First National Bank (FNB), and Figure 2 shows the mortgage spread defined as the prime rate of interest minus the actual interest rate on new mortgage loans. Both graphs suggest that credit conditions were eased from 2003 to 2008, followed by a sharp contraction associated with the Global Financial Crisis (GFC). After a modest recovery, somewhat earlier for the loan-to-value indicator than for the mortgage spreads indicator, there was renewed tightening until about 2016. From 2017, the two graphs diverge, with the loan-to-value ratio trending upwards while the spread narrows.

One can raise questions about how to interpret LTV data averaged over all mortgages (rather than applying to first-time buyers). In the US, data on loan-to-value ratios for all mortgages show much less of an association with the evolution of credit conditions than do data on LTVs for first-time buyers, Duca *et al.* (2016). This is because, in the long housing market upswing from the late 1990s to 2006, many repeat buyers were able to use the increased equity on their previous homes to moderate the degree of leverage they needed to take on to move up the housing ladder. Hence, average LTVs in the US show far less of a rise than LTVs for first-time buyers, many more of whom will have been credit-constrained. In South Africa, credit constraints are prevalent among a greater fraction of all types of mortgage borrowers, and hence it is plausible that the *average* LTV is a more reliable indicator of credit conditions than in the US. Nevertheless, the rise in the average LTV from 2017 is somewhat puzzling. It is possible that with the low volumes of transactions after 2017, lenders were being stricter about credit scores and income checks and so were able to offer higher LTVs. A plausible alternative explanation is that, with the rise of stock market valuations, the ability to use pension assets as security for mortgages, as is allowed in South Africa, enabled lenders to offer larger loans that permitted lower initial cash deposits from borrowers.

3. A House Price Model

3.1. Theoretical background - Inverse Demand Approach for Derivation of a House Price Equation[61]

There are two theoretical frameworks for deriving an aggregate house price equation: the house price-to-rent approach, based on asset market arbitrage, and the inverse demand approach that comes from consumer theory. The house price-to-rent approach requires well-functioning rental markets, where many households are at the choice threshold between renting and becoming owner-occupiers, and good historical data on rents. Such data are absent for South Africa. The alternative framework inverts the demand for housing services, treating the housing stock as predetermined. The resulting inverted demand equation implies that real house prices are driven by the user cost of housing, real incomes, and the housing stock. Kearl (1979) first fully articulated this approach, followed by Hendry (1984), Poterba (1984), and DiPasquale and Wheaton (1994). In practice, changes in mortgage borrowing constraints importantly alter the effective demand for housing and hence house prices, see Dougherty and Van Order (1982), Meen (1990, 2001), Muellbauer and Murphy (1997),

[61] This section follows the exposition in Duca, Muellbauer and Murphy (2021).

Anundsen and Heebøll (2016), and Favara and Imbs (2015), inter alios. Critical to implementing this framework, which we adopt below, are good estimates of the housing stock and income.

The demand for housing services is assumed to be proportional to the housing stock. The latter is fixed in the short run, and prices are solved by inverting the demand function. In a simple log-linear approximation, the log of demand, h_{ijt}, of household i in area j (considered in isolation) is:

$$ln\ h_{ijt} = -\alpha_i\ ln\ hp_{jt} + \beta_i\ ln\ y_{it} + z_{it} \qquad (1)$$

where *hp* is the real house price in location j, *y* is real income and *z* denotes other demand shifters including the user cost of housing.[62] The own-price elasticity of demand is –*α*, and the income elasticity is *β*. As Muellbauer and Murphy (1997) note, equation (1) may be derived from an explicit multi-period utility maximisation problem where there are two goods - housing services and a composite consumption good (see Dougherty and Van Order (1982), for example). Then y is a measure of permanent income or some combination of physical and financial wealth and current and future real income. Equation (1) also omits transactions costs such as real estate agent and legal fees and property transaction taxes, which in some countries can amount to 8% or more of the house price. The literature on investment with lumpy adjustment costs suggest that households will adjust their demand for housing in a discrete manner, when some thresholds are crossed. However, since both households and the housing stock are heterogeneous, aggregate behaviour is likely to be smooth (Bertola and Caballero 1990). For example, if income or another demand shifter rises, marginal households near the threshold where benefits equal transaction costs, are pushed over the threshold and will transact, raising demand. Equation (1) is static but, in practice, the response of house prices to demand shocks is likely to be drawn-out since housing transactions take time and generally entail time-consuming search, (Wheaton 1990 and DiPasquale and Wheaton 1994)[63].

[62] The price of housing services equals the product of the user cost of housing times house prices.

[63] Product heterogeneity and time-consuming search can make the sales time for a house long, variable, and difficult to attribute to demand swings or randomness (Chinloy 1980, Haurin 1988, and Wheaton 1990). In this environment, rapid price adjustments may not be rational (see Quigley 1979, Rothschild 1981, and Stull 1978).

Solving for housing prices, *hp*, involves aggregating the micro-demands to a market demand schedule and inverting this to obtain a solution for average house prices at location *j*:

$$\ln hp_{jt} = (\beta \ln y_{jt} - \ln h_{jt-1} + z_t)/\alpha \qquad (2)$$

where y_{jt} is average real income in location *j*, h_{jt-1} is last period's housing stock in area *j*, z_t is the average of other demand shifters, and α and β are averages of micro-parameters. This stylistic representation omits the lags resulting from transactions costs, and the demand equation (1) ignores household location choices.

Demand for housing at location *j* depends on both current residents and on those living elsewhere who opt to relocate to *j*. Aggregating to the national level, these local relative considerations tend to wash out. However, international relativities can matter, particularly for major cities where internationally mobile households tend to locate (Englund and Ioannides 1997).

An advantage of the inverted demand approach is that it is well grounded theoretically, unlike many 'ad hoc' approaches. In addition, there are strong priors for key long-run elasticities, such as the 'central estimates' in Meen (2001). For example, time-series estimates of the income elasticity of demand often find that β is near 1, in which case the income and housing-stock terms in equation (2) simplify to log income per property, i.e., $\ln y - \ln h$. However, the income elasticity of *house prices*, given the stock, is β/α, which typically notably exceeds 1 since the own-price elasticity of demand for housing, α, is below 1 in absolute magnitude. Forecasts of house prices from this approach need to model construction or residential investment (e.g., DiPasquale and Wheaton 1994) as well as to forecast income, interest rates and credit availability.

The demand shifters (*z*) include the user cost of housing, demography and credit availability. As housing is durable, intertemporal considerations suggest that expected or permanent income and user costs are important. The latter considers that durable goods deteriorate, but may appreciate in price, and incur interest and tax costs. Absent transaction costs and credit constraints, and tax deductibility of mortgage interest, user costs are usually approximated as:

$$uc = i + t_p + \delta + \sigma - \Delta HP^e / HP \qquad (3)$$

where i is the nominal mortgage interest rate, t_p is the property tax rate, δ is the deterioration rate, σ is a (possibly time-varying) risk premium, and $\Delta HP^e/HP$ is the expected nominal rate of appreciation. The formulation for the property tax rate assumes a tax rate fixed in the short-run and continuous revaluation to current prices of the house on which the tax is charged. If this is not the case, it is preferable to make a separate allowance for the tax rate outside the user cost term. The derivation of equation (3) assumes houses are traded every period. However, as DiPasquale and Wheaton (1994) stress, the expected appreciation term should reflect planned holding periods, as transactions costs impede trading, and so should not just refer to very short-run appreciation.

The user cost is not the only channel through which interest rates affect housing demand. Kearl (1979) notes that typical mortgages stabilise nominal payments. For credit-constrained households, cash-flows matter so that the debt-service ratio affects demand. Moreover, the debt-service-to-income ratio (DSTI), along with LTV and DTI ratios, is used by lenders to set loan terms and decide whether or not to lend. Thus, as nominal mortgage rates fall, one of the lending criteria becomes less binding, thereby increasing credit supply[64]. The implication is that nominal, as well as real, mortgage interest rates are likely to affect housing demand and therefore house prices in countries where the debt-service ratio is a key lending criterion.

The user cost, first formulated for consumer durable goods by Cramer (1957), regards the durable good only as a consumption item. However, the structure and land components of housing are also major stores of value that compete with other assets. This means that part of the demand for housing comes from its role as part of a wealth portfolio, implying that relative returns and risks for other assets also affect housing demand. The relevance of low returns on other assets versus strong house price appreciation is particularly high in the current period of lower bond yields. It also means that the positive effect of income growth expectations on housing demand—and hence on house prices that comes from thinking of housing purely as a consumption good—could be reversed if a major motive is the saving motive. Indeed, Campbell (1987)

[64] In France, regulatory DSTI caps strengthen the effect of nominal interest rates (Chauvin and Muellbauer (2018)).

highlights how saving could rise in anticipation of future income declines.

The user cost term in equation (3) does not account for how leverage affects the relative returns to buyers using mortgages, see Muellbauer and Murphy (1997). Leverage amplifies returns and risks, implying that the coefficient of the user cost term in a house price equation should depend on how much leverage lenders provide to home-buyers, as measured by the LTV, and hence the general state of mortgage credit conditions.

In addition to such portfolio considerations, the availability of home equity withdrawal creates a potential third source of demand for housing, in addition to the standard demand for a durable good and the portfolio demand. In countries, such as South Africa and the U.S., with easy access to home equity loans, the role of housing as collateral for borrowing, gives households with positive housing equity a means of overcoming credit constraints that would otherwise prevent or raise the cost of borrowing.

3.2. Empirical Evidence from Other Countries

We begin by summarising the evidence from applying the systems approach using the 'Latent Interactive Variable Equation System' (LIVES) methodology described above in a range of individual countries. LIVES has been used to model house prices, mortgage debt and other variables in Chauvin and Muellbauer (2018) for France, and in Geiger *et al.* (2016) for Germany. In each case, a six-equation system was modelled, and the other variables included were consumption, non-mortgage debt, liquid assets and permanent income. Similarly, Muellbauer *et al.* (2015) used the LIVES approach to model house prices, mortgage debt and consumption in Canada.

Empirical findings for Germany and Canada for house prices are broadly in line with those for France. Some of the main findings for France of Chauvin and Muellbauer (2018) give useful context for the South African evidence shown in our models below.

The house price equation has a very strong long-run solution with a quarterly speed of adjustment of around 0.12 (t=13). Mortgage credit conditions are crucial: if MCCI is omitted the speed of adjustment collapses and few long-run coefficients make sense.

The elasticity of real house prices with regard to the nominal mortgage rate is -0.38 (t=-12) and seems to be quite stable. In France's fixed mortgage rate market, where lenders focus strongly to keeping the debt-service ratio below a ceiling of around 40%, the importance of the nominal mortgage rate makes particular sense. The elasticity with regard to user cost varies significantly with MCCI and is around -0.035 at the peak value of MCCI. The elasticity of house prices with respect to income per house is 2 and looks to be fairly stable. Assuming an income elasticity of demand for housing of 1, which can be accepted, this implies that the price elasticity of housing demand with regard to average house prices is -0.5, which is in line with studies surveyed by Meen (2001). The elasticity of real house prices with regard to log permanent/current income is around 0.5. There are also significant demographic effects from the ratios of children and pre-retirement adults to the total number of adults.

Multi-country empirical evidence on the determination of house prices comes from Cavalleri *et al.* (2019). They apply equilibrium correction models to real house prices in 23 countries. Taking averages across countries, they find an average elasticity of response in the long-run of real house prices to income per house of 1.8 and an average response to the real mortgage rate of -0.3. They did not check for the influence of the nominal mortgage rate and do not attempt to control for variations in credit conditions. Demography is represented only by the log of population. Estimates of the average speeds of adjustment are not reported, but are probably quite low, given the omissions in the specification chosen for estimation.

Finally, we turn to central bank models. Many major central banks have semi-structural econometric models comparable with the SARB's Core model, several of which include house price equations. For example, the influential FRB-US model adopts the house price/rent approach based on asset pricing theory. However, the weaknesses are that there are no controls for credit conditions and the speed of adjustment is only 0.012 per quarter, implying almost no role for the adjustment of house prices to rents. In Australia's core model, MARTIN, the house price equation, like that in FRB-US, is also based on the asset price arbitrage approach. In the long-run, real house prices depend on the rent index and on a real interest rate, taking no account of varying credit conditions. The short-term dynamics does include a calibrated effect from the change in the nominal mortgage interest rate. The speed of adjustment is a remarkably low 0.02, a clear sign of omitted variables. A new policy model from the Bank of France, Lemoine *et al.* (2019) assumes that real house prices are governed by a simple autoregressive process with two lags, to the exclusion of all economic variables.

The new ECB model, ECB-BASE, for the whole Eurozone, does adopt the inverse demand approach, unlike the models at the FRB, RBA and Bank of France. This model finds an elasticity with regard to income per house of around 1, lower than suggested by other studies, and a strongly significant user cost effect. There are no controls for the nominal interest rate, credit conditions and demography, and the speed of adjustment is a low 0.036 per quarter, suggesting omitted variables. One can also question the choice of aggregating data over countries with such diverse credit institutions and house price dynamics, likely to give rise to measurement biases from implausible restrictions.

The Netherlands central bank model DELFI 2.0 is different from most models in assuming a long run solution for log nominal house prices as a linear function of the log nominal mortgage stock. The dynamics includes lagged growth in the mortgage stock and changes in interest rates and the speed of adjustment is 0.04. For the model as a whole, in which house prices also influence consumption and residential investment, much then depends on the equation for mortgage credit. This includes three proxies for credit conditions amongst the explanatory variables: an S-shaped linear trend (a proxy for the gradual loosening of bank lending standards in the 1990s), the ECB's Bank Lending Survey (available from the end of 2002 onwards), and the banking sector's leverage ratio.

3.3. Econometric Specification for House Prices

We follow the general specifications of Chauvin and Muellbauer (2018). An equilibrium correction framework is adopted, in which adjustment to the long-run solutions implied by theory takes time. Given the theory background set out in section 2 above, the long-run solution for the house price equation is an inverted log-linear demand function, where real house prices, *rhp*, are determined by household demand, conditional on the lagged housing stock.

$$\ln rhp_t = h_{0t} + h_{1t} \ln prime_t + h_{2t} \ln user_t + h_3(ln(y_t/hs_{t-1}) + h_{4t}E_t \ln (y_t^p/y_t)) + h_5\, demog_t + h_6 LA_{t-1}/y_t + h_7 IFA_{t-1}/y_t \quad (4)$$

In this equation, the intercept term, h_{0t}, captures shifts in demand, which should increase with mortgage credit conditions, represented by an index MCCI. The nominal mortgage rate is *prime,* and user cost, measuring interest rates minus expected appreciation, is *user.* Both effects should be negative, and potentially could vary with MCCI. The coefficient h_3, for the log ratio of income to the housing stock[65], is expected to be positive, and from the theory is measuring minus the inverse of the price elasticity of demand for housing, see above. The coefficient h_{4t} captures the relative effect of permanent to current income. The sign is ambiguous as there are offsetting influences. Standard demand for housing as a consumption item would suggest a positive coefficient as in a consumption function. But portfolio and collateral demand for housing, as a way of saving for the future, imply the opposite sign: more optimistic income expectations should reduce the demand for this store of value[66]. In principle, either influence could vary with mortgage credit conditions MCCI. The remaining terms represent the effects of demography and liquid and illiquid financial assets relative to income.

The role of demography is potentially mixed. On the one hand, the proportion or changes in the proportion of households in the younger, first-time buyer age groups could be a factor influencing house prices, mainly derived from housing demand as a consumption good. However, the portfolio demand for housing among middle aged and pre-retirement households is likely to be high. This suggests that the proportion of households in this age group could also be a positive factor for house prices. In principle, demography and the income distribution should interact, as the purchasing power of the different demographic groups, as well as their size, should be relevant. In practice, lack of data typically makes this impossible to implement. The different components of portfolio wealth could also have dual roles: other things being equal, higher wealth would increase the consumer good demand for housing. However, higher financial wealth would tend to diminish demand for housing as a store of value.

[65] This formulation imposes the constraint that the income elasticity of demand for housing is 1.
[66] Note that house price expectations are already embodied in the user cost term.

South Africa is unusual in the liberal way in which mortgage market regulations permit pension wealth to be used as part collateral for house purchase. This implies that increases in pension wealth, for example, by extending pension coverage, or because of the appreciation of financial assets, could increase the demand for housing and hence of house prices.

The long-run relationship in equation (4) is embedded in an equilibrium correction form. Conventionally, this would imply that the dependent variable is the change in $\ln rhp_t$, with the lagged deviation between the LHS and the RHS of equation (4), as a key driver, together with changes in the other regressors and potentially changes in other variables such as the inflation rate, employment and the exchange rate. However, while the long-run relationship is formulated in real terms, implying no money illusion in the long run, 'nominal inertia' is often found in short-run dynamics, for example, because of lags in perceptions of the price level. Reformulating the dynamic relationship with the change in the log of the *nominal* house price index as the dependent variable would imply a coefficient of 1 on the current inflation rate on the RHS *if* market participants were fully aware of the current price level and were able to make decisions in 'real' terms. In practice, the hypothesis that the coefficient on the current inflation rate is 1 when the dependent variable is change in the log of the *nominal* house price index is strongly rejected: the empirical evidence is for a coefficient of zero. This implies that a more parsimonious form of the equilibrium correction equation is with the change in the log of the *nominal* house price index as the dependent variable.

3.4 Data

The data used in this paper are defined in Table 1, where summary statistics are also presented.

The dependent variable is the change in the log of the nominal house price index. In the long-run equilibrium relationship for the log of real house prices, the key drivers are the log user cost, the log income per house measure, a measure of the rate of property taxes, and two proxies for mortgage credit conditions (moving averages of the mortgage rate spread and the LTV, the latter from the FNB). User cost (see Figure 3) is defined as the prime rate of interest divided by 100, minus annual house price appreciation experienced over the previous 16 quarters, plus a constant proxying a risk premium and transactions costs. Experimentation with different periods suggested this 4-year measure of past appreciation gave the best fit and most stable results, see discussion below. The log of user cost enters directly and also in an interaction term with the moving average of the lagged LTV. The availability of higher leverage from a higher LTV is expected to increase the relevance of expected appreciation of house prices – as captured in the user cost - in the demand for housing. Income per house (see Figure 4) is measured as real household disposable income divided by the previous quarter's housing stock, both from the National Accounts. Recent changes in the housing stock are likely to be endogenous because residential investment rises with higher house prices. To address the problem of a downward endogeneity bias in the estimated housing supply effect, the log income per house is lagged by two quarters. The property tax rate (see Figure 5) is measured as the local government tax revenue from tax rates charged on housing, divided by housing wealth. As the tax revenue data are volatile, a four-quarter moving average is used. The short-run dynamics are measured by changes in five variables: the lagged log house price index, the mortgage rate, the rate of consumer price inflation measured by the four-quarter change in the log consumption deflator, the log real exchange rate and log employment.

3.5 Results

The starting point was the dynamic specification, with an equilibrium correction form and a rich lag structure, of equation (4). The dependent variable is the change in the log of the *nominal* house price index, as explained above.

It soon became evident that no significant role could be found for the balance sheet variables. Reducing the general formulation to a parsimonious form resulted in the estimates shown in Table 2. Column 1 of Table 2 provides estimates for the period 2000:Q4 to 2020:Q1, ending just before the pandemic struck. The long-run solution comes through strongly with a quarterly speed of adjustment of 0.171. This is a little higher than estimates of the well-specified equations for the other countries reviewed in Section 3.2. There is thus very strong evidence of co-integration. The log income per house term is strongly significant; with a coefficient of 0.268 relative to a speed of adjustment of 0.171, this implies a long-run income elasticity for house prices of around

1.6. As explained in Section 3.1, the inverse of minus 1.6, namely -0.625, therefore gives the average price elasticity of the demand for housing in South Africa. By contrast, Chauvin and Muellbauer (2018) estimate a price elasticity of around -0.5 for France, and Geiger *et al.* (2016) of around -0.8 for Germany. The less elastic finding for France than South Africa may arise from the highly Paris-centric nature of the French economy, suggesting greater spatial substitution possibilities in South Africa. However, economic activity in South Africa is arguably more concentrated than in Germany, consistent with fewer spatial substitution possibilities than in Germany.

The log user cost term, entering both singly and in interaction with a moving average of the LTV, is also strongly significant. The four-year memory of house price appreciation in the construction of the term is similar to that found for the US in Duca *et al.* (2011, 2016). This implies strong persistence of past appreciation, and hence, the considerable risk of house prices overshooting beyond fundamentals. The interaction effect is strongly significant suggesting that higher leverage raises the salience of the house price expectations embodied in user cost. As explained in Section 3.1, the property tax rate is not incorporated in the construction of user cost, as would be implied by simple theory, see equation (3), but is highly significant. In South Africa, property valuations, on which local rates are based, tend to lag considerably behind current house prices, violating the theory assumption in equation (3). The coefficient of the tax rate is therefore lower than that implied by a simple tax-adjusted user cost, in which taxes are a given fraction of current values. Neither a nominal mortgage rate nor a real mortgage rate was significant in the long-run solution, see equation (4). The real interest concept represented by user costs seems to capture appropriately the long-run effect of interest rates. However, both nominal and real rates matter in the dynamics, see below. The absence of a long-run nominal interest rate effect in South Africa, as found for France, suggests the possibility that lenders pay less attention to debt-service ratios than in France, where there are mostly fixed rate mortgages. With interest rates so unpredictable in South Africa, and largely floating rate mortgages, lenders are more likely to pay attention to the debt-to-income ratio, credit scores and the loan-to-value ratio. In Germany, France and the US, the real interest rate is fully captured by the user cost and also does not appear in the long-run solution directly.

The final term in the long-run solution is the interest rate spread, entering as a lagged three-quarter moving average. This proves highly significant, confirming that credit conditions play an important part in driving house prices in South Africa. This effect of credit conditions supplements that from the interaction effect of the LTV with user cost, discussed above.

In the short-term dynamics, lagged house price appreciation in the previous quarter captures a momentum effect additional to that embodied in the user cost term. The negative short-term effects of interest rates on house prices is captured by four-quarter changes in the nominal prime rate of interest and in the real prime rate, though the nominal effect is far greater. Changes in nominal rates have important cash-flow implications in a floating rate environment; changes in real rates are more forward-looking and reflect affordability, taking inflation into account. The lagged annual rate of inflation has a negative coefficient, perhaps because it signals a temporary loss of real income if nominal wages are sticky or expectations of tighter monetary policy. While income growth proved insignificant, shocks to employment, measured by the quarterly rate of acceleration of employment in the previous two quarters appear to have significant positive effects on house prices.

The effect of changes in the real exchange rate can be summarised in a positive effect from the eight-quarter change in the log real exchange rate. This implies that periods of exchange rate appreciation tend to be followed by higher house prices, probably because of stronger capital inflows associated with such appreciation. Some of this foreign capital presumably found its way into the property market[67]. However, the estimated coefficient clearly declines towards the end of the period, and by the end of 2016, this exchange rate effect has vanished completely. To explore this further, we modelled this transition by interacting the exchange rate term with a smooth transition dummy that is one before 2015 and smoothly falls to zero at the end of 2016, and for the rest of the period to 2020. This can be interpreted as a loss of interest of foreign buyers[68] in South African property perhaps associated with rising problems of power

[67] This has interesting implications for how to interpret exchange rate effects on foreign buyers of housing and hence on aggregate house prices in South Africa. On the face of it, rather than increasing purchases after a period of strong currency depreciation when housing looks cheaper in foreign currency terms, foreign buyers seem to have been more influenced by the momentum effect of currency appreciation and the associated economic optimism. The lack of a connection between a weak exchange rate and higher demand by foreign buyers in Cape Town found by Georg and Davids (2019) and presented at the SARB Biennial conference is consistent with this interpretation of the macro-evidence.

[68] The absence of a similar effect in the equation for *domestic* mortgage debt is consistent with the interpretation

shortages, increases in violent crime, and the perception that the state capture scandal and rising government debt to GDP have undermined the future stability of the country. When using the interaction term, the real exchange rate coefficient becomes stable, contrasting with column 2 where the interaction term is missing, and the coefficient is substantially lower.

The diagnostic statistics for the estimates are all satisfactory, including the Chow test for parameter stability and tests for autocorrelation and heteroscedasticity of the residuals. Columns 3 and 4 of Table 2 provide robustness tests. Column 3 gives estimates for 2000:Q4 to 2014:Q4, which are very similar to those in Column 1. This suggests that the transitional dummy from 2015 captures the shift in the exchange rate effect quite accurately. Column 4 examines the consequence of omitting *the interaction term* of the log user cost with the moving average of LTV. The fit is worse, and the estimated speed of adjustment is lower, but the general story is robust. With this specification, the level of the LTV is insignificant but the lagged change in the LTV makes a small contribution. Finally, replacing the log user cost by the level of user cost gives very similar results, with only a slight deterioration in the equation standard error from 0.00458 to 0.00469. The log linear functional form is marginally superior. These are remarkably low standard errors, of the order of one half of one percent of house prices. Apparently, the banks which provide house price indices engage in smoothing of the data, which would contribute to the low equation standard errors. This smoothing probably also helps explain the high coefficient, of close to 0.6, on the previous quarter's change in the log of house prices.

Summary

The evidence from the data is thus for a plausible and well-fitting house price equation for South Africa. It yields important insights into the powerful transmission from interest rates and credit conditions in the mortgage market to house prices. Both mortgage spreads and loan to value ratios appear to be relevant as features of credit conditions. Moreover, there is evidence of a memory of up to four years regarding house price appreciation by many housing market participants in forming expectations of future appreciation. This implies that a series of positive shocks to housing demand can feed back positively on housing demand and on house prices, so extending boom

---

of a shift in *foreign* demand associated with the momentum effect of recent currency appreciation.

conditions, and potentially causing house prices to overshoot relatively to their fundamentals. This has clear implications for risks to financial stability and stabilisation policy.

4 A Mortgage Borrowing Model[69]

4.1 Theoretical background

There is not a single, simple theoretical model that underlies the demand for housing. Clearly, the demand for mortgages is strongly linked to the demand for housing, which implies that there is also no single, simple theoretical model behind this demand. However, while some home buyers are cash buyers or buyers with so much wealth that the mortgage represents only a small part of the purchase price, the demand for mortgages tends to be dominated by the subset of potential buyers with less wealth. Younger first-time buyers are likely to be prominent, suggesting that the proportion of the population in this age group is likely to be a factor. Moreover, to model the mortgage stock, or the flow of new mortgage lending, the credit supply side is crucial. All lenders use screening rules, such as limits on leverage as represented by loan-to-value ratios, and affordability criteria as represented by debt-service or debt-income ratios, as well as checks on the credit worthiness of individual households, to allocate credit. This implies that credit conditions, a proxy for shifts in credit availability other than that represented by the standard mortgage interest rate, need to be a key feature of a model of the mortgage stock.

Given the link to demand for housing, a key issue for modelling the demand for mortgages is the average price of housing. For those committed to a home purchase, higher house prices suggest the need to borrow more, though some buyers might be forced into lower quality housing. This would imply a positive effect from house prices onto the mortgage stock. A second reason to expect such an effect is that existing home-buyers, considering trading up in the market, will have more equity in the market and so be able to achieve a cheaper loan at a lower low-to-value ratio or, if previously at an LTV constraint imposed by a lender, be able to buy a more expensive home. However, there is a potential argument pointing in the opposite direction, which comes

[69] This section draws on Chauvin and Muellbauer (2018).

from a shift in the 'extensive margin', i.e., by reducing the pool of potential first-time buyers able to enter the market at all, when lenders demand substantial down-payments to obtain a mortgage. As a result of a rise in average house prices relative to the incomes of potential first-time buyers, fewer of such buyers will have saved enough to offer the (substantial) minimum down-payment necessary and will therefore remain renters in the interim.

Turning to the role of mortgage interest rates, the above discussion of the demand for housing emphasised that the nominal mortgage rate was likely to be important in countries where lenders focus on the debt-service ratio as a lending criterion, as well as the real rate represented by the user cost of housing. The nominal mortgage interest rate should be even more relevant for mortgages than for house prices as affordability in terms of short-term cash-flows is not only a concern for mortgage lenders but also one for borrowers: defaulting on a mortgage and losing one's home is damaging both for lenders and borrowers. If the mortgage stock model is partly driven by the level of house prices, and that, in turn, is strongly influenced by the user cost of housing, it is quite possible that there is no *direct* effect from user cost on the demand for mortgages but only the indirect effect via house prices. However, the real mortgage interest rate based on expectations of consumer price inflation may well be relevant for mortgage demand as a measure of the longer-term servicing cost of debt.

4.2 Empirical Evidence from Other Countries

Desirable properties for a mortgage stock equation are a well-determined long-run solution for the mortgage stock and a moderate speed of adjustment (given the long-run nature of mortgage debt). From Chauvin and Muellbauer (2018), the mortgage stock equation for France has a speed of adjustment a little under 0.08 (t=16). The mortgage credit conditions index, MCCI, enters both directly (with a t-ratio of 12) and in interaction with the log house price to income ratio (with a t-ratio of 6). Given log house prices to income and the other independent variables, the nominal mortgage rate is highly significant, as in the French house price equation. There are no significant direct effects from user cost or from a real interest rate. However, by conditioning on the log house price to income ratio, there is in effect an *indirect* user cost effect, as well as the *indirect* effect of nominal interest rates that operates via house prices. Demography has a similar role to that in the house price equation. The hypothesis can be accepted that the income elasticity of the mortgage stock is 1.

Few central bank policy models, except for the Dutch model, have an equation for the mortgage stock. For example, neither the FRB-US nor the ECB-BASE model has an equation for mortgage debt - or indeed for household debt. Because they rely on net worth to drive consumption, these models depend on an equation which updates net worth every quarter by net disposable income minus consumption and minus residential investment, and a revaluation adjustment. This does not permit an explicit role for credit conditions. In the French model of Lemoine *et al.* (2019), there is no role for household wealth or debt, and therefore no model for these, and hence no role for credit conditions. The Australian model MARTIN includes an equilibrium correction equation for household debt extended by banks. In the long-run, household debt is proportional to the value of the housing stock, and also depends on the real mortgage interest rate.

4.3 Econometric Specification for Mortgage Debt

Next, we turn to a specification for modelling mortgage debt. As explained above, the demand for mortgage debt is driven by the demand for housing. Higher house prices should increase the demand for mortgages for the reasons explained in Section 4.1, though with the proviso that some potential first-time buyers might be priced out of the market. A very general formulation of the long-run solution that corresponds to the economic arguments above is as follows:

$$\begin{aligned} \ln\ rmdebt_t = {} & m_0 + m_1 \ln y_t + m_2 MCCI_t + m_{3t} \ln user_t + m_{4t}\, rprime_t + m_{5t} \ln prime_t \\ & + m_{6t} E_t \ln (y_t^p / y_t) + m_{7t} \ln (hp_t / y_t) + m_8 demog_t + m_{9t} ln(LA_t/y_t) \\ & + m_{10t} \ln (nmdebt_t / y_t) + m_{11} \ln (IFA_t / y_t) \end{aligned} \tag{5}$$

Here, $rmdebt$ is per capita mortgage debt in real terms, i.e., nominal debt divided by the consumer expenditure deflator, and y is per capita real household disposable income. If the income elasticity of mortgage debt, m_1, is one, the dependent variable can be reformulated as the log of the mortgage debt to income ratio. $MCCI$ is an indicator of credit conditions in the mortgage market; $user$ measures user costs as previously explained; $rprime_t$ is the real prime rate of interest; $prime$ is the nominal prime rate of interest; y_t^p/y_t is the ratio of permanent to current per capita real household disposable income; hp/y is the ratio of the real house price index to per capita real household disposable income; $demog$ is a demographic indicator; LA/y is the ratio of liquid assets to income, and $nmdebt/y$ and IFA/y, the corresponding ratios for non-mortgage debt and illiquid financial assets.

Credit market liberalisation could impact in several ways on these long-run relationships as indicated by time subscripts on several parameters. In principle, the strength of the effects of user cost and real interest rates rmr_t is likely to increase with credit liberalisation, making m_{3t} and m_{4t} more negative for example, while nominal interest rates may have less impact, making m_{5t} less negative[70]. The impact of income expectations could also vary with shifts in credit liberalisation, for example causing an upward shift in m_{6t} with increasing *MCCI.* Higher house prices relative to income should increase demand for mortgages but this could increase further if liberalisation relaxed the down-payment constraint, hence shifting up m_{7t}. Demography and asset to income ratios are represented in the next four terms in equation (5). Generally, a higher ratio of liquid assets may indicate greater availability of liquidity to fund mortgage deposits, but with easier credit access, that could become less relevant. A higher level of non-mortgage debt relative to income reduces the ability of households to take on mortgage debt and may also make lenders more cautious about mortgage lending. It is possible that, when mortgage credit conditions are more relaxed, this negative effect becomes somewhat less pronounced. In practice, in short samples, empirically identifying such interaction effects can be very demanding. Nevertheless, testing for such possibilities is advisable.

Embedding the long-run relationship in equation (5) in an equilibrium correction form suggests that the dependent variable is the change in $\ln\, rmdebt_t$. Reformulating the dependent variable as the change in the log of per capita mortgage debt in *current prices* would imply that on the RHS of the dynamic relationship would appear the current inflation rate with a coefficient of 1. In practice, the data strongly reject this implication, instead finding a zero effect from the current inflation rate. The implication

---

[70] This would be the case if mortgage market liberalisation was mainly about easing loan-to-value constraints. However, if it more concerned relaxing debt-to-income or debt service ratio constraints, m_{5t} might become more negative.

is that, while there is no money illusion in the long run, so that only real variables or ratios of nominal variables, matter, this is not so in the short run, as we found for the dynamics of the house price equation. This implies formulating the dependent variable as the change in the log of per capita mortgage debt in current prices for a more parsimonious form of the dynamic relationship.

4.4 Data

The data used in this paper are defined in Table 1, where summary statistics are also presented. The dependent variable is the change in the log of nominal mortgage debt per capita. In the long-run equilibrium relationship for the log of nominal mortgage debt to income, the key drivers are, in principle, the LTV, the log of the nominal prime rate, the log of the user cost, the real prime rate (Figure 6), the ratio of permanent to current income, the log ratio of house prices to income (Figure 7), population growth, and the logs of liquid assets to income, of illiquid assets to income, and of non-mortgage debt to income (Figure 8). LTV (Figure 9) and user cost were defined in Section 3.4. In the version estimated below, we do not include permanent income because this would require a separate equation – see discussion in the concluding section, Section 6. Demography is captured by the annual population growth rate. The three balance sheet variables are from the SARB's household balance sheet, which originated from the methods of Aron *et al.* (2006), Aron and Muellbauer (2006) and Aron *et al.* (2008).

The short-run dynamics are measured by changes in five variables: the lagged log house price index, the mortgage rate, the rate of consumer price inflation measured by the four-quarter change in the log consumption deflator, the log of mortgage debt and log employment.

4.5 Results

For two reasons, it is even more difficult to model mortgage debt in South Africa over the relatively short period from 2001 to 2020, than to model house prices. First, the theory background is even more eclectic than for house prices, where the inverse demand approach gives clear guidelines. Secondly, the mortgage debt-to-income ratio has only one turning point over this period, whereas the real house price index has two, which is more informative for empirical work. It is important, therefore, to have priors in mind to avoid selecting spurious, but well-fitting, representations of the data. Higher house prices relative to income, should, with access to credit, generate a higher demand for mortgages[71]. Therefore, as in Chauvin and Muellbauer (2018), the house price to income ratio should be one of the explanatory variables in the mortgage equation. Another point is that the mortgage stock evolves relatively slowly, although new advances are likely to be very sensitive to economic conditions. Research on France and the UK could suggest orders of magnitude for the speed of adjustment for plausible specifications of an equilibrium correction model for the South African mortgage stock. Chauvin and Muellbauer (2018) find a speed of adjustment of 0.077 in a quarterly equilibrium correction model for the mortgage stock in France, while Fernandez-Corugedo and Muellbauer (2006) find a speed of adjustment of 0.061 in the UK.

The dependent variable in the dynamic formulation is the change in the log of the nominal per capita mortgage stock, as explained earlier. In principle, the long-run solution for the log mortgage debt-to-income ratio could contain both log user cost and its interaction with LTV, levels of real and nominal interest rates, the level of LTV, the property tax rate, the mortgage rate spread, the log house price to income ratio and its interaction with credit conditions as represented by the level of LTV, and several balance sheet to income ratios and related possible interactions. However, by conditioning the long-run solution for the log mortgage debt-to-income ratio on the log house price-to-income ratio, the main long-run drivers for the latter should plausibly become irrelevant. On a sample of under 20 years, identifying all these effects empirically is too ambitious, even with priors on the expected signs. From a range of specifications, it was soon apparent that, indeed, three of the long-run drivers of house prices were not significant in the mortgage stock equation. These are the log of user cost and its interaction with LTV, the property tax rate and the mortgage rate spread (albeit there is a trace of a positive effect in the last). However, the lagged level of the LTV was also significant in all specifications. The real interest rate is highly significant as a measure of the cost of mortgage debt service, entering most parsimoniously as a two-quarter moving average. Among the balance sheet to income ratios, the level of the log non-mortgage debt-to-income ratio stood out in a variety of different dynamic

[71] With tight lending conditions, a rise in house prices relative to income could exclude sections of the population from obtaining a mortgage, and so have a negative effect on aggregate mortgage lending. Chauvin and Muellbauer (2018) find evidence for such a pattern in French data.

specifications. This is unsurprising for, in South Africa, non-mortgage debt has at times accounted for almost half of total household debt. In principle, households already servicing non-mortgage debt should find it harder to incur the additional debt burden of a mortgage. This would suggest a negative balance sheet effect on the demand for mortgages, as is confirmed by the data.

The results of the reduction of the dynamic equation to a parsimonious form, are shown in the first column of Table 3. The speed of adjustment is a plausible 0.055, not far from that found in the UK. The log of the house price to income ratio is highly significant. However, unlike in the house price equation where it enters only through an interaction term, the lagged level of LTV, expressed as a moving average, is very significant in the mortgage stock equation. The mortgage spread, in a lagged moving average form, is not quite significant, and is omitted in the results reported in the second and third columns. Among the balance sheet to income ratios, only the log non-mortgage debt-to-income ratio is significant.

Turning to the short-term dynamics of the estimated equation, the growth of mortgage debt in the previous three quarters picks up slight persistence in the dynamics. The current change in the mortgage spread – a proxy for easier credit conditions – and growth in real per capita income have positive effects on mortgage debt growth, as do recent positive shocks in employment measured by the rate of acceleration. Population growth also enters as an acceleration effect in log population, implied by testing down from a more general form. As annual population growth is itself a strongly persistent, i.e. an I(1) variable, the change in the growth rate makes the measure I(0)[72]. An impulse dummy adjusts for an outlier in 2002:Q3.

The diagnostics are generally satisfactory. The Chow test for parameter stability is very satisfactory and Column 3, shows that estimates to the end of 2014 are very close to the full sample estimates in Column 1.

[72] It is possible that, since the fall in population growth from the late 1990s is linked to the AIDS epidemic, the variable is picking up the negative effect of the epidemic on mortgage demand.

Summary

As in the house price equation, the evidence is that interest rates and credit conditions have powerful direct effects on the mortgage stock as well as indirect effects through the level of lagged house prices. The relative roles of the mortgage spreads and LTVs is somewhat different from the house price equation. For mortgages, the direct effect of the level of LTVs is greater than for house prices, while spreads have only temporary effects. The extrapolative element of expectations of house price appreciation embedded in the lagged house price to income ratio implies that mortgage debt, like house prices, can overshoot fundamentals. High levels of mortgage debt relative to income can pose risks for financial stability, and for risks of sharp downturns in consumer spending if interest rates were to rise.

5 A Residential Investment Model

5.1 Theoretical background

Residential investment, comprising a significant and volatile part of GDP, is an important channel for monetary policy transmission. Further to this, an equation for residential investment potentially serves two additional functions in an econometric policy model. First, if housing wealth is one of the drivers of the consumption function, as is the case in the SARB's Core model, an equation is required for the acquisition of housing assets by households. This acquisition would be captured largely by residential investment since most of such investment is in the form of home improvements or home purchases by households. Second, a residential investment equation is needed to endogenise the housing stock, which is an important driver of house prices in the house price equation, see Section 2. This is because the housing stock itself arises from the accumulation of past residential investment.

The simple theory of a profit-maximising firm in a competitive market suggests that profits of a house builder depend on the sales prices of houses built relative to the costs of construction. Given lags in construction, sales occur several quarters after construction begins and this should affect the timing of observations of prices and costs. House builders need capital to build, which suggests a role for interest rates as a measure of financing costs. While house prices are driven by demand, in the short run, house prices tend to adjust to demand with a lag, as we saw above. This suggests that short-term demand shocks should affect construction volumes.

5.2 Empirical Evidence from Other countries

Research on residential investment has been reviewed by Duca *et al.* (2021). An important recent study for the OECD, by Cavalleri *et al.* (2019), covers 25 countries – including South Africa[73], and updates an earlier study by Caldera and Johansson (2013). The key driver in this research is the ratio of house prices to an index of building costs[74], which for many countries is well proxied by the price deflator for residential investment. Countries vary a great deal in the supply elasticity of residential investment. For example, their estimate for the U.S. is that a 1% increase in real house prices leads eventually to a 2.8% increase in the volume of residential investment. The figure is under 1% in Belgium, France, Germany, Italy and the UK, while the estimate for South Africa is 1.09% (their Table B6). Their evidence thus suggests that the house price link to residential investment is stronger in South Africa than it is in Europe, but weaker than it is in the US. The study of Cavalleri *et al.* (2019), who find that more habitable land per head, greater ease of construction (proxied by the past expansion of built-up area) and less land-use restrictiveness all boost the price elasticity of housing supply, is suggestive in explaining cross-country differences. Hence there may be important structural and procedural/planning differences between countries affecting monetary transmission, realised via housing markets.

The OECD study is an important contribution, but its limited short-term dynamics probably do not fully capture timing differences between the effects of house prices and construction costs. The study also omits interest rate effects, which could bias estimates. Since house prices are sticky, short-term demand shocks influence residential investment directly without being mediated through prices, as noted above. Proxies for such demand shocks could help capture the short-term dynamics in residential investment. These proxies need to be based on the *changes* in demand drivers, such as income, interest rates and employment. Note that the long-term

---

[73] This is a single equation where the period covered is not clear but ends in 2017.

[74] To be precise, the model is formulated in terms of the log real house price index and the log real construction deflator. For several countries, the coefficients on the two are approximately equal and of opposite sign, so that the two terms can be combined into a single log price ratio of house prices to the construction cost deflator.

demand drivers are already captured by the level of house prices, which enter the residential investment expressed as a ratio to construction costs.

Duca *et al.* (2021) argue that future research in this area needs to take account of the major structural break caused in countries such as the U.S. and Ireland by the GFC. Much productive capacity, all the way down the supply chain, was lost in these countries. The construction industry became more concentrated as many smaller building firms went bankrupt when cash flows and the value of their land banks collapsed. This suggests that post-crisis, monetary transmission via the housing market will have altered. However, as the building industry in South Africa proved more resilient to the GFC than in the countries above, a structural break in the residential investment equation is less likely to be relevant. Situating South Africa in the international spectrum of the mechanism connecting residential investment with house prices and possibly other drivers would be informative. It would establish the magnitude of the transmission channel between monetary policy, house prices and residential investment.

5.3. Econometric Specification for Residential Investment

As explained in Section 5.1, the demand for residential investment is driven by the ratio of selling prices to costs of construction and by financing costs, and by demand shocks. The formulation of the long-run form of residential investment equation is as follows:

$$ln\ inv_t = h_0 + h_1 \ln gdp_t + h_2 \ln (hp_t / hc_t) + h_3 rprime_t + h_4\, demog_t \qquad (6)$$

Here $hinv$ is per capita residential investment in constant prices, gdp is per capita real GDP, hp is the house price index, hc is the deflator for residential construction, $rprime$ is the real prime rate of interest, and $demog$ is a demographic indicator. If the parameter h_1 equals one, the dependent variable can be expressed as the log of the investment to GDP ratio.

The dynamic form of the equation is of the equilibrium correction type, with the dependent variable the change in the log of per capita residential investment in constant prices. Apart from changes in the elements of the long-run solution, other variables in change form are the log of real per capita household disposable income, in nominal and real prime rates of interest, the inflation rate, and log population. As the data on residential investment are not seasonally adjusted, seasonal dummies were included.

5.4. Data

The data used in this paper are defined in Table 1, where summary statistics are also presented. The dependent variable is the change in the log of the residential investment in constant price terms. In the long-run equilibrium relationship for the log of the residential investment, the key drivers are the log of real GDP, the log of the ratio of house prices to construction costs (see Figure 10) and the real prime rate. The short-run dynamics are measured by annual changes in five variables: the lagged log house price index, the log of construction costs, the log of population (see Figure 11), the log of real personal disposable income per capita, and the surprise terms in the prime rate and inflation.

5.5. Results

After testing for more general lag structures in the dynamic specification, it became apparent that many could be summarised by four-quarter changes. The motivation for using four-quarter changes in the formulation of the dynamics of the equation arises from lags in the process of house building. The U.S. Census Bureau publishes data for the average lag between the granting of building permits and building starts, which is typically of the order of two months; between starts and completions, the average lag ranges from seven to 14 months, depending on the type of housing[75]. It seems likely that a lag of around four quarters would represent the typical experience in South Africa, and this four-quarter simplification of the lag structure was confirmed by testing. Furthermore, residential investment appeared to have moderated from about 2017 by more than predicted by the other explanatory variables, probably connected with a worsening economic and political situation. A smooth transition dummy with a negative coefficient reflects this phenomenon, moving from zero at the end of 2016 to 1 at the end of 2018.

[75] See the website: https://www.census.gov/construction/nrc/lengthoftime.html

The results for the residential investment equation over three different sample periods are shown in Table 4. The estimated equation easily passes Chow tests for parameter stability for all three samples. The long-run solution is well-specified and with a high speed of adjustment for all three samples, lying in a narrow range of around 0.3, with a t-ratio around 9 for the longest sample. This is evidence of strong co-integration amongst the long-term variables. The key supply elasticity, given by the ratio of the coefficient on the log house price to building cost ratio relative to the coefficient on log residential investment to GDP ratio, lies in a narrow range around 0.95 for all three of the sample periods shown, with a high individual t-ratio. The high precision of this estimate gives confidence in the strength of the transmission channel, for example, from interest rates and credit conditions via house prices to residential investment. These findings are not far from those of the OECD study discussed above, which finds a speed of adjustment of 0.35 for South Africa and a supply elasticity of 1.09. This places the supply elasticity for South Africa above that of most European countries but below that of the US. Our slightly lower estimate of the supply elasticity is probably the result of including more controls. Excluding such controls, likely to be correlated with house prices, is likely to bias up the effect attributed to house prices.

It turns out that the population growth rate is an I(1) variable – that is, a trending variable, see Figure 2 - and it proves important for the long-run solution. Figure 11 shows the annual growth rate of the population, which fell notably after the end of the 1990s, probably because of the AIDS epidemic and perhaps emigration. The numerically large coefficient for this variable in the estimated equation is surprising if it were to be attributed to a pure population growth effect. But population growth changes also reflect shifts in the structure of the population. Small changes in population growth can therefore be linked to economically significant changes, for example in the composition of the population by age or income (on which we lack data) or shifts in the economic environment such as productivity growth. As Figure 11 illustrates, population growth is trending with only one substantial turning point in the relevant period. This could lead to a spurious regression problem. However, the stability of the estimated coefficient over different samples, even those ending in the early 2000s, is reassuring on this point.

Turning to the dynamics, lags between the decision to start building and completion suggests that the relevant price ratio is formulated as the price for which a dwelling can be sold relative to the cost of construction four quarters earlier. With this formulation, growth rates of house prices and of the construction deflator become insignificant. The close cyclical correspondence between the log ratio of residential investment to GDP and this formulation of log relative prices can be seen in Figure 10. Gaps between the two are accounted for by lags in adjustment, and by the other drivers included in the equation: population growth, income growth, changes in interest rates and the dummy for the post-2017 decline in the economic environment. The ‘surprise’ terms in interest rates can be parsimoniously represented by the four-quarter change in the nominal prime rate of interest, see Figure 12. There is a much smaller effect from the four-quarter change in the real prime rate of interest, typically with a t-ratio below 2 and hence not shown in Table 4. The strong and very significant coefficient on the change in interest rates is further evidence of the direct monetary policy channel of transmission to residential investment (in addition to the indirect effects via house prices discussed above).

Figure 13 illustrates annual growth in per capita real disposable income, showing the greater stability of income after inflation targeting was adopted in 2001, reflecting a decline in inflation volatility. The exception, is of course, the pandemic in 2020 with record falls in income growth.

The specification includes impulse dummies for 1979:Q2, 1988:Q3, 1988:Q4, and 1997:Q4. As noted above, a dummy is included for the deteriorating economic environment from 2017 onwards[76]. As the data are not seasonally-adjusted, seasonal dummies are included for the first quarter and the second quarter from 2007 onwards, when the seasonal pattern in the data shifted. The four-quarter lag in the rate of change of residential investment may also be picking up an evolving change in seasonality.

[76] This dummy has a smooth transition form, set at 0 before 2017 and transitioning to 1 from the beginning of 2017 to the end of 2018. The dummy represents factors such as the adoption of the land expropriation policy by the ANC in December 2017 (widely-discussed beforehand), growing emigration, the growing crisis in electricity generation by Eskom, the growth in government debt to GDP and the associated risks to South Africa’s international credit ratings, the ‘state capture’ scandal of the Zuma government, and widespread problems of delivery of public services at lower levels of government. A discussion of some of these factors is found in Aron and Muellbauer (2020).

Summary

To summarise, we have evidence of a model of residential investment with a surprisingly stable structure, fitting data back to 1978. The key driver of residential investment relative to GDP is the relative price of houses to construction costs. This implies that the powerful effect of interest rate and credit conditions on house prices transmits to this volatile component of aggregate demand. While there is no evidence of direct interest rates effects in the long-run solution, there are powerful short-term effects of changes in prime rates of interest on residential investment. There is evidence of a moderation in residential investment as population growth fell with the AIDS epidemic and from 2017, probably associated with worsening economic and political prospects.

6. Conclusion and further Development

This paper has developed econometric models for house prices, the mortgage stock and residential investment in South Africa, complementing our previous research on consumption in South Africa in 2013. These models have important implications for the understanding of monetary transmission via credit and housing markets and for assessing potential risks to financial stability. The starting point is the house price equation, as house prices are a major driver both of mortgages and of residential investment, as well as having a significant influence on consumption.

First, the evidence from the data is for a plausible and well-fitting house price equation for South Africa. It is based on the 'inverted demand principle', where the price is determined by demand that varies relative to the existing housing stock. The estimated equation yields important insights into monetary transmission in South Africa. There is a powerful transmission from interest rates and credit conditions in the mortgage market to house prices. Both mortgage spreads and loan-to-value ratios (LTVs) appear to be relevant proxies for credit conditions. House price expectations relative to mortgage rates determine 'user cost', which is a key driver of housing demand. There is evidence of a memory of up to four years regarding the expectations of house price appreciation by housing market participants. This implies that a series of positive shocks to housing demand can feed back positively onto housing demand and onto house prices, so extending boom conditions. This can potentially cause house prices to overshoot relative to their fundamentals, as seems to have occurred in 2007-8. Such overshooting has clear implications for risks to financial stability and is relevant when designing stabilisation policy. From our evidence of a shift in the effect of changes in the exchange rate, there was an apparent waning influence of foreign home buyers on house prices in South Africa from about 2015. Given the linkages from interest rates to the exchange rate, this suggests there may have been a shift in the effect of monetary policy on house prices after 2015 in South Africa.

Second, for the mortgage stock, as for the house price equation, the evidence is that interest rates and credit conditions have powerful effects. The relative direct effects of mortgage spreads and LTVs is somewhat different on the mortgage stock than in the house price equation. For mortgages, the direct effect of the level of LTVs is greater than for house prices, while spreads have only temporary effects. However, since a key driver of the mortgage debt to income ratio is the level of house prices to income, there are large *indirect* effects of interest rates and credit conditions on the mortgage stock via the house prices to income ratio. Also, the extrapolative element of expectations of house price appreciation, embedded in the house price ratio, has an *indirect* effect. This implies that mortgage debt, like house prices, can overshoot fundamentals. High levels of mortgage debt relative to income can thus pose risks for financial stability. There may also be risks of sharp downturns in consumer spending if interest rates were to rise.

Estimates for both the house price and the mortgage stock equations are limited by the historical span of data on LTVs and on mortgage spreads, which begin around 2000. In particular, there is only one turning point in the series for mortgage debt to income, from 2001 to just before the arrival of the pandemic in 2020, making robust identification of parameter estimates difficult. Hence, the model for mortgage debt is necessarily provisional. Nevertheless, the model is very consistent with evidence from other countries.

Third, for residential investment, in contrast, house prices capture almost all the relevant long-run information on demand. We were therefore able to find a stable relationship for data back to 1978, despite the many structural changes and shocks experienced by the South African economy. The key driver of residential investment relative to GDP is the relative price of houses to construction costs. This finding is consistent with international evidence from an important OECD study, Cavalleri *et al.* (2019). The implication for monetary transmission is that the powerful effect of interest rate and credit conditions on house prices in South Africa also transmits to this rather volatile component of aggregate demand, residential investment. While there is no evidence of direct interest rates effects in the long-run solution, there are powerful short-term effects of changes in the prime rate of interest on residential investment. There was an apparent moderation in residential investment as population growth fell with the AIDS epidemic and from 2017, probably associated with worsening economic and political prospects.

An important area for further development of these models is a more explicit treatment of expectations, particularly of interest rates and of income. In all three equations, the prime rate of interest is a key variable. In South Africa, the prime rate moves one-for-one with the policy rate, the repo rate, producing a clear link between policy decisions and macroeconomic consequences. Since our evidence shows that house prices, mortgages and residential investment are all highly sensitive to interest rates, it seems likely that interest rate expectations would affect private sector behaviour. In the current version of the equations, expectations are implicit in the lag structure, for example lags in interest rates and inflation.

A useful development would be to make explicit the private sector's perception of the SARB's policy rule, as in Aron and Muellbauer (2002), but formulated with a one-year-ahead forecasting equation for the prime rate. While interest rate expectations over a longer horizon are probably relevant, the one-year-ahead outlook is a useful starting point. The fitted value of the expected change in the prime rate could then be incorporated in the above behavioural equations for house prices, mortgage stock and residential investment (and potentially consumption). The advantage of an explicit model is better to understand monetary policy transmission by separating out the expectations channel from other effects. This would also allow incorporation of shifts in the policy rule, meeting the Lucas Critique, i.e., to take account of the consequent shift in parameters. An example is in analysing data before and after the advent of inflation targeting in 2001. To the extent that forward guidance could play a role in the SARB's monetary policy, an explicit expectations channel should also be helpful in designing such policies.

Another expectations variable is that for future growth in household disposable income. We have argued above, that for house prices and mortgages, such income expectations are probably less relevant than they were shown to be for consumption (Aron and Muellbauer, 2013). The reason is that housing is not only a consumption good, but it is also an asset, and mortgages provide the mechanism for acquiring that asset. When people become more pessimistic about future income, they tend to save more if possible. They may also increase liquidity, and thereby raise mortgage demand, through withdrawing equity when lenders make it easy to do so. These tendencies may offset, at least in part, the concerns people may have about servicing their debts if income growth diminishes. Practically, it is an empirical question as to the net outcome of these conflicting effects of permanent income relative to current income - which is the most relevant measure of income growth expectations. In Aron and Muellbauer (2013), income growth expectations were not significant in an equation for *total* household debt. However, as mortgage debt is far more long-term than is non-mortgage debt, this finding may not hold for mortgage debt.

Finally, regarding credit conditions, Chauvin and Muellbauer (2018) and Duca and Muellbauer (2013) argue for a 'latent variable approach', the 'Latent Interactive Variable Equation System (LIVES), to model credit conditions, both for mortgage and non-mortgage markets, and the consumption function. The evidence in this paper suggests that data available since 2000 on loan-to-value ratios and mortgage spreads in South Africa would be excellent candidates to drive an indicator of credit conditions in the mortgage market. To extend the estimation period backward, the credit conditions indicator before 2000 could be defined using dummy variables. In principle, Deeds Office data on LTVs could be extended backwards before 2000. As suggested in Aron et al. (2020), it may also be possible, using the Deeds Office data on mortgage transactions, to approximate first-time buyer LTVs. These are likely to be a more accurate measure of credit conditions since first-time buyers are more likely to be credit constrained. Credit conditions indicators, including the mortgage spread and LTVs, provide the key channel by which shifts in macro-prudential policy settings and data on the asset position of banks, including impaired assets, can feed into an econometric policy model. The models presented in this paper should therefore make a useful contribution to future development of the SARB's Core model and especially the version extended with links between banking and the real economy (see our assessment in Aron and Muellbauer (2022a)).

References

Anundsen, Andre K., and Christian Heebøll. 2016. "Supply Restrictions, Subprime Lending and Regional US House Prices." *Journal of Housing Economics* 31 (1): 54-72.

Aron, J. and J. Muellbauer. 2002. "Estimating Monetary Policy Rules for South Africa", in Norman Loayza and Klaus Schmidt-Hebbel (eds) "Monetary Policy: Rules and Transmission Mechanisms", Series on Central Banking, Analysis and Economic Policies, Volume 4, Central Bank of Chile, pp. 427-475. ISBN # 956-7421-09-9.

Aron, J., J. Muellbauer and J. Prinsloo. 2006. "Estimating household-sector wealth in South Africa." *Quarterly Bulletin*, South African Reserve Bank, June, 61-72.

Aron, J. and J. Muellbauer. 2006. "Estimates of Household Sector Wealth for South Africa, 1970-2003." *Review of Income and Wealth* (International Association for Research in Income and Wealth) 52: 2 285-308, June.

Aron, J., J. Muellbauer and J. Prinsloo. 2008. "Estimating the Balance Sheet of the Personal Sector in an Emerging Market Country: SA, 1975-2005." Chapter 10 in James B. Davies (ed.), *Personal Wealth from a Global Perspective*, UNU-WIDER Studies in Development Economics, Oxford: Oxford University Press, October 2008, pp. 196-223.

Aron, J. and J. Muellbauer. 2013. "Wealth, credit conditions and consumption: evidence from South Africa." *Review of Income and Wealth* 59 (S1): S161–S196, October.

Aron, Janine, Greg Farrell and John Muellbauer. 2020. "Financial Stability and the Housing Market in South Africa." Mimeo. Nuffield College, University of Oxford, February 2020.

Aron, J. and J. Muellbauer. 2022a. "Enhancing Financial Stability and Monetary Analysis in the Core Model of the SARB." Mimeo. Nuffield College, University of Oxford, February 2022.

Aron, J. and J. Muellbauer. 2022b. "Non-Performing Loans and Arrears in South Africa: a Scoping paper for Future Model Development." Mimeo. Nuffield College, University of Oxford, February 2022.

Bertola, Guiseppe and Ricardo Caballero. 1990. "Kinked Adjustment Costs and Aggregate Dynamics." *NBER Macroeconomics Annual 1990*, 237-96. Cambridge: National Bureau of Economic Research.

Campbell, John Y. 1987. "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis." *Econometrica* 55 (6): 1249-73.

Cavalleri, Maria Chiara, Boris Cournède and Ezgi Özsöğüt. 2019. "How Responsive are Housing Markets in the OECD? National Level Estimates", OECD Economics Dept. working paper 1590.

Cerutti, Eugenio, Jihad Dagher, and Giovanni Dell'Ariccia. 2017. "Housing Finance and Real Estate Booms: A Cross-Country Perspective." *Journal of Housing Finance* 38: 1-13.

Chauvin, Valerie and John Muellbauer. 2018. "Consumption, household portfolios and the housing market in France." *Economie et Statistique/Economics and Statistics* 500-501-502: 157-178, plus online complement.

Chinloy, Peter T. 1980. "An Empirical Model of the Market for Resale Homes." *Journal of Urban Economics* 7 (3): 279-92.

Cramer, Jan S. 1957. "A Dynamic Approach to the Theory of Durable Consumer Goods." *Review of Economic Studies* 25 (2): 87-96.

de Jager, S., R. Ehlers, K. Mojapelo and P. Pienaar. 2021. "Short-term impacts and interaction of macroprudential policy tools." Working Paper Series WP/21/20, South African Reserve Bank.

DiPasquale, Denise and William C. Wheaton. 1994. "Housing Market Dynamics and the Future of Housing Prices." *Journal of Urban Economics* 35 (1): 1-27.

Dougherty, Ann and Robert Van Order. 1982. "Inflation, Housing Costs and the Consumer Price Index." *American Economic Review* 72 (1): 154-164.

Duca, John V., John Muellbauer, and Anthony Murphy. 2011. "House Prices and Credit Constraints: Making Sense of the U.S. Experience." *Economic Journal* 121 (552): 533–551.

Duca, John V. and John Muellbauer. 2013. "Tobin LIVES: Integrating Evolving Credit Market Architecture into Flow of Funds Based Macro-Models." ECB working paper 1581, published in A Flow-of-Funds Perspective on the Financial Crisis Volume II, edited by Bert Winkler, Arne Van Riet, and Peter Bull, 11-39. New York: Palgrave-Macmillan, 2014.

Duca, John V., John Muellbauer, and Anthony Murphy. 2016. "How Mortgage Finance Reform Could Affect Housing." *American Economic Review* 106 (5): 620–24.

Duca, J V, J Muellbauer and A Murphy (2021), "What Drives House Price Cycles? International Experience and Policy Issues", *Journal of Economic Literature* 59(3): 773-864.

Englund, Peter and Yannis M. Ioannides. 1997. "House Price Dynamics: An International Empirical Perspective." *Journal of Housing Economics* 6 (2):119-136.

Favara, Giovanni and Jean Imbs. 2015. "Credit Supply and the Price of Housing." *American Economic Review* 105 (3): 958-92.

Fernandez-Corugedo, E. and J. Muellbauer. 2006. "Consumer credit conditions in the United Kingdom." Bank of England Working Papers 314, Bank of England.

Geiger, Felix, John Muellbauer, and Manuel Rupprecht. 2016. “The Housing Market, Household Portfolios, and the German Consumer.” European Central Bank Working Paper 1904, European Central Bank, Frankfurt.

Georg, Co-Pierre and Davids, Allan. 2019. “The Cape of Good Homes: Exchange Rate Depreciations, Foreign Demand and House Prices.” Presented at the SARB Biennial Conference, October 2019.

Geiger, Felix, John Muellbauer, and Manuel Rupprecht. 2016. “The Housing Market, Household Portfolios, and the German Consumer.” European Central Bank Working Paper 1904, European Central Bank, Frankfurt.

Haurin, Donald. 1988. “The Duration of Marketing Time of Residential Housing.” AREUEA Journal 16 (4): 396-410.

Hendry, David F. 1984. “Econometric Modelling of House Prices in the United Kingdom.” In Econometrics and Quantitative Economics, edited by David F. Hendry and Kenneth F. Wallis. Oxford: Basil Blackwell.

Kearl, James R. 1979. “Inflation, Mortgages, and Housing.” *Journal of Political Economy* 87 (5): 1115-38.

Lemoine, M., Turunen, H., Chahad, M., Lepetit, A., Zhutova, A., Aldama, P., Clerc, P., and Laffargue, J.-P. (2019), ‘The FR-BDF Model and an Assessment of Monetary Policy Transmission in France’, Working Paper Series No. 736, Banque de France.Meen Geoffrey. 1990. "The Removal of Mortgage Market Constraints and the Implications for Econometric Modelling of UK House Prices." *Oxford Bulletin of Economics and Statistics*, 52 (1): 1-24.

Meen, Geoffrey. 2001. *Modeling Spatial Housing Markets: Theory, Analysis and Policy*, Norwell, MA: Kluwer Academic Publishers.
Muellbauer, John and Anthony Murphy. 1997. “Booms and Busts in the UK Housing Market.” *Economic Journal* 107 (445): 1701–27.

Muellbauer, John, Pierre St. Amant and David Williams. 2015. “Credit Conditions and Consumption, House Prices and Debt: What Makes Canada Different?” Bank of Canada, Staff Working Paper 2015-40. https://www.bankofcanada.ca/wp-content/uploads/2015/11/wp2015-40.pdf.

Poterba, James. 1984. “Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach.” *Quarterly Journal of Economics* 99 (4): 729-52.

Quigley, John M. 1979. “What Have We Learned About Urban Housing Markets?” In *Current Issues in Urban Economics*, edited by Peter M. Mieszkowski and Mahlon R. Straszheim, 391-429. Baltimore: Johns Hopkins University Press.

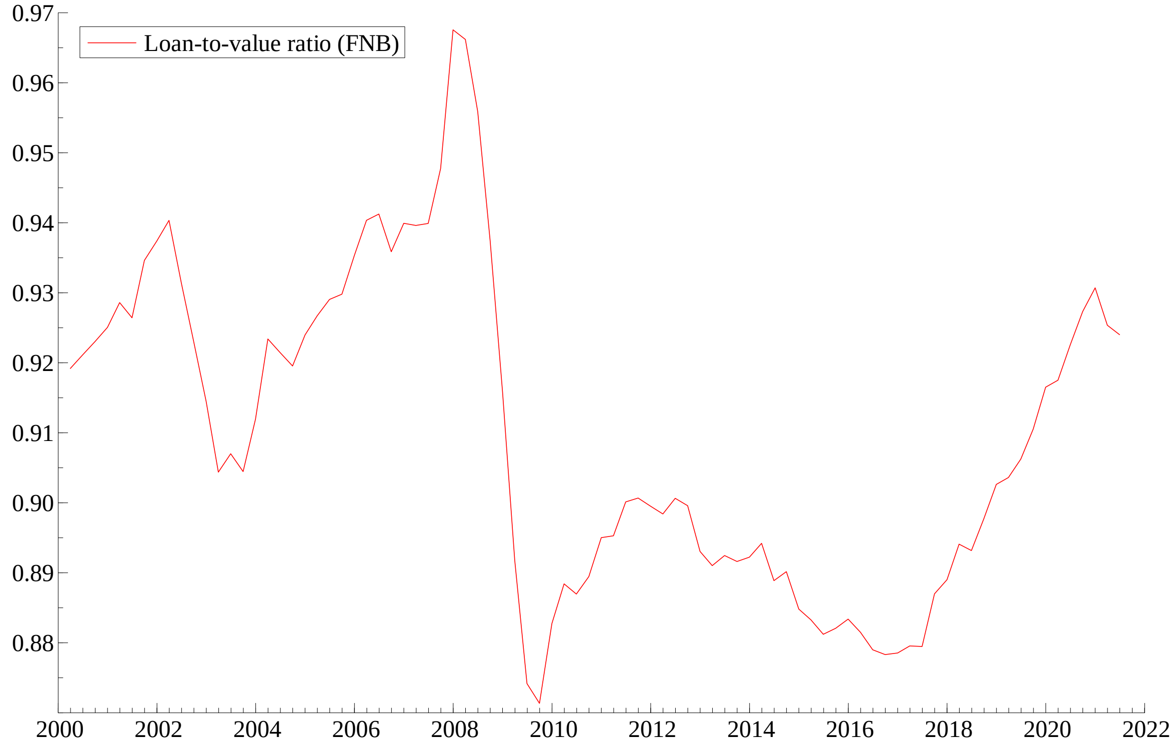
Rothschild, Michael. 1981. “Searching for the Lowest Price When the Distribution of Prices is Unknown.” *Journal of Economic Theory* 25 (1): 152-62.

Smal, D., C. Pretorius, and N. Ehlers. 2007. “The Core Forecasting Model of the South African Reserve Bank,” SARB Working Paper WP/07/02, June.

Stull, William J. 1978. “The Landlord’s Dilemma: Asking Rent Strategies in a Heterogeneous Housing Market.” *Journal of Urban Economics* 5 (1): 101-15.

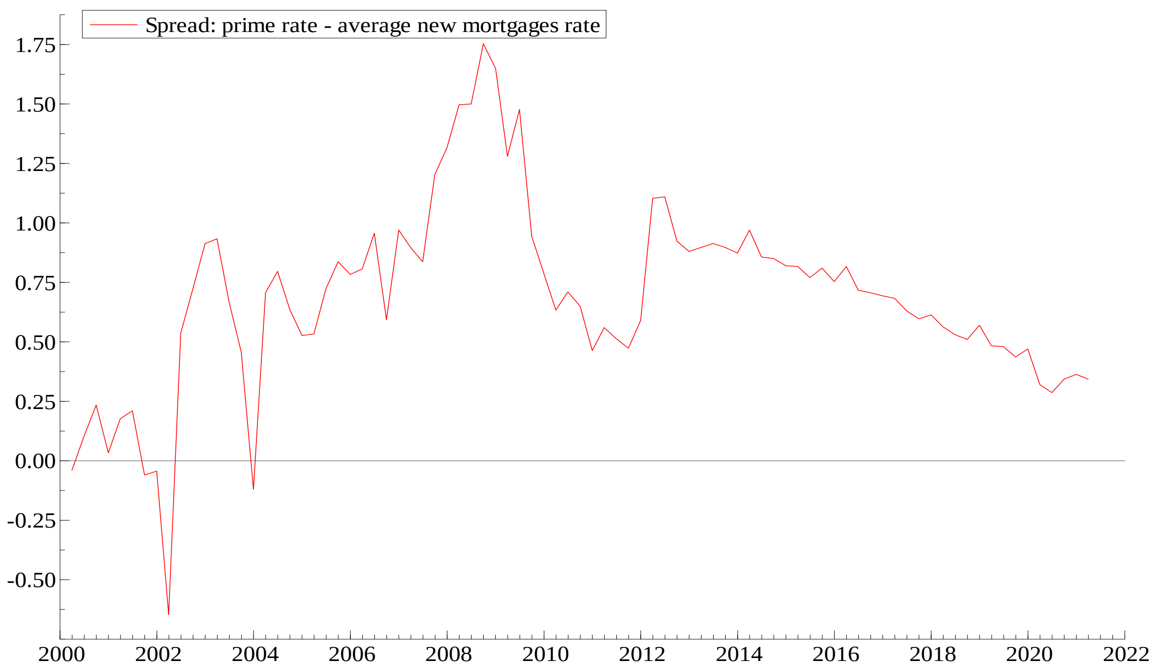
Wheaton, William C. 1990. “Vacancy, Search, and Prices in a Housing Market Matching Model.” *Journal of Political Economy* 98 (6): 1270-92.

Figure 1: CCI proxy: the loan-to-value ratio for all types of mortgages



Source: FNB compilation from Deeds Office data, South Africa.

Figure 2: CCI proxy: prime rate of interest minus average rate on new mortgages spread



Source: The average mortgage interest rate is BAT9612M, SARB, from 2001Q1. Data for 2000 interpolated by the authors.

Figure 3: Log of real house price and log user cost

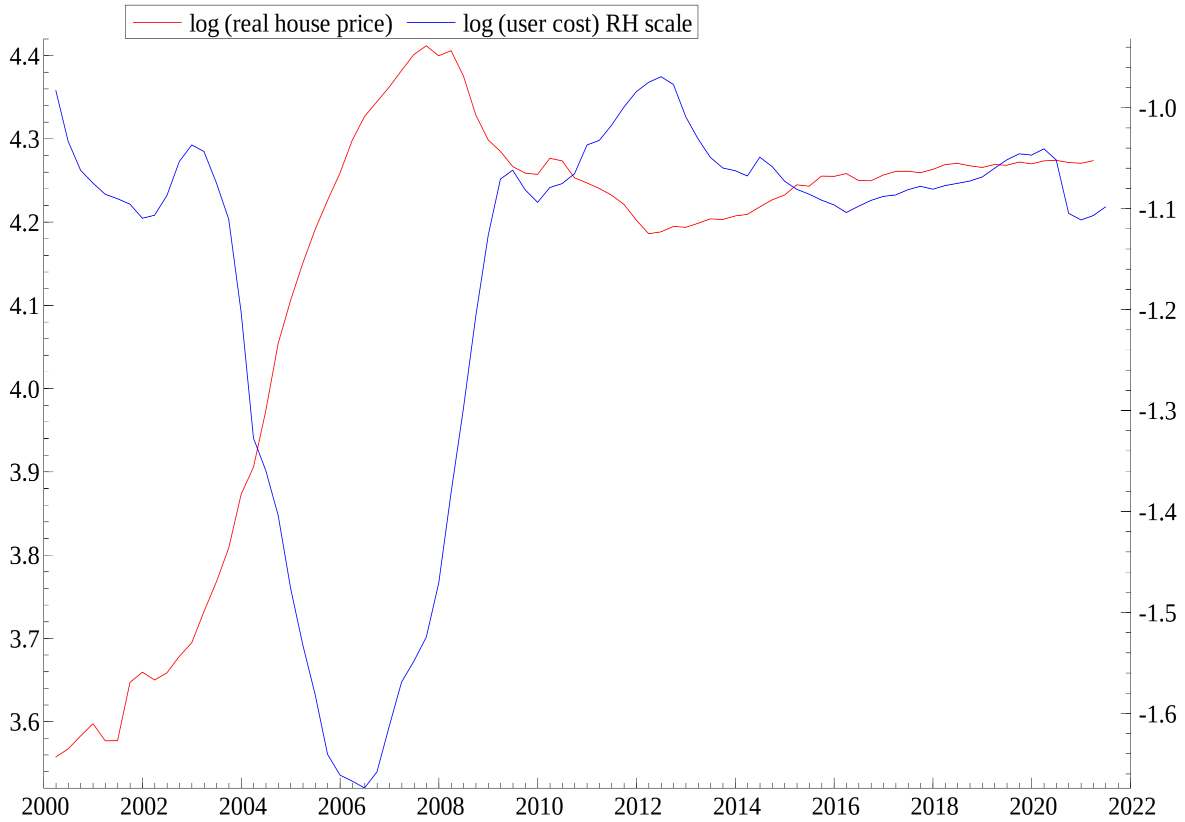


Figure 4: Log of real house price and income per house

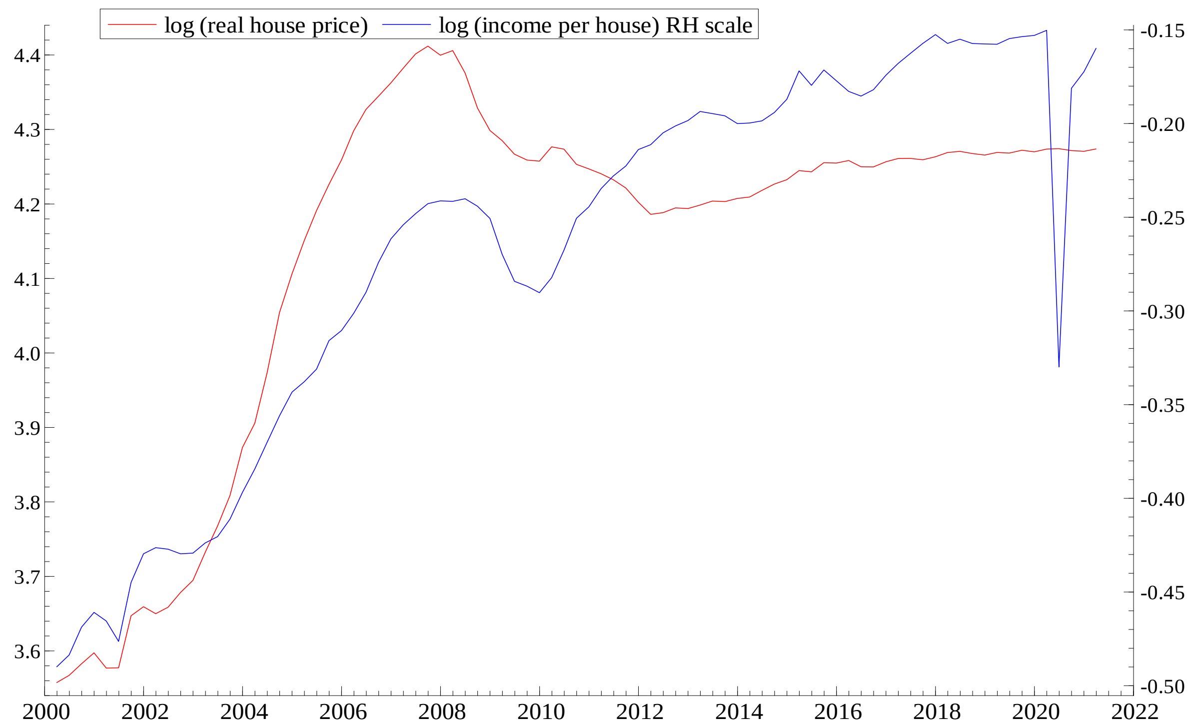


Figure 5: Log of real house price and property tax rate

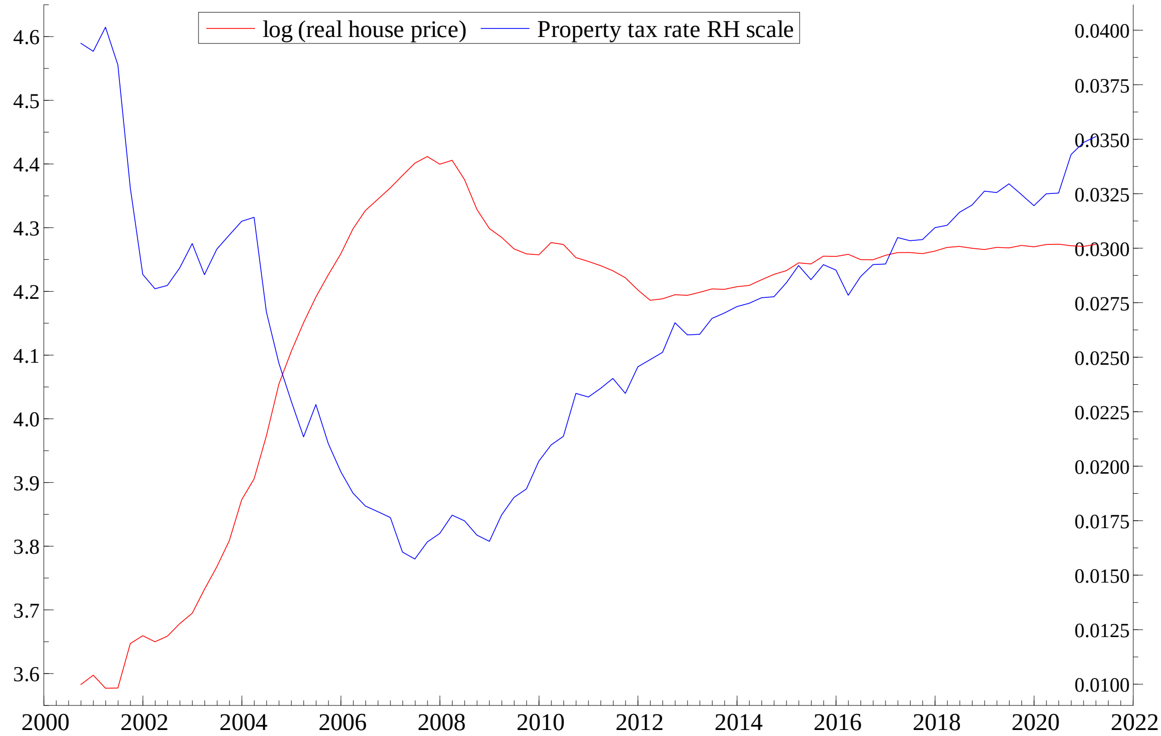


Figure 6: Log mortgage stock to income ratio and real interest rate

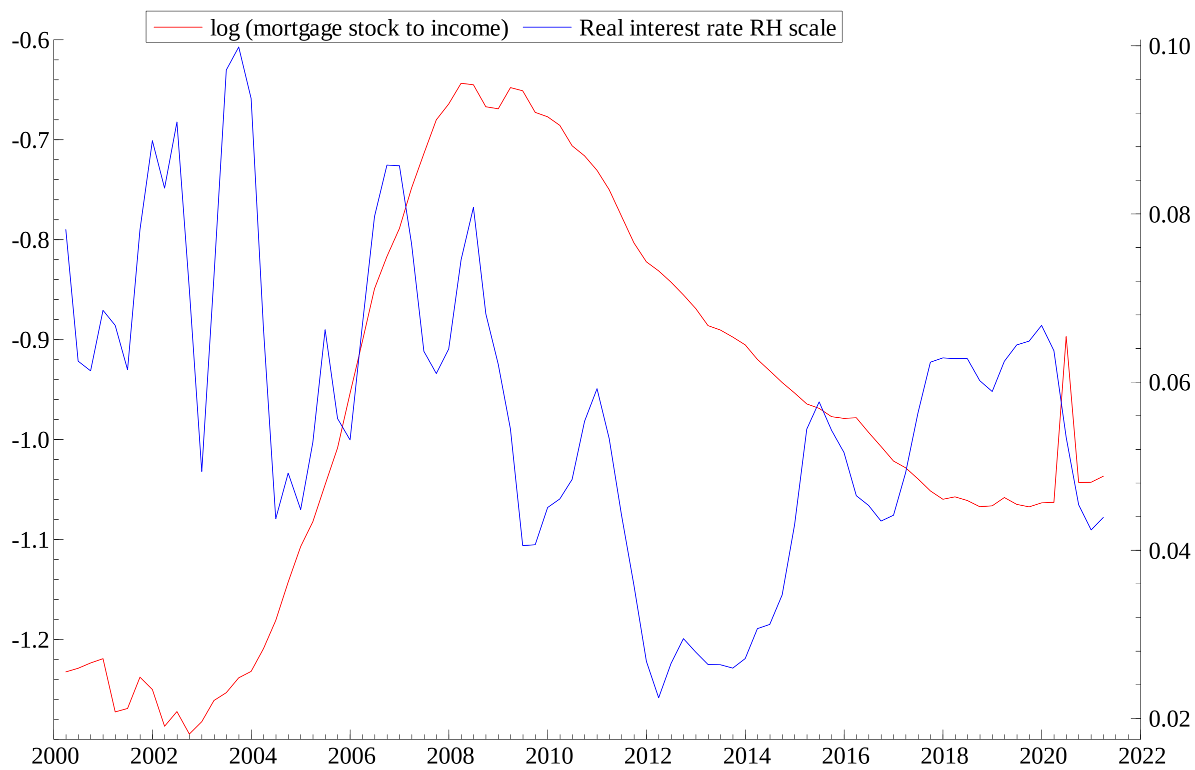


Figure 7: Log mortgage stock to income ratio and log house price to income ratio

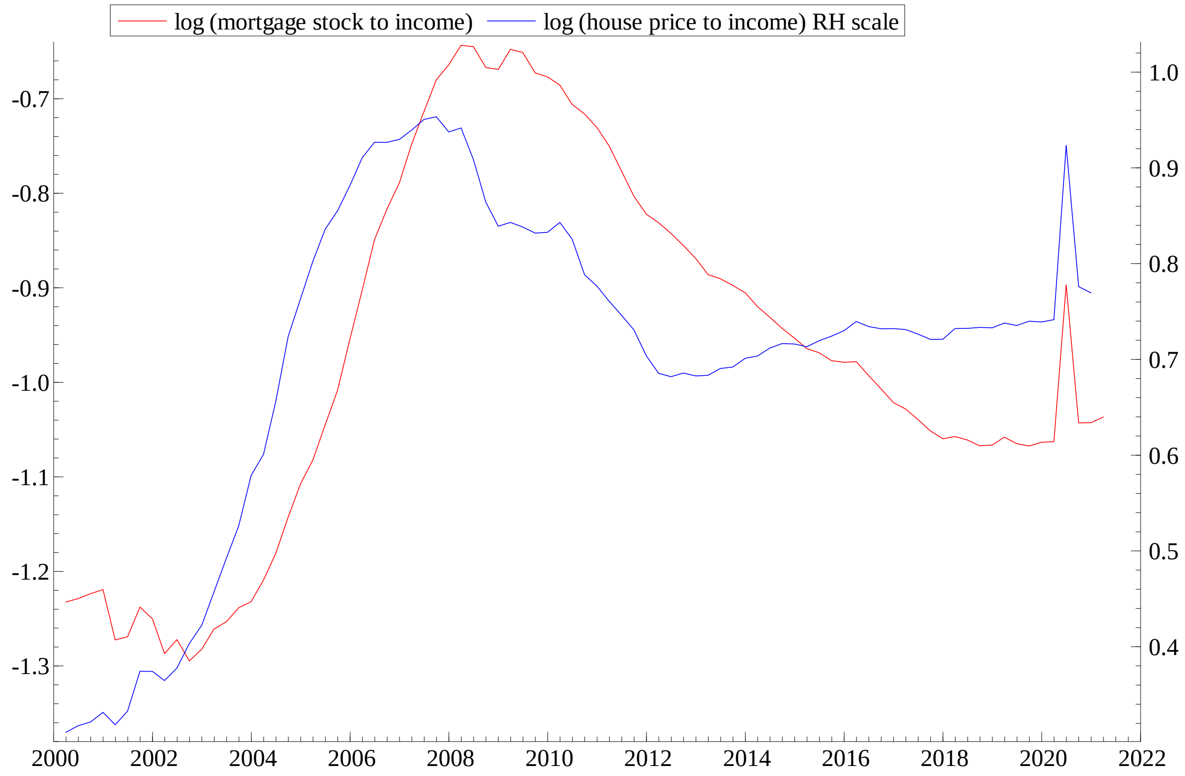


Figure 8: Log mortgage stock to income ratio and log non-mortgage stock to income ratio

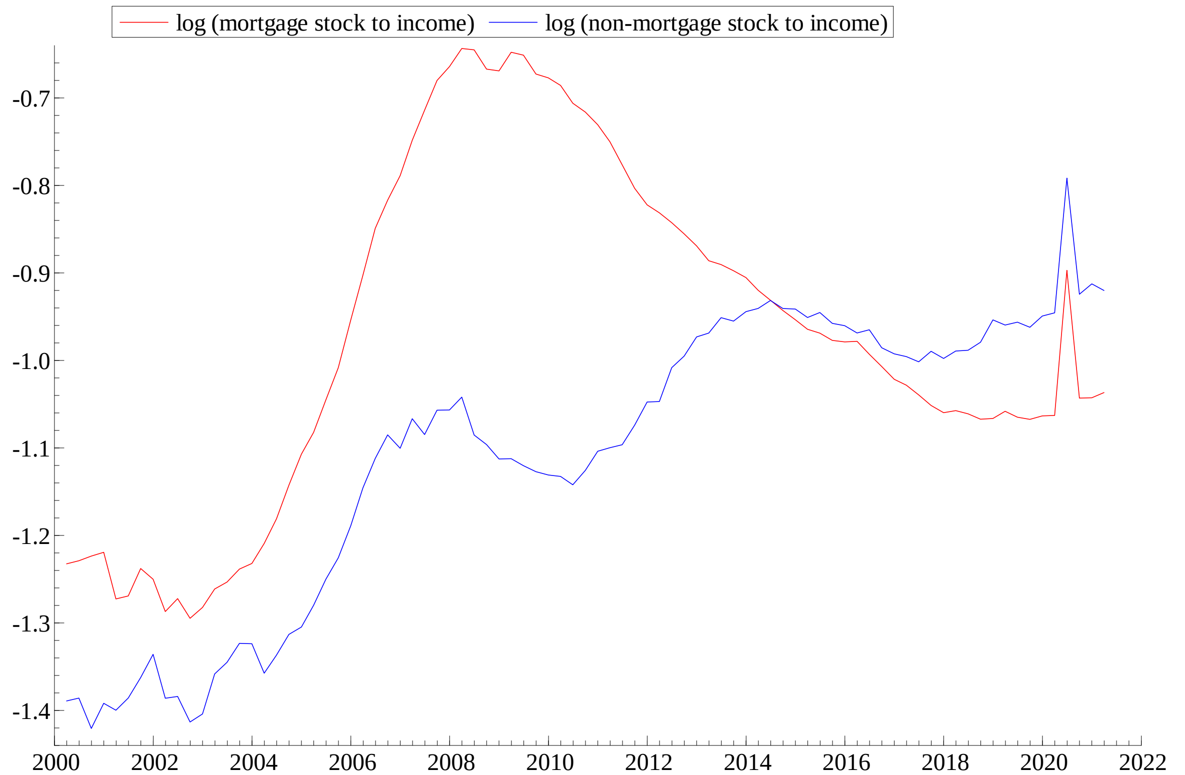


Figure 9: Log mortgage stock to income ratio and loan-to-value ratio

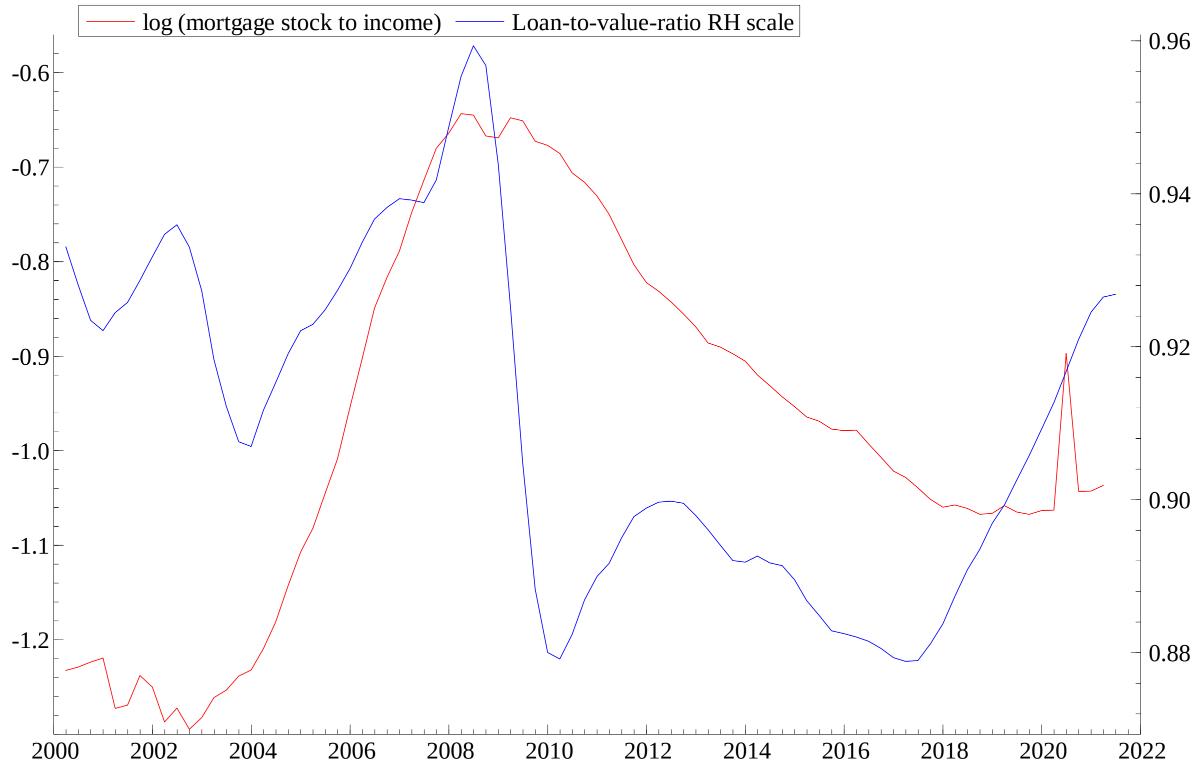


Figure 10: Log residential investment to GDP and log house price index to construction costs

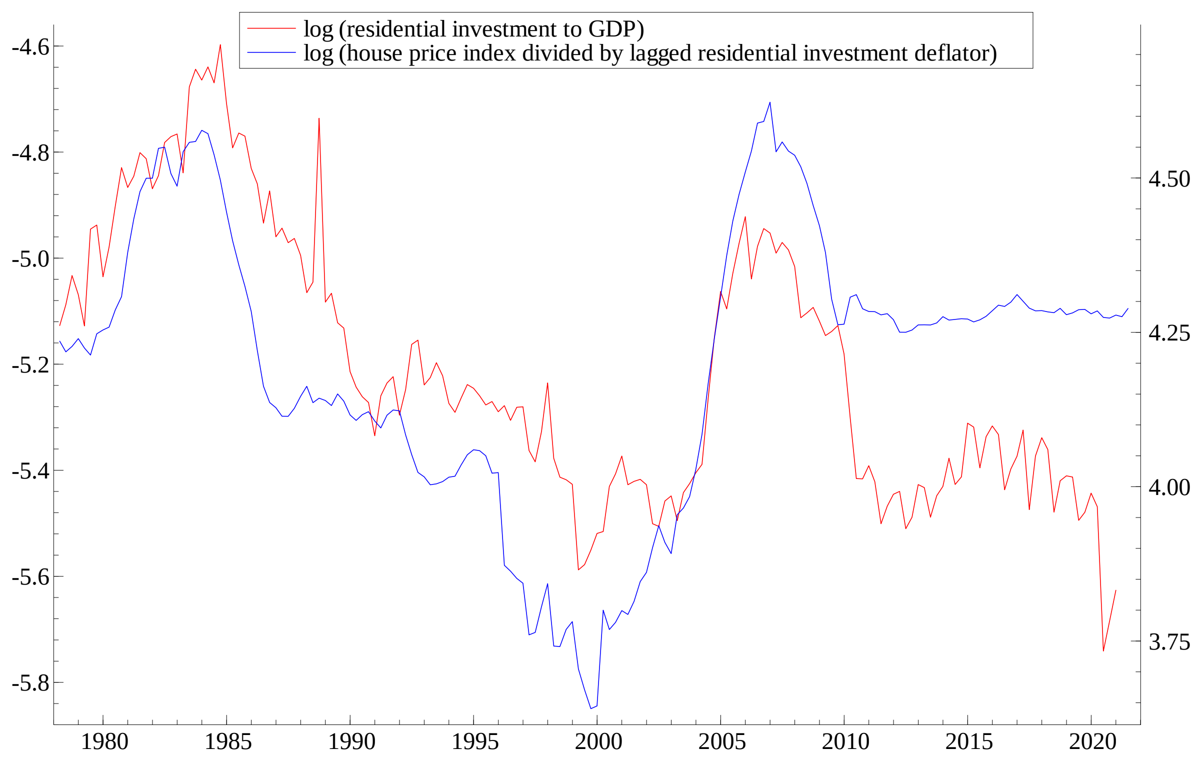


Figure 11: Log residential investment to GDP and annual population growth

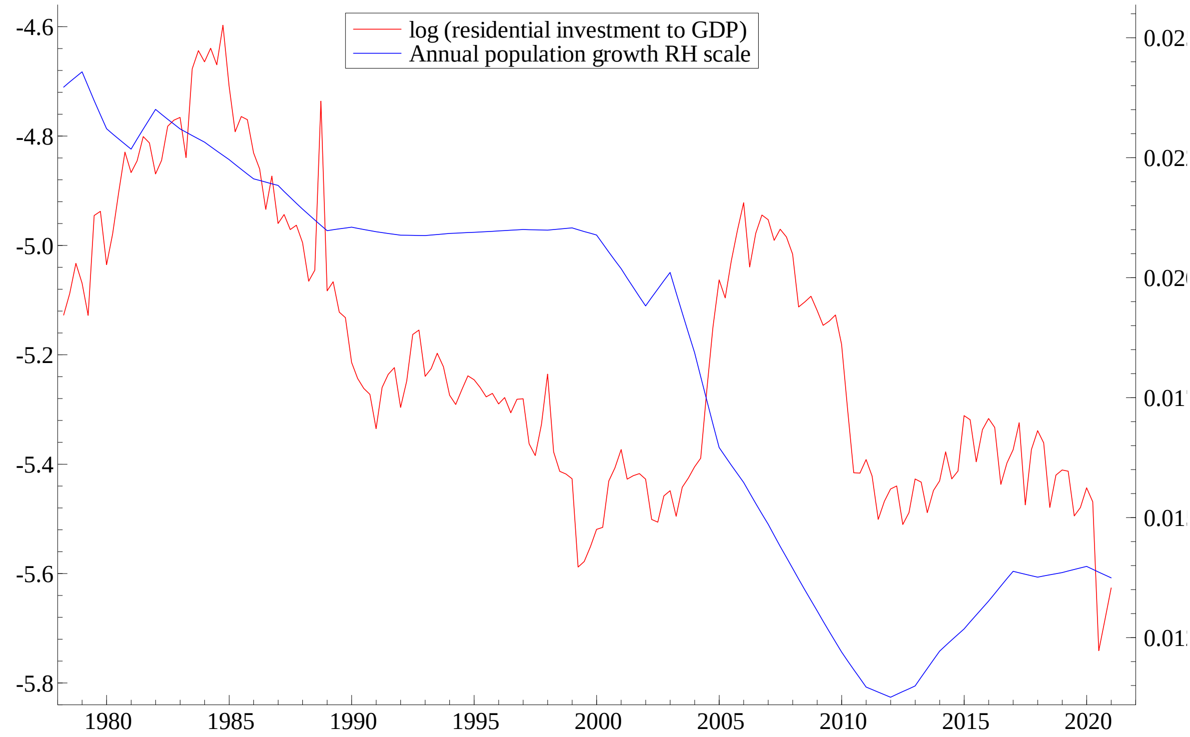


Figure 12: The four-quarter change in the prime rate of interest/100

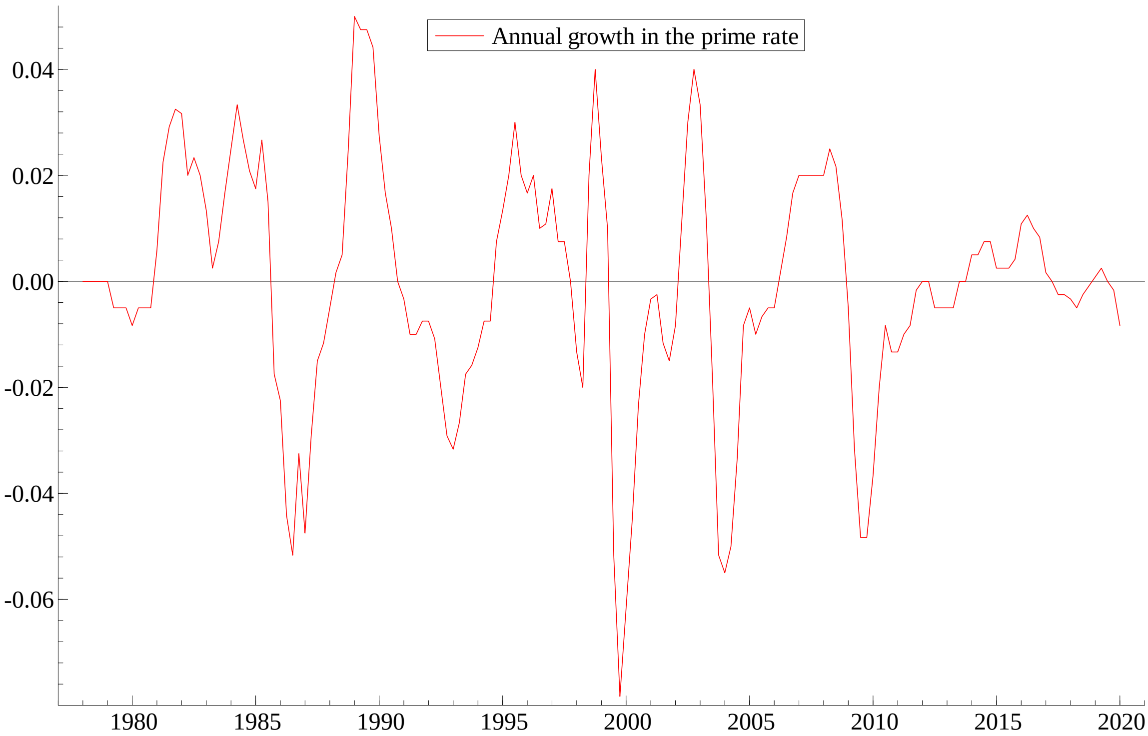


Figure 13: Real income growth

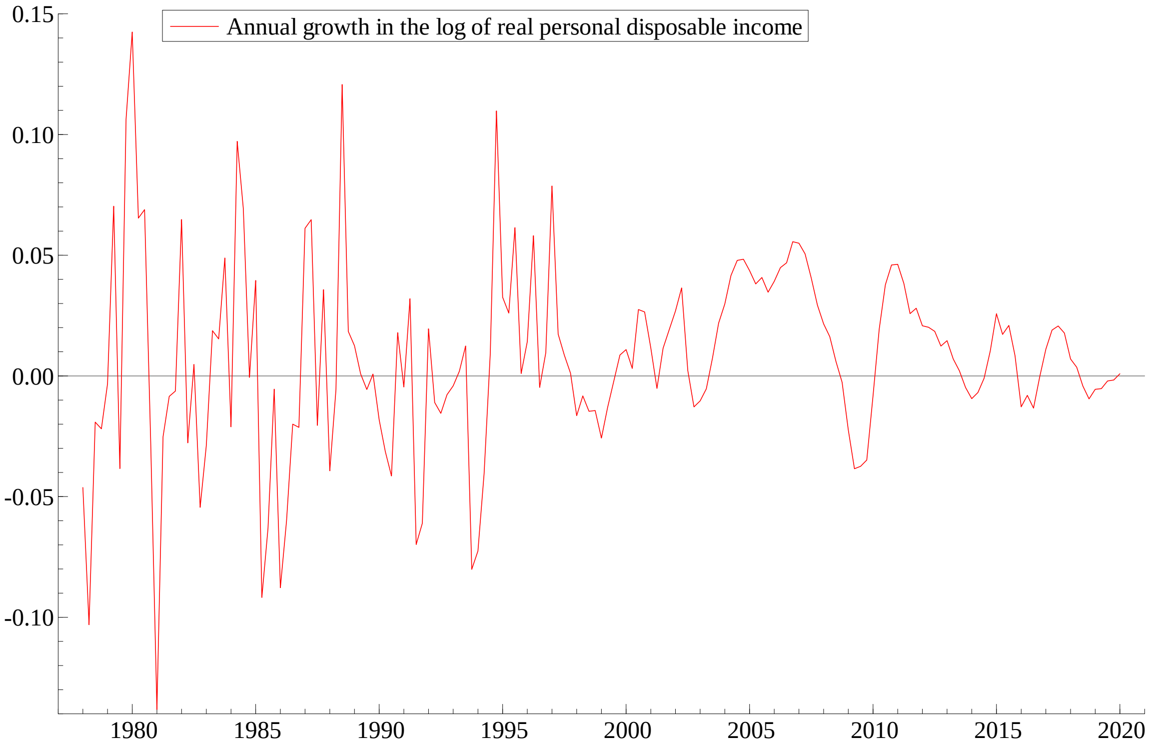


Table 1: Data definitions and sources

Variable	Definition	Mean	Std. deviation	Minimum	Maximum	Data source
House Price Equation						
DEPENDENT VARIABLE						
Δ log (nominal house prices)	Quarterly change in the log of the house price index. This is an average of several indices of house prices provided by mortgage lenders.	0.0226	0.0204	-0.0130	0.0804	SARB
INDEPENDENT VARIABLES						
log (real house prices)	The log of the house price index divided by the consumer expenditure deflator.	4.15	0.228	3.58	4.41	SARB
log (user cost)	The user cost is defined as follows: User cost= (prime rate $_{t-1}$)/100 -lagged house price appreciation + constant (proxying risk premium and transactions cost. Lagged hp appreciation: Δ_{16} log (house price) $_{t-1}$)/4. Constant = 0.3, see Section 3.3.	-1.17	0.204	-1.67	-0.969	Quarterly Bulletin and SARB
Interaction term: LTV (ma3) × log (user cost)	Loan-to-value ratio for all residential mortgages, derived from the Deeds Office data (a three-quarter moving average) multiplied by the log of user costs, as defined above. Both terms are de-meaned before being interacted.	-0.00268	0.00505	-0.0176	0.00122	FNB, Quarterly Bulletin and SARB
Log (income per house)	The log of the ratio of the real per capita household disposable income to the housing stock measure from the National Accounts (per capita, in constant prices and lagged one quarter).	-0.261	0.0962	-0.476	-0.150	Quarterly Bulletin

Variable	Definition	Mean	Std. deviation	Minimum	Maximum	Data source
mortgage rate spread (ma3)	Prime rate minus the average interest rate on new mortgage loans (all expressed as a three-quarter moving average).	0.00721	0.00354	-0.0025	0.0163	SARB
property tax rate (ma4)	The local government revenue from property taxes on residential property divided by the value of total housing wealth from the household balance sheets (expressed as a four-quarter moving average).	2.29	0.486	1.46	3.64	SARB
Δ_4 log (consumer expenditure deflator)	Annual change in the log of the consumer expenditure deflator.	0.0558	0.0192	0.0223	0.118	Quarterly Bulletin
Δ_4 prime rate	Annual change in the prime rate divided by 100.	-0.00252	0.0185	-0.055	0.04	Quarterly Bulletin
Δ_4 real prime rate	Annual change in the prime rate divided by 100 minus the annual change in the log of the consumer expenditure deflator.	-0.00028	0.0226	-0.0574	0.0518	Quarterly Bulletin
Adjusted Δ_8 log (REER)	The adjusted two-year change in the log of the real effective exchange rate. A rise is a Rand appreciation. The adjustment multiplies the variable by (1 minus a smoothed transition dummy), see Section 3.4. The dummy is zero until 2014Q4; a smooth rise to 1 over two years from 2015 to the end of 2016; then 1.	-0.0113	0.134	-0.297	0.323	Quarterly Bulletin
$\Delta\Delta$ log (employment)	Quarterly acceleration in the log of employment.	7.63E-05	0.00776	-0.0315	0.0243	Quarterly Bulletin
Δ LTV	Quarterly change in the loan-to-value ratio for all residential mortgages, derived from the Deeds Office data.	-7.1E-05	0.00684	-0.0240	0.0198	FNB

Variable	Definition	Mean	Std. deviation	Minimum	Maximum	Data source
Debt equation						
DEPENDENT VARIABLE						
Δ log (mortgage debt per capita)	Quarterly change in the log of mortgage debt (in current prices) divided by population.	0.0199	0.0234	-0.00549	0.0810	Quarterly Bulletin
INDEPENDENT VARIABLES						
log (mortgage debt to income ratio)	The log of mortgage debt (in current prices) divided by income (in current prices).	-0.948	0.192	-1.29	-0.643	Quarterly Bulletin
log (house price to income ratio)	The log of the house price index divided by per capita nominal household disposable income.	0.730	0.143	0.365	0.953	Quarterly Bulletin and SARB
log (non-mortgage debt to income ratio)	The log of non-mortgage debt (in current prices) divided by income (in current prices).	-1.10	0.147	-1.413	-0.931	Quarterly Bulletin
real prime rate (ma2)	Prime rate divided by 100 minus the annual change in the log of the consumer expenditure deflator (all expressed as a two-quarter moving average).	0.0563	0.0187	0.0225	0.0999	Quarterly Bulletin
LTV (ma4)	Loan-to-value ratio for all residential mortgages, derived from the Deeds Office data (expressed as a four-quarter moving average).	0.908	0.0227	0.879	0.959	FNB
mortgage rate spread (ma4)	Prime rate minus the average interest rate on new mortgage loans (all expressed as a four-quarter moving average).	0.00743	0.00331	-0.00135	0.0160	SARB
Δ mortgage rate spread	Quarterly change in the above spread.	0.0000147	0.00251	-0.00603	0.0118	SARB

Variable	Definition	Mean	Std. deviation	Minimum	Maximum	Data source
Δ log (real household disposable income)	Quarterly change in the log of per capita real household disposable income.	0.00384	0.00755	-0.0177	0.0283	Quarterly Bulletin
Δ_4 log (consumer expenditure deflator)	Annual change in the log of the consumer expenditure deflator.	0.0548	0.0189	0.0223	0.118	Quarterly Bulletin
$\Delta_4\Delta_4$ log (population)	Annual acceleration in log population.	-0.000327	0.000727	-0.00198	0.000725	Quarterly Bulletin
$\Delta\Delta$ log (employment)	Quarterly acceleration in the log of employment.	0.0000186	0.00779	-0.0315	0.0243	Quarterly Bulletin
Dummy 2002Q3	Impulse dummy.					Constructed
Residential Investment equation						
DEPENDENT VARIABLE						
Δ log (residential investment per capita)	Quarterly change in the log of residential investment (in constant prices) divided by population.	-0.00146	0.0684	-0.343	0.317	Quarterly Bulletin
INDEPENDENT VARIABLES						
log (residential investment to GDP ratio)	The log of residential investment (in constant prices) divided by GDP (in constant prices).	-5.20	0.245	-5.59	-4.60	Quarterly Bulletin
log (house prices) $_{t-1}$ –log (construction costs) $_{t-5}$	The log of house prices less the log of construction costs a year previous, measured as the residential investment deflator.	4.19	0.234	3.64	4.62	Quarterly Bulletin and SARB
Δ_4 prime rate	Annual change in the prime rate of interest.	-0.000493	0.0222	-0.0783	0.0500	Quarterly Bulletin
Δ_4 log (real household disposable income)	Annual change in the log of per capita real household disposable income.	0.00752	0.0395	-0.138	0.142	Quarterly Bulletin
Δ_4 log (population)	Annual change in log population.	0.0184	0.00415	0.0113	0.0243	Quarterly Bulletin
Seasonal Q1	Seasonal					Constructed

Variable	Definition	Mean	Std. deviation	Minimum	Maximum	Data source
Smoothed transition dummy (2017 to 2018)	Zero until 2016Q4; a smooth rise to 1 in 2018Q4; then 1.					Constructed
Seasonality shift from 2007 for Q2	Zero before 2007; Q2 seasonal from 2007 onwards.					Constructed
Dummy 1979Q2	Impulse dummy.					Constructed
Dummy 1988Q3	Impulse dummy.					Constructed
Dummy 1997Q4	Impulse dummy.					Constructed

Table 2: House Price Model Results

Dependent variable: Δ log (nominal house prices) $_{t}$	*2000:4 to 2020:1*		*2000:4 to 2020:1*		*2000:4 to 2020:1*		*2000:4 to 2014:4*	
	Eq. 1		*Eq. 2*		*Eq. 3*		*Eq. 4*	
	Coefficient	*t-Statistic*	*Coefficient*	*t-Statistic*	*Coefficient*	*t-Statistic*	*Coefficient*	*t-Statistic*
constant	0.789	8.1	0.696	7.0	0.686	7.0	0.856	7.4
log (real house prices) $_{t-1}$	-0.171	-8.2	-0.151	-7.2	-0.155	-7.1	-0.184	-7.5
log (user cost) $_{t-1}$	-0.0264	-2.3	-0.0252	-2.0	-0.0555	-5.6	-0.0242	-1.7
Interaction term: LTV (ma3) $_{t-1}$ × log (user cost) $_{t-1}$	-1.56	-3.4	-1.26	-2.7	-	-	-1.80	-3.1
log (income per house) $_{t-2}$	0.268	7.0	0.229	6.0	0.228	6.0	0.284	6.4
mortgage rate spread (ma3) $_{t-1}$	1.32	4.1	1.28	3.8	1.46	3.8	1.49	4.0
Property tax rate (ma4) $_{t-1}$	-0.0165	-5.2	-0.0149	-4.5	-0.0166	-4.8	-0.0191	-4.7
Δ log (house prices) $_{t-1}$	0.589	9.0	0.627	9.1	0.516	7.1	0.591	7.7
Δ_4 log (consumer expenditure deflator) $_{t-1}$	-0.106	-2.2	-0.117	-2.2	-0.132	-2.6	-0.111	-1.9
Δ_4 prime rate $_{t}$	-0.191	-4.0	-0.185	-3.6	-0.105	-2.3	-0.212	-3.6
Δ_4 real prime rate $_{t}$	-0.0548	-2.0	-0.0598	-2.0	-0.0642	-2.1	-0.0570	-1.8
Adjusted Δ_8 log (REER) $_{t}$	0.0377	5.4	0.0278	4.3	0.0304	4.4	0.0384	4.9
$\Delta\Delta$ log (employment) $_{t-1}$	0.175	2.0	0.160	1.7	0.145	1.6	0.210	2.0

$\Delta\Delta$ log (employment) $_{t-2}$	0.170	1.9	0.150	1.6	0.106	1.1	0.267	2.1
Δ LTV $_{t-1}$	-	-	-	-	0.213	1.7	-	-
Equation standard error	0.00458		0.00489		0.00488		0.00504	
Adjusted R-squared	0.950		0.943		0.943		0.951	
Durbin-Watson	1.90		1.90		1.69		2.01	
Breusch/Godfrey LM: AR/MA4	*p* = [.095]		*p* = [.371]		*p* = [.080]		*p* = [.073]	
Chow test	*p* = [.148]		*p* = [.109]		*p* = [.152]		*p* = [.461]	
Breusch-Pagan het. Test	*p* = [.457]		*p* = [.911]		*p* = [.354]		*p* = [.780]	

Notes: Estimation performed in TSP 5.0 of Hall and Cummins.

‡ Corresponds to an unadjusted Δ_8 log (REER), i.e., it indicates the drop in the estimated effect without adjustment.

Table 3: Mortgage Debt Model Results

Dependent variable: Δ log (mortgage debt per capita) $_t$	*2001:3 to 2020:1*		*2001:3 to 2020:1*		*2001:3 to 2014:4*	
	Eq. 1		*Eq. 2*		*Eq. 3*	
	Coefficient	*t-Statistic*	*Coefficient*	*t-Statistic*	*Coefficient*	*t-Statistic*
constant	-0.517	-5.9	-0.546	-6.2	-0.628	-5.6
log (mortgage debt relative to income) $_{t-1}$	-0.0546	-6.0	-0.0514	-5.6	-0.0413	-2.2
log (house price to income ratio) $_{t-1}$	0.0781	4.1	0.0915	5.1	0.0887	3.6
log (non-mortgage debt relative to income) $_{t-1}$	-0.0937	-5.4	-0.102	-5.9	-0.123	-4.7
Real prime rate (ma2) $_t$	-0.374	-5.5	-0.385	-5.6	-0.493	-4.3
LTV (ma4) $_{t-2}$	0.389	4.9	0.412	5.2	0.497	4.7
mortgage rate spread (ma4) $_{t-1}$	0.657	1.9	-	-	-	-
Δ log (mortgage debt per capita) $_{t-1}$	0.131	1.9	0.119	1.7	0.0978	1.2
Δ log (mortgage debt per capita) $_{t-2}$	0.163	2.6	0.137	2.2	0.164	2.1
Δ log (mortgage debt per capita) $_{t-3}$	0.207	3.2	0.195	3.0	0.208	2.6
Δ mortgage rate spread $_t$	1.58	5.5	1.47	5.1	1.34	3.9
Δ log (real household disposable income) $_t$	0.552	4.4	0.510	4.0	0.655	3.7
Δ_4 log (consumer expenditure deflator) $_{t-1}$	-0.363	-5.4	-0.369	-5.4	-0.412	-4.5
$\Delta_4\Delta_4$ log (population) $_{t-2}$	10.3	4.3	11.2	4.6	14.0	4.1
$\Delta\Delta$ log (employment) $_{t-1}$	0.219	2.1	0.231	2.2	0.246	1.9

ΔΔ log (employment) $_{t-2}$	0.353	3.1	0.355	3.1	0.467	2.9
Dummy 2002Q3	-0.0395	-5.9	-0.0425	-6.4	-0.0430	-5.4
Equation standard error	0.00537		0.00549		0.00614	
Adjusted R-squared	0.947		0.945		0.943	
Durbin-Watson	2.08		1.95		2.03	
Breusch/Godfrey LM: AR/MA4	*p* = [.165]		*p* = [.244]		*p* = [.407]	
Chow test	*p* = [.340]		*p* = [.164]		*p* = [.116]	
Breusch-Pagan het. Test	*p* = [.108]		*p* = [.086]		*p* = [.553]	

Notes: Estimation performed in TSP 5.0 of Hall and Cummins.

Table 4: Residential Investment Model Results

Dependent variable: Δ log (residential Investment per capita) $_t$	*1978:1 to 2020:1*		*1994:3 to 2020:1*		*1978:1 to 2007:4*	
	Eq. 1		*Eq. 2*		*Eq. 3*	
	Coefficient	*t-Statistic*	*Coefficient*	*t-Statistic*	*Coefficient*	*t-Statistic*
constant	-2.99	-8.8	-3.22	-6.5	-3.05	-7.7
log (residential investment to GDP ratio) $_t$	-0.304	-8.9	-0.329	-6.9	-0.308	-7.5
log (house prices) $_{t-1}$ – log (construction costs) $_{t-5}$	0.283	8.1	0.303	5.8	0.287	7.4
Δ log (residential Investment per capita) $_{t-4}$	0.261	5.8	0.284	3.7	0.252	4.9
Δ_4 prime rate $_{t-1}$	-0.985	-6.6	-0.907	-5.3	-1.09	-6.7
Δ_4 log (real household disposable income) $_{t-1}$	0.327	4.1	0.193	1.3	0.370	4.0
Δ_4 log (real household disposable income) $_{t-5}$	0.289	3.6	0.278	2.0	0.317	3.5
Δ_4 log (population) $_{t-1}$	12.6	7.5	13.8	4.8	13.7	5.7
Seasonal Q1 $_t$	-0.026	-3.8	-0.034	-3.7	-0.0282	-3.4
Smoothed transition dummy (2017 to 2018)	-0.042	-2.9	-0.047	-3.3	-	-
Seasonality shift from 2007 for Q2	-0.065	-5.1	-0.066	-4.9	-	-
Dummy 1979Q2 $_t$	0.119	3.2	-	-	0.118	3.0
Dummy 1988Q3 $_t$	0.321	8.6	-	-	0.320	8.2

Dummy 1988Q3 $_{t-1}$	-0.250	-6.3	-	-	-0.253	-6.0
Dummy 1997Q4 $_{t}$	0.124	3.3	0.120	3.5	0.125	3.2
Equation standard error	0.0365		0.0328		0.0382	
Adjusted R-squared	0.714		0.681		0.719	
Durbin-Watson	1.98		1.65		2.02	
Breusch/Godfrey LM: AR/MA4	p = [.975]		p = [.116]		p = [1.00]	
Chow test	p = [.907]		p = [.592]		p = [.891]	
Breusch-Pagan het. Test	p = [.042]		p = [.613]		p = [.018]	

Notes: Estimation performed in TSP 5.0 of Hall and Cummins.

South African Reserve Bank

Financial Stability Department

Enhancing Financial Stability and Monetary Analysis in the Core Model of the SARB

Abstract

Central bank models without a well-articulated credit channel and links between the financial sector and the real economy may misrepresent the timing and profile of monetary policy transmission and risks to financial stability. This paper draws on international literature to propose improvements to the Core econometric policy model of the South African Reserve Bank (and a recent extension adding a banking sector). Real estate plays a major role in the monetary transmission mechanism, but the Core model misses the crucial elements. There is almost no role for credit conditions, and no explicit role for house price expectations. Empirical evidence on expectations of house price appreciation suggests a potential for overshooting of house prices (currently not captured in the model) and of mortgage debt, followed by painful corrections. We propose a forward-looking approach to incorporate income expectations in the consumption equation through modelling permanent income (currently not in the model). The consumption equation should include changing credit conditions and relax the net worth restriction on household wealth to capture the separate impacts of housing wealth, illiquid and liquid assets, and debt. The long-run solution of the house price equation should improve by incorporating the supply side, with explicit roles for credit conditions and house price expectations, as well as interest rates. House prices transmit strongly into mortgage debt. We propose replacing the single aggregate household debt equation by separate equations for mortgage debt and non-mortgage debt and capturing not just a direct interest rate effect but also the indirect effects of interest rates via house prices and shifting credit availability. House prices also transmit strongly to residential investment, and we propose including a

residential investment equation, currently missing in the Core model. We propose adjustments to the banking sector equations to enhance the understanding of linkages with macro-prudential policy. Improving the database on commercial real estate, closely tracking loan-to-value ratios and credit spreads in the mortgage market and modelling the consequences of changing credit availability for consumption, debt, house prices and investment, should enhance the understanding of financial stability risks in South Africa.

Acknowledgements: We thank Shaun de Jager, Pieter Pienaar and Riaan Ehlers for helpful discussions and a critical review. We are grateful to the following persons at the SARB for their advice, particularly on data: Wian Boonzaaier, Elriette Botes, Rashad Cassim, Shaun de Jager, Karen Kuhn, Danie Meyer, Caswell Monyela, Mpho Moloto, Lesego Morope, Khumbudzo Muneri, Susana Paulse, and Bart Stemmet. Outside the SARB, we thank David Aikman, Greg Farrell, John Loos and Siphamandla Mkhwanazi (FNB), Hendrik Nel, Johan van den Heever and Ben Smit.

Contents

Executive Summary

Our paper recommends improvements to the SARB's Core model to increase its relevance for macro-prudential stress testing and for setting monetary policy. Since the publication of the Core model in 2007, there has been further model development but no updated publication. The most recent published version, see De Jager *et al.* (2021), adds a banking sector and expands the linkages that can be influenced by macro-prudential policy between the banking system and the real economy. This is not the version used by the Monetary Policy Committee. Our suggested model improvements apply to both versions of the Core model.

Our comments apply to the consumption function, the house prices equation, the mortgage stock equation, and suggest including a residential investment equation, currently absent. Improving the treatment of the Capital Adequacy Ratio (CAR) and the Loan-to-Value Ratios (LTVs) in the new banking sector part of the model is a further aspect.

A *well-specified model for consumption* (given that it comprises about 60% of GDP) is crucial for understanding monetary transmission and financial stability. The consumption equation is discussed in Section 2. We recommend three improvements. First, a more explicit treatment of income expectations for consumption. Second, to relax the highly restrictive 'net worth' assumption on household wealth to capture the different impacts of housing wealth, illiquid and liquid assets, and debt. And third, to introduce credit conditions (which vary over time) into the equation.

On a more explicit treatment of expectations, we propose using a forward-looking approach to incorporate income expectations through modelling permanent income. In contrast to using the text-book concept of permanent income, which uses a very low discount rate for future income, our proposal reflects the more limited horizons of real-world households with a more realistic discount rate. The weight that households place on expected income as compared to current income needs to be estimated empirically.

On relaxing the 'net worth' assumption, which applies *equal weights* to the different components of wealth, we regard it as crucial that the very *different weights* of these components should be estimated separately. For example, cash in a bank deposit is clearly more ‘spendable’ than an illiquid financial asset such as a pension and hence will have a far bigger impact on consumption. Moreover, housing is a consumption good as well as an asset with an important collateral role for access to credit and should be distinguished from financial assets.

On introducing credit conditions, these vary over time and need to be controlled for in consumption functions. This is because the asymmetric information between lenders and borrowers means that lenders impose collateral requirements for mortgage borrowing and use a wide range of screening devices to reduce the risk of bad loans, and these are far from constant over time.

Our recommendations for the *house price equation* are explained in Section 4. The long-run solution should be improved in three ways: by incorporating the supply side, bringing in an explicit role for credit conditions, and introducing house price expectations. Drawing on Aron and Muellbauer (2022a), we note evidence that South African housing market participants extrapolate past house price changes over several years in forming expectations of appreciation (as in the US). This can lead to the *overshooting* of house prices and of mortgage debt, followed by painful corrections. Neither are currently captured in the model.

House prices transmit strongly into mortgage debt, an effect which is missing in the current model. Improvements toward a household mortgage debt equation are discussed in Section 5. We suggest the current single aggregate household debt equation (driven by bank credit extension) in the Core model be replaced by *separate equations for mortgage debt and for non-mortgage debt*, as these are driven by different factors. Monetary policy transmission should not be confined to an interest rate effect in the household debt equation. Instead, house prices and shifts in credit availability should be explicitly incorporated into the proposed mortgage debt equation. Interest rates feed strongly into house prices, and both direct and indirect effects of interest rates on mortgage debt are important in monetary transmission.

House prices also transmit strongly to residential investment. We propose the inclusion of a *residential investment equation* (currently only aggregate investment is modelled), and this is discussed in Section 6. A candidate equation, given in in our partner paper, Aron and Muellbauer (2022a), is remarkably stable back to the late 1970s, despite many shocks and structural changes in the economy. This equation captures a further important indirect effect on aggregate demand of interest rates via house prices, currently missing in the Core model.

'Oven-ready equations'. The empirically-estimated equations for house prices, mortgage debt and residential investment in our partner paper, Aron and Muellbauer (2022a), which incorporate the above recommendations, could relatively straightforwardly be introduced into the Core Model, better to interpret monetary transmission, with important implications also for financial stability. For the consumption function, we suggest in Sections 3 and 8, some simplifications of our earlier work on consumption (Aron and Muellbauer, 2013), for ease of implementation. Monetary policy transmission is stronger in the Aron-Muellbauer (2013) consumption function than in the MPRU-Core or Core models. Not only is there a direct effect of interest rates on consumption, also present in the Core models, but strong, indirect effects via housing wealth, illiquid financial assets and permanent income.

Finally, the paper proposes, in Section 7, several corrections and adjustments to the four banking sector equations in the 2021 published version of the Core model. While the introduction of a banking sector with four new equations linking the capital adequacy ratio (CAR) to credit extension and spreads is an important step forward, quite a few improvements are needed to make clear how macro-prudential policy transmits to the real economy through credit extension. In very brief summary: (i) Controls should be added in the two credit extensions equations in order to properly interpret the coefficient on changes in the CAR. (ii) We recommend replacing the long-term spread on outstanding mortgages with the spread on new mortgage loans. (iii) The two interest rate spreads should be treated as credit conditions indicators in their own right, and hence as drivers of the equations for credit extension. (iv) The average loan-to-value ratio for mortgages needs to be modelled explicitly. (The details are given in Section 7.)

1. Introduction

The SARB maintains two econometric models of the economy. One is the Quarterly Projection Model (QPM), described in Botha *et al.* (2017). This model assumes that the economy evolves around an underlying, well-defined equilibrium path, which cannot be influenced by monetary policy. The model type is a 'gap model' and focuses on four important deviations from the equilibrium path: the output gap; the real exchange rate gap; the real interest rate gap; and the inflation gap (defined as deviation of inflation from target)[77]. The model is designed to explain the main elements of inflation dynamics in South Africa, but it has a relatively simplistic account of the drivers of aggregate demand. Of key relevance to the objective of this paper, there is *no explicit role* for housing and credit. As the SARB acknowledges, such a highly aggregated model needs to be complemented by more disaggregated models that help explain the components of aggregate demand.

The second of the SARB's models is such a disaggregated model, the 'Core' econometric model. The revised Core model includes important extensions to the 2007 version, and the new Macro-prudential (MPRU)-Core model now has a banking sector (De Jager *et al.*, 2021). Links have been strengthened between credit extension by banks and aggregate demand, with recent credit growth appearing in several equations, for example, in equations for consumption and investment. There are links from share prices and house prices to the gross assets of households, in new equations for these two variables. Household net worth is measured as gross assets, i.e., financial assets plus physical assets (mainly housing wealth), minus household debt. There are now separate equations for gross assets and for household debt. The revised model does not distinguish between mortgage debt and non-mortgage debt for households. As part of the banking sector equations needed to model bank balances sheets, there are new equations for long-term and short-term credit extension to the private sector, the former mainly comprised of mortgages.

[77] Gap models of this kind were developed at the Bank of Canada in the 1990s. They have been widely used at other central banks and by the IMF. Survey-based inflation expectations play an important role in the model and the formal analysis incorporates implicit expectations of output, the exchange rate and short-term interest rates.

Although progress has been made to articulate the credit channel, the current version of the SARB's MPRU-Core econometric model needs further development to better capture the channels of direct and indirect transmission of monetary policy. In the financial stability context, such model development would also improve the modelling of the macroeconomic transmission of shocks and potential macro-prudential policy decisions, and the resulting feedbacks. Integrating macroeconomic feedbacks in the design of stress tests of the financial system is high on the agenda of financial regulators such as the ECB. Improving the household and real estate sectors of the MPRU-Core model would make an important contribution to the analysis of risks to financial stability and the design of policy responses.

We have a detailed discussion below, in Sections 2 to 6, of how to improve the treatment of the household sector, drawing on international literature. Real estate plays a crucial role in the monetary transmission mechanism in South Africa, but this is only weakly captured by both Core and MPRU-Core models, which likely therefore misrepresents the timing and profile of monetary transmission. There is almost no role for credit conditions and no role for house price expectations. Income expectations are a potential additional channel through which interest rates and asset prices can affect expenditure decisions. We propose a forward-looking approach to incorporate income expectations through modelling permanent income (currently not in the model). In addition to income expectations, the consumption equation should include credit conditions, and relax the net worth restriction on household wealth to capture better the impacts of housing wealth and of debt. The house price equation should address three key omissions to improve the long-run solution: incorporating the supply side and bringing in an explicit role for credit conditions and house price expectations. House prices transmit strongly into mortgage debt and residential investment. We suggest the single aggregate household debt equation be replaced by equations for both mortgage and non-mortgage debt, to bring in house prices and credit conditions, explicitly to capture not just an interest rate effect but also shifts in credit availability. A residential investment (construction) equation (currently absent) could straightforwardly be included. We find empirically that housing market participants extrapolate past house price changes over several years in forming expectations of appreciation (as in the US). This can lead to the overshooting of house prices, as observed in 2007-8 (currently not captured in the model), and of mortgage debt, followed by painful corrections.

Such an improved model would be a useful framework for discussion at both Monetary Policy and Financial Stability Committees, and for interactions between them. In Section 7, we discuss aspects of the banking sector model to enhance the characterisation of links between bank balance sheets, regulatory and macro-prudential policy, and consequences for credit extension, housing markets and consumption.

The critique of aspects of the Core model in this paper does not necessarily mean that the MPC lacks awareness of the channels of transmission involving housing. Every Quarterly Bulletin, Inflation Report and Financial Stability Review discusses housing and mortgage market developments, including, recently, surveys of lending attitudes by banks[78], which are indicators of credit conditions. But improving the output of the Core model regarding monetary transmission and financial stability will strengthen the analytical hand of the MPC, and indeed of the Financial Policy Committee (FPC) that is responsible for macro-prudential policy decisions.

2. Improving the Consumption Function

Consumer expenditure in South Africa is of the order of 60% of GDP. As such it goes without saying that using the best practices internationally to model consumption is important for understanding both monetary transmission and financial stability. Comprehensive surveys of an older literature on consumption functions include Muellbauer and Lattimore (1995), and Muellbauer (1994) for a less technical account. Cooper and Dynan (2016) survey the literature on wealth effects in consumption functions and Muellbauer (2020) critically assesses consumption functions in the current policy models of major central banks.

The SARB's consumption equation in the Core model could benefit from several improvements. Constructive additions would be to introduce forward-looking income expectations, as do several central bank policy models, to disaggregate net worth into several different measures of wealth, including housing wealth, and to introduce the effects of changing credit conditions, which is especially important for financial stability.

[78] For the survey see: https://www.coefs.org.za/research/working-papers/south-african-bank-lending-practices-survey-wp-2018-03/

2.1 Some Specification Issues for a Consumption Equation

2.1.1 Theory Background to the Consumption and Model

The basic, aggregate, life-cycle/permanent income consumption function of Friedman-Ando-Modigliani has the form:

$$c_t = \gamma^* A_{t-1} + \omega^* y_t^P \tag{1}$$

where real, per capita consumption, c, depends on permanent, real, per capita, non-property income,[79] y^p, and the real, per capita level of net wealth, A, and γ^* and ω^* are parameters. Permanent income, y^p, is defined as the constant flow of income that corresponds to the present value of expected future income streams. Equation (1) captures in a specific form a basic comprehension of life-cycle budget constraints. A household wanting to sustain consumption will realize that not all of its assets can be spent now without damaging future consumption, and that future income has a bearing on sustainable consumption. Estimating this consumption function requires devising and estimating an income forecasting model to generate permanent non-property income, see discussion in Section 3.

Since consumption and income tend to grow exponentially, formulating the consumption function in logs has advantages. The log approximation of equation (1) is:[80]

$$\ln c_t = \alpha_0 + \ln y_t + \gamma A_{t-1}/y_t + \ln\left(y_t^P/y_t\right) \tag{2}$$

where $\gamma = \gamma^* / \omega^*$ and $\alpha_0 = \log \omega^*$[81]. The log ratio of permanent to current income

[79] Non-property income is the relevant income concept in highly stylised text-book life-cycle models where property income is defined by the rate of return on the single asset assumed in such models, and the asset level is a choice variable. Non-property (labour plus transfer) income, omits dividends and interest earned on wealth. One can therefore interpret equation (1) to say that consumption depends on permanent non-property income and on permanent property income proportional to the net wealth term.

[80] See Aron *et al.* (2012).

[81] One important advantage of equation (2) is that it avoids the log assets formulation employed in many studies of consumption. The log formulation gives a poor approximation of the marginal propensity to consume out of assets when asset levels are low, as they are for many households, especially in emerging economies. It is also a poor approximation when disaggregating net worth into several components since the log function is not additive.

$\ln\left(y_t^p / y_t\right)$ reflects expectations of income growth.

A dynamic specification of the static form, for instance to introduce habits or adjustment costs, implies a partial adjustment form of equation (2). If real interest rates are variable, by standard consumption theory, the real interest rate r_t enters the model with the usual interpretation of inter-temporal substitution and income effects. The model can be extended to include a measure of income uncertainty, θ_t. These considerations suggest the following generalisation of the canonical permanent income model of consumption in equation (2):

$$\Delta \ln c_t \approx \lambda(\alpha_0 + \alpha_1 r_t + \alpha_2 \theta_t + \ln y_t + \alpha_3 E_t \ln\left(y_t^p / y_t\right) + \gamma A_{t-1} / y_t - \ln c_{t-1}) + \varepsilon_t \quad (3)$$

where λ measures the speed of adjustment of consumption to its long-run equilibrium level.

A first modification relaxes the present value formulation of permanent income, to allow for uncertainty concerning future income and liquidity constraint, reflected in a higher discount rate than a market real rate of interest. In practice, with aggregate data it is difficult to forecast income beyond about three years except by reversion to a trend. Shorter horizons are suggested if households anticipate future credit constraints, according to the buffer-stock theory of saving explained in Deaton (1991). Precautionary behaviour also generates buffer-stock saving, as in Carroll (2001a,b), where it is argued that plausible calibrations of micro-behaviour can give a practical income forecasting horizon of about three years. This horizon was originally suggested by Friedman in his application of the permanent income hypothesis to aggregate consumption data.

A second important modification is that the formulation of aggregate assets, A, in equation (3) needs to be split up into liquid and illiquid types of assets, each with different 'spendibilities', i.e., allowing different weights for the different types of assets. There are several reasons that strongly support allowing this disaggregation of assets in empirical models of consumption. Housing wealth differs fundamentally from financial assets since it gives shelter (i.e., it has utility value) as well as having an asset value. Moreover, with credit constraints, housing wealth has a vital collateral role, see Muellbauer (2007) or Aron *et al.* (2022a) for further discussion. A third reason is that illiquid financial assets, which are subject to asset price volatility, and pensions, also subject to trading restrictions, have different and weaker effects on consumption from liquid financial assets[82] and debt. Muellbauer (2020) notes that the great majority of central bank policy models retain the net worth restriction, which ignores these differences between the various balance sheet components.

A third modification is to address the fact that variations in household access to credit may potentially induce time variation in key parameters of the consumption function. Because of asymmetric information, lenders use screening devices such as credit scores and evidence of borrowers' income, and, for secured lending, especially for housing, collateral requirements, to reduce the risk of bad loans. As their willingness to lend increases, given changes in their capital base, the cost of funds, industry structure and regulatory constraints, lenders tend to relax the stringency of their lending conditions, with a corresponding impact on household demand, including for housing, and hence on house prices. This is why variations in credit conditions need to be controlled for in specifying the household sector in policy models, though rarely included in central bank policy models, as pointed out by Muellbauer (2020).

These considerations suggest the following 'credit-augmented' version of the Friedman-Ando-Modigliani consumption function:

$$\begin{aligned}\Delta \ln c_t \approx \lambda\,(&\alpha_{0t} + \alpha_{1t} r_t + \alpha_{2t}\theta_t + \alpha_{3t} E_t \ln\left(y_t^p / y_t\right) + \gamma_1\, NLA_{t-1}/y_t \\ &+ \gamma_2\, IFA_{t-1}/y_t + \gamma_{3t}\, HA_{t-1}/y_t + \ln y_t - \ln c_{t-1}) \\ &+ \beta_{1t}\Delta \ln y_t + \beta_{2t}\Delta nr_t\left(DB_{t-1}/y_t\right) + \beta_{3t}\Delta\theta_t + \varepsilon_t\end{aligned} \tag{4}$$

The speed of adjustment is given by λ, and the γ parameters measure the marginal propensity to consume (*mpc*) for each of three types of assets. The net worth to income ratio has been disaggregated into liquid and illiquid elements: *NLA/y* is the ratio of liquid assets minus debt to non-property income[83], *IFA/y* is the ratio of illiquid financial assets

[82] Otsuka (2006) has formalised a model in which trading costs for illiquid assets imply a higher 'spendibility' for liquid assets.
[83] This could be further split into separate ratios to income of liquid assets and debt.

to non-property income, and *HA/y* is the ratio of housing wealth to non-property income, all in real terms. The term, $\Delta nr_t\left(DB_{t-1}/y_t\right)$ measures the cash flow impact on indebted households from changes in nominal rates, where *nr* is the nominal interest rate on debt, *DB*.[84] The evidence from several countries is that the change in the unemployment rate is a good proxy for income uncertainty, θ_t , or for a shift in income uncertainty. The term in the log change of current income allows for the empirical possibility that some households' spending growth follows current income growth more closely than is implied by equation (2). This could be the result of some households taking current income growth as an indicator for their future income growth. Equation (4) embodies the most basic life-cycle model (i.e., equation (2)) as a special case[85]. Finally, the time variation in some of the parameters is captured by their time subscripts, and is induced by shifts in credit availability, as discussed below.

The credit channel for monetary transmission is reflected in the consumption function through the different *mpc*s for net liquid assets, housing and illiquid assets; through the cash flow effect for borrowers via nominal interest rates; and, by allowing for possible parameter shifts in several variables stemming from credit market liberalisation. Credit market liberalisation potentially should: (i) raise the intercept α_0 , which implies a higher level of $\ln\left(c/y\right)$, mainly because of reduced required saving for a housing down-payment – a direct effect of liberalisation; (ii) make the real interest rate coefficient, α_1 , more negative as scope for inter-temporal substitution of consumption rises; (iii) should lower α_2 and β_3 on the uncertainty effects, because easier credit reduces concerns with income uncertainty, though higher debt levels could cancel this tendency; (iv) raise α_3 by increasing the scope for the impact of expected income growth by relaxing the borrowing constraint; (v) increase the *mpc* from housing wealth, γ_3 , given the greater access to home equity loans; (vi) lower the current income growth effect, β_1 , because there will be fewer credit-constrained households depending mainly

[84] Recent research has highlighted the importance of the cash-flow channel in monetary transmission in floating interest rate environments, see Muellbauer (2020), p.516-517, for a review. The evidence is that heavily-indebted households, faced with a drain on their cash-flows because of higher debt-service costs, reduce their spending by more than savers increase theirs, in response to higher interest income on their liquid deposits.

[85] Note that $\lambda = 1,\ \alpha_{1t} = \alpha_{2t} = 0,\ \gamma_1 = \gamma_2 = \gamma_{3t},\ \beta_{1t} = \beta_{2t} = \beta_{3t} = 0$ and $\alpha_{3t} = 1$ are the restrictions which result in equation (2). Equation (4) also encompasses (i.e., is more general than, but has as a special case) equation (3).

on their current income; and (vii) lower the cash flow impact, β_2 , of a change in the nominal rate since refinancing might become easier.

With a measurable indicator of the degree of credit market liberality, a credit conditions index (*CCI*), it would ideally be possible to make each potentially time-varying parameter a linear function of the *CCI and* test these hypotheses about time variation. However, in practical applications only a few of these interaction effects are likely to be empirically identifiable. It is possible that the financial conditions index (FCI), available since 2000, and used by the SARB in monitoring risks to financial stability, might have useful information content for measuring credit conditions[86].

Finally, equation (4) satisfies long-run homogeneity in income and assets: that is, doubling both, doubles consumption. The long run coefficient on *ln y* is set to 1, as in equation (2), and hence it is not being estimated. Then the income endogeneity issues highlighted in Hall (1978) cease to be of concern for the measurement of the long-run income effects. Concerning the asset to income ratios, these are dominated by the movements of volatile, lagged asset prices, so that the endogeneity of income is in practice largely irrelevant. The change in log income, $\Delta \ln y_t$, will be endogenous, and may be estimated with a slight bias but with little impact on the long-run solution.

2.1.2 An Estimated Consumption Equation for South Africa

The most complete consumption function corresponding to the above theoretical developments, is found in Aron and Muellbauer (2013)[87], analysing quarterly data for 1971-2005. We used the balance sheet estimates of disaggregated wealth data developed in Aron and Muellbauer (2006) and Aron *et al.* (2006, 2008), which work was later adopted and adapted for ongoing used by the SARB[88].

There were two data challenges. The first concerned the theoretically-preferred measurement of current income, y , in South Africa. The theoretical measure of real,

---

86 This hypothesis can be tested empirically by checking for financial conditions index (FCI) effects in the consumption function, house price and mortgage stock equations.

87 An earlier version of a model of this kind for South Africa was reported in Aron and Muellbauer (2000a), with the consumption equation for the household sector being reported in greater detail in Aron and Muellbauer (2000b).

88 Johan Prinsloo, formerly Head of National Accounts at the SARB, and who worked with us, was important in realising the adoption by the SARB of our household balance sheet estimates for use in the Core model.

per capita, non-property income measure, y , consists of tax-adjusted income from paid and self-employment, and transfers from the government. Matching theoretical concepts with the National Accounts in practice can be difficult[89]. A second measurement issue concerned developing a proxy for the change in the unemployment rate, a possible indicator for $\Delta\theta$, the measure of changes in uncertainty. South African data on the unemployment rate are considered unreliable. The rate of growth of e*mployment* proved a useful alternative proxy (with the opposite sign).

We now summarise key empirical findings from Table 3 of Aron and Muellbauer (2013), corresponding to equation (7) of that paper, a specific empirical version of our equation (4) above. The speed of adjustment, λ, was estimated at 0.45 per quarter, and strongly significant. In the long-run solution, there was powerful evidence that the long-run ratio of consumption to income increased with greater access to credit. There was a strongly significant negative effect of the real prime rate, measured as a four-quarter moving average. There was a significant effect of the fitted log ratio of permanent to current income, capturing income expectations, with evidence that income expectations became more important with the easing of credit conditions. Estimates of the coefficient on the ratio of net liquid assets ranged from 0.11 to 0.16 for different samples, while those for the coefficient on illiquid financial assets (measured as a four-quarter moving average) ranged from 0.022 to 0.028. The implication is that liquid assets are far more 'spendable' than illiquid assets and debt has far more negative effects on consumption than the restrictive the net worth formulation would have implied. The coefficient on the housing wealth to income ratio was found to be zero at the lowest level of credit conditions, but close to 0.1 at the peak of credit access. Thus, this evidence suggests that housing wealth does not act like a 'classical' wealth effect as in equation (1), but strongly supports the collateral interpretation of the 'housing wealth effect'[90]. Apart from this housing wealth interaction effect, and the interaction effect with log permanent to current income, none of the possible interaction effects in the long-run solution shown in equation (4) were found to be significant.

[89] Tax-adjusted measures of non-property and of property income are not directly available in National Accounts data, see Aron and Muellbauer (2013) on possible approximations.

[90] The collateral interpretation says: the effect of higher housing wealth is to increase consumption by allowing more borrowing through increased collateral and also equity withdrawal from housing wealth, see Aron *et al.* (2012) and Berger *et al.* (2018). The latter present an optimising model of a household facing collateral constraints and lumpy transactions costs, with a collateral effect of house prices on consumption, and where the size of the effect increases as the down payment constraint is relaxed. This implies that the housing wealth effect on consumption varies with credit conditions.

In the short-term dynamics, the annual growth rate of employment (+), the rate of inflation (-), and the lagged change in log consumption (-), were found relevant, where signs are given in parentheses. The interpretation of employment growth is as a proxy for income uncertainty, as noted above. The rate of inflation could also be interpreted as a proxy for income uncertainty as a high rate of inflation tends to be associated with greater volatility of real income. Finally, the negative coefficient on lagged consumption growth indicates a correction for 'over-spending' in the previous quarter, which can arise especially for durable goods, where a recent acquisition reduces replacement demand.

A crucial aspect to estimating this consumption function is to obtain an estimate of the credit conditions index CCI. This was obtained by joint estimation of an equation for consumption and total household debt, using dummy variables, selected with prior restrictions on periods when documented episodes of credit liberalisation occurred, e.g., see Figure 1. The estimated index in two variants is shown in Figure 1. As noted above, CCI has both an intercept effect, shifting up the average propensity to consume, and interaction effects (especially with the housing wealth to income ratio).

2.2. How the SARB treats Consumption in the Core Model

The South African Reserve Bank's Core forecasting model, see Smal *et al.* (2007) for an earlier version, uses an equilibrium correction model linking log consumption with log personal disposable income, log net worth and the real interest rate, using data from 1985 to 2005. This is an important advance on earlier models for South African consumption, which all omitted the role of assets. However, the (commonly-made) assumption that all components of wealth have the same effect on consumption – implicit in the restrictive net worth measure – runs counter to economic theory. As discussed above, housing is a consumption good as well as an asset. Thus, inter-temporal consumption theory implies that a rise in house prices, unlike a rise in the stock market prices, has both an income and substitution effect, and a wealth effect on consumption, see Aron *et al.* (2012). Moreover, liquid assets are necessarily more spendable than, say, pension wealth.

However, the consumption function in the Core model does poorly in capturing the impact of household debt on consumption and the time-varying impact of credit conditions on consumption via the housing collateral channel. The *long-run solution* for consumption has several short-comings. It does not consider permanent income explicitly, nor does it allow for shifts in credit conditions. It uses the *aggregate concept* of net worth as the only way in which household balance sheets and asset prices can affect consumption[91]. The net worth restriction implausibly implies that there will be *identical* effects on consumption of a 100 Rand increase in liquid assets, illiquid financial assets (such as pensions) and housing wealth, and of a 100 Rand decrease in debt. But it is very important to separate out these effects, as we have argued above. The estimated speed of adjustment in the equation to the long-run equilibrium, after short- to medium-run perturbations, is very low, at 0.11 per quarter. A low speed of adjustment is a typical symptom of omitted variables in the model, and we have indicated the key omissions above. Indeed, the estimated speed of adjustment is a quarter of that estimated in the consumption function for South Africa in Aron and Muellbauer (2013), which uses a comprehensive equilibrium correction model and disaggregates net worth. As well as disaggregating household balance sheets, the consumption analysis in Aron and Muellbauer (2013) incorporates a credit conditions index, which influences the effect of housing wealth on consumption, and includes a role for income expectations via permanent income.

2.3. Our Proposals to Improve Modelling of Consumption

Monetary policy transmission is stronger in the Aron-Muellbauer (2013) consumption function than in the MPRU-Core or Core models. Not only is there a *direct* effect of interest rates on consumption, also present in the Core models but a *strong, indirect* effects via housing wealth, illiquid financial assets and permanent income. The fact that the indirect interest rate effect via housing wealth increases when credit conditions are loose has the policy implication that it is particularly dangerous to lower interest rates in a credit boom. Since the effect of interest rates is so large, when there is a crisis, relaxing monetary policy has powerful effects. Moreover, its effects feed through relatively quickly, given a high speed of adjustment. The Aron-Muellbauer (2013) consumption function shows considerable heterogeneity for the components of net worth: it implies around -0.11 to -0.16 for debt, and, at the peak of credit availability

---

91 The model does incorporate the lagged rate of change of private credit in the *short-term* dynamics, which brings in some influence of credit conditions. However, this will not capture longer-term shifts in the supply of credit and will not differentiate credit demand side from supply side influences.

(since it varies with credit conditions), almost 0.1 for housing wealth. In a boom, housing wealth rises strongly, but so does household debt. This makes the household sector vulnerable to a contraction of credit conditions and fall in house prices. The equation makes this vulnerability clear, which is of considerable importance to financial stability policy.

By contrast, in the SARB's Core model, the indirect interest rate effect operates through total net worth. Net worth in the model depends on changes in the JSE stock market index and in house prices, rather than on levels of these variables. There are no interest rate effects in the model for the JSE index and the indirect interest rate effect via house prices is only transitory and does not vary with credit conditions. Since, in the Core model, the marginal propensity to spend out of net worth is around 0.04, this implies coefficients of -0.04 for debt and +0.04 for housing wealth. This understates the potential vulnerability of the household sector to a contraction of credit conditions and fall in house prices, especially as consumption reacts so slowly to shocks.

The 2013 Aron-Muellbauer model uses a single credit conditions index, CCI, applying to *total* household debt and to consumption. However, the evidence from Chauvin and Muellbauer (2018) and Geiger *et al.* (2016) is that the effects of shifts in credit conditions on non-mortgage debt and mortgage debt are different. Two CCIs are needed: one for mortgage debt and one for non-mortgage debt. Moreover, they find that the effects of both these CCIs on consumption are different. South Africa is likely to resemble France and Germany in this respect. Figure 2 plots the ratios to household disposable income of the two types of household debt. It appears that, as in France, there was a greater easing in the 1980s of credit conditions in the non-mortgage debt market than for mortgages.

This suggests modelling the two types of debt separately. Chauvin and Muellbauer (2019) and Geiger *et al.* (2016) apply the 'latent variable' method[92] to estimate separate credit conditions indices for each type of debt in a six-equation model for: consumption, both types of debt, liquid assets, house prices and permanent income.

[92] This is the 'Latent Interactive Variable Equation System (LIVES), see Duca and Muellbauer (2013). Here the 'latent variable' is function of dummy variables, which appears in multiple behavioural equations. Known changes in financial architecture and regulation provide priors that influence the selection of dummies.

In any case, separate models are needed for the two types of debt as other drivers differ for them too, and these have different implications for financial stability.

In principle, a similar technique could be applied to South African data. However, there may be a simpler alternative. In our new paper on models for house prices, mortgage debt and residential construction, Aron and Muellbauer (2022a), we find that two indicators of mortgage credit conditions, available back to 2000, are important. These are the mortgage spread defined as the difference between the prime rate and the average rate on *new* mortgage loans, see Figure 3 and the loan-to-value ratio on new mortgage loans, derived from Deeds Office transactions data, see Figure 4. For non-mortgage debt, or more generally, short-term loans to households, data on the spread between prime (or repo) rate and the average rate on new loans may well encapsulate useful information on credit conditions in the market for non-mortgage debt, i.e., credit cards, overdrafts, personal loans and loans for cars and other consumer durables. An empirical model for non-mortgage debt would confirm whether such an indicator has a plausible time profile. Unfortunately, such data are available only from 2014. However, it may well be that the data from banks on the average short-term lending rate on outstanding loans, which go back to 1996, could be used to construct a spread measure useful for capturing shifts in credit conditions for non-mortgage debt. For mortgage debt, the spread between the average rate on *outstanding* loans and prime, currently used in the MPRU-Core model, is less satisfactory than the spread on new loans. Average rates on outstanding mortgages tend to lag behind rates on new mortgages and are therefore less 'current'. Moreover, as interest receipts on outstanding loans fall when payment arrears rise, the spread defined on outstanding loans perversely would tend to fall too, should credit conditions tighten in response to worsening payment defaults. It is therefore far less satisfactory as an indicator of credit conditions in the mortgage market.

Rather than follow the latent variable approach in Aron and Muellbauer (2013), it should be possible to estimate a single equation for the consumption function back to around 2001, using the above empirical indicators of credit availability to track shifts in the average propensity to consume out of income. These indicators for credit conditions in the mortgage market should potentially be interacted with other variables as in equation (4). If interactions with the housing wealth term prove significant, this would confirm the findings in Aron and Muellbauer (2013) for South Africa and in Aron *et al.* (2012) for the US and UK. In the house price and credit boom of 2005 to 2008, when easier credit fed into higher house prices, the consumption impact of house prices was *amplified* by the easy mortgage credit availability of that period. Given that consumption comprises over 60% of aggregate demand, this aspect of the financial accelerator would help explain the strength of the economic boom of that time, as well as its sensitivity to conditions in credit and housing markets. In the GFC, the sudden contraction of credit conditions probably helps explain the sharp decline in the consumption to income ratio (after a prior steep rise during the boom period), see Figure 5.

However, in such a relatively short sample, estimating from 2001, it may be hard to obtain precise empirical results. There could be a case for calibrating some of the coefficients on the different wealth components, particularly as evidence from other countries supports findings reported in Aron and Muellbauer (2013). The evidence suggests a coefficient of the order of 0.12 to 0.15 on the ratio of net liquid assets (liquid assets minus debt) to income; of the order of 0.02 to 0.025 on the ratio to income of illiquid financial assets; and variable coefficients ranging from 0.04 to 0.09, depending on credit conditions, on the ratio to income of housing wealth. This would be preferable to assuming that only the ratio to income of net worth matters, though it would be simple to test by comparing the fit of the alternative formulations.

Another simplification of the Aron-Muellbauer approach would be to use conventional disposable household income instead of trying to approximate the non-property income concept called for by highly stylised textbook theory. At the Fed, the FRB-US model develops the permanent income concept for three different types of income, including property income, all of which are relevant for consumption. Such a complication would be out of place in the context of South Africa, given the need to keep the size of the Core models at a reasonable level.

3. Introducing a Permanent Income Forecasting equation

Forward-looking expectations are not modelled in the Core model. Income expectations are a potential additional channel through which interest rates and asset prices can affect expenditure decisions. We propose a forward-looking approach to incorporate income expectations through modelling permanent income, which is a particularly important component of the consumption function.

3.1. Some Specification Issues for an Income Forecasting Equation

The difference between log permanent income and log current income, $\ln\left(y_t^p/y_t\right)$, used in equations (2) to (4) is effectively a rate of expected income growth. This ratio can be closely approximated by an expression in logs of expected future non-property income (or alternatively, of conventional disposable income, in the simplification suggested above)[93]:

$$\ln\left(y_t^p/y_t\right) = \left(\sum_{s=1}^{k} \delta^{s-1} E_t \ln y_{t+s}\right) \Big/ \left(\sum_{s=1}^{k} \delta^{s-1}\right) - \ln y_t \quad (5)$$

Permanent income, defined as the constant amount of real income that corresponds to the present discounted value of expected future income streams, in the absence of uncertainty will use a discount factor, δ, of 1/(1 + r), where the discount r is a real rate of interest for an horizon k. In the presence of income uncertainty and liquidity constraints, however, the discount rate, r, will be higher because less weight is placed on the more uncertain distant future.

Muellbauer and Lattimore (1995) argue that a discount rate of the order of 20% per annum is appropriate, given income uncertainty and liquidity constraints, based on micro-evidence[94]. On quarterly data, where the discount rate then is 5%, this suggests that the quarterly δ might be 0.95.

There are several ways of generating forecasts of permanent income. The Federal Reserve model, FRB-US, offers two options: model-consistent forecasts, and forecasts from a small VAR for income and its drivers. The former is very complex to implement and fits less well than the latter, as reported in Brayton *et al.* (1997). A third alternative is the one adopted in Aron *et al.* (2012), Aron and Muellbauer (2013), Chauvin and Muellbauer (2019), and several other papers. This involves the direct estimation of a

---

93 Equation (5) is also equivalent to a weighted moving average of forward-looking income growth rates.

94 As well as being used by FRB-US model of the Federal Reserve, such a high discount rate is consistent with the empirical micro-estimates by Hausman (1979) and Warner and Pleeter (2001) of discounts for future cash-flows used in practical household decision-making.

reduced form equation for $\ln\left(y_t^p / y_t\right)$. One uses actual values of income up to ten years ahead to construct y^p, to generate the dependent variable in this reduced form equation. When using a quarterly discount factor of 0.95, a ten-year horizon gives an adequate approximation for the income horizon used in the construction of permanent income[95].

The disadvantage is that with a ten-year horizon one needs income forecast data that extend 10 years beyond the estimation period for an up-to-date estimate of the parameters of the reduced form. For example, in 2021, the sample period would otherwise have to end in 2011. This would be less of a problem is there were no structural breaks. However, structural breaks in the income process make it particularly desirable to use up-to-date estimates. As an alternative, income forecasts from a variety of sources such as the IMF or Oxfordeconomics.com, or indeed the SARB itself, could be substituted for the missing future income in order to estimate the equation.

3.1.1. Examples of Estimated Income Forecasting Equations

We derived a forecasting model for South Africa for the rate of growth of real, per capita, disposable, non-property income, $\log(yperm / y)$, based on equation (5), in Aron and Muellbauer (2013). Since we used a long sample of data, estimating from 1971 to 2005, the significant regime changes in South Africa during the 1980s had to be taken into account. Examples were the move to new operating procedures for monetary policy and internal financial liberalisation, both likely to have shifted monetary transmission. Political crises entailed the increasing international isolation of South Africa, reflected in diminished trade and finance, while its mineral dependency as a primary exporter gave an important role to terms of trade shocks in determining income growth. The long-run changes in productivity growth expected in an economy subject to such regime changes were captured with split trends[96]. By incorporating such shifts, the consumption function including the income growth forecasts should be robust to

---

95 With an even higher quarterly discount factor of 0.9 per quarter, a four-year horizon would be adequate.

96 In Aron and Muellbauer (2013), see Table 1, the following three split trends were used in the income forecasting equation: Split trend 1984 = zero before 1984 and 1, 2, 3, . . . thereafter; Split trend 1990 = zero before 1990 and 1, 2, 3, . . . thereafter; and Split trend 1994 = zero before 1994 and 1, 2, 3, . . . thereafter. The included split time trends reflect a slowdown beyond 1984 stemming from the 1985 debt crisis, and faster growth after the release of Nelson Mandela in 1990Q1 and the democratic elections in 1994Q2, following which capital flows increased.

the Lucas critique (Lucas, 1976).

The model took the following form:

$$\ln(yperm_t / y_t) = \alpha_0 + Split_t + \alpha_1 \ln y_t + \sum_{i=2}^{n} \alpha_i X_{i,t} + \sum_{i=1}^{n}\sum_{s=0}^{k} \beta_{is} \Delta X_{i,t-s} + \varepsilon_t \qquad (6)$$

where y_t is measured by real, per capita, disposable, non-property income; $Split_t$ are split trends reflecting the evolution of the capacity of the economy to produce and to sustain per capita personal incomes; and the X_i for i=2...n are explanatory variables (note that the dynamic terms in the explanatory variables, ΔX_i , include $\Delta ln\ y$, with a coefficient of β_1), for a maximum lag length of k[97]. The explanatory variables for the general formulation included: the level of real interest rates and changes in nominal interest rates, the government surplus to GDP ratio, capacity utilization (as a proxy for the unemployment rate), terms of trade, a measure of trade openness, the real exchange rate, the growth rate of OECD industrial production capturing trading partners, domestic credit growth, real house prices and a real stock market price index[98].

Equation (6) is in effect an equilibrium correction formulation, where the long-run solution is given by:

$$\ln yperm = -(\alpha_0 + Split \ + (1+\alpha_1)\ln y + \sum_{i=2}^{n} \alpha_i X_i) \qquad (7)$$

To construct log permanent income using equation (5), a 40-quarter horizon was adopted with a quarterly discount factor of 0.95, equivalent, as noted above, to an annual discount rate of about 20%. We used actual data on personal per capita income to 2010:Q4 and assumed a quarterly per capita income growth rate of 0.6% thereafter. The evidence showed a well-fitting parsimonious equation for the following key drivers: moving averages of the real prime rate of interest, changes in the nominal prime rate,

---

97 Note that k here is different from the horizon measure in equation 5.

98 A dummy indicator based on prescribed liquid asset requirements for commercial banks was included, see Aron and Muellbauer (2002), to capture the changing sensitivity of income growth to interest rates as the monetary policy regime changed. For samples beginning after 1986, this complication could be avoided.

the gold terms of trade and a real house price index, and split trends. For France, Chauvin and Muellbauer (2019) estimated a broadly similar model for permanent income. Key drivers, along with split trends, were moving averages of real interest rates, terms of trade as measured by the real oil price, the real exchange rate, and the real stock market share price index. With income on a per capita basis, the ratio of the working population to the total population was an important determinant. For Italy, de Bonis *et al.* (2020) find that the real interest rate, the real exchange rate, an index of competitiveness and the ratio of the labour force to the total population are key drivers. In both these papers, split trends, especially given the structural regime shift of the global financial crisis (GFC) and its aftermath, play an important role.

3.2. How the SARB treats Income Forecasting in the Core Model

Forward-looking expectations are not explicitly modelled in the Core models.

3.3. Our Proposals to Improve Modelling of Income Forecasting

We recommend that a similar approach is used to estimate equation (6) with data for real per capita household disposable income for South Africa. Estimating from after 1985 would avoid the necessity of introducing some of the parameter shifts in Aron and Muellbauer (2013). The evidence from the earlier model suggests that clear candidates for the explanatory variables are moving averages of the real prime rate, changes in nominal prime rates, the log real house price index, and a measure of the terms of trade. The terms of trade could be defined by international prices of a basket of commodity exports to international prices of a basket of imports, or an approximation to the concept. It would be important to test for a real exchange rate effect, as found for European countries. It is not obvious that this measure of competitiveness would operate in a similar way to the European countries. Negative sentiment about the medium-term economic outlook in South Africa could depreciate the exchange rate, though the short-term effects in improving competitiveness would benefit growth. Hence, the real exchange rate could present ambiguous signals for forecasting income growth. It may be worth testing forward-looking consumer confidence indices from survey data for their informational content in forecasting income growth. The downside is that they would then need to be modelled for simulation purposes.

An econometrician, with the advantage of hindsight, can, at the appropriate points in time, introduce split trends for important shifts in the growth rate of income. The ratio of permanent to current income is meant to represent the view of well-informed market participants at the time. They will typically not immediately be aware that a shift in a growth rate of income has occurred. Therefore, the fitted value of a model that incorporates that shift will not represent peoples' actual expectations at the time. A more realistic way of incorporating the shift, therefore is adopted by Chauvin and Muellbauer (2019) and de Bonis *et al.* (2020), for France and Italy, respectively, where they assume that there is a gradual learning process over two years about such shifts. This has the realistic implication that households tended to be over-optimistic before the negative trend shift in income and in income expectations induced by the GFC (see Chauvin and Muellbauer (2019) for details on how this is embodied in estimates of permanent income).

Suggestions of split trends for South Africa are a split trend from around 1994, as used by Aron and Muellbauer (2013) to track the improvement in income growth after the democratic election and the end of Apartheid; a split trend from 2009 for the GFC; another from around 2015 for the worsening economic performance associated with the Zuma presidency, electricity supply constraints and fiscal fragility; and one from 2020 because of the pandemic.

4. Improving the House Price Equation

4.1. Some Specification Issues for a House Price Equation

The inverse demand approach to deriving a house price equation is based on the idea that while the demand for housing depends on real house prices, income and other demand shifters, the housing stock is relatively fixed in the short run, while house prices are highly endogenous. This suggests inverting the demand function to make real house prices the dependent variable, driven by demand factors relative to the pre-existing, that is, the lagged, housing stock. The long pedigree of the inverse demand approach to modelling house prices, back to Kearl (1979), is traced in Duca *et al.* (2021). A separate residential construction equation, see below, explains the pre-existing housing stock as the result of cumulative investment. In our partner paper, Aron and Muellbauer (2022a), we apply these ideas, with background in the international literature, to derive new econometric models of house prices and residential investment.

The evidence from the data is for a plausible and well-fitting house price equation for South Africa based on this inverted demand principle. The key demand drivers include income relative to the housing stock, credit conditions for mortgages, interest rates, the expected house price appreciation and the rate of property taxation charged by local governments. The estimated elasticity of house prices with respect to income, given the housing stock, is 1.6. This is in line with evidence from other countries, see discussion in Aron and Muellbauer (2022a).

The estimated equation yields important insights into *monetary transmission* in South Africa. There is a powerful transmission to house prices from interest rates and credit conditions in the mortgage market. Both mortgage spreads and average loan-to-value ratios (LTVs) appear to be relevant proxies for credit conditions. House price expectations relative to mortgage rates determine the 'user cost', which is a key driver of housing demand. There is evidence of a memory of up to four years regarding the expectations of house price appreciation by housing market participants (similar to the empirical findings in US house price models, see Duca *et al.* (2021)). This implies that a series of positive shocks to housing demand feeds back positively onto housing demand and onto house prices, so extending boom conditions. This could potentially cause house prices to overshoot relative to their fundamentals, as seems to have occurred in South Africa in 2007-2008. Such overshooting has clear implications for risks to financial stability and is relevant when designing stabilisation policy.

Our evidence is of a shift in the effect of changes in the exchange rate on house prices. This suggests there may have been a weakening of capital inflows entering the housing market after 2015, which previously were driven by a momentum effect from exchange rate appreciation. Given the linkages from interest rates to the exchange rate, this suggests there may have been a shift in the effect of monetary policy on house prices in South Africa after 2015.

4.2. How the SARB treats House Prices in the MPRU-Core Model

The macro-prudential version of the Core model includes a model for real house prices. The long-run solution is driven by real per capita GDP and a real rate of interest, defined by the rate on outstanding mortgages minus the expected inflation rate two years ahead. The real rate is only just significant. In the dynamics, the current growth rate of real GDP and the lagged growth rate of long-term credit (mainly mortgages) in constant prices have powerful effects. Lagged house price appreciation, which one might have expected to be relevant as an aspect of extrapolative expectations, does not appear in the equation; nor does the LTV ratio or mortgage credit spread. Three highly significant impulse dummies are needed to explain the data in a relatively short 2001-2015 estimation period, though the speed of adjustment is a creditable 0.09. To capture the role of interest rates and credit conditions, and possible tendencies of house prices to overshoot fundamentals, much has to rest on the drivers of the stock of long-term credit but as we shall see below, these mechanisms are also lacking in the long-term credit equation of the MPRU-Core model.

4.3. Our Proposals to Improve Modelling of House Prices

Our estimated equation from Aron and Muellbauer (2022a) could easily be taken up in the Core models as a "ready-made"[99]. Two more variables would then be added to the MPRU-Core model system: the mortgage spread and the average LTV ratio. These would provide a formal link between bank balance sheets and both monetary and macro-prudential policy, consistent with the objectives of development of the MPRU-Core model to incorporate macro-prudential policy settings by De Jager *et al.* (2021). The current version of the model has an equation for the spread on mortgage rates on *outstanding* mortgages, which is explained by the capital adequacy ratio of banks and the average LTV ratio. In our house price model, the spread on *new* mortgages proves more relevant than the spread on outstanding mortgages as an indicator of current credit conditions and for driving house prices. The dominance of new lending in credit growth suggests that the spread on new mortgages is a more appropriate concept, as discussed above in Section 2.3. The MPRU-Core model also

[99] A further potential improvement would be to check for a role for income growth expectations as represented by the log ratio of permanent to current income, discussed above.

lacks an equation for the LTV ratio, though this plays an important role in the model. The LTV ratio should not be treated as exogenous or as set by policy, as explained in detail in Section 7 below. The LTV ratio should be modelled as a market outcome, driven by market conditions and the banks' balance sheets, and influenced by regulation, including macro-prudential policy settings.

5. Focusing on Mortgage Debt

5.1 Some Specification Issues for a Mortgage Debt Equation

In contrast to the vast literature on consumption, little systematic econometric work exists on household debt, either for mortgage or for non-mortgage debt, see the reviews in Fernandez-Corugedo and Muellbauer (2006) and in Meen (1990). The canonical rational expectations-life cycle model of the representative consumer has little to contribute to understanding the determination of aggregate household debt. In that model there is only a single asset, so that the life-cycle model can explain only the evolution of aggregate net wealth. In practice, consumers have multiple motives for holding debt, and these differ for mortgage debt and non-mortgage debt (consisting of credit card debt, overdrafts and personal loans and finance to acquire consumer durables such as cars and furniture). Even for mortgages, there are several potential motives. Most obviously, one motive is to acquire a roof over one's head, i.e., housing as an important consumption item. Easy access to equity withdrawal in South Africa, using housing as collateral, means that another motive for home ownership and hence acquiring a mortgage, is the buffer stock role of housing equity. Additional mortgage borrowing backed by housing collateral can support spending in the event of a short-term need for cash, for example, because of an income drop or a medical emergency. Housing is also a key component of wealth that can help support consumption in retirement. These multiple motives suggest that no simple theoretical model can adequately explain the demand for mortgages. Nor is it entirely clear what the impact of income growth expectations should be, as the consumption aspect of housing suggests a positive effect, while the saving aspect – acquiring housing as an asset – suggests the opposite.

In our partner paper, Aron and Muellbauer (2022a), we set out an eclectic model of potential determinants of mortgage demand. One of the key drivers is likely to be the level of house prices as, other things being equal, higher house prices require larger mortgages[100]. The cost of credit and credit availability- the ease of access to credit- are also obvious drivers. From a *monetary transmission* respect, the evidence suggests that for the mortgage stock, as for the house price equation, interest rates and credit conditions have powerful effects. The relative direct effects of mortgage spreads and LTVs is somewhat different on the mortgage stock than in the house price equation. For mortgages, the direct effect of the level of LTVs is greater than for house prices, while spreads have only temporary effects. However, since a key driver of the mortgage debt to income ratio is the level of house prices to income, there are large *indirect* effects of interest rates and credit conditions on the mortgage stock via the house prices to income ratio. Also, the extrapolative element of expectations of house price appreciation, embedded in the house price ratio, has a further *indirect* effect. This implies that mortgage debt, like house prices, can overshoot fundamentals. High levels of mortgage debt relative to income can thus pose risks for *financial stability,* which is relevant when designing stabilisation policy. There may also be risks of sharp downturns in consumer spending if interest rates were to rise, as mortgage debt is an important driver of consumption, see Section 2 above.

Estimates for both the house price equation and the mortgage stock equation are limited by the historical span of data on LTVs and on mortgage spreads. These begin around 2000. In particular, there is only one turning point in the series for mortgage debt to income, between 2001 and early 2020, making robust identification of parameter estimates difficult. Hence, the model for mortgage debt is necessarily provisional. Nevertheless, the model is highly consistent with evidence from other countries, see Aron and Muellbauer (2022a) for discussion.

[100] With limited access to credit, however, there is the possibility that higher house prices relative to income may lock some aspirant purchasers out of the market.

5.2 How the SARB treats Mortgage Debt in the Core Model

The new Core model includes an equation for what is mainly mortgage debt issued by banks, 'long term bank claims on the private sector'. This private sector-wide coverage is somewhat different from our measure, which comes from the household balance sheets and is therefore for households only and includes non-bank sources of finance. Nevertheless, the two series are likely to be closely correlated, so if this dependent variable is retained, then key features of our specification are likely to be relevant. In the MPRU-Core model, the long-run solution for long-term credit issued by banks is quite weakly determined, driven by the lagged long-term credit to GDP ratio, with an (only just) significant estimated coefficient, and an (insignificant) effect from the nominal effective rate of interest on outstanding long-term credit. The short-term drivers consist of growth of GDP and the M3 measure of the money stock and the lagged change in a moving average of banks' capital adequacy ratio, see Section 7 for further discussion. Total household debt is modelled as a homogenous relationship to private sector credit extension and enters the consumption function as a component of net worth. There is no equation for the mortgage debt of households as a separate item.

5.3 Our Proposals to Improve Modelling of Mortgage Debt

Whichever measure of mortgage debt is used, whether it is household mortgage debt or long-term bank credit (combining mainly household and business mortgage debt), the key elements of the equation we have provided in Aron and Muellbauer (2022a) could be used in the MPRU-Core model. As noted above, these include powerful effects from the level of house prices, the level of the real prime rate of interest and credit conditions, especially from the lagged LTV. As for the house price equation, it would be advisable to check if there is a significant effect from income growth expectations, measured by the log ratio of permanent to current income.

6. Introducing Residential Construction

6.1 Some Specification Issues for a Residential Investment Equation

An important paper by Caldera and Johansson (2013) estimates housing supply (new construction) elasticities for 21 OECD countries in a common specification. They model quarterly real residential investment and real house prices using separate log-linear equilibrium correction models in a Seeming Unrelated Regression (SUR) set-up. Long-run construction depends on real house prices, construction costs, and demography (which affect the incentive to build), and the short-run relationship includes lagged changes in these drivers. An updated version of this approach in Cavalleri *et al.* (2019) extends the country coverage and includes an equation for South Africa. It suggests a very strong long-term relationship between residential investment and the ratio of house prices to construction costs, measured by the deflator for residential construction from the National Accounts. The estimated speed of adjustment of 0.35 for South Africa confirms the solidity of this long-run finding. This equation is an excellent starting point for more detailed work.

In Aron and Muellbauer (2022a), we were able to find a stable relationship for data back to 1978, despite the many structural changes and shocks experienced by the South African economy. The key driver of residential investment relative to GDP is the relative price of houses to construction costs, with an elasticity around unity. This finding is consistent with international evidence, as well as that for South Africa, from the OECD study, Cavalleri *et al.* (2019). The implication from our model for monetary transmission is that the powerful effect of interest rate and credit conditions on house prices in South Africa also transmits to residential investment, which is a rather volatile component of aggregate demand. While there is no evidence of interest rates effects in the long-run solution for residential investment, there are powerful short-term effects of changes in the prime rate of interest. There was an apparent moderation in residential investment as population growth fell with the AIDS epidemic, and also from 2017, probably associated with worsening economic and political prospects.

The dynamic specification of the equation is quite simple. It suggests that the level of the house price index relative to construction costs four quarters earlier captures the timing of the price effects well. This is consistent with a lag of about a year between deciding to start construction and selling the completed building.

6.2 How the SARB treats Residential Investment in the Core Model

There is no equation for residential investment in the Core model for this important further link in the chain of monetary transmission. Yet residential investment is strongly influenced by house prices, providing another transmission channel for monetary policy, and typically with considerable lags. These lags potentially matter for policy because they signal that a strong rise in house prices is grounds for early tightening of policy. The *aggregate* investment equation in the MPRU-Core model includes a small role for the long government bond yield (ten-year bond), but not for the SARB's policy rate, or the mortgage interest rate, or indeed for any asset prices. The main driver is GDP, subtracting corporate tax revenue. However, while there is a short-term transmission channel via the lagged rate of growth of broad private sector credit extension, the estimated effects of the policy rate on this are actually quite moderate. Thus, if the main expenditure components of GDP are only mildly affected by the prime rate, which moves in line with the repo rate, the main instrument, then monetary policy has apparently only a *weak* effect on aggregate investment. The evidence from our own research contradicts this conclusion from the Core model, not only for the substantial residential investment channel for monetary policy via housing and related credit markets, but also for the consumption channel, and hence more generally for GDP.

6.3 Our Proposals to Improve Modelling of Residential Investment

The new residential investment equation from Aron and Muellbauer (2022a) is essentially 'oven ready' for inclusion in the SARB Core model. It does, however, require another equation, namely for the residential investment deflator. This should not be so hard to model, as one would expect wages and material costs to be the main drivers. Generally speaking, macro-econometric models find it hard to find stable relationships for aggregate private investment. It is sometimes argued that investment is driven by volatile 'animal spirits' or that aggregate investment depends on hard-to-model profit or growth expectations. It is reassuring, that at least for the residential investment component, there is well-fitting and stable relationship for South Africa. It may be the case that the removal of this component from aggregate investment will make it easier to find a coherent model for the rest of investment.

7. Improving the treatment of the Capital Adequacy Ratio (CAR) and LTVs in the new banking sector part of the model[101]

The motivation for including a banking sector in the Core model is to introduce channels through which the regulatory settings of macro-prudential policy could influence credit pricing of long-term and short-term interest rate spreads and hence credit growth. These channels might operate through the capital adequacy ratio (CAR)[102], the liquidity coverage ratio (LCR)[103], the net stable funding ratio (NSFR)[104], and, potentially, by placing ceilings on loan-to-value or debt-to-income ratios for mortgages or altering their risk weights. However, neither the LCR nor the NSFR plays an explicit role in the behavioural equations of the revised Core model. Only the capital adequacy ratio, CAR, is formally introduced as a driver in four new equations for credit extension and interest spreads. Risk weights on LTVs have not changed since 2008 and there are no LTV ceilings, but should this position change, further adaptations to the model would be required to include ceilings and changing risk weights.

In the *two new credit extension equations* in the Core model, credit extension by banks depends, *inter alia*, on *the change* in the CAR in the previous four quarters (the relevant coefficient is calibrated rather than estimated). One interpretation of the potential effect of CAR in these equations is that an increased CAR from raising safe assets and/or reducing risk-weighted assets, could be reflecting an *increase* in risk aversion by banks themselves, and hence a tightening of credit conditions with reduced credit extension (implying a negative coefficient in the equation). Other interpretations stem from various changes in the regulatory environment, but the sign, though mostly negative,

[101] This section benefitted from discussion with David Aikman, Director of the Qatar Centre for Global Banking & Finance at Kings College London.

[102] The CAR is defined as the ratio of bank capitalisation to risk-weighted assets. Bank capitalisation includes the sum of allocated, qualifying, common equity Tier 1 capital and reserve funds, additional Tier 1 capital and reserve funds, and Tier 2 capital and reserve funds. The measure of risk-weighted assets is applied to short-term and long-term bank credit extended to the private sector and other bank investments and bills, with different weights on each. In South Africa, the regulatory minimum for the CAR, MINCAR, is currently set at 10.5%.

[103] The LCR is defined as the ratio of high-quality liquid assets to short-term money balances.

[104] The NSFR is modelled as long-term money balances plus bank capitalisation plus a fraction of short-term money balances, divided by a weighted combination of long-term and short-term claims on the private sector by the banking sector.

in rare situations may be positive. However, the occurrence of loan losses can potentially weaken or even reverse the negative sign.

Hence, the possible conflation of the above various effects with potentially different signs into one coefficient, means the coefficient cannot be reliably interpreted. Controls need to be added to the equation, therefore, for regulatory changes and for loan losses.

In order to explore this question in more detail, and to suggest appropriate controls for a better interpretation of the role of the CAR, consider the following *five situations* regarding changes in the risk and lending environment.

1. A bank becomes concerned about the risk environment and voluntarily chooses a higher capital buffer. (This could also occur if the punishments for breaching the buffer become more severe.) The bank is likely to increase retained earnings and reduce risky lending. In this case, we expect the CAR and hence the buffer above the regulatory minimum, (CAR – MINCAR), to increase, leading to weaker lending and higher spreads. This fits with the negative sign of the change in CAR in the current credit extension equations of the Core model.

2. There is an unexpected regulatory change, an increase in MINCAR. One would then expect a bank to want to restore its desired buffer above the minimum and to tighten credit conditions to achieve this. We expect to observe CAR rising (perhaps with a lag) and the capital buffer, (CAR – MINCAR), instantly falling, before recovering, so that lending falls and spreads increase. This also fits with the negative sign of the change in CAR in the current version of the credit extension equations. *It would be advisable to include an explicit role for changes in MINCAR in the equations for credit extension to control for the circumstances of this case.*

3. A tightening of another regulatory buffer like the CCyB or LCR could plausibly increase the observed CAR and the CAR buffer, and lead to weaker lending and higher spreads. This is similar to cases 1 and 2, with a negative sign. *Ideally, it would be advisable to include controls for changes in other regulatory buffers (though, as noted above, CCyB and LCR are not currently reflected in the model).*

4. There are also more complex regulatory situations. Consider a crisis with much uncertainty where the regulator forces the bank to raise its capital ratio by issuing fresh equity capital (as, for example, in the US Supervisory Capital Assessment Program of 2009). Then we might expect a rise in the CAR, but that credit conditions would loosen and lending increase. This will give the *opposite* sign for the change in CAR. *This special case could be handled in the credit equations with an appropriate dummy variable*. If instead the regulator leaves the choice of how to increase the CAR to the bank, a tightening of credit conditions is more likely (as seen in the EBA stress tests during the euro crisis). Then, the situation will resemble cases 1, 2 and 3, with a negative sign.

5. Finally, there is the case of an *unexpected* increase in realised losses/write-offs. Such losses will reduce profits and hence the equity capital of the bank, and banks will respond with tighter credit conditions. The CAR will fall, as will the capital buffer, (CAR- MINCAR), followed by a reduction in lending and an increase in spreads. This means there is a *positive* association between the fall in the CAR and the fall in credit extension. *Here, an additional control could be included in the equations for credit extension, an indicator for worsening loan losses. One candidate could be the change in the ratio of impaired loans to gross outstanding loans, as published in the Financial Stability Review*[105]*.*

We are not claiming that, over available historical data, all these cases arose or can be identified. However, it is important to check the relevance of the suggested additional controls. It is also important to be aware that, in the future, some of these special situations may arise, and their effects may need to be calibrated before enough data makes estimation feasible.

Turning to the *two new spreads equations* in the MPRU-Core model, their relevance for the CAR lies in a second mechanism by which the CAR influences credit growth, namely through interest rates. The interest rates used in the new banking equation for credit extension are the effective lending rates, measured as averages of interest rates

---

[105] The Core model includes a Memo item for an equation for impaired loans measured in real terms. It does not allow for important shifts in the data in 2008, and in 2018 and beyond, due to shift from Basel I to Basel II in 2008, and the switch in the accounting standard from IFRS 39 to IFRS 9 in January 2018. This equation could certainly be improved, see Aron and Muellbauer (2022b) for discussion. This measure could potentially play a significant role in the equations for credit extension and interest rate spreads.

on outstanding loans, not as averages on new loans. These average rates on outstanding loans are lagging indicators compared to interest rates on new loans, and they have other problems as discussed in Section 2.3. It would be greatly preferable to use effective rates on new loans, available back to 2001, for long-term credit extension. For short-term credit extension, data for the effective average rates on new loans are available only from 2014, so the less preferred average measure on outstanding loans has to be used.

In the MPRU-Core model, two lending spread equations are used to connect the effective rates on outstanding loans with the prime rate, which proxies the cost of capital, and in South Africa moves one-for-one with the repo rate. The lending spread, or prime rate minus actual lending rate, can be interpreted as a risk price or a proxy for the (un)willingness of banks to take on risk. In the past 20 years, the actual mortgage rates extended have generally been below the prime rate, so that when the mortgage spread becomes less negative, or turns positive, this indicates a tightening of credit conditions.

In the MPRU-Core model, the *spread for short-term lending* depends *positively* on the *level* of CAR (expressed as a lagged four-quarter moving average) and depends negatively on the output gap. The *long-term credit spread* also depends *positively* on the *level* of CAR (expressed as a lagged four-quarter moving average) and depends negatively on a lagged four-quarter moving average of the loan-to-value ratio, LTV (averaged over all types of mortgages observed in the mortgage market). Both effects of the level of the CAR are calibrated rather than estimated. However, the expected sign of the relationship between the level of the CAR and the spreads is ambiguous and not necessarily positive, again requiring the addition of controls in these two spread equations.

The discussion of five cases above suggested that should be a negative association between *changes* in the CAR and the growth of credit, given additional controls for special circumstances. Consistent with this reasoning, a high *level* of the CAR could indicate high levels of risk aversion by lenders. On the other hand, a high *level* of the CAR above MINCAR could plausibly signal there is scope for a subsequent *reduction* of the CAR by easing credit conditions. There is thus again an ambiguity about the role of the CAR in these equations. Furthermore, using the level *exclusively,* instead of the level as well as the change, could give the wrong information, because this would miss the dynamic response to the adjustment in the CAR. Better would be to include *both* the level *and* the change of the CAR in the interest rate spread equations to control for circumstances when a high level of CAR is followed by a fall in the CAR, indicating a relaxation of credit conditions. Also, the additional controls discussed above would need to be included, as for the credit extension equations.

The spread is not treated as a credit conditions indicator in its own right in the MPRU-Core model, i.e., it is not directly used to drive the equations for credit extension. It only serves as a link between the lending rates and the prime rate. We believe this to be a significant omission and that the equations for credit extension would improve by including an explicit role for the two spreads.

Finally, we turn to the role of the average LTV ratio in the MPRU-Core model equations for the long-term mortgage spread. Here a higher level of the lagged LTV ratio reduces the spread. In the MPRU-Core model, as there is no equation for it, the *average* LTV ratio is assumed to be exogenous. This means it is implicitly treated as a policy instrument. But this cannot be correct[106].

Macro-prudential policy can set ceilings on levels of LTV ratios, and these ceilings will influence the average LTVs observed in the market, but policy cannot set the *average* LTV ratio. This is because the average LTV ratio has heterogenous components, and it can vary over time independently of macro-prudential controls. For example, repeat buyers in the housing market will have existing housing equity that is influenced by how much appreciation they have experienced since they acquired their existing home. If they have had high appreciation, they are likely only to require lower LTV ratios on their next purchase. Another example is that observed LTV ratios in South Africa are influenced by the fraction of borrowers taking advantage of the rules allowing part of a person's pension assets to be used as collateral. Then, a rise in average pension assets relative to average house prices, would be expected to increase the average LTV, where the measure of the LTV ratio does not include pension collateral. We conclude, therefore, that the average LTV ratio needs to be modelled explicitly, taking

[106] Analogously, while the SARB can *set* the repo rate, it cannot *set* the 10-year government bond yield. The latter will be *influenced* by the policy rate, but also by a host of other factors such as the international rate environment, inflation expectations, SA's credit rating and the government debt to GDP ratio.

into account factors such as lagged house price appreciation, the average pension value relative to house prices, the mortgage interest rate spread, and macro-prudential policy settings. The empirical findings of such a model will be helpful in interpreting data on the average LTV as part of monitoring potential risks to financial stability.

7.1 Summary of our Proposals to Improve Modelling of the effects of CAR

While the introduction of a banking sector with four new equations linking the CAR to credit extension and spreads is an important step forward, there are quite a few improvements to make clear the implications of how macro-prudential policy transmits to the real economy through credit extension.

Controls should be added in the two credit extensions equations in order to properly interpret the coefficient on changes in the CAR, in the form of changes in the MINCAR, for changes in other regulatory buffers like CCyB and LCR, a dummy variable for the event of the regulatory case 4 above (which has not so far occurred in South Africa), plus an indicator of bad loans, such as changes in the ratio of impaired loans to gross outstanding loans. The two interest rate spreads, after replacing the long-term spread with the spread on new loans, should be treated as credit conditions indicators in their own right, and hence as drivers of the equations for credit extension.

In the two equations for the interest rate spread, *both* the level *and* the change of the CAR should be included, to control for circumstances when a high level of CAR is followed by a fall in the CAR, indicating a relaxation of credit conditions. Also, the same controls would need to be included, as for the credit extension equations. The average LTV ratio needs to be modelled explicitly, taking into account factors such as lagged house price appreciation, the average pension value relative to house prices, the mortgage or long-term credit interest rate spread, and macro-prudential policy settings.

8. Conclusions

This paper recommends improvements to the SARB’s Core model to increase its relevance for macro-prudential stress testing and for setting monetary policy. Since the publication of the Core model in 2007, there has been further model development but no updated publication. The most recent published version, see De Jager *et al.* (2021), adds a banking sector and expands the linkages that can be influenced by macro-prudential policy between the banking system and the real economy. Our suggested model improvements apply to both versions of the Core model.

A *well-specified model for consumption* (given that it comprises about 60% of GDP) is crucial for understanding monetary transmission and financial stability. We have suggested three improvements. First, a more explicit treatment of income expectations for consumption. Second, to relax the highly restrictive 'net worth' assumption on household wealth to capture the different impacts of housing wealth, illiquid and liquid assets, and debt. And third, to introduce time-varying credit conditions into the equation.

On a more explicit treatment of expectations, we propose using a forward-looking approach to incorporate income expectations through modelling permanent income. In contrast to using the text-book concept of permanent income, which uses a very low discount rate for the future, our proposal reflects the more limited horizons of real-world households with a more realistic discount rate. The weight that households place on expected income as compared to current income needs to be estimated empirically.

On relaxing the 'net worth' assumption, which applies *equal weights* to the different components of wealth, we regard it as crucial that the very *different weights* of these components (namely, liquid assets, (minus) debt, illiquid financial asset like pensions, directly-held equity and housing wealth) should be estimated separately. For example, cash in a bank deposit is clearly more 'spendable' than a pension and hence will have a far bigger impact on consumption. Moreover, housing is a consumption good as well as an asset, and should be distinguished from financial assets.

On introducing credit conditions, these need to be controlled for in consumption functions. This is because asymmetric information between lenders and borrowers means that lenders impose collateral requirements for mortgage borrowing and use a wide range of screening devices to reduce the risk of bad loans, and these are far from constant over time.

For the *house price equation,* we suggest the long-run solution should be improved in three ways: by incorporating the supply side, bringing in an explicit role for credit conditions, and introducing house price expectations. Drawing on Aron and Muellbauer (2022a), we note evidence that South African housing market participants extrapolate past house price changes over several years in forming expectations of appreciation (as in the US). This can lead to the *overshooting* of house prices as occurred in 2007-8, and of mortgage debt, followed by painful corrections. Neither are currently captured in the model. House prices transmit strongly into mortgage debt, an effect which is missing in the current model.

Turning to the household mortgage debt equation, we suggest the current single aggregate household debt equation (driven by total bank credit extension) in the Core model be replaced by *separate equations for mortgage debt and for non-mortgage debt*, as these are driven by different factors. House prices and shifts in credit availability should be explicitly incorporated into the proposed mortgage debt equation. Interest rates feed strongly into house prices, and both direct and indirect effects of interest rates on mortgage debt are important in monetary transmission. An alternative approach could be to continue to drive household debt with long-term and short-term credit extension from banks, and incorporate house prices and shifts in credit availability, as well as interest rate effects in the bank credit extension equations.

House prices also transmit strongly to residential investment. We propose the inclusion of a *residential investment equation* (currently only aggregate investment is modelled. A candidate equation is given in in our partner paper, Aron and Muellbauer (2022a), which is remarkably stable back to the late 1970s, despite many shocks and structural changes in the economy. This equation captures a further important indirect effect of interest rates on aggregate demand via house prices, currently missing in the Core model.

The empirically-estimated equations for house prices, mortgage debt and residential investment in our partner paper, Aron and Muellbauer (2022a), which incorporate the above recommendations, could relatively straightforwardly be introduced into the Core and MPRU-Core models, better to interpret monetary transmission, with important implications also for financial stability. For the consumption function, we have suggested some simplifications of our earlier work on consumption (Aron and Muellbauer, 2013), for ease of implementation. Instead of following the latent variable approach, it should be possible to estimate a single equation for the consumption function back to around 2001, using the empirical indicators of credit availability to track shifts in the average propensity to consume out of income.

We are well aware of the trade-off between model size and the feasibility of simulating a model for scenario forecasting and for policy analysis. The new equations introduce further variables. This would require further equations for the average loan-to-value ratio, the property tax rate and the residential investment deflator, and an equation for a redefinition of mortgage rate spreads. Alternatively, assumptions would need to be made about the future trajectory of the additional variables in forecasting exercises. However, well-fitting equations with strong long-run solutions are likely to generate more robust simulations even in a somewhat larger model. Importantly, a convincing economic story would greatly enhance the usefulness of the modelling exercise for interpretation of simulations with the Core or MPRU-Core model. Examples include a better understanding of the credit cycle and of the housing market channel of monetary policy transmission.

The paper also proposes several adjustments to the four banking sector equations in the 2021 published version of the Core model. While the introduction of a banking sector with four new equations linking the CAR to credit extension and spreads is an important step forward, quite a few improvements are needed to make clear how macro-prudential policy transmits to the real economy through credit extension. These can be briefly summarised as follows (see Section 7 for details). Controls should be added in the two credit extensions equations in order to properly interpret the coefficient on changes in the CAR. We recommend replacing the long-term spread on outstanding mortgages with the spread on new mortgage loans. The two interest rate spreads should be treated as credit conditions indicators in their own right, and hence as drivers of the equations for credit extension. The average LTV ratio needs to be modelled explicitly.

We previously suggested improving the database on commercial real estate (Aron *et al.*, 2020). This, together with the close tracking of loan-to-value ratios and credit spreads in the mortgage market emphasised in this paper, and modelling the consequences of their changes for consumption, debt and investment should further enhance the understanding of risks to financial stability in South Africa.

The current nature of the risks to financial stability for South Africa, in the context of the lingering global pandemic, climate change challenges, major supply constraints, limited fiscal capacity and stagflation, are rather different from what they were in the mid-2000s. The supply side of the Core model will require enhancement to adapt to these changed current circumstances. While in the near future, the probability of a credit and house price boom on the scale that occurred in the 2000s is vanishingly small, similar dynamic processes operate in recessions as well as in booms.

Hence, if a well-designed empirical policy model has the ability to capture well the dynamic processes that operated in the past, this will help reduce the risks of misinforming policy-makers facing such an uncertain future.

References

Aron, J. and J. Muellbauer. 2000a. “Personal and Corporate Saving in South Africa,” *World Bank Economic Review*, 14(3), 509–44.

Aron, J. and J. Muellbauer. 2000b. "Personal and Corporate Saving in South Africa.” Centre for Economic Policy Research, London, Working Paper Series no. 2656, December.

Aron, J. and J. Muellbauer. 2002. “Interest Rate Effects on Output, Evidence from a GDP Forecasting Model for South Africa,” *IMF Staff Papers*, 49, 185–213, November.

Aron, J., J. Muellbauer and J. Prinsloo. 2006. “Estimating household-sector wealth in South Africa.” *Quarterly Bulletin*, South African Reserve Bank, June, 61-72.

Aron, J. and J. Muellbauer. 2006. "Estimates of Household Sector Wealth for South Africa, 1970-2003." *Review of Income and Wealth* (International Association for Research in Income and Wealth) 52: 2 285-308, June.

Aron, J., J. Muellbauer and J. Prinsloo. 2008. "Estimating the Balance Sheet of the Personal Sector in an Emerging Market Country: SA, 1975-2005." Chapter 10 in James B. Davies (ed.), *Personal Wealth from a Global Perspective*, UNU-WIDER Studies in Development Economics, Oxford: Oxford University Press, October 2008, pp. 196-223.

Aron, J. and J. Muellbauer. 2013. “Wealth, credit conditions and consumption: evidence from South Africa.” *Review of Income and Wealth,* 59 (S1): S161–S196, October.

Aron, J., G. Farrell and J. Muellbauer. 2020. “Financial Stability and the Housing Market in South Africa.” Mimeo, South African Reserve Bank.

Aron, J. and J. Muellbauer. 2022a. “New Econometric Models for South African House Prices, Mortgage Debt and Residential Investment.” Mimeo. Nuffield College, University of Oxford, February 2022.

Aron, J. and J. Muellbauer. 2022b. “Non-Performing Loans and Arrears in South Africa: a Scoping paper for Future Model Development.” Mimeo. Nuffield College, University of Oxford, February 2022.

Aron, J., J. Duca, J. Muellbauer, A. Murphy, and K. Murata. 2012. “Credit, Housing Collateral and Consumption, Evidence from the U.K., Japan and the U.S,” *Review of Income and Wealth*, 58(3), 397–423.

Attanasio, O., A. Leicester, and M. Wakefield. 2011. “Do House Prices Drive Consumption Growth? The Coincident Cycles of House Prices and Consumption in the UK,” *Journal of the European Economic Association*, 9(3), 399–435, June.

Berger, D., V. Guerrieri, G. Lorenzoni and J. Vavra. 2018. “House Prices and Consumer Spending.” *Review of Economic Studies*, 85(3), 1502-1542.

Brayton,F. Flint, E. Mauskopf, D. L. Reifschneider, P. Tinsley and J. Williams. 1997. “The role of expectations in the FRB/US macroeconomic model”, *Federal Reserve Bulletin*, 83, issue Apr, 227-245.

Caldera, A. and A. Johansson. 2013. “The Price Responsiveness of Housing Supply in OECD Countries.” *Journal of Housing Economics* 22 (3): 231-49.

Carroll, C. 2001a. “Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation),” *Advances in Macroeconomics*, 1(1), article 6, 1–32.

Carroll, C. 2001b. “A Theory of the Consumption Function with and without Liquidity Constraints,” *Journal of Economic Perspectives*, 15(3), 23–45.

Cavalleri, M.C., B. Cournède and E. Özsöğüt. 2019. “How Responsive are Housing Markets in the OECD? National Level Estimates”, OECD Economics Dept. working paper 1590.

Chauvin, V. and J. Muellbauer. 2018. “Consumption, household portfolios and the housing market in France.” *Economie et Statistique/Economics and Statistics* 500-501-502: 157-178, plus online complement.

Cooper, D. and K. Dynan. 2016. "Wealth Effects and Macroeconomic Dynamics," *Journal of Economic Surveys*, 30(1), 34-55.

de Bonis, R. D., Liberati, J. Muellbauer and C. Rondinelli. 2020. "Consumption and wealth: new evidence from Italy," Temi di discussione (Economic working papers) 1304, Bank of Italy, Economic Research and International Relations Area.

de Jager, S., R. Ehlers, K. Mojapelo and P. Pienaar. 2021. “Short-term impacts and interaction of macroprudential policy tools.” Working Paper Series WP/21/20, South African Reserve Bank

Deaton, A. 1991. “Saving and Liquidity Constraints,” *Econometrica*, 59(5), 1221–48.

Duca, J. V. and Muellbauer, J. 2013. “Tobin LIVES: Integrating Evolving Credit Market Architecture into Flow of Funds Based Macro-Models.”, ECB Working Paper No. 1581, also in B. Winkler, A. van Riet, and P. Bull (eds), *A Flow-of-Funds Perspective on the Financial Crisis Volume II—Macroeconomic Imbalances and Risks to Financial Stability*, Palgrave Macmillan, 2014.

Duca, J.V., J. Muellbauer and A. Murphy. 2021. “What Drives House Price Cycles? International Experience and Policy Issues,” *Journal of Economic Literature,* 59(3): 773-864.

Fernandez-Corugedo, E. and J. Muellbauer. 2006. "Consumer credit conditions in the United Kingdom." Bank of England Working Papers 314, Bank of England.

Geiger, F., J. Muellbauer, and M. Rupprecht. 2016. “The Housing Market, Household Portfolios, and the German Consumer.” European Central Bank Working Paper 1904, European Central Bank, Frankfurt.

Hausman, J. 1979. “Individual Discount Rates and the Purchase and Utilization of Energy Using Durables,” *Bell Journal of Economics*, 10(1), 33–54, Spring.

Kearl, J. R. 1979. ‘Inflation, Mortgages, and Housing.’ *Journal of Political Economy,* 87(5), 1115-38.

Lucas, R. E. 1976. “Econometric Policy Evaluation, a Critique,” *Carnegie-Rochester Conference Series Supplement to the Journal of Monetary Economics*, 1, 19–46.

Meen, G. P. 1990. “The Measurement of Rationing and the Treatment of Structural Change in the UK Mortgage Market,” *Journal of Applied Econometrics*, 5(2), 167–87.

Muellbauer, J. 2020. "Implications of household-level evidence for policy models: the case of macro-financial linkages," *Oxford Review of Economic Policy*, 36(3), 510-555.

Muellbauer, J. 2007. “Housing, Credit and Consumer Expenditure,” in *Housing, Housing Finance, and Monetary Policy, Proceedings of the Federal Reserve Bank of Kansas City Symposium*, Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, 26–334, 2007.

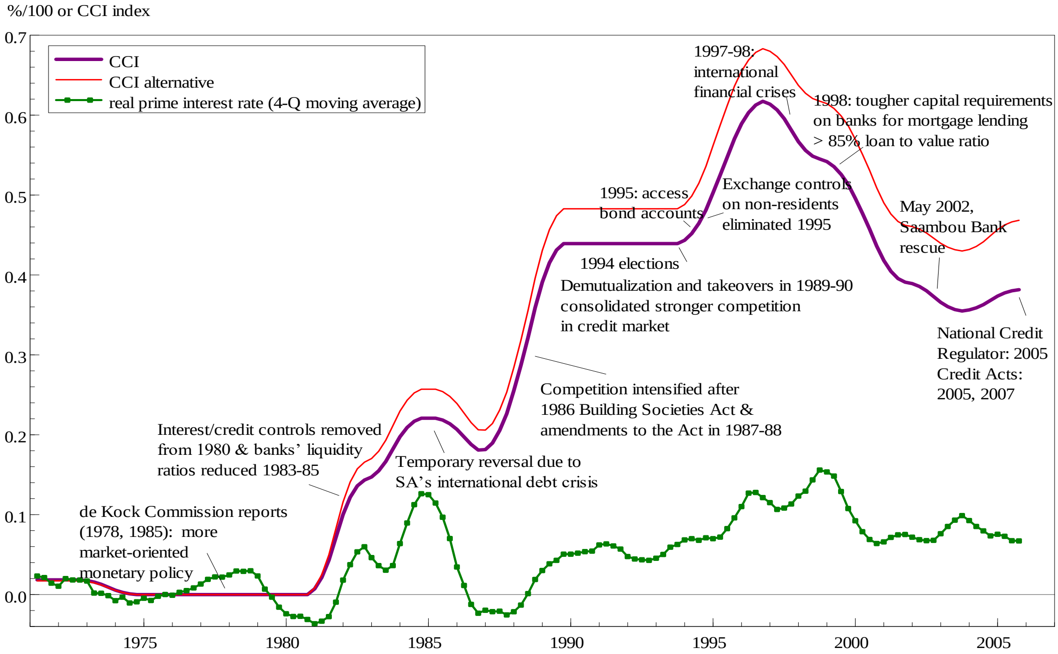
Muellbauer, J. 1994. “The Assessment: Consumer Expenditure,” *Oxford Review of Economic Policy*, 10 (2), 1–41, https://doi.org/10.1093/oxrep/10.2.1.

Muellbauer, J. and R. Lattimore. 1995. “The Consumption Function: A Theoretical and Empirical Overview.” in M.H. Pesaran and M. Wickens (eds), Handbook of Applied Econometrics, Blackwell.

Smal, D., C. Pretorius, and N. Ehlers. 2007. “The Core Forecasting Model of the South African Reserve Bank,” SARB Working Paper WP/07/02, June.

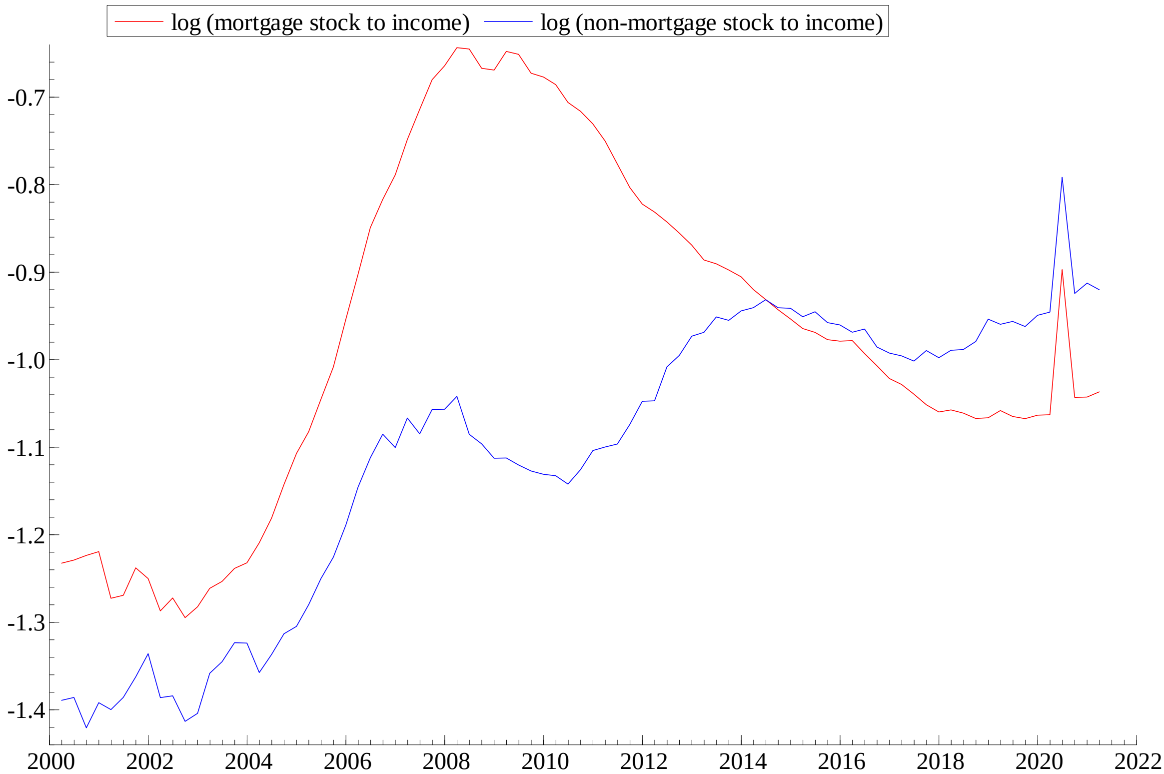
Warner, J. T. and S. Pleeter. 2001. “The Personal Discount Rate, Evidence from Military Downsizing Programs,” *American Economic Review*, 91(1), 33–53.

Figure 1: Estimated Credit Conditions Index for South Africa and the real interest rate



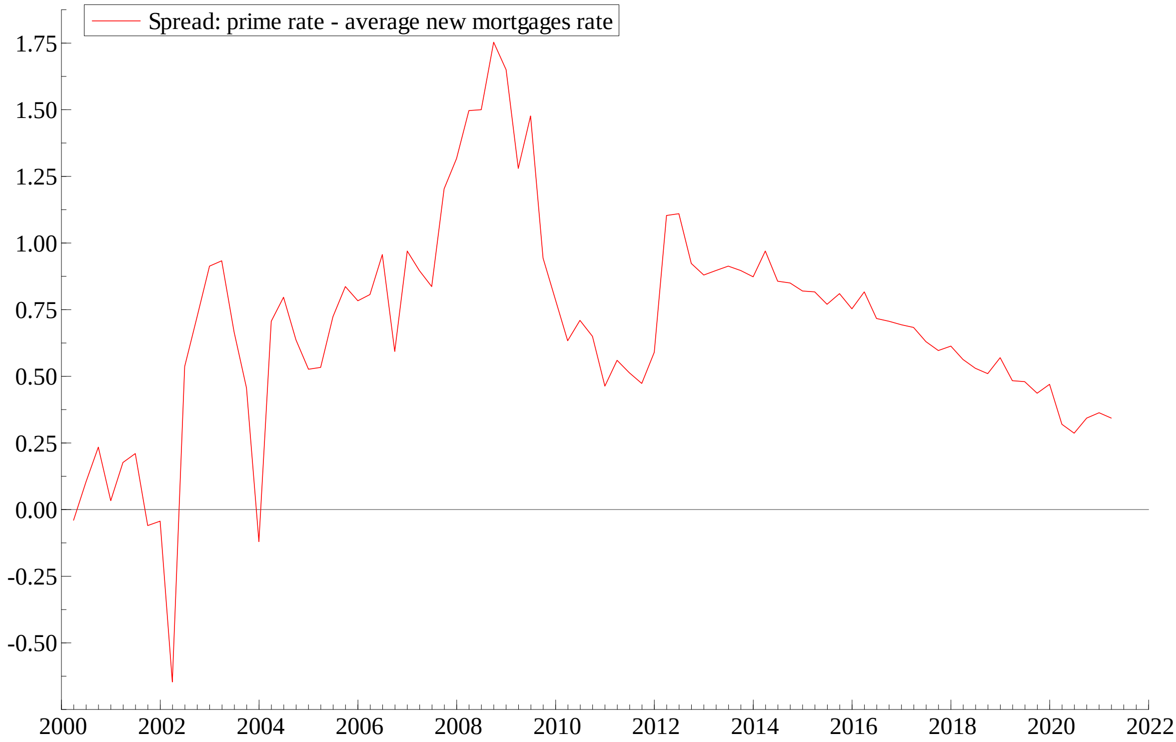
Source: Aron and Muellbauer (2013).

Figure 2: Log mortgage stock to income ratio and log non-mortgage stock to income ratio



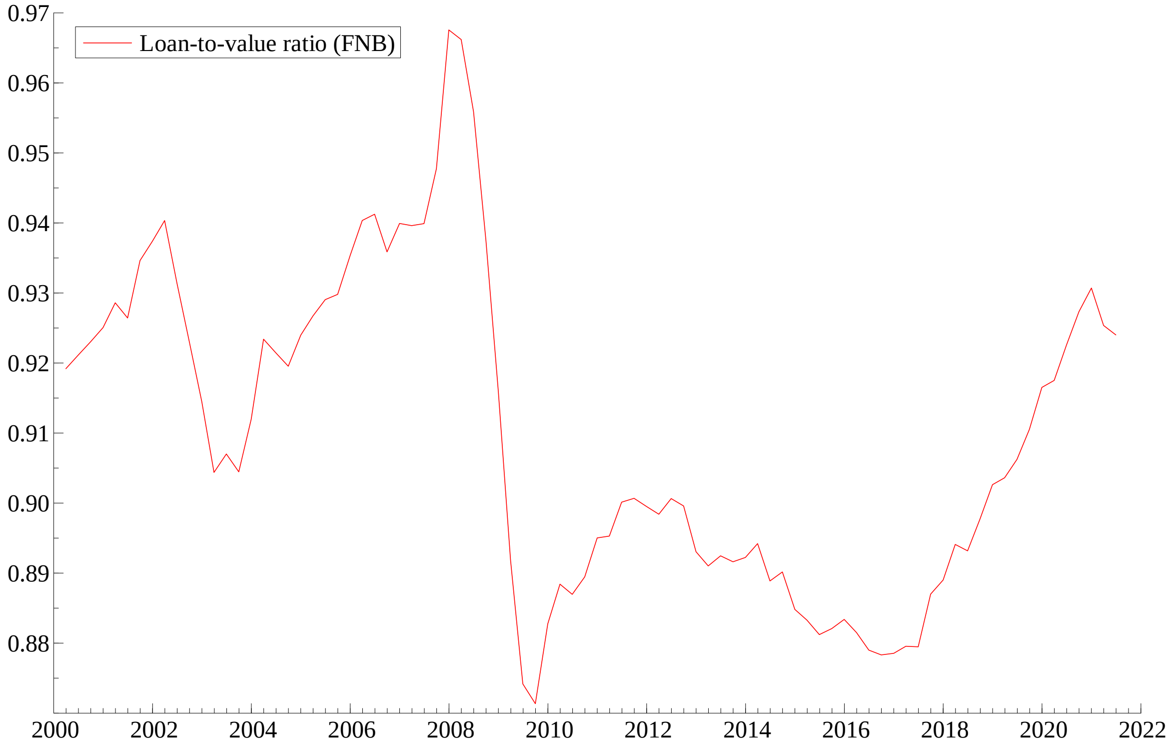
Source: SARB data.

Figure 3: Spread between prime and the effective mortgage rate on new loans



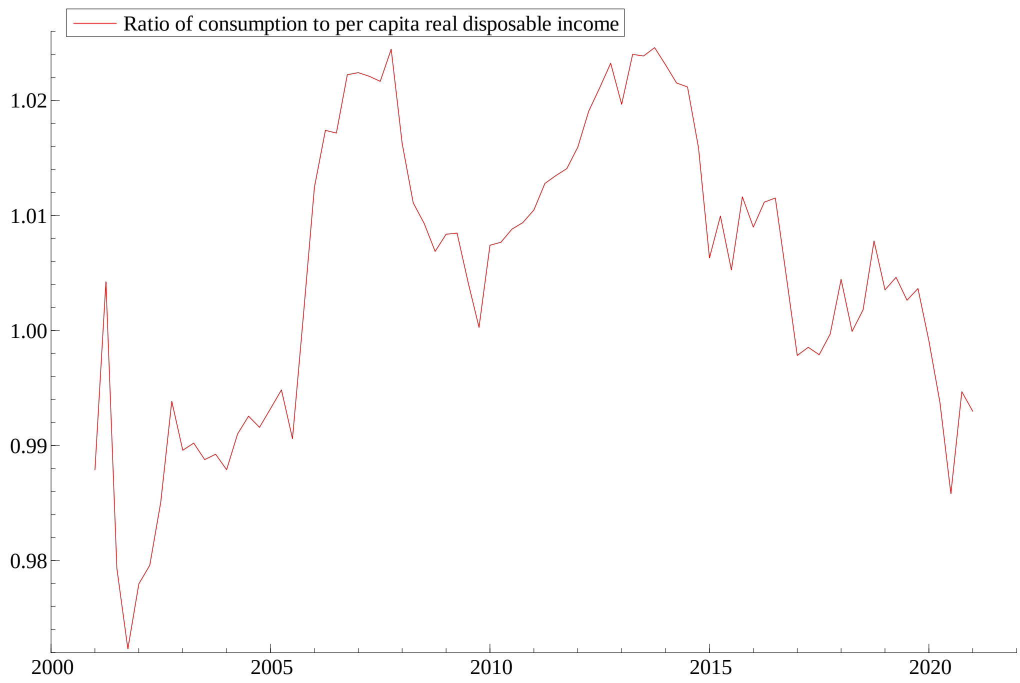
Source: The average mortgage interest rate is BAT9612M, SARB, from 2001Q1. Data for 2000 interpolated by the authors.

Figure 4: Loan-to-value ratio for mortgages from Deeds Office data



Source: FNB compilation from Deeds Office data, South Africa.

Figure 5: Consumption to disposable income



Source: Constructed from Quarterly Bulletin data.

TOPICAL BRIEFING: No. 7

Financial Stability Department

South African Reserve Bank

Non-Performing Loans in South Africa: a Scoping paper for Future Model Development

Abstract

Elevated non-performing loans (NPLs) are a recurrent characteristic of banking crises, with an important *two-way connection* between credit conditions and NPLs. Modelling such credit risk indicators could be highly relevant for informing monetary and macroprudential policy in SA, by strengthening model linkages between the financial sector and real economy. This paper surveys international literature on identifying, measuring and modelling NPLs. SA is not immune from the inconsistency of concepts across countries and jurisdictions, and within countries between different institutions and across time. Our clarifying typology details the evolution of NPL concepts in SA, heavily affected by regulatory definitional changes. We propose how pre- and post-2008 data on three different NPL concepts might be joined to permit an analysis of data from 2001 on reasonably consistent definitions (the volatile period, 2001-2007, can help draw robust insights). The SARB publishes a time series for the related credit risk indicator: 'credit impairments', a loan loss provision, not an NPL concept. We develop a *new* empirical model for the ratio of 'credit impairments' to gross loans and advances from 2001. Major drivers are the ratios of mortgage debt and house prices to income, credit conditions in the previous three years measured by credit spreads, and the GDP growth rate in two prior years. We allow for regulatory data breaks in 2008 and 2018. For potential NPL models, we expect similar key drivers with different relative weights in the long-run solution, and different short-run dynamics. We further propose panel studies of NPLs for individual banks and loan classes, even if only post-2008. Before institutional memory is lost, we urge a concerted effort from the Prudential

Authority and Financial Stability Department to publish time series data for at least two of the NPL measures for 2001-2007, and from 2008, with clear documentation, to benefit in-house and more general modelling efforts.

Acknowledgements: We are extremely grateful to Esté Nagel (Prudential Authority) and Hugh Campbell (Financial Stability Department) for their advice. We also thank the following persons at the *SARB* for their discussion and assistance: Vafa Anfari, Shaun de Jager, Danie Meyer, Mpho Moloto, Susana Paulse, Irene Peter, and Pieter Pienaar; and from the *FNB*: John Loos and Siphamandla Mkhwanazi; and from *71point4*: Illana Melzer. We thank an anonymous reviewer at the SARB for comments.

Contents

Executive Summary

This paper explores the international literature on NPLs to establish the scope for an analysis of NPLs by the SARB. It should be a priority to clarify definitional issues concerning NPLs and related proxies in South Africa, as well their driver variables, and to design models for early warnings systems for NPLs and for their use in the Core macro-model for a better understanding of the financial stability linkages.

Elevated levels of non-performing loans (NPLs) are a recurrent characteristic of banking crises. Banking crises are typically preceded by poor quality of lending, excessive credit growth and high levels of leverage. The value of non-performing loans, low and stable in boom periods, can rise sharply when the crisis breaks. Rising NPLs raise funding costs for banks, damaging their efficiency and profitability. As banks apply tougher lending criteria for firms and households, a credit crunch may follow with falling GDP or stagnant economic growth.

Modelling NPLs, particularly as part of the Core model, is highly relevant for informing both monetary and macroprudential policy in South Africa. Even without a major crisis, a period of easy credit conditions, resulting in lax lending criteria, can create financial vulnerability among borrowers and potentially among lenders, particularly if followed by an economic downturn. Then, rising NPLs will amplify the economic cycle.

This paper surveys international literature on identifying, measuring and modelling NPLs. There is inconsistency of concepts across countries and jurisdictions, and within countries between different institutions and across time (e.g. see Bholat *et al.* (2018)). South Africa is not immune, with different NPL concepts heavily affected by regulatory definitional changes.

We present a clarifying typology detailing the evolution of NPL concepts in South Africa. After detailed investigation of the data and definitions, we propose how pre- and post-2008 data on different NPL concepts might be joined to permit an analysis of data back to 2001 on reasonably consistent definitions. Including the volatile period of 2001-2007 could help draw robust economic insights.

The three possible NPL concepts are the ratios to total loans and advances of: (i) defaults; (ii) 90-day overdue loans; and (iii) impaired advances. Of these, the default ratio concept, available since 2008, is probably the most suitable to be taken back on a *consistent* basis to 2001 – although the implementation requires further investigation. On the 90-day overdue concept, we judge that breaks in the data in 2008 and later, as banks successively switch from the standardised to the Internal Ratings Based (IRB) approach for measuring credit quality, would complicate the linkage of data for a consistent series back to 2001. However, we recommend that the linkage be attempted on a bank-by-bank basis to provide a second NPL concept for modelling. The third measure, impaired loans, published since 2008, is the least satisfactory candidate as a potential NPL measure. It is strongly affected by the accounting switch in 2018, and pre-2008 proxies may be more elusive.

In general, inconsistency of concepts makes it harder to draw firm conclusions from empirical studies, whether from country panels, time series for individual countries or bank-specific panels. Nevertheless, our review of empirical studies points to the relevance of rates of economic growth in reducing NPLs, and interest rates and the unemployment rate in raising NPLs, amongst a range of macroeconomic drivers, as well as relevance for bank-specific and non-financial corporate drivers. However, few studies include real estate-connected drivers such as mortgage debt-to-income and house price-to-income ratios. Again, few consider non-linear or asymmetric relationships in the data.

There is only one published time series for a credit risk indicator for South Africa back to the 1990s, which is not for an NPL concept, but rather for the stock of 'credit impairments', a loan loss provisions concept. We have developed a *new* empirical model for the 'credit impairments' to gross loans and advances ratio. The major drivers are the ratios of mortgage debt-to-income and house prices-to-income, credit conditions in the previous three years (measured by credit spreads), and the growth rate for GDP in the previous two years. It is important to allow for breaks in the data from definitional changes in 2001, 2008 and 2018. High levels of impairments are typically followed by weak current credit conditions measured by spreads and loan-to-value ratios.

Our findings for this loan loss provisions model underline the two-way connection between credit conditions and credit risk for South Africa. We anticipate that the key drivers in the long-run solution would be similar for prospective NPL models for South Africa. A similar model for credit impairments and/or for an aggregate NPL concept should be an important part of the Core model. Credit risk indicators are likely to affect credit pricing and credit extension by banks and will thus improve linkages in the model between the financial sector and the real economy. Further, comparing results for NPL and loan loss provisions models would illuminate questions about the pro-cyclicality of provisioning.

We set out the different possible dependent variables for *disaggregated* NPL series for South Africa and have reviewed the data availability of potential bank-specific, corporate and macro-drivers of these NPLs. We discuss the scope for *bank panel studies* of NPL data, including of sectoral data from 2008. We also recommend comparisons with NCR data for households and to exploit data from credit reference bureaus on NPLs by vintage of the loan; we suggest adaptation by the SARB of a micro-simulation approach developed to analyse the profitability of mortgages, Melzer and Hayworth (2018), for analysing mortgage defaults and the scale of potential losses.

We recommend that the SARB publish time series data for at least two of the NPL measures for the 2001-2007 period and 2008 onwards, with clear documentation. A special background paper should give transparent methodological detail on joining, using a bank-by-bank basis, the different time segments for the various NPL concepts.

In summary, for the *default ratio* NPL measure, for all banks (regardless of whether they use the standardised or IRB approaches), the pre-2008 data require joining with the post-2008 data. For the *90-day overdue ratio* NPL measure, for banks using the standardised approach, the pre-2012 data require joining with the post-2012 data; for IRB banks, data from the last month of their using the standardised approach needs to be joined to the first month that they use the IRB approach. Thereafter, the aggregate NPL time series data, with appropriate qualifications and explanations of methodology, should be routinely published.

1. Introduction

Loans support household and firm investment and spending, and may be collateralised or unsecured, cover different maturity profiles, and require payment of interest and potentially penalties, as part of the contractual obligations. In good times, these assets of the lenders support their profitability, while releasing cash flow for business debtors, and bridging the lifetime budget constraint for household borrowers seeking housing, for example. However, negative shocks may convert loans to non-performing loans (NPLs) that are in or close to default when debtors fail to meet the contractual obligations of the loan. For example, an NPL can be defined as a loan upon which the debtor has not made scheduled payments for at least 90 days. Elevated levels of NPLs are a recurrent characteristic of banking crises (Bholat *et al.* (2018). Ari *et al.* (2019), studying 88 banking crises in 78 countries since 1990, found that for over 80% of crises, the NPL levels exceeded 7% of total loans, and for almost half the crises, the NPL levels more than doubled relative to the pre-crisis period. Moreover, these NPL levels persisted well beyond the crisis peak, and for a third of cases, exceeded 7% of total loans seven years later. The effect of rising NPLs is to undermine bank balance sheets, curtail credit growth, and impede post-crisis output recovery.

Non-performing loans ballooned globally in the last two decades, in consequence of the Global Financial Crisis (GFC) and the European sovereign debt crisis, see Table 1. The pandemic era is expected to exacerbate the problem of NPLs worldwide, as government fiscal support and various regulatory forbearance measures, such as rental and mortgage payment moratoria and eviction bans, are withdrawn (Kasinger *et al.*, 2021). South Africa, too, has seen a jump in NPLs in the pandemic, rising especially sharply for specific sub-sectors and for households' unsecured loans, see Section 2. Yet, even recently, Ari *et al.* (2019) pronounced that 'we know little about the patterns of NPL build-up and the factors that affect NPL resolution'.

It would be useful to draw lessons from cross-country analyses of NPLs and their drivers, and to use panel and time series studies within banking sectors for individual countries to incorporate NPL and other credit risk indicators into macro-models and early warning systems. However, a serious consideration is that the criteria for classifying NPLs or impaired loans across countries vary not only across jurisdictions and lenders, but also within lenders across time (Bholat *et al.*, 2018). The goal of promoting a harmonised NPL definition across countries has been promoted by guidelines from the IMF (2005), the European Banking Authority (ECB, 2017), and the Basel Committee on Banking Supervision (BCBS, 2017). Nevertheless, these discrepancies and variations necessarily make comparative analyses unreliable, across countries or regions, and even analyses of different banks' asset quality within a single country may be compromised. Data constraints further limit the lessons that can be drawn as many countries have inadequate or missing data, especially on lending quality.

The nomenclature for 'problem loans' is wide in general, and this issue is discussed in Section 3. South Africa is no exception, using three different terminologies for "non-performing loans" since the 1990s. The different NPL definitions may not coincide and should not be used inter-changeably[107]. The data are mainly sourced from the Prudential Authority (formerly the Bank Supervision Department of the SARB), collected as regulatory data. All three of the terms have been affected by definitional changes over time.

The first NPL concept is *'impaired advances'*. Loans are impaired when the amount expected to be repaid falls below the contracted value carried on bank's balance sheet: "impaired advances are advances in respect of which a bank has raised a specific impairment and includes any advance or restructured credit exposures subject to amended terms, conditions or concessions that are not formalised in writing", Financial Stability Review (2021-1). The *specific impairments*, or loan loss provisions (LLPs), are an accounting deduction representing the difference between the contracted repayments and the banks' most current estimate of what they will receive. The reporting of "impaired advances" followed South Africa's implementation of Basel II with effect from 1 January 2008. The data are collected via bank survey forms as assessed by the banks themselves, embodying discretionary thresholds when banks evaluate categories of default. The SARB publishes a time series for the specific impairments (provisions) in respect of loans and advances which it terms: 'credit impairments'. The credit impairments are loan loss provision concept, and not an NPL concept.

[107] The sources and definitions of the different NPL data reported in the *Financial Stability Review* should always be precisely given.

The terminology used in South Africa does not always coincide exactly with terms used by the BCBS or other international agencies. For example, the conventional usage of the NPL term 'impaired advances' is as an accounting concept (see Bholat *et al.*, 2018). However, in South Africa prudential guidance is overlaid onto the accounting definition using qualitative criteria. Hence, the set of loans encompassed by South Africa's term, 'impaired', will exceed those of the pure accounting definition of 'impaired'. This overlay is the typical procedure globally, see also Baudino (2018).

The second NPL concept is *overdue loans,* currently defined as all exposures overdue for more than 90 days and where the recovery thereof was considered to be doubtful, expressed as a percentage of on-balance sheet exposures. Before a change in banking regulations implemented in 2001, data were collected from banks on 'overdue advances' classified into months overdue categories such as 0-1, 1-3 and more than 3 months overdue. Overdue advances were reported in the Annual Bank Supervision reports from 1994 to 2007. These quarterly data apply to the different credit products such as mortgage loans and instalment finance. From 2008, quantitative information on 90-day overdue loans has been required from the banks following the Internal Ratings Based (IRB) approach, see below. Following bank regulatory changes in 2012, bank reporting forms have required quantitative information on 90-day overdue loans and advances from the banks following the standardised approach to credit risk reporting, see below.

The third and final NPL concept is *monthly 'default ratios',* constructed by those banks with permission to use the Internal Ratings Based (IRB) models for credit quality assessment from 2008, see BCBS (2001a). The number of authorised banks has expanded from 4 to 5 banks since 2008. The remaining banks use the standardised approach to credit risk rating from 2008, see BCBS (2001b), and also construct default ratios. These use the sum of the three credit risk buckets: 'sub-standard', 'doubtful' and 'loss', see Appendix, to define 'default'. The aggregate default ratios for all banks have been shown at least from 2010 in the *Financial Stability Review*, as a total, and for retail and corporate sectors.

An important distinction is between stock and flow measures. NPL is a stock concept and the NPL ratio is defined as the stock of (some concept of) NPLs divided by the underlying total of loans outstanding. Analogously, the loan loss provision ratio is the stock of loan loss provisions divided by the underlying total of loans outstanding. However, flows into and out of these stock measures are also likely to be informative. Ferrari *et al.* (2021) make the point that data on *flows* of *new* loan loss provisions and non-performing loans can sometimes be more informative than stocks or changes in stocks. Write-offs (when a loan is considered unrecoverable) and reversals (when a loan previously classified as non-performing is reclassified as performing) can distort the picture provided by the stock data. For example, large write-offs may result in a decrease in the stock measures even when the flow of new provisions and non-performing loans has increased. In their study of Belgian data, Ferrari *et al.* (2021) conclude that the stock ratios better reflect the sensitivity of credit risk to macroeconomic variables than do the flow measures, but that no single ideal credit risk measure exists.

This paper explores the international literature on NPLs to establish the scope for an analysis of NPLs by the SARB. It should be a priority to clarify definitional issues concerning NPLs and related proxies in South Africa, as well their driver variables, and to design models for early warnings systems for NPLs and for their use in the Core macro-model for a better understanding of the financial stability linkages. Section 2 links rising NPLs to the credit cycle and places South Africa within the context of global trends. Section 3 clarifies the global definitional discrepancies and the international attempts to standardise NPL identification. We present a clarifying typology detailing the evolution of NPL concepts in South Africa. After detailed investigation of the data and definitions, we propose how pre- and post-2008 data on the three different NPL concepts might be joined to permit an analysis of data back to 2001 on reasonably consistent definitions (including the volatile period of 2001-2007 to help draw robust economic insights). In Section 4, we explore the empirical literature to assess the methods and variables used to model and forecast NPLs, and, in some cases, with attempts to mitigate the definitional problems.

NPLs are not currently formally modelled in the SARB's Core model[108]. Stress testing exercises for banks currently use post-2018 data for individual banks (since the International Financial Reporting Standard (IFRS) 9 accounting standard became

[108] There is an equation for 'credit impairments' (a loan loss provisions concept not an NPL concept) as a Memo item, see review in Section 5.1.

effective at the SARB under IFRS accounting standards) for bank asset-class-level analyses using calibrated models. In Section 5, we derive a *new* equation for the ratio of 'credit impairments' to gross loans and advances extended by banks. We find that lending conditions in earlier years, and recent debt-to-income and house price-to-income ratios play an important role in driving this concept of an aggregate credit risk ratio and confirm that changes in definitions in 2008 and 2018 resulted in quantifiable jumps in the data. Our findings for this model of loan loss provisions underline the two-way connection between credit conditions and credit risk. A similar model for credit impairments and/or for an aggregate NPL concept should be an important part of the Core model. We anticipate that the key drivers in the long-run solution would be similar for prospective NPL models for South Africa. Credit risk indicators are likely to affect credit pricing and credit extension by banks and will thus improve linkages in the model between the financial sector and the real economy. Further, comparing results for NPL and loan loss provisions models would illuminate questions about the pro-cyclicality of provisioning.

In Section 6, we suggest a 'suite of models' approach to improve the macro-prudential side of the Core model of the SARB, and to allow a better understanding of the credit cycle with banking sector level models and early warning forecasting models of NPLs. We set out the different possible dependent variables for disaggregated NPL series for South African and review the data availability of potential bank-specific, corporate and macro-drivers of these NPLs and model selection methods. We discuss the scope for bank panel studies of NPL data, including of sectoral data from 2008. Section 7 concludes.

2. Models and the Role of NPLs in the Credit Cycle and Financial Stability

Banking crises are typically preceded by poor quality of lending, excessive credit growth and high levels of leverage. When the crisis breaks, the value of non-performing loans, often fairly stable in boom periods, can rise very sharply. Where the potential loan losses have been under-provisioned for, this negative non-linear relationship of NPLs with the credit cycle can result in a sudden, large, deleterious impact on financial and economic stability. This, in turn, has negative consequences for the banking system's ability to provide financing to the real economy. Rising NPLs drive up the funding costs for the associated banks through higher borrowing costs, loan loss provisions and legal and administrative costs, and lower interest income, damaging efficiency and profitability, and weakening regulatory capital. If, as typically happens, credit costs to borrowers then rise and banks use tougher lending criteria for firms and households, this may lead to a credit crunch and falling or stagnant economic growth. The solvency of both banks and borrowers may be at stake, with damaging feedbacks onto bank and firm share prices with liquidations, and onto house prices with repossessions. Further negative feedbacks onto the economy may stem from the spending constraints of the indebted households and firms. Financial sector interconnectedness in the economy may be large enough to cause systemic risk. There is thus a two-way connection between credit conditions and NPLs.

The empirical evidence confirms that there are important macro-financial linkages in crisis recovery, associated with NPLs. Ari *et al.* (2019), analysing a new dataset of over 80 banking crises, tracked NPL ratios across banking crises[109]. High levels of unresolved NPLs were linked with more severe recessions after crises, and output was lower than in crises with low NPLs. Moreover, the high NPL ratios in a third of cases persisted well beyond the crisis. This research points to the potential usefulness of early warning systems on the trends for NPLs and using a more forward-looking, ‘expected loss’ model for the associated loan loss provisioning, to reduce the dire economic effects of shock-induced crises. While a timely recognition of NPLs is important, regulatory definitional changes have still not given banks the right incentives for early NPL recognition and loss-provisioning (Bholat *et al.* (2018), Kasinger *et al.* (2021). This is discussed in the next section.

Table 1 presents the ratio of non-performing loans as a percentage of total gross loans[110] for a range of countries, though as noted in the introduction, definitions may not be entirely consistent either between countries or over time. Table 1 shows that the credit quality of loan portfolios across most countries, fairly stable prior to the GFC, fell sharply from 2007-8, with increasing NPLs. This deterioration was uneven across countries. In the US, about three-quarters of the total loan portfolios of banks was from real estate lending; asset quality fell through holdings of real estate loans, exposure to

---

[109] An updated version of this paper is forthcoming in 2021, Ari *et al.* (2021).

[110] Bholat *et al.* (2018) suggest that this ratio may reward leverage: “a more leveraged bank would show a higher denominator and therefore a lower NPL ratio in situations where it has the same number of NPLs as a bank with lower leverage, even though overall risk of failure may be higher in a highly leveraged bank, since by definition it would have a lower capital buffer”.

mortgage-backed securities and credit derivatives based on these securities. US and Western European banks with exposure to US residential and commercial mortgage-backed securities saw considerable asset quality deterioration in the GFC and its aftermath, though NPL figures thereafter declined. Italy, Ireland, Greece and other countries on the periphery of the Eurozone were caught in a sovereign debt crisis that reflected structural problems in the Eurozone. The trigger was Greece's revelation in 2009 that its fiscal deficit more than twice exceeded the previously reported figure, raising spreads dramatically on sovereign debt in countries with poor fiscal and competitive positions, such as Greece, Italy and Portugal. This generated a 'doom loop' between sovereign debt and the banking system, prolonging problems that initially came to light in the GFC and the recession that followed. These countries experienced persistent double-digit NPL ratios, especially in countries where competitiveness and economic growth were slowest to recover.

South Africa, too, has seen growing NPLs. Using first the definition based on the annual ratio of total impaired advances to gross total loans and advances, this ratio fell from a peak of 5.94% in 2009 to 2.84 in 2017, rising again to 5.18% in 2020, see Table 1. Figure 1 shows total impaired advances to gross total loans and advances by type of bank from 2016. NPLs at the major banks, Standard Bank Group Ltd., FirstRand Ltd., Absa Group Ltd. and Nedbank Group Ltd., have increased over the last five years, a time of record high unemployment and electricity blackouts[111]. Part of the jump in 2018, probably of the order of 10%[112], is related to a redefinition of impaired advances under the new accounting standard of the IRFS 9, applied from January 2018. The greater part of NPLs cover the household sector. Figure 2a uses the 90-day default measure of NPLs. Although debt-service costs fell in 2020 and 2021, the ratio of the value of overall household 90-day-overdue loans to total outstanding household loans exceeded 6% in these years. Of this ratio, the portion for *secured* household debt reached 4.9% in July 2020, declining moderately to 4.4% by February 2021; the ratio for unsecured household debt continued to increase, reaching a six-year peak of 12.4% in February 2021, see Figure 2a. The spread between loan

[111] According to S&P Global Market Intelligence (2021), Standard Bank's NPL ratio was 6.21% in 2020, up from 4.22% a year earlier and nearly double that of 2016. FirstRand's NPL ratio rose 109 basis points in 2020 to 5.04%, while Nedbank's rose 189 basis points over the same period. Credit losses were considerable in 2020. Standard Bank's net income halved year to year to ZAR13.2 billion, FirstRand's declined 41% to ZAR18.2 billion, Absa's dropped 59% and Nedbank's fell by 71%.

[112] Communication, Financial Stability Department, SARB. The research reported in Section 5.2 below shows a related jump in 'credit impairments', a loan loss provisions measure.

performance on secured and unsecured debt points to growing economic difficulties for part of the population in the Covid-19 era, and to pressures on smaller, less diversified bank lenders. Mortgages comprise the bulk of bank credit to households, averaging over 60% as a share of household credit between 2001 and 2011, and just under 60% since then. Mortgage 90-day-overdue loans to total outstanding mortgage loans rose sharply in the first half of 2020, but appear to have stabilised in 2021, see Figure 2b. Sectoral figures on NPLs defined by default ratios, see Figure 3, are available for the aggregate of IRB reporting banks only.

In the Covid-era, commentators are exploring scenarios for the likely profile of NPL growth. In several countries, apart from eventual retraction of the fiscal support given to firms and households, withdrawal of regulatory moratoria on forbearance in mortgage repayments and tenancy rental payments could compound NPLs. Kasinger *et al.* (2021) point to three characteristics unique to the Covid-crisis: the differential impact across industrial sectors, the huge intervention and fiscal support at the industry level, and exceptional uncertainty affecting expectations for longer-run outcomes. The last of these may also affect the trend of NPLs as banks may adopt a wait-and-see approach before recognising such loans, and possibly then will maintain less efficient lending. This outcome could be compounded by lengthy forbearance episodes. These authors echo others in urging pro-active NPL management, with planning for loss provisions and appropriate incentives to restructure vulnerable firms and banks. They argue that a realistic assessment of current loan values for the early identification and recognition of NPLs on bank balance sheets would be encouraged by IFRS 9 standard accounting rules or similar, and stress tests. Early recognition of NPLs could promote the development of secondary loan markets giving more transparent loan quality information, to help reduce the ultimate capital losses for banks from NPLs.

South Africa has been included in the Fitch (2021) study of emerging market countries in the Europe, Middle East and Africa region in which banks have utilised moratoria programmes and of likely risk to asset quality and rising NPL ratios as forbearance is withdrawn. Loan moratoria programmes have had the highest utilisation in Georgia, Hungary and Nigeria (between 40% and 60% of loans were subject to moratoria at peak levels) followed by South Africa and the UAE (each at close to 20%). The risk of asset quality deterioration is viewed as most significant in Nigeria, but material in

Turkey, Georgia, Hungary, South Africa, the UAE and Qatar, in the set of countries studied.

3. Issues concerning NPL Classification

Prior to the GFC, beginning with the Basel I agreement in 1988, there was progress in the harmonisation and international comparability of claims on banks. However, far less attention had been paid to the standardisation of loan classification of the asset side of banks' balance sheets and the definition of non-performing loans. Since the crisis there have been consecutive endeavours by regulatory and multilateral organisations towards a harmonisation of the NPL definition across jurisdictions. Nevertheless, there is currently no universal standard for NPLs across countries (Bholat *et al.*, 2018). The analysis of Ari *et al.* (2019), amongst other studies, demonstrates that reliable and comparable NPL data are crucial for NPL monitoring and evidence-based NPL resolution polices.

3.1 Clarifying the different definitions of NPLs and their variation over time

The timely identification of 'problem loans' or 'non-performing loans'[113] helps to ensure that the stock of these is recognised on bank balance sheets. The nomenclature describing 'problem loans' is wide, examples being non-performing loans, impaired loans, restructured loans, delinquent loans, past-due loans, and defaulted assets. One reason for the different terminology is that there are multiple players with different perspectives in the system. There is an accounting perspective, governed by the accounting rules of a particular jurisdiction, requiring a provision to be made if a loan cannot be fully recovered, when the loan is reclassified from a performing to an impaired loan in the financial statement of the bank. There is a regulatory perspective, governed by the supervisory and prudential rules of a particular jurisdiction, that may require additional equity to be held for loans that are non-performing. Then there is a broader economic perspective, used by central banks and institutions like the ECB, the

[113] Baudino *et al.* (2018) distinguish between the terms, "non-performing assets", "non-performing loans" and "non-performing exposures": "Of the three, non-performing loans is the narrowest concept, as it refers only to problem loans, but is the term most commonly used in the academic literature as well as among market participants. Nonperforming exposures is typically the widest concept, and it includes loans, debt securities and certain off-balance sheet exposures, but may exclude certain asset classes, such as foreclosed collateral. In some jurisdictions that provide a definition of non-performing assets, they include various asset classes such as foreclosed collateral."

BIS and the IMF, which may combine and extend features of both of the above two definitions.

Traditionally, the identification of problem assets and the associated calculation of losses have been subject both to accounting principles and prudential oversight (Baudino, 2018): whether and when an exposure is deemed to be 'non-performing' is typically not clear-cut but requires judgement by banks and regulators, based on both quantitative and qualitative factors.

The *accounting concept* is 'impaired loans', but accounting frameworks are not globally harmonised either, as for instance, there is a divergence between the US Generally Accepted Accounting Principles (US GAAP).and IFRS jurisdictions. 'Impaired loans' are those in respect of which a bank has raised a 'specific impairment' or 'loan loss provision' (LLP), capturing the difference between the expected repayment and the larger contracted value. The accounting rules affect how bad loans are identified, disclosed and provided against. The IFRS standard, consistent across all industries and not just banking, is currently used by 166 jurisdictions including the European Union and South Africa (International Accounting Standards Board (IASB))[114]. The IAS 39 standard applied up until January 2018, when the new IFRS 9 standard on provisioning was implemented[115]. Under the IAS 39, impaired assets were governed by an 'incurred loss' model so that impairment was recognised only when a loss had actually occurred, and expected losses, even if likely, were not taken into account in the definition. Under the new IFRS 9 standard, provisions are based on forward-looking expectations, and governed by a three-stage model. Stage 1 (performing) and Stage 2 (under-performing) and Stage 3 (non-performing) categories replace the impaired and unimpaired categories of IAS 39, see the comparison by classification and provisioning requirements in Table 2a. Stage 3 of the IRFS 9 is similar to the impaired classification[116] of the IAS 39. The three-stage method differentiates credit quality, as well as the method for calculating the loan loss provision, which under

114 See webpage: https://www.ifrs.org/groups/international-accounting-standards-board/

115 The IFRS 9 was implemented by banks at different months in 2018 as the implementation date was based on the institutions' particular financial years.

116 Baudino *et al.* (2018, Table 1) states that any one or more of the following suggests evidence of 'credit impairment' under IFRS 9 and IAS 9: Significant financial difficulty of the borrower. A breach of contract such as default or past-due event. The lender has granted the borrower a concession due to the borrower's financial difficulty. It is probable that the borrower will enter bankruptcy. The disappearance of an active market for that financial asset because of financial difficulties. The purchase or origination of a financial asset at a deep discount that reflects the incurred credit losses.

IFRS 9 covers all credit exposures. Baudino *et al.* (2018, p.6) clarifies that the definition of ‘impaired’ has remained unchanged (in the sense that Stage 3 of IFRS 9 corresponds to impaired under IAS 39), but that IFRS 9 requires a more granular assessment of credit risk than under IAS 39.

Baudino *et al.* (2018) suggests that because accounting standards are principle-based, and hence applicable to all industries, and not only banks, the complexity of identifying and managing credit risk typically requires more detailed guidance than only from the applicable accounting standard. They make a highly informative comparative typology for the US, Europe, Latin America and some Asian countries that *maps* the *regulatory concept* of NPLs with the accounting concept of ‘impaired’. This is reproduced for their selection of countries in Table 2b. Credit exposures are classified into ‘risk buckets’ (the most common being: Normal, Special Mention (or Watch), Substandard, Doubtful and Loss), based on criteria developed by the various prudential regulators, related to the loan classification scheme of the Institute of International Finance (see Table 3 and Krueger (2002)).

Baudino *et al.* (2018) argue that irrespective of whether a *formal regulatory* definition of NPLs was adopted (as for example the EU countries did in 2014, of which see more below), or a more *informal regulatory* definition for NPL identification, the regulatory frameworks will in general tend to *encompass* the accounting definition of “impaired” within a broader set of NPLs. This is because qualitative criteria, e.g., the Unlikely to Pay criteria, classify exposures as NPLs that otherwise might be considered as ‘unimpaired’ or ‘performing’ under the locally applicable accounting frameworks. This is indicated in Table 2b by the non-performing (blue) segment by regulatory definitions exceeding in size the impairment segment by the accounting definition.

There are further distinctions between a *broader economic concept* of NPLs (sometimes called NPEs, see BCBS (2017) and European Banking Authority (EBA) (2014, 2017)), the prudential concept of defaulted loans, and the accounting concept of impaired loans. Bholat *et al.* (2018) present a detailed documentation of the global progression towards greater harmonisation of the economic concept through recommendations and guidance issued by a range of international institutions, both before and after the GFC.

They suggest that the harmonisation effort began with the Basel Committee's work including the credit risk calibration based on bank's internal risk models using the IRB methodology as part of the Basel II framework. They suggest that a definition of default was thereby established[117], and notes its similarity to the later EBA definitions of non-performing. In 2014 the EBA clarified its definition of non-performing exposures[118], for a concept broader than NPLs: non-performing is defined as material exposures greater than 90 days past-due, and/or where the debtor is assessed as unlikely to pay its credit obligations in full without realisation of collateral, regardless of the existence of any past-due amount or of the number of days past due. The subsequent ECB (2017) guidance clarifies the application of the 'unlikely to pay' condition and the management and monitoring of forbearance, loan write-off and collateral valuation. In ECB (2017, Figure 2, p.48) it is illustrated how this supervisory definition of NPEs encompasses both the prudential definition of default (European definition under the CRR) and the accounting definition of impaired (by the IAS 39 standard of the time) - itself encompassed by the default definition, as indicated above. All impaired loans and all defaulted loans are necessarily NPEs. However, NPEs can also encompass exposures that are not recognised as impaired or as defaulted in the applicable accounting or regulatory framework. Differences between impaired and NPEs concerns the extent to which the (automatic) 90 days past due cut-off used in NPEs is not used for impaired. Differences between defaults and NPEs concerns the cure period to exit NPE, exposures greater than 90 days past due preventing exit of NPE and features of forborne treatment.

Guidelines by the BCBS (2017) on prudential treatment of problem assets has also helped harmonise definitions for 'non-performing' and 'forborne' exposures, including entry and exit criteria, using both quantitative and qualitative considerations. The BCBS definition combines three concepts: it includes all exposures that are considered

---

[117] BCBS (2004): *Default is defined as* where an obligor is 90 days past due **or** is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realising security. Indicators of unlikeliness to pay include: the bank puts the credit obligation on non-accrued status; the bank makes a charge-off or account-specific provision resulting from a significant perceived decline in credit quality subsequent to the bank taking on the exposure; the bank sells the credit obligation at a material credit-related economic loss; the bank consents to a distressed restructuring of the credit obligation where this is likely to result in a diminished financial obligation caused by the material forgiveness, or postponement, of principal, interest or (where relevant) fees; the bank has filed for the obligor's bankruptcy or a similar order in respect of the obligor's credit obligation to the banking group; or the obligor has sought or has been placed in bankruptcy or similar protection where this would avoid or delay repayment of the credit obligation to the banking group. (Note that the accounting concept 'non-accrual loans' does not exist under the IFRS.)

[118] Exposures cover all debt instruments (loans, advances and debt securities) and off-balance sheet exposures (loan commitments, financial guarantees and other revocable and irrevocable commitments) excluding trading exposures and off-balance sheet exposures except held for trading exposures, Bholat *et al.* (2018).

defaulted under the Basel II framework; all exposures that are impaired under applicable accounting standards (this equates to ‘stage 3’ of the IFRS 9 provisioning model, Bholat *et al.* (2018)); and all other exposures that are not defaulted or impaired but are material exposures that are more than 90 days past due or where there is evidence that full repayment is unlikely. This NPE definition also encompasses a broader range of exposures[119] than considered as “impaired” under accounting standards. This is because it has a qualitative “unlikely to pay” criteria without an equivalent in accounting frameworks and includes designation of NPE status on a debtor basis, the specific rules to exit the NPE category and the minimum repayment period for forborne NPEs to return to performing status, not explicitly observed under relevant accounting standards (Baudino *et al.* (2018). The guidelines specify that collateralisation does not influence past due Status and should not be considered in the classification of an exposure as non-performing. For countries following the IMF or European reporting standards, there has been a convergence to the UN System of National Accounts definition of NPLs[120].

There is thus considerable heterogeneity in the definition of an NPL across regulatory jurisdictions and systemically important banks and firms, and the aggregation of data may introduce measurement errors and biases. With no common standard for categorising loans applied internationally, it is difficult to compare asset quality and draw lessons from cross-country analyses. Especially in view of the mandatory implementation of the IFRS 9 accounting standard on loan loss provisioning from January 2018, it is argued that harmonising the cross-border definitions of NPLs toward a universal categorisation would facilitate the development of comparable indicators for assessing asset quality (Bholat *et al.*, 2018). These authors also argue that the IFRS 9 entails greater *discretion* in the determination of provisions from NPLs[121]. The greater level of estimation entailed may cause divergences in the recognition of impairments, and hence divergences across banks’ balance sheets. It potentially increases the divergences of loan loss provisions across jurisdictions, and even within banking sectors of a single country.

[119] The definitions apply to all credit exposures from on balance sheet loans, debt securities, and other items due, and off-balance sheet items, such as loan commitments and financial guarantees.

[120] “A loan is non-performing when payments of interest or principal are past due by 90 days or more, or interest payments equal to 90 days or more have been capitalized, refinanced, or delayed by agreement, or payments are less than 90 days overdue, but there are other good reasons (such as a debtor filing for bankruptcy) to doubt that payments will be made in full.” (UN, 2008; Bholat, 2018; IMF, 2005).

[121] These authors also examine the strategic choices and trade-offs banks face under the IFRS 9 provisioning rules.

3.2 How different are the NPL definitions globally? Insights from the empirical literature

Since this paper draws on the international literature about the drivers of NPLs, it is important that comparisons concerning the dependent variables be like-for-like. Yet, as noted above, there are wide differences in the NPL concepts across different jurisdictions and even differences within countries due to discretionary thresholds and imprecisely-defined concepts. Measurement errors are further introduced into time series and panel data by the changing definitions for international standards, an example being the introduction of the IFRS 9. Thus, when assessing the sign and importance of NPL drivers with cross-sectional or panel data for different countries' NPLs, or within a country, using time series or panel data for different banks' NPLs, the measurement biases in the dependent variable must be considered.

Several empirical studies, from surveys (e.g., BCBS, 2017) and from cross-country and cross-bank tabulations of definitions (e.g., Bholat *et al.* (2018), Baudino *et al.* (2018) and Barisitz (2011, 2013a, 2013b)), have confirmed that there are considerable differences both across and within countries, and across systemically important banks, in NPL definitions.

Barisitz (2011, 2013b) identifies biases in NPL definitions for ten CESEE countries: Bulgaria, Croatia, Czech Republic, Hungary, Poland, Romania, Russia, Serbia, Slovakia and Ukraine, while Barisitz (2013a) focuses on nine Western European countries: Austria, Finland, France, Germany, Ireland, Italy, Portugal, Spain and the United Kingdom. The majority of these countries classify loans as non-performing when one of two (primary) elements is present: the principal or interest is 90 days or more past due and/or there is 'well-defined weakness of loan or borrower'. However, the interpretation of 'well-defined weakness' is not precisely defined within and across jurisdictions and open to different interpretations. For the CESEE countries, credit quality categories proposed by the Institute for International Finance, see Table 3, are used in addition, of which 'substandard, doubtful, loss' help categorise NPLs; but these categories are mostly not applied in the nine Western countries (Barisitz, 2013a). The actual interpretation of these categories differs perceptibly in practice across the CESEE countries (Barisitz, 2011). There are also secondary elements that affect NPL identification across all the above countries which may result in upward or downward biases. These include the treatment of replacement (restructured) loans, whether collateral and guarantees or other types of security are considered when classifying credit quality, recording the total loan or only part of the loan as an NPL, and the treatment of multiple loans to one borrower (whether to take the 'customer view' and to downgrade all loans if any are classified as NPLs, or just the one loan if taking a 'product view'). With the aim of improving international comparability, the author documents the likely direction of biases from the primary and secondary definitions as tabulated in these papers[122].

A broader tabulation, for the G20 group of countries, including South Africa, was done by Bholat *et al.* (2018). Although they too find that for countries following the IMF or European reporting standards, there is a convergence to the UN System of National Accounts definition of NPLs, they state that *no definitions across these countries are quite the same*. They point to variable categories in loan classification schemes across countries, and the flexible nature of definitions which leaves scope for discretion by firms in interpretation in practice, for example imprecision in the meaning of 'other good reasons'. They also tabulate the heterogeneity in classifying NPLs for global systemically important banks. These authors conclude that cross-country and cross-firm comparisons are considerably complicated by the diverse practices, interpretations and definitions. They suggest that the accurate aggregation of NPLs is probably impossible. Similarly, Baudino *et al.* (2018) reach the conclusion that there are considerable differences across jurisdictions in applicable accounting standards, which are exacerbated by divergent prudential frameworks that govern NPA identification and measurement, across selected Asian, Latin American and Caribbean countries, as well as the United States and European countries. Meaningful comparisons of credit quality metrics are thus difficult to make.

The Basel Committee on Banking Supervision surveyed Thailand and the 27 member jurisdictions of the Basel Committee concerning the regulatory frameworks and supervisory practices across jurisdictions for problem loans, see BCBS (2017). They also surveyed industry practices using a questionnaire and case studies for 39 banks from these jurisdictions. Combined with a literature review, the findings were that there

[122] For example, Finland has upward and a downward bias with regard to different aspects of NPL categorisation with some cancelling out. Austria and Germany have small downward biases; larger downward biases are present in Portugal and the U.K., while Italy has a larger upward bias.

were no consistent international standards for categorising problem loans. Practices varied widely across the jurisdictions, and there were multiple definitional layers within jurisdictions. Terms differed by accounting versus regulatory frameworks, and terms were not consistently defined or reported. Half the examined jurisdictions had established local/national supervisory definitions that categorised assets differently from those used in the accounting framework. They referred to a significant influence of local accounting, regulatory, legal or tax standards resulting in different criteria for including loans in particular categories. The result was that categories carrying the *same name* in different jurisdictions or different banks mostly did not refer to loans with the *same degree* of creditworthiness. For the surveyed banks, they discovered that the internal categorisation systems could be highly idiosyncratic, as when based on the Internal Ratings Based (IRB) models. In sum, the Basel Committee identified multiple layers of credit risk categorisation, those used for banks' internal credit risk categorisation, those used for regulatory and supervisory credit risk categorisation, and those used in the accounting frameworks for financial statements. They found that similar loans fell into different categories in various jurisdictions, but they did note equivalence in some cases.

Moreover, widely-used commercial sources for NPL data, such as *The Banker* and *Bankscope*, report different measures of NPLs, and hence different representations of balance sheet health (Bholat *et al.*, 2016). Using its own survey sent to the top 1,000 bank holding companies on a global basis and cross-checked against publicly disclosed data, *The Banker* reports the ratio for NPLs to Gross Total Loans, where NPLs are defined as all loans that are overdue for longer than 90 days. *Bankscope*, which covers 29,000 private and public banks globally over more than 15 years, reports impaired loans sourced from banks' annual reports and accounts, which are all loans that have a specific impairment against them. Bholat *et al.* quote *Bankscope*'s caveat: "'there is no conformity to defining impaired loans, both across country and intra-country' because all accounting standards 'are vague in their definition of when a loan is impaired' and because 'management discretion can change from one year to the next within a particular bank'". Bholat *et al.* contrast from the two sources these data, and the loan loss provisions to gross loans (respectively, loan loss reserves to gross loans), and the over/under-provisioning' ratios for G-SIBs. They conclude that the absence of a common benchmark means that different policy conclusions can be reached - such as whether banks were adequately provisioned at the onset of the GFC in different jurisdictions.

3.3 A clarifying typology for the evolution of NPL concepts in South Africa

To assess the possibilities for *consistent* estimation of NPL models, for enhancing the financial stability linkages in the Core model and for early warning forecasting, the changing definitions of the different NPL variables, their span, their possible disaggregation, must be clarified.

South Africa has used several different terminologies for “non-performing loans” since the 1990s. The terminology used in South Africa does not always coincide exactly with terms used by the BCBS or other international agencies, for example, this is the case for ‘impaired advances’, see below. South Africa has three types of bank-level NPL data (over varying dates): impaired advances (related to an accounting definition); 90-day-overdue, 180-day-overdue or past due (typically 60-day-overdue) loans (a prudential/supervisory concept); and default ratios from Internal Ratings Based (IRB) models for authorised banks (following the implementation of Basel II). Authorisation of banks was progressive, beginning in 2008, so that more banks have been added to the default ratio series over time (numbering five banks as of 2021).

In Table 4, we create a *typology* of the definitions, definitional changes, date spans and possible disaggregation for South Africa’s NPL data, as well as for a loan loss provisions concept, credit impairments. The credit impairments measure is the *stock* of loan loss provisions accumulated over time. Most of these data are sourced via bank survey forms[123] from the Prudential Authority (formerly the Bank Supervision Department of the SARB). As regulatory return data, all three NPL concepts, and the credit impairments, have been subject to definitional changes from changes in regulation from time to time. The main regulatory changes are indicated in Table 4.

One important focus of this paper is the possibility of constructing consistent series for time series analysis of NPLs; this is discussed in detail in Section 3.4 and summarised

[123] Bank survey forms: the *aggregated DI and BA returns* data for the South African banking sector can be found at the SARB websites: BA200 from June 2008 onwards on the website (but from January, 2008, communication Prudential Authority); BA210 available from March 2019 on the website, (but from March, 2008, communication Prudential Authority); and for January, 1994 to December, 2007, D1500, available on the web.

in Table 4. The current data can be accessed from 2008 from the Prudential Authority, SARB (but some detailed asset classes are only available from 2012 due to a change in regulation). The data are largely complete for large and publicly-listed domestic banks covering over 90% of bank assets, but less complete in the 'other bank' data, which include smaller and less diversified domestic banks whose loans tend to be dominated by unsecuritised lending, and branches of foreign banks.

A monthly time series for 'credit impairments' from 1991 (September) is published in the Quarterly Bulletin[124]. This series was used from 2007:Q3 by the macro-models' team for an 'NPL equation' listed as a Memo item in the revised Core model (De Jager *et al.*, 2021), though in fact this is not an NPL concept – see Section 5.1. for a review of the equation. We present a new empirical model for the ratio of 'credit impairments' to gross loans and advances from 2001 in Section 5.2.

The main NPL definitions are as follows. *Impaired* advances are the total value of the advances in respect of which banks have raised a specific impairment (provision) and includes any advance or restructured credit exposures subject to amended terms, conditions or concessions that are not formalised in writing. A specific provision means any *impairment, allowance or provision* made against losses on a debt that has been specifically identified as bad or doubtful (an incurred event). Impaired advances are expressed as a percentage of on-balance-sheet loans and advances.

The conventional usage of the term 'impaired advances' is as an *accounting concept* (see Bholat *et al.*, 2018; EBA, 2017; BCBS, 2017). The term in South Africa refers to *more* than the accounting definition, however: it overlays prudential guidance onto the accounting definition using qualitative criteria. Hence, the set of loans encompassed by South Africa's term, 'impaired', will exceed those of the pure accounting definition of 'impaired'. This overlay is the typical procedure globally, see also Baudino (2018).

Impaired advances' are collected from bank survey forms, and are assessed by the banks themselves, embodying discretionary thresholds, when banks evaluate categories of default. The reporting of "impaired advances" followed South Africa's

[124] The monthly series for 'credit impairments' is entitled: "Assets of banking institutions: specific provisions in respect of loans and advances", with code: KBP1123M, sourced, since 2008, from the PA's BA900 survey which covers the balance sheets of private banks.

implementation of Basel II with effect from 1 January 2008[125]. Since the reported *loan loss provisions* back to the 1990s were set against *some definition* of an *impaired loan,* it is possible that a reasonably consistent series for the underlying impaired loans, with some definitional changes, could be constructed at least to 2001. Data availability and continuity for 'impaired advances' are discussed further in Section 3.4.

There are two other 'non-performing loans' concepts, but each with some differences. As noted in the Introduction, *overdue loans* are currently defined as all exposures overdue for more than 90 days and where the recovery thereof was considered to be doubtful, expressed as a percentage of on-balance sheet exposures[126]. Before a change in banking regulations implemented in 2001, data were collected from banks, via the bank forms above-mentioned, on 'overdue advances' classified into months overdue categories such as 0-1, 1-3 and more than 3 months overdue. Overdue advances were reported in the *Annual Bank Supervision* reports from 1994 to 2007[127]. These quarterly data apply to the different credit products such as mortgage loans and instalment finance. Following bank regulatory changes in 2012, bank reporting forms have required quantitative information on 90-day overdue loans and advances from the banks following the standardised approach to credit risk reporting, see below. From 2008, quantitative information on 90-day overdue loans has been required from the banks following the Internal Ratings Based (IRB) approach, see below. Data availability and continuity for overdues are discussed further in Section 3.4.

---

[125] The term is first used at the SARB in 2008 in the *Bank Supervision Annual Report* of that year, and in the second of the two *Financial Stability Reviews* of 2008. The monthly data based on this definition go back to 2008.

[126] According to Chapter 7 of Government Gazette No. R. 1029, 12 December 2012, the definition of an "overdue amount" is given (p.1212) as:
(a) an overdraft facility includes an amount due by a person who has exceeded an advised limit or has been advised of a limit smaller than the current outstanding amount;
(b) an amount payable in instalments or in relation to bills issued in a series, includes the full amount not yet written off, outstanding under the transaction concerned, including, in the case of an amount payable in instalments, such instalments not yet due and penal interest, if any, incurred in respect of overdue amounts, but excluding, in the case of an amount payable in instalments or of bills issued in a series, interest not yet due, if-
(i) the relevant account has vested in the hands of a third party for collection; or (ii) the debtor has become subject to an administration order, has surrendered his estate, has entered into a compromise with his creditors, has been put under judicial management, is wound up or sequestrated or has been declared insolvent; or (iii) the reporting institution considers recovery of the debt for any reason doubtful or has identified the debt as a non-performing debt;
(c) an amount not payable in instalments, including an overdraft facility, includes-
(i) any amount the recovery of which the reporting institution for any reason considers doubtful; (ii) any amount in respect of which the reporting institution has identified the debt concerned as a non-performing debt; or (iii) the full amount, not yet written off, outstanding under the transaction concerned if any of the circumstances contemplated in subparagraph (i) or (ii) of paragraph(b) become applicable.

[127] The earliest *Bank Supervision Annual Report* on the web is for 1994. Havrylchyk (2010) reports regressions based on "overdues" at least from 1994, see Section 5.2. Because of the amendment of regulations, the pre- and post-2001 data on overdues are not quite comparable, as the 2001 Annual Report explains. Strictly speaking, the 'credit risk buckets' introduced for bank reporting in 2001 (e.g., see IIF definition of these in Table 3), do not exactly match the different month overdue quantitative timing concepts, due to the qualitative prudential features overlaying these.

Finally, *defaulted exposures* are reported by those banks that were authorised to use the internal ratings-based (IRB) approach from 2008, see BCBS (2001a; 2004), to calculate their minimum regulatory capital for credit risk, using the 90-day definition. The default ratio is defined as the ratio of defaulted exposure as a percentage of total exposure at default. These cover private households and several non-financial corporate sub-sectors, like real estate. In 2008, four banks were first authorised to use basic or more advanced versions of this method, and the number has since expanded to five banks; the remaining banks use the standardised approach to credit risk rating from 2008, see BCBS (2001b). Default ratios are also constructed for the banks following the standardised approach to credit risk reporting. These use the sum of the three credit risk buckets: 'sub-standard', 'doubtful' and 'loss', see Appendix, to define 'default'. The aggregate default ratios for all banks have been shown at least from 2010 in the *Financial Stability Review*, as a total, and for retail and corporate sectors. Since all the banks followed a similar approach to credit risk rating from 2001 to 2007, the different segments of *total* default rates may be able to be linked over time, making default rates potentially the most promising of the three NPL measures for a consistent historical series, see next section.

As noted in the Introduction, the credit risk indicators here considered are stock ratios, but data on new flows into these stock measures can provide additional information beyond that contained in changes in stocks. Current and historic bank reporting forms ask for information on write-offs and debt recoveries that could, combined with changes in stocks, be used to generate flow measures with potentially useful credit risk information.

3.4 A proposal for connecting NPL concepts in South Africa for continuous data series

There are effectively four credit risk indicators for which South African time series data exist or could be assembled, see Section 3.3 and Table 4. For modelling aggregate indicators, it is important to include data for the volatile period from 2001 up to 2007, for drawing robust economic insights. Aggregate data post-2008 may not give a long enough time span for robustness in time series analyses, though panel data analyses for individual banks and loan classes may be informative from a modelling perspective, compensating for the lack of historical information.

We propose how pre- and post-2008 data on three different NPL concepts might be joined to permit an analysis of data back to 2001 on reasonably consistent definitions. The only published time series for a credit risk indicator for SA is for 'credit impairments', a loan loss provision, which is not an NPL concept. *Credit impairments* is published back to 1991 under the code KBP 1123M. These data have been affected by at least three definitional changes, see Table 4, in 2001 with a change in bank regulations to take account of credit quality in impaired loans, in January 2008 with the implementation of Basel II and the IRB basis, and in January 2018 with the move from the IAS 39 to the IFRS 9 accounting standard. The post-2008 credit impairments data need to be compared with the pre-2008 'specific provisions' data from the DI500 banking form and reported in the Banking Supervision reports, as significant data discrepancies between the pre- and post-2008 data have been noted by us. We understand that the inclusion of 'general provisions', from 2008, accounts for much of the jump in the credit impairments data[128]. General provisions are those against performing loans and against those loans in early arrears (or with a significant increase in credit risk)[129]. Given that the 'impaired loans', see below, are defined as those against which a 'specific provision' has been made, ideally the 'general provisions' should be separated out from the credit impairment series for improved modelling purposes.

Overdue loans

Turning to the three broad measures of non-performing loans, one hope for an NPL concept that *could* be free of these regulatory shifts is a 90-day *overdue loans* concept. Havrylchyk (2010) implies that 90-day overdue data are available from 1998 to 2008 for three loan classes: mortgages, instalment finance and other loans and advances, but does not give a source. The annual Bank Supervision Reports, which began to be published on the web in 1994, provide an age-analysis (0-1 months, 1-3 months, and >3 months overdue) from 1998 but the 2000 Report is the last to provide these data. Before the year 2001, the *total* overdue loan measures published in the Banking Supervision reports refer to 90 days plus.

[128] The pre-2008 data are defined as 'specific provisions', while from January 2008 they include 'specific and general provisions', see Table 4.

[129] Communication from the Prudential Authority: although pre-dating the IFRS 9 implementation, these two 'general provisions' categories *notionally* correspond to IFRS 9 Stages 1 and 2, respectively, whereas the 'specific provisions' corresponds to the IFRS 9 Stage 3.

From January 2001, new bank regulations were implemented which took account of credit quality in impaired loans, see Table 4. In the 2001 Banking Supervision Report, there is a significant step-change down in the total overdue loans reported (see Figure 30, p.62 of the report), and the report states that only a *narrower* category of loans, classified as 'doubtful' and 'loss', qualify as 'overdue' for reporting purposes. This classification does not correspond precisely to a days-overdue definition. The corresponding data for the new reporting structure are shown on the published DI500 form (with the final of these forms dated December-2007) with categories: 'standard/current, special mention, sub-standard, doubtful and loss'. However, these categories do not translate exactly into simple 'days-overdue' concepts because there is a significant qualitative aspect as part of the definition of credit quality (see excerpts in the appendix). Regulatory changes in 2012[130], require data under the standardised approach to be collected showing what *proportion* of each category comprised loans on a per days concept, like 'greater than 90 days'. These types of detailed overdue per days data were also collected in 2008 to 2011 for IRB-authorised banks but not for banks using the standardised method. The banking forms that underlie the aggregate data published in the DI500 spreadsheets and in the annual Banking Supervision Reports, reveal that the 'credit risk buckets' used between 2001 and 2007 were not supplemented by requiring the quantitative fraction of loans in each risk bucket to be provided[131]. This means that there is a gap in the 90-day overdue data between 2001 and 2007, and for the banks reporting on the standardised approach, also until 2011[132]. But if an approximation could be made between the 90-day overdue measure and the credit risk buckets, it is possible that a longer series approximating the 90-day overdue measure could be constructed. However, this is uncertain[133]. The graphs of overdue ratios for households published in the FSRs indicate no jump in 2018, suggesting that these data may be immune from the change in the accounting concept in 2018[134].

[130] See 12-Dec-2012, Government Notice no 1029.
[131] Unlike in the BA200 forms used from 2012.
[132] Quarterly data on "special mention" "substandard", "doubtful" and "loss" are available on banks reporting on the standardised approach for the period March 2008 to December 2011 (communication, Prudential Authority).
[133] The Annual Supervisory Report for 2000 indicates for 2000Q4, R11,189 million in overdue mortgages, of which 1,377 were up to one month overdue, 1,322 were one to three months overdue, and 8,491 were three or more months overdue. The DI500 spreadsheet for the beginning of the month of January 2001 reports for mortgages, R5,320 million in the 'special mention' credit risk bucket, 4,888 in 'sub-standard', 4,399 in 'doubtful' and 5,125 in the 'loss' bucket. It is not obvious what combination of the last three risk buckets could approximate the 90-day overdue measure but is seems likely that the sum of 'doubtful' and 'loss' is closer than other summations.
[134] The accounting and the regulatory definitions appear to have coincided at this time (Prudential Authority communication).

The default ratio

For another NPL concept, the *default ratio,* data are available from 2008, but can be linked quite closely at the aggregate and bank-by-bank level for continuous series back to 2001. We conclude that default ratios are likely to be the best of the alternative concepts for harmonising information over time. For 2001 to 2007, default ratios can be defined for all banks by summing the three credit risk buckets: 'sub-standard', 'doubtful' and 'loss' to define 'default'. For banks continuing on the standardised approach after 2008, there should be no jump in the data in 2008. For the larger banks that switched to the IRB method (four of them in 2008, and now five banks), the default rates defined by the IRB method can be linked with the preceding standardised method. By comparing, bank by bank, the last month on the *standardised approach* with the first month of the *IRB approach*, it should be possible to link the two sets of data. Indeed, by looking at monthly trends just before and just after the switch it should be possible to calculate adjustment factors. Carrying out this bank-by-bank exercise would be very useful.

There is apparently no jump in the default ratio in 2018 when the new accounting rules came in, according to Figure 47 of the November 2020 *Financial Stability Review*, while credit impairments and 'impaired advances' jump sharply. This suggests that the default ratio, constructed as above on a bank-by-bank basis will provide the most consistent aggregate NPL data back to 2001. Having a consistent NPL concept is extremely valuable for modelling purposes. Moreover, from 2008, the default ratio can be constructed for different loan classes for additional insight.

For mortgages it may be possible to link another two data sources. There are the 90-day overdue series from 2007:Q3 from the National Credit Registry (NCR) data base on household mortgages. While the data on loan quality categories from the DI500 source back to 2001 are available for all mortgages, they will be dominated by household mortgages. It may be possible to link some aggregate of the loan quality categories for mortgages from the DI500 source with the NCR data, particularly as there is an overlap between the two in 2007:Q3 and 2007:Q4. However, the two series would be only approximately comparable.

Impaired loans

The third possible NPL measure is the accounting concept of *impaired loans*, on which the loan loss provisions/credit impairments (captured in the series KBP1123M) are based. This terminology was introduced in 2008. To try to extend impaired loans data back to 2001, requires linking the published impaired loans data from 2008 (defined on the Basel II definition of 90-day overdue plus unlikelihood of payment, see Section 3.1) with an aggregate of the credit risk categories or 'buckets' in the DI500 data. From 2008, when the banking form switched to the BA200 returns, to 2017, there appear to be temporally-consistent aggregate data with the revised concept under Basel II of what is an impaired loan. However, the further revision in 2018 with IFRS 9 has affected the impaired loans measure, with a large increase reported between December 2017 and January 2018. It appears, therefore, that while an aggregate series could be compiled for these impaired loans back to 2001, switches in data in 2008 and 2018 would have to be corrected for. This suggests that the impaired loans data are likely to be the least satisfactory for temporal consistency of the three measures potentially available for South Africa. We would also expect the long-run drivers, after correcting for shifts in definitions in 2008 and 2018, to be similar to those found for the credit impairments ratio in Section 5.2, though the short-run dynamics could differ somewhat. The BA200 data only give the aggregate of impaired loans with no breakdown by sector or loan type, though, of course bank by bank data should be available to the SARB.

Summary

In summary, we recommend that the SARB publish time series data for at least two of the NPL measures for the 2001-2007 period and 2008 onwards, with clear documentation. A special background paper should give transparent methodological detail on joining, using a bank-by-bank basis, the different time segments for the various NPL concepts. For the *default ratio* NPL measure, for all banks (regardless of whether they use the standardised or IRB approaches), the pre-2008 data require joining with the post-2008 data. For the *90-day overdue ratio* NPL measure, for banks using the standardised approach, the pre-2012 data require joining with the post-2012 data; for IRB banks, data from the last month of their using the standardised approach needs to be joined to the first month that they use the IRB approach. Thereafter, the aggregate NPL time series data, with appropriate qualifications and explanations of methodology, should be routinely published.

4 The Drivers of NPLs: Insights from the International Literature

NPLs are an important ingredient in the financial accelerator in which worsening credit conditions amplify adverse shocks to the economy. General equilibrium models with a financial accelerator go back to Bernanke and Gertler (1989), Bernanke, Gertler and Gilchrist (1996, 1999) and Kiyotaki and Moore (1997). In Bernanke, Gertler and Gilchrist, fluctuations in the external finance premium of firms can lead to an amplification of business fluctuations. In Kiyotaki and Moore, the need to collateralise debt on the value of land leads to cyclical fluctuations in the price of land, which can amplify the economic cycle. However, none of these papers feature loan defaults except in a fairly rudimentary form, through a costly state verification mechanism. There are partial equilibrium models of aspects of the financial accelerator, for example of bank runs, see Diamond and Dybvig (1983) for the seminal paper, and of network effects, see Allen and Gale (2000). Links between NPLs and household decisions are made by Lawrence (1995) in a two-period model of household choice with a default option, which, if exercised, restricts subsequent access to credit. Such models suggest that households facing higher unemployment risk tend to have higher rates of default. There is a closely related literature on mortgage defaults. Kau *et al.* (1992) develop an option pricing model of the default decision, while Vandell (1995) questions the 'ruthless default' implications of that literature. This led to the double-trigger model of the default decision, see Elmer and Seelig (1998) for an early exposition, in which either or both cash-flow problems and a negative equity position motivate the default decision[135].

In empirical work, comparative analyses of the drivers of NPLs have been compromised by the lack of a universal standard to classify non-performing loans across countries. This applies both across countries and regions, and for panel studies of banks within a single country, as documented in Section 3. Studies of the dynamics of NPLs are limited by the number of available years of data, but often also by inadequate data in earlier years. Both panel and time series studies are challenged by

[135] In the UK housing context, using aggregate and county court regional data, arrears and repossessions were modeled following the double-trigger model in Aron and Muellbauer (2016; 2011).

the evolution of accounting and prudential standards of NPL classification so that the dependent variable is not consistently defined.

The empirical determinants of NPLs are loosely drawn from the financial accelerator and lifetime consumption theories. Our aim in this section is to extract the relevant independent variables used to model various constructs of NPL definitions (such as peak NPLs, or percentage of NPLs above a threshold, or time to peak NPLs) and to identify any consistent patterns across the most convincing diverse studies. We are not aiming to conduct a thorough critical survey of these studies, and we note the drawbacks of poor data comparability in comparative studies and the exclusion of relevant drivers, and possible asymmetries and non-linearities in most studies.

4.1 A Benchmark Cross-sectional Analysis: Macro-, Banking- and Corporate-determinants

A comprehensive paper by Ari *et al.* (2019) provides a useful benchmark against which other studies can be compared. They construct a new dataset on NPLs for 78 countries from 1990, covering 88 banking crises, and reporting NPLs for an 11-year window, three years before and seven after the crisis[136]. The authors source annual NPL data from IMF's Financial Soundness Indicators (FSI), and where data are missing, they use hand-collected data from IMF Staff Reports or from the official statistics of the national authorities or other national sources. The authors attempted to adjust for NPL definition differences across data sources to ensure consistency within countries[137]. Across countries, however, the same concerns about poor comparability of the data for NPLs remain (Section 3).

Ari *et al.* (2019) present predictor models that use pre-crisis independent predictor variables, measured as averages or cumulative changes over the five years prior to the crisis, with constructed dependent NPL variables, dated on or after the crisis date. Regressions are conducted for five constructs of NPLs (and some variation of these)

[136] Two earlier but related datasets are used by Laeven and Valencia (2013, 2018), that cover only peak NPLs during banking crises; and Balgova et al. (2017), where the data are criticized as being inferred mainly from the commercial provider Bankscope, concentrating on larger banks only and hence not fully representative and with poor coverage in earlier years.

[137] When extending the data from a more prioritized source using a less prioritized source, they multiplicatively rescale the less prioritized source to match the more prioritized source in the first overlapping year. They also only combine data sources when their definitions are consistent, and the data discrepancy is minor.

on three sets of independent variables, sourced from the literature. The five constructed dependent variables and the three sets of predictor variables are presented in Table 5[138]. They use a form of general-to-specific selection to select the most informative combination of predictors for each NPL metric in each case: this is the machine learning approach called "post rigorous least absolute shrinkage and selection operator" ("post-r-lasso"; Belloni *et al.*, 2012; Belloni and Chernozhukov, 2013)[139].

The results of the Lasso statistical selection exercise are as follows[140], which potentially suggests a reduced set of indicators for NPL risk monitoring. Higher NPLs (*dependent variable 1*) are predicted by a lower pre-crisis GDP per capita (proxying for weak institutional strength) and a higher corporate debt-to-assets ratio (where higher corporate leverage reflects weaker corporate sector conditions); and, for the alternative less stringent definition of elevated NPLs, see Table 5, by a rise in domestic credit to private sector. The peak (elevated) NPLs (*dependent variable 2*) are higher in countries with weaker banking (there is a rise when the bank return on assets falls) and a shorter corporate debt maturity (capturing a weaker corporate sector); and, for the alternative definition, the peaks are reduced if there is exchange rate depreciation against the USD (reflecting the competitiveness hypothesis, see Table 6).

The time relative to the crisis start year defined by Laeven and Valencia (2013; 2018)[141] to the NPL peak (*dependent variable 3*) is reduced by a depreciation of the exchange rate against the USD and also by abandoning an exchange rate peg prior to the crisis, interpreted as reflecting the cushioning effect of floating exchange rates in facilitating adjustment or reflecting a timelier policy response overall. The time to peak is also reduced by a higher pre-crisis GDP growth (suggesting better debt management capacity for banks and debtors) and a lower pre-crisis government debt-to-GDP ratio, interpreted as capturing fiscal space. An increase in domestic credit growth (reflecting

---

[138] The candidate predictor data are standardized to Z-scores (i.e., zero mean and unit standard deviation across banking crises).

[139] The Lasso (least absolute shrinkage and selection operator) regression analysis method aims to enhance the prediction accuracy and interpretability of the resulting statistical model, by requiring the sum of the absolute value of the regression coefficients to be less than a fixed value, which forces certain coefficients to zero, thereby excluding them. Castle et al. (2020) argue that Lasso struggles with negative correlations, as negatively correlated variables need to enter jointly as they may not matter much individually. This is also a problem for step-wise regression.

[140] Subject to qualifications re data comparability, possible omitted variables and non-linearities, and the short-comings of Lasso methodology.

[141] It is not clear from Ari *et al.* (2019) how the start year was chosen.

the adverse effect of credit booms) lengthens time to peak, as does higher longer debt maturity. Combined with regression 2, this implies that short-term corporate debt leads to a more rapidly reached peak NPL, and to higher peak NPLs. An increase in unemployment lowers the time to peak[142].

The *final two dependent variables* reflect how soon, relative to the start year of the crisis, NPLs are resolved, and the likelihood of NPL resolution, defined by a dummy equal to 1 if the NPL ratio is reduced below 7% within 7 years after the start of the crisis. Material to both are lower pre-crisis government debt and lower credit growth. The likelihood of NPL resolution (*dependent variable 5*) is also higher with higher growth, after exchange rate depreciation, and with unemployment increase (interpreted as due to the pressure to resolve the debt sooner) and promoted by a high bank non-interest-income-to-total-income ratio, interpreted as proxying for profitability and good management. It is lowered by a higher pre-crisis current asset to liability ratio (suggesting that liquid assets held by borrowers reduce banks' incentives to write off debt). With the alternative definition, defined by a dummy equal to 1 if the NPL is reduced by at least 25% relative to the peak within 7 years, resolution likelihood is promoted by a higher corporate debt-to-asset ratio (reflecting weaker corporate sector conditions). The latter could mean that when companies are in trouble, there is greater pressure for timely resolution.

The authors demonstrate therefore that better ex-ante macroeconomic, institutional, corporate, and banking sector conditions and policies may assist in reducing NPL vulnerabilities during a crisis. These authors also examine in a dynamic context whether post crisis output is affected, respectively, by elevated NPLs, and by the resolution of elevated NPLs. They find output to be on average lower in crises with elevated NPLs and in countries with unresolved NPLs. This suggest that elevated and unresolved NPLs are associated with more severe post-crisis recessions.

---

[142] In principle, these results are hard to interpret. A longer time to the peak may suggest a worse outcome or a better outcome depending on how high the peak is. A long time to a low peak might be a sign of a resilient economy and banking system. A fast time to a high peak may indicate a severe crisis then resolved by aggressive intervention e.g., by the fiscal authority or by an internationally-organised bailout or a rapid improvement in the international economic environment.

4.2 An Overview of Past Work from a Meta-study and a Survey

There is an overlap between the determinants considered in the reviews below, but the Ari *et al.* (2019) single study has a larger set of variables than most.

A *meta-study* of NPL ratio determinants by Macháček *et al.* (2017) is reported here, though with some scepticism of its usefulness. These authors chose 37 studies, all prior to 2014, of which seven predated the GFC, two thirds (24) investigated an individual country (as a time series model or a bank-level panel data model), and the remainder (13) used cross-country panel data models. About two thirds of the studies apply dynamic model specifications (with at least one lag of the dependent variable, the non-performing loans ratio). The studies were selected to contain at least one of the following of the five most common macroeconomic determinants of the non-performing loans ratio: *real economic growth*, the *interest rate*, *inflation*, the *exchange rate* and *unemployment*. For three variables, the empirical evidence is in line with the theoretical assumption: on the effect of real economic growth, a majority of (significant) coefficients had a negative value (real growth increases the ability of debtors to pay off debts); a majority found the effect of interest rates to be positive (debt service costs of loans with variable rates are especially sensitive to an increase in interest rates); and similarly, a majority found unemployment to have a positive effect. But for inflation and exchange rates, the results were ambiguous, with equal numbers of (significant) coefficients having both signs, reflecting potentially offsetting theoretical effects. However, the meta-study compares different periods, countries, definitions and methodologies, with obvious room for specification errors. Indeed, in an analysis of the discrepancies across the studies, factors such as data specification, estimation method, number of countries and observations included in the model played a significant role.

A more comprehensive review from the finance literature, Naili and Lahrichi (2020) – henceforth 'The Review', explored 69 studies published between 1987 and 2019 in 40 peer-reviewed journals, with over 70% concerning the period post-GFC. Methods range from simple regressions to dynamic panel models. Two-step general methods of moments was used in 44% of the reviewed papers, while the rest mostly relied on Pooled Ordinary Least Square or Two-Stage Least Square models. The determinants of NPL ratios from the studies covered are included in our Table 6 (this table had to be constructed from the body of the Review, as the empirical conclusions were not made precise in a tabular form). The first five of the variables were also considered in the meta-study. The variables considered in the predictor models of Ari *et al.* (2019), and the resultant signs from the Lasso statistical selection study, are given in the final column[143].

We summarise from the meta-study, the Review and the comprehensive study of Ari *et al.* (2019) whether there is consensus on the relevance and direction of the effect of the independent variables covered in the table. The advantage is to contrast a *generalised view* from the two reviews, which between them cover about 100 studies (the overlap is only seven studies), with the more specific findings from Ari *et al.* (2019). The outcomes from the two reviews, however, must be treated cautiously. They do not properly assess the quality of the studies but merely quote the findings. They give little information about comparability of the dependent variable NPL measures and methodologies, including how forward-looking, i.e., predictive, were the different studies[144]. Further, different methods are conflated, and different countries or different set of banks, and institutional features, and they cover differing periods.

Several <u>macro-economic factors</u> were considered. *GDP growth* is used to indicate the state of the business cycle, with bad loans increasing during slowdowns (reflecting asset prices and employment worsening). *Unemployment* captures the difficulties borrowers with uncertain income face in servicing their debts. *Inflation* has ambiguous results. On the one hand, inflation erodes real incomes, reducing the ability to service debt obligations, especially with variable interest rate loans. On the other, inflation erodes the value of outstanding debt, and debt is more sustainable if wages exceed or keep pace with inflation. Many find no significant inflation effect, e.g., in CESEE and EU countries. A rising *policy interest rate* raises the lending rate, constraining borrowers' ability to repay, for floating rate loans. *Exchange rate depreciation*, measured both in real and nominal terms, worsens banks' loan quality as unhedged borrowers with foreign currency-denominated debt face raised debt servicing costs in

---

[143] To recall, the five dependent variables of Ari *et al.* (2019) are elevated NPLs (*Dep 1*), the peak NPLs as a percentage of total loans (*Dep 2*), the time to reach the NPL peak (*Dep 3*), the time to resolve NPLs (*Dep 4*), and the likelihood of resolution within 7 years (*Dep 5*). Three sets of variables are used, the first in *Set 1* comprises macro-variables, then in *Set 2* appended by banking variables, which are in *Set 3* appended by non-financial firm/industry variables. The most complete general-to-specific results therefore would be when *Set 3* is included.

[144] That said, the Table 6 results for the Nailu and Lahrichi (2020) review are apparently based on the more cited of the referenced studies.

local currency. Without the mismatch but with high export volumes, a depreciation could favour loan quality through the growth and competitiveness channel. An adverse *public debt position* worsens NPLs with governments having less fiscal space to intervene and in the extreme, a sovereign debt crisis can signal a bank crisis. Banks' creditworthiness is impacted by poor a sovereign rating, curtailing lending which limits refinance. Several studies examine the *institutional framework* with the hypothesis that credit quality is improved where there is little corruption, sound regulatory frameworks and accountability. At least one study found that corruption promoted NPLs.

Comparing the above macro-economic factor findings with Ari *et al.* (2019), the same result is found, that if (*pre-crisis) GDP growth* is higher, this reduces the time to the NPL peak (*Dep 3*) and increases the likelihood of NPL resolution (*Dep 5*). Similarly, high unemployment reduces the time to the peak NPL (*Dep 3*) but increases the likelihood of resolution (*Dep 5*) - interpreted as due to the pressure to resolve the debt sooner. However, neither the *inflation rate* nor *interest rates* were selected by the Lasso statistical model. (Nominal) *exchange rate depreciation* or *abandoning an exchange rate peg* prior to the crisis the reduces the time to reach the peak (*Dep 3*), interpreted as reflecting the facilitating effect of floating exchange rates in adjustment, and by the same token increase the likelihood of resolution (*Dep 5*). However, the appendix of Ari *et al.* (2019) with an alternative specification for the dependent variable suggests that depreciations and floating exchange rates also predict lower peak NPLs[145]. There is also correspondence with the general findings for higher *(pre-crisis) government-debt-to-GDP ratio,* which increases the time to the peak NPL, reflecting less fiscal space, increases the time to resolve NPLs (*Dep 4*) and reduces the likelihood of resolution (*Dep 5*). Ari *et al.* (2019) use higher *GDP per capita* to proxy for institutional strength which reduces the probability of elevated NPLs and concurs with related findings in Table 6.

Neither the meta-study nor the Review considered *private credit extension*, as did Ari *et al.* (2019). Instead, the Review examines loan growth as a banking sector variable, see below. Private credit extension features strongly in most of the models of Ari *et al.*, with the findings that a rise in domestic credit to private sector elevates NPLs (*Dep 1*),

[145] The Appendix, page 19, gives a *negative* sign for an exchange rate depreciation in Panel B, but the text in the main paper says "Note, however, from Panel B, that depreciations and floating exchange rates *do not* (our italics) predict lower peak NPLs, possibly due to currency mismatch-associated losses in firms and banks." However, for this alternative definition of the peak NPL, the coefficient is not significant at the 5% level.

lengthens the time to the peak NPLs (*Dep 2*); lengthens the time for NPLs to be resolved (*Dep 4*), and reduces the likelihood of NPL resolution (*Dep 5*).

None of the above three studies considered the housing market. The US is a useful model for South Africa in this respect, where the housing market and associated changes in house prices are also likely to be an important NPL determinant. In the US studies of Ghosh (2017)[146], as for Beck *et al.* (2013), changes in the housing price index are included as a potential macro-determinant. Rises in house prices are expected to reduce NPLs, especially for the real estate sub-sector NPL. The mechanism through which this operates is via a wealth channel, since rising house prices raise property wealth, helping borrowers cope with unexpected adverse shocks or to refinance their mortgages by boosting the value of their housing collateral. Ghosh (2015, 2017) confirm the fall in NPLs with higher house prices for both real estate NPLs and individuals' NPLs, capturing the countercyclical nature of these types of loans and the effect of house prices on collateral values.

Of the <u>bank-specific variables</u>, on *bank capitalisation*, several studies, e.g., Klein (2013) and Makri *et al.* (2014) in their panel of European nations, find that a high capital adequacy ratio (CAR) reduces NPLs, with the interpretation that with more capital at risk, banks are likely to engage in prudent lending with adequate loan screening. Two prominent studies on the US, Ghosh (2017) and Ghosh (2015), however, find the opposite result, that managers in banks that are highly capitalised may engage in a liberal credit policy leading to rising NPLs.

Studies on *bank size* offer no clear-cut evidence. "Too big to fail" banks may take excessive risk, but potentially, large-sized banks with modern risk management systems and procedures may be better able to conduct risk screening of loans. On *bank efficiency*, cost inefficiency appears to be linked with poor management and worsening NPLs, according to several studies. "Skimping" on good loan quality with apparent cost efficiency, may appear in burgeoning NPLs in the future. Shocks which raise NPLs at the same time decrease apparent efficiency because extra costly managerial operations will be required. On *bank performance or profitability* several prominent studies (e.g., Louzis *et al.* (2012), Ghosh (2015) in the US, and Makri *et al.*

[146] Ghosh (2017) defines the NPL ratio is defined as the sum of total loans and leases past due 90 days or more and non-accrual loans, divided by total (gross) loans.

(2014) in Europe), find that higher profitability reduces NPLs, as there is less inclination for excessive risk-taking with resultant higher quality loan portfolios. But the opposite sign is reached in studies where banks game the market to conceal bad loans by a liberal credit policy for current gain, but with subsequent escalation of NPLs. On rapid *loan growth*, the results concur with those of Ari *et al.* (2019) on private credit extension. Rapid loan growth is linked to riskier lending behaviour (Keeton and Morris, 1987), through adverse selection, inappropriate managerial incentives and reduced screening standards in boom periods, worsening credit quality. The short-term easing of credit quality promotes short-term profits at the expense of heavy future losses. On *bank diversification*, a limited literature suggests greater diversification increases risk from inexperience and lack of comparative advantage, leading to business failures and worsening NPLs.

Evidence on *executive compensation* and a direct link with NPL levels is limited; various studies contend that while higher executive compensation could increase managerial risk-taking in risky banks, more option-based compensation would likely increase managerial risk aversion. *Managerial overconfidence* may reduce risk aversion, which can be detrimental, though possibly temporarily rewarding in a boom period. US evidence using the stock options proxy for overconfident management, found an association with excessive risk-taking, relaxed lending standards and increased bank leverage resulting in high levels of NPLs (Ho *et al.*, 2016). There are few studies on *corporate social responsibility*, but the evidence tends to suggest CSR-banks incur lower NPLs. Similarly, there are few studies directly on *ownership identity* and NPLs, but there is some evidence that state ownership is linked to high risk-taking and poor performance while institutional ownership reduces NPLs. On *ownership concentration*, the results are conflicting. Dispersed ownership has been linked with poor incentives for proper monitoring of banks, but also more risk taking by controlling interests prioritizing personal interest. There is evidence from the US, China and global studies that concentrated ownership reduces bank risk exposures, and significantly reducing NPLs. Other cross-country studies find poorer performance, lower cost efficiency, higher risk-taking and worse loan quality. On *banking industry concentration*, the outcomes in the literature appear to be ambiguous. Some evidence supports the competition-fragility hypothesis of Keeley (1990), which suggests that competition increases NPLs since the ensuing lower profit margins decrease the discounted net value of banks and raise their risk tolerance. Wang (2018) finds for US aggregate data that the level of future NPLs increases in competitive markets. A different view supports raised incentives for managers to behave prudently in a competitive environment to improve the perception of good risk management for regulators and investors (Jiménez & Saurina, 2005; Ozili, 2019).

Comparing the above findings with Ari *et al.* (2019), similar findings are achieved for increasing *bank efficiency* which is linked with a more rapid resolution of NPLs (*Dep 4*). On the other hand, the opposite result from the Naili and Lahrichi (2020) overview is found for *bank diversification*: measured by a high bank non-interest-income-to-total-income ratio, and interpreted as proxying for profitability and good management, this promotes the likelihood of NPL resolution (*Dep 5*). Though included as a variable, the *banking industry concentration* measures were not selected by the Lasso statistical model.

The two reviews summarised in Table 6 did not consider non-financial corporate sector variables. In brief, Ari *et al.* (2019) find that a higher pre crisis *debt-to-assets* ratio (capturing corporate leverage and reflecting weaker corporate sector conditions), both drives up NPLs (*Dep 1*) and increases the likelihood of timely resolution of NPLs (*Dep 5*), by their definition. A higher share of *short-term debt in total debt* (capturing corporate debt service capacity and reflecting a weaker corporate sector), both raises the peak NPLs (*Dep 2*) and shortens the time to the peak (*Dep 3*). A higher pre-crisis *current-asset-to-liability* ratio reduced the probability of timely resolution of NPLs (*Dep 5*). Neither the share of foreign assets in total assets nor an alternative corporate debt service capacity measure (EBIT to total interest expense ratio) were selected by the Lasso statistical model.

4.3 Shortcomings and omissions in surveyed work on NPL drivers and loan loss provisions

It is notable that none of the models briefly surveyed above introduced non-linearities into the predictors of NPL ratios, such as asymmetries in the business cycle or in GDP growth, weighting more greatly the falls in growth. Cross-country panel studies with fixed effects do not always consider carefully enough heterogeneneity of slope coefficients between countries. These could be handled with interaction effects for institutional differences such as the exchange rate regime or fixed rate versus floating rate mortgage markets or the depth of the corporate bond market. Moreover, few studies incorporate the full set of drivers recommended for modelling 'growth at risk' by the IMF in Prasad et al. (2019). These consist of three underlying aggregates and the credit-to-GDP gap. These aggregates attempt to capture respectively household sector vulnerabilities, corporate sector vulnerabilities, and housing market imbalances. The measures capturing household and corporate sector vulnerabilities are aggregated from indicators that capture leverage, debt servicing capacity, and indebtedness. Housing market imbalances are aggregated from indicators that measure imbalances from multiple aspects, including house price dynamics, construction activity, inventory and sales, mortgage activity, and household financial strength. The relevance of such drivers can vary across countries, for example, with rates of owner-occupation, leverage, the structure of the financial system, and whether home equity withdrawal is readily available. The growth at risk approach uses quantile regressions which give more weight to periods with probabilities of low or negative growth. NPLs are likely to have a non-linear relationship with growth, with high NPL values particularly associated with recessions, especially ones associated with financial crises. A linear predictive model for NPLs, therefore, is implicitly designed to put more weight on forecasting recessions accurately, than on forecasting variations in normal positive growth rates. Therefore, one should expect similar predictive variables to be relevant in forecasting NPLs using conventional methods as in the growth at risk models based on quantile regressions.

5 Insights from a New Model for Loan Loss Provisions in South Africa

The SARB appears to use data on credit risk indicators in two ways, from a modelling perspective. The recent update to the Core model introducing links with macro-prudential policy, De Jager *et al.* (2021), presents a Memo equation (which plays no part in the Core model) for what it calls NPLs, but which in fact is a simple model for 'credit impairments'[147], a loan loss provisions concept, and not an NPL concept. Second, micro-level stress-testing at the SARB uses a 'bottom-up' approach. The data are approached from the accounting perspective using the "impaired" concept under the IFRS 9, from 2018 onwards. The exercise is performed by the six major banks covering 92% of assets. Individual banks perform stress testing of their

[147] The dependent variable uses the monthly series KBP1123M from the Quarterly Bulletin (entitled: "Assets of banking institutions: Specific provisions in respect of loans and advances"), see Table 4.

portfolios and these individual results are aggregated to give total credit losses over three years[148]. The models are calibrated and not estimated because of difficulties with historical data and linkages across changing definitions. For each bank, average over-the-cycle transition matrices for each asset class, showing how assets evolve between different states over a 12-month period, are used to project opening balances forward, including the derivation of aggregate credit losses for each bank. The transition matrices can be adjusted with respect to a macroeconomic index that captures differences in the macroeconomic environment according to different scenarios run with the Core model.

Ozili and Outa (2017) review the recent academic and policy literature on bank loan loss provisioning, giving some limited attention in their section 6 on the determinants of bank provisioning.

In Section 5.1, we review earlier work on modelling loan loss provisions for South Africa, as a prelude to introducing a new model of this credit indicator for its insights into potential NPL models. Some early work on a 'top down' approach to stress testing for South Africa was done in 2010 by Havrylchyk (2010) using a by-bank analysis with panel data for quarterly loan loss provisions. We also review the Memo equation of De Jager *et al.* (2021) for credit impairments.

In Section 5.2, we present a *new* empirical model for the ratio of 'credit impairments' to gross loans and advances from 2001. Major drivers are the ratios of mortgage debt and house prices to income, credit conditions in the previous three years measured by credit spreads, and the GDP growth rate in two prior years (with important data breaks in credit impairments in 2008 and 2018). We anticipate the key drivers in the long-run solution in potential NPL models (from 2001) would be similar, with different relative weights on the drivers and short-run dynamics.

[148] At the SARB, stress testing began in 2008 with bottom-up and top-down stress tests as part of the IMF FSAP/Article IV Consultation, but mainly as an IMF exercise. In 2012, the SARB ran its own bottom-up stress test, and in 2014, again in collaboration with the IMF FSAP, a stress test of the South African banking sector, including both bottom up and top-down components. On its own, and after an IMF technical assistance mission, the SARB ran a common scenario stress test focused on solvency between Q4 2015 and Q2 2016. Participating banks conducted bottom-up stress tests, while the SARB conducted a complementary top-down stress test focused on solvency to validate and benchmark the results (*FSR* H1 2016). There is now a small stress testing division at the SARB to run top-down and bottom-up tests biennially. Liquidity testing is being added to solvency testing.

5.1 Earlier work on South African loan loss provisions

Havrylchyk (2010) uses a bank panel approach for modeling a measure of credit risk in South Africa as a function of macroeconomic variables. Her random effects model for credit risk during 2001-2008 uses bank-reported quarterly loan loss provisions against sectoral loans for the five biggest banks[149]. But for robustness she also tests 'overdue loans' as the dependent variable (an NPL concept), for which data were available to her from 1994 through 2008, though with several definitional changes. Her model is:

$$\textit{credit risk}_{it} = \alpha 1 + \alpha 2\ [\textit{business environment}]_t + \alpha 3\ \textit{prices}_t + \alpha 4\ \textit{interest rate}_t + \alpha 5\ [\textit{household}]t + \alpha 6\ [\textit{external}]_t + \varepsilon_{it}$$

The measures of credit risk are made a function of: the business environment (real GDP growth, real growth in gross fixed capital formation and the change in All-share index); the cost of borrowing in real or nominal terms; inflation excluding the property component; household sector indicators (such as the growth in nominal property prices, in real consumption, the debt-to-income ratio, the change in the employment index and the change in the wage index); and external factors (such as the change in a commodity price index or the gold price, in oil prices, in the real effective interest rate and in the terms of trade).

Credit risk models were presented for total loans, for household loans (the largest share in banks' portfolios consist of mortgages at flexible interest rates), and then sectoral models for mining, electricity, transportation and residual sectors. All explanatory variables were lagged by one year, so that she is effectively producing a one-year ahead forecasting model. Since *changes* in property prices were non-stationary, she explains that they were included in first differences, hence in the form of the rate of acceleration of the level of house prices. This may be problematic as it is unclear that the dependent variable being explained is necessarily non-stationary[150].

[149] Absa Bank Ltd, FirstRand Bank Ltd, Investec Ltd, Nedbank Ltd and The Standard Bank of South Africa Ltd, which constituted 92% of the total banking assets in June 2008.

[150] Even if the dependent variable was stationary, it is possible that several non-stationary explanatory variables could be jointly cointegrated so that their linear combination could be stationary. Therefore, individual non-stationarity is not necessarily a ground for exclusion.

Certainly the credit impairments measure, which is the same as Havrylchyk's loan loss provisions, that we model in Section 5.2, is highly non-stationary.

The loan loss provisions models for total loans and for household loans were driven by real GDP growth (-), real interest rates (+), inflation (-/+), acceleration of property prices (-) and the real effective exchange rate appreciation (-), with the signs she expected shown in parenthesis. During an upswing, individuals can more easily service loans because of higher wages and lower unemployment. Higher interest rates increase the repayment burden for borrowers with flexible interest rate contracts, making them more likely to default. A large share of loans to individuals are mortgages, so that rising house prices enhance borrower wealth, facilitating loan repayment. Since most domestic loans in South Africa are denominated in domestic currency, exchange rate fluctuations might be expected to have little direct effect on credit losses, though they may be an indicator of economic sentiment, with appreciation being associated with more positive growth expectations and a period of stronger capital inflows. Her model suggests that the main drivers of credit risk are high interest rates and especially declining property prices, in an environment where most borrowers have loans with flexible interest rates. An increase in the price of gold lowers loan loss provisions for the mining sector. Higher oil prices reduce profitability in the electricity sector, leading to higher defaults and hence higher loan loss provisions.

The related model for 'overdue loans' based on the 90-day overdue concept, which is a non-performing loans concept, fits less well over the entire longer period, though appears to fit better over a short period from about 2000, which she attributes to definitional changes that exaggerated the size of overdue loans pre-1996, see her Table 3. However, she does not give a source for her data. It seems likely that she used one or more of the 'credit risk buckets', e.g., the sub-standard plus doubtful categories, reported for mortgages on the CI500 forms as a proxy for 90-day overdue loans.

Since the publication of the Core model in 2007, Smal *et al.* (2007), there has been further model development but no updated publication. The most recent published version of the Core model, see De Jager *et al.* (2021), adds a banking sector and expands the linkages that can be influenced by macro-prudential policy between the banking system and the real economy. This version includes an equation estimated from 2007 onwards for ‘credit impairments’ as a Memo item - although it is misleadingly labelled as non-performing loans when it is in fact a measure of loan loss provisions, see Table 4. The model is shown in Box 1, as an error correction model for real credit impairments, deflating by the CPI and driven in the long-run by real disposable income and the prime rate. It would be more usefully modeled as a ratio to gross loans and advances, as in the sectoral by-bank panel model of Havrylchyk (2010): the fraction of loans that go bad is the relevant concept for the dependent variable, not the level of bad loans. This equation makes no allowances for the regulatory breaks in the data, and this omission will bias all the estimated coefficients. There are no linkages to mortgages markets or housing.

Box 1: The current treatment of *credit impairments (i.e. loan loss provisions)* in the SARB’s Core model – a Memo item (wrongly labelled Non-performing loans). The series code: KBP1123M.

B.5 Non-Performing loans

Real non-performing loans are credit impairments of banks in respect of loans and advances, where debtors have not made the scheduled payments for a specified period of time. Here, non-performing loans are estimated as a function of real personal disposable income of households, and the prime interest rate. Sample = 2007Q3 - 2019Q1.

$$\Delta \log\left(\frac{NPL}{Cpitd}\right) = \left[-\beta_1 * \left(\log\left(\frac{NPL_{(-1)}}{Cpitd_{(-1)}}\right) - log\left(Pdinc6_{(-1)}\right)\right) + \beta_2 * \left(Primei_{(-2)}\right)\right]$$
$$+ \beta_3 + \beta_4 * \Delta\, log(Pdinc6) + \beta_5 * Dum16q2 + \varepsilon_{NPL}$$

β_1 = 0.15075 ; β_2 = 0.00660 ; β_3 = -0.35228 ; β_4 = -1.5 ; β_5 = -0.15432

Where:

- NPL = Non-performing loans - Credit impairments in respect of loans and advances
- $Cpitd$ = Targeted headline consumer price index
- $Pdinc6$ = Real personal disposable income (Constant 2010 prices)
- $Primei$ = Prime overdraft rate
- $Dum16q2$ = Dummy for irregular data (2016Q2 = 1)

5.2 Insights from a new model for the ‘credit impairments’ to gross loans and advances ratio

We argued in Aron and Muellbauer (2022b) that credit risk measures - such as the credit impairments ratio or an NPL ratio - should play a functional role in the Core model, for example in equations for credit extension by banks and interest rate spreads. We develop a *new* empirical model for the ratio of 'credit impairments' to gross loans and advances from 2001.

While this series is subject to definitional shifts, modelling loan provisioning is relevant for inclusion in the Core model since an increase in provisions set aside reduces, other things being equal, the amount of credit banks can extend and probably affects their pricing of credit. One can even make an argument that loan provisioning i.e., credit impairments, is more immediately relevant for modelling credit extension than the underlying NPL data.

Table 4 details the regulatory definitional changes for credit impairments. The definition for credit impairments changed in 2008 with the implementation of Basel II, and in 2018 with the switch from IAS 39 to IFSR 9, the latter phased in gradually. Havrylchyk (2010) argues that there was another definitional change in 2001, when her series on loan loss provisions begins. Indeed, see Table 4, amended Regulations relating to Banks were implemented on 1 January 2001, which altered the definition and classification of overdue accounts. This suggests that any modelling of the series needs, at the very least, to include dummy variables to capture the 2008 and 2018 shifts in definitions and exclude pre-2001 data. Our preference is to model quarterly data on a credit impairments ratio, named CIR, which we measure as credit impairments (from the monthly series KBP1123M) divided by total gross loans and advances issued to the private sector by banks[151].

We have developed an equilibrium correction model for the CIR, for the period 2002:Q1 to 2020:Q1, including the relevant dummies to capture the 2008 and 2018 shifts in definitions. The data used in this paper are defined in Table 9, where summary statistics are also presented. We follow a general-to-specific model selection strategy, adding several innovative measures to the set of variables most often used in the empirical studies reviewed in Section 4. Most studies test for the effects of GDP per capita and GDP growth, interest rates, the inflation rate, the exchange rate and the unemployment rate. We expect that growth rates of employment are a more robust indicator of the state of the labour market in South Africa than the unemployment rate, which includes many individuals with little attachment to formal labour markets. We

[151] The sectoral by-bank panel model of Havrylchyk (2010) also uses such ratios.

include rates of change of employment, per capita GDP, per capita real household disposable income, and of the real exchange rate, and levels of real per capita GDP and household income, interest rates and inflation in the general specification.

To this selection of potential drivers, we add debt-to-income ratios, the house price-to-income ratio, the rate of growth of total bank credit and a proxy for credit conditions based on mortgage interest spreads. Household debt is a major component of overall bank credit extension and usually accounts for an even larger share of impaired loans. This suggests that the ratios to income of mortgage debt and of non-mortgage debt of households should be included in the relevant set of drivers. The higher are these debt-to-income ratios, the more likely are households to have difficulty servicing their debts. However, because mortgages are secured by the value of houses, the higher are average house prices relative to income, the lower should be the credit impairments ratio, for given debt-to-income ratios.

Of course, these are broad generalisations for aggregate data. The right-hand side tails of the distributions of debt-to-income are where the most vulnerable households are to be found. The quality of lending in recent years has an impact on the size of the vulnerable tail of the debt-to-income distribution, and it is therefore likely to be an important driver of the credit impairments ratio and of associated NPL ratios. There is the potential to proxy lending quality with a measure of credit conditions. In Aron and Muellbauer (2022a), we found that the mortgage spread defined as the prime rate minus the average effective mortgage rate on new mortgage loans was a good measure of credit conditions in South Africa: high values of the spread are associated with easy credit conditions. Figure 5 graphs this spread, showing increasingly high values in 2005 to 2008, followed by a sharp contraction in the financial crisis, followed by a partial recovery and a renewed decline.

We turn to the dummies needed to capture changes in the scope of CIRs as defined here. The first of these is DUM2007, which is 1 up to 2007:Q4 and then 0, to capture the switch to the treatment of credit impairments under Basel II. An evolutionary approach was adopted towards the internal ratings based (IRB) approach to provisioning, see Basel (2001) and PA Reports (2007, 2008), with a further step implemented in 2018. The second dummy therefore is DUM2017, which is 1 up to 2017:Q4 and then 0, to capture the implementation of the new accounting standard, IFRS 9. But as this change was intended to be phased in, we also include the four-quarter moving average of DUM2017, for a more gradual transition. A fourth dummy, DUM2016Q1 is 1 up to 2016:Q1 and then zero. It captures a reclassification by a bank of their impairments data[152]. These four dummies are all highly significant, and the results are reported in Table 10.

The estimated effects of the dummies can be used to construct an *adjusted* credit impairments ratio, ACIR, that corrects for the shifts in definitions in 2008 and 2018, to make the measure correspond to the current definition of credit impairments. These adjustments increase the scale of CIR before 2008 and again before 2018. In other words, had the current (post-2019) definition been used, reported CIRs in earlier years would have been substantially higher, as shown in Figure 6 which compares the unadjusted and adjusted CIRs[153]. Interestingly, the estimated model suggests that the rise in measured credit impairments in 2018 was almost entirely a figment of the change in definition, rather than a substantive change (see Box 3, Financial Stability Review, April 2018 and p.8, Financial Stability review, November 2018).

In the final reduction to a parsimonious from of the equation, there is evidence that strong growth of real per capita GDP in the previous two years reduces the credit impairment ratio and evidence of a marginal effect from the previous year's appreciation of the real exchange rate. Both the log mortgage debt-to-income ratio and the log house price-to-income ratio proved highly significant, confirming the interpretations discussed above. Moreover, long lags in the proxy for mortgage credit conditions, namely moving averages of the mortgage interest spread, also prove extremely important. They suggest that lax lending criteria over the previous three years are a major cause of high levels of CIRs. Recent experience of power outages tends to raise the CIR[154]. Probably this is because of the impact of outages on the cash flow of businesses and hence their ability to service debt. Impulse dummies for outliers, +1 in 2004:Q4 and -1 in 2005:Q1, were also included, but with no impact on the clearly determined long-run solution. The speed of adjustment is accurately estimated at 0.7,

[152] It appears that one of the banks reclassified credit impairments for its unsecured lending to households, so reducing the aggregate CIR, see *Financial Stability Review*, April 2018.

[153] It is also possible to adjust the CIR differently by converting the post-Basel II data to the pre-2008 definition of credit impairments, which is somewhat less useful from a current policy perspective.

[154] We measure the difference between electricity generated in the current quarter and 'normal' generation measured as the average over the previous eight quarters. The lagged three-quarter moving average of this measure has a significant negative effect on the CIR.

providing strong evidence of cointegration, given controls for regulatory shifts, between the CIR, the log ratios of debt-to-income and house price-to-income, and mortgage spreads, which are all I(1) variables. Parameter stability tests suggest excellent stability and other specification tests for heteroscedasticity and autocorrelation of the residuals are all satisfactory.

Omitting the marginally significant exchange rate effect results in the power outage measure losing significance, resulting in the simpler equation shown in column 2 of Table 10. Estimating this version for the period after 2008:Q1 gives results shown in column 3 of Table 10, confirming the stability of the parameter estimates. However, it is no surprise that the unprecedented rise in CIR in the pandemic that began in 2020:Q2 cannot be explained by the model.

Further insight into the long-run relationship is obtained by graphing the adjusted CIR against the fitted contributions of these three sets of drivers, see Figure 7. Figure 7 shows that a decline in the CIR from 2002 to 2008 was mainly driven by the rise in the house price-to-income ratio, to which the large fall in interest rates in 2003 and the later easing of credit conditions contributed, and the extent of the decline was only partially offset by the upward trend in mortgage debt-to-income and easy credit conditions in the previous three years. From 2004 to 2007, particularly strong growth of GDP over the previous two years, contributed to the decline in CIR. The house price-to-income ratio reversed after 2008, and the growth rate of GDP reversed sharply, partly because of rising interest rates after 2006, which together with the sustained high level of mortgage debt-to-income and the aftermath of loose lending conditions, drove the CIR sharply higher to a peak in 2010. In the following years, a further decline in the house price-to-income ratio raising the CIR, was offset by a fall in the mortgage debt-to-income ratio and stricter lending criteria. Eventually, a gradual rise in the house price-to-income ratio, a continued decline of mortgage debt-to-income and a decline in credit conditions led to falls in the CIR.

While the aggregate credit impairments ratio refers to all loans and advances, the empirical results appear to emphasise household mortgage debt and house prices both relative to household income. Levels and rates of change of the ratios to GDP of either total loans and advances or total bank credit extended to the private sector proved insignificant. However, it is plausible that the mortgage spread measure is a good proxy for lending conditions across the economy and therefore captures the general build-up of debt, and not only for household mortgages. Moreover, bad loans on bank lending to house builders and other businesses in the real estate sector are likely to be affected by the house price-to-income ratio.

The credit impairments ratio, as we have seen, is sensitive to regulatory shifts in provisioning that occurred with Basel II and the shift in accounting rules from IAS39 to IFSR9. It is also likely, that because of differences in internal procedures at different banks, CIRs may not be ideally comparable across banks. It would therefore be desirable to model NPL measures such as the ratio of impaired loans to gross loans and advances, default ratios, the ratio of 90-days overdue loan and advances to gross loans and advances, or the default ratios (see Table 4). We turn to possible models for the NPL ratios next.

6 A Suite of Models Approach for NPLs and other Credit Risk Indicators in South Africa

Published descriptions of the SARB Core model can be found in De Jager *et al.* (2021), and for an earlier version without a banking sector, in Smal (2007). A summarised commentary on the macro-financial linkages in the Core model is given in our partner paper, Aron and Muellbauer (2022a). This paper also discusses the integration of credit risk indicators into the banking block of the model.

We have drawn on Sections 3 and 4 to examine for South Africa the availability of data for NPLs and related concepts (by various measures and disaggregation) as *dependent* variables, and the available data for the different possible macro-, bank-specific, industrial- and household *drivers* of aggregated and disaggregated credit risk indicators.

In Table 7 possible dependent variables for credit risk indicator measures are presented, both aggregated and disaggregated by sector. This table should be read in conjunction with Table 4, giving the definitional changes over time for the indicators. In principle, this could allow an analysis of the NPLs for all (large) banks in a time series from 2008, or for individual (large) banks in a panel analysis or repeated-cross-sectional analysis. The household sector could be analysed as a time series for all lending, for securitised or for unsecuritised lending, and for pure mortgage lending. Default ratios for the non-financial corporate sector as a whole and distinguished by sub-sector, could be analysed by time series methods. Potentially commercial mortgages could also be examined separately by time series. There is scope for further disaggregation of some dependent NPL variables within a banking panel analysis.

In Table 8, a range of different possible macro-, bank-specific, industrial- and household-drivers of aggregated and disaggregated NPLs and other credit risk indicators for South Africa are tabulated. These are linked with the discussion from the literature in Section 3. The definitions and frequency and span of the data are given.

In Section 6.1., we consider the scope for modelling NPLs within the Core model, given the available data and drivers, drawing on lessons from Section 5.2. In Section 6.2, we suggest a suite of possible models at different degrees of disaggregation for NPLs. We explore implications for forecasting models for credit risk indicators, or an early warning system (EWS) approach, which could assist the stress testing exercise by helping to design alterative macro-scenarios. We also examine other avenues for improving the data towards better modelling of NPLs and other credit risk indicators in South Africa. We suggest that a micro-simulation approach developed by analysts at the Centre for Affordable Housing Finance in Africa to analyse the profitability of mortgages could be adapted by or for the SARB to analyse mortgage defaults and the scale of potential losses.

6.1 Improving the Core model by including a model for NPLs

It would be desirable to apply a broadly similar model to the one developed for the credit impairments (loan loss provisions) ratio to NPL ratios, if data on these could be extended, at least approximately, to cover a similar period from 2001 to the present. In this way, the two-way linkage between credit conditions and NPLs could be formalised, capturing in the Core model a critical element in business cycle dynamics in South Africa.

Credit conditions are an integral part of what drives the business cycle, for example, affecting the dynamics of house prices and affecting consumption and residential investment. For South Africa, Aron and Muellbauer (2022a) show that mortgage credit conditions as proxied by the spread between the prime rate of interest and the effective rate on new mortgage loans is one of the key drivers of house prices, and hence indirectly of mortgage debt and residential investment. The evidence in Aron and Muellbauer (2013) on consumption and debt suggests that credit conditions have a direct effect on consumption in South Africa, as well as indirect effects via the collateral role played by housing wealth and the partly offsetting negative effects of higher debt.

Credit conditions have a two-way connection with credit risk indicators: a period of easy credit conditions, resulting in lax lending criteria, tends to create financial vulnerability among borrowers and potentially among lenders, particularly if followed by an economic downturn. Then, rising NPLs and other credit risk measures result in a reduced ability and willingness of banks to extend credit, resulting in tighter credit conditions that amplify the downturn in the economy. This dynamic process is exactly what occurred in South Africa in the period 2004 to 2014. In the CIR model, Section 5.2, we have quantified the link running from previous credit conditions to credit risk.

We now consider the reverse connection from credit risk indicators to current credit conditions. Evidence of the effect of variations in NPLs on mortgage credit conditions comes from Chauvin and Muellbauer (2019), see Figure 4. This plots the independently-estimated credit conditions index for the French mortgage market against the NPL ratio for the French banking system. It shows the downturn in credit conditions with the rise in NPLs after 1990, the subsequent recovery as NPLs declined at the end of the 1990s, rising to a peak just before the financial crisis, and then once more declining as NPLs rose again. Chauvin and Muellbauer (2019) document the effects of the mortgage credit conditions index on consumption, mortgage debt and house prices in France.

As noted above, the model for credit impairments developed in Section 5.2 helps illustrate for South Africa the two-way relationship between credit risk measures and credit conditions. Loose lending conditions sustained over several years contributed to a later rise in the credit impairments ratio. The reverse relationship is apparent in Figure 8, which plots the adjusted CIR against the contemporaneous mortgage spread (measured as a centered three-month moving average). From 2004 to around 2014, there is a mirror image between the two, with the CIR leading the mortgage spread. When the CIR was falling, lenders felt able to loosen lending, but when it rose, lenders were more restrictive. After 2014 both measures appear to decline. It may be that in a weak mortgage market, with transactions volumes sharply lower than in the boom years, lenders boosted profit margins by reducing the gap between the prime rate and effective mortgage rates (i.e., by increasing the margin between the base rate and actual lending rates). The stress testing exercises carried out by the SARB, and the desire to have a substantial buffer between capital ratios and the minimum capital adequacy ratio, may have contributed to a declining gap between the prime rate and effective mortgage rates.

The loan-to-value ratio reveals another aspect of the mortgage lending conditions. Figure 9 plots the adjusted CIR against the average loan-to-value ratio (provided by the FNB and based on Deeds Office data). There is again a mirror image pattern from 2004 to 2014, though it seems to be closer to simultaneous than in Figure 8. This may be explained by the sensitivity of LTVs to *current* house price developments. Valuers employed by mortgage lenders are likely to use more optimistic valuations in a rising market and more cautious valuations when prices are stagnant or dropping back and transactions volumes shrink. When the housing market began to turn in 2008, the loan levels that the lenders were prepared to advance fell back relative to housing values, so that LTVs fell. At the same time, bad loans increased sharply, reinforcing lender caution on lending criteria. The LTV graph shows a rise from 2017, which is slightly puzzling. It is possible that with tighter loan conditions (for example, requiring higher credit scores), tighter mortgage spreads and low transactions volumes, competing lenders used higher LTVs to gain market share. The absence of macro-prudential regulation on LTVs, by contrast with the greater attention to capital adequacy ratios, may also have been a factor.

Turning to modelling NPL ratios for South Africa in the Core model, having the volatile period of 2001 to 2007 in the sample is important for empirical identification of the key relationships. We anticipate that the key drivers in the long-run solutions would be similar but the relative weights on the different drivers and the short-run dynamics could differ. However, given the probable need to patch pre-2008 data to later data from a different source, including a shift dummy for the 2008 shift may be necessary, depending upon which NPL concept is selected for modelling. Comparing the estimated equations for the credit impairments and for NPL ratios and their drivers could throw light on the important question of whether banks are slow in provisioning, whether potential lags in provisioning change with the business cycle, and whether they have been affected by the regime and accounting shifts in 2008 and 2018[155].

One potential NPL measure is the quantitative 90-day overdue ratio, for which, as explained earlier, data has been gathered since 2008 from IRB banks and from 2012 for the other banks. As discussed earlier, and summarised in Tables 4 and 7, these data can be linked with earlier data on credit risk buckets: from 2001 to 2007 for IRB banks, and from 2001 to 2011 for banks following the standardised approach. It looks as though these days-overdue data are likely to be immune from the 2018 shift to the IFSR 9 accounting standard.

A second potential NPL measure is the default ratio measured since 2008. As explained earlier, this is the most promising of the three potential NPL measures, when linking data from 2001 on a bank-by-bank basis, as the banks successively switched from the standardised approach to the IRB approach (from 2008 onwards).

Another NPL measure is *impaired loans* for which aggregate data are available from 2008. As discussed in Sections 3.3 and 3.4 and Table 4, it may be possible to construct a proxy for the period 2001 to 2007 based on the 'credit risk buckets' in the DI500 spreadsheets to approximate the aggregate impaired loans concept. Large shifts in the data in 2008 and 2018 would need to be controlled for, making this measure the least satisfactory of the three. Again, the empirical form of the credit impairments equation discussed in Section 5.2 would be the obvious starting point for modelling an NPL ratio defined this way.

6.2 A suite of possible models at different degrees of disaggregation for NPL

It would be desirable to undertake further investigations for data on NPL or credit impairments ratios for different sectors of the non-financial corporates as well as households, and for secured and unsecured loans. As discussed in Section 5.1, and Table 7, from 2008 data are available for the NPL concept defined as default ratios for

[155] The FSR in May 2021 comments reassuringly that loan loss provisions in the pandemic have broadly matched defaulted exposures.

private households and non-financial corporate sub-sectors, like real estate, constructed by those banks with permission to use internal models for credit quality assessment.

However, while the historical span is short, the possibility exists of carrying out a bank-by-bank panel study for different loan classes, such as households and number of sectoral splits for lending to businesses. This kind of disaggregation is likely to be informative. Table 8, based on an extensive literature review, suggests key measures of bank-specific regressors, as well as more general macroeconomic data for such a panel study. With such a large selection of potential drivers, one has to be careful that model selection methods to find parsimonious formulations begin with general specifications, but where it is possible to impose plausible economic priors in the selection process[156]. The automatic selection methods in Doornik and Hendry's Autometrics software can be very useful in this context, see Hendry and Doornik (2014). In a forecasting context, as Hendry and Clements (2004) show, pooling forecasts over different forecasts often improves the robustness of out-of-sample forecasts.

The bank returns on credit risk, forms DI500 from 2001 to 2007, and forms BA200 from 2008 onwards are filled in by individual banks. Therefore, the SARB has at its disposal, at least quarterly, bank-specific time series data on credit risk for the periods 2001-2007 and 2008 onwards. Based on these data, separate bank panel studies could be run for the two periods, though we suspect that for harmonised default ratios constructed as explained above, it should be possible to examine bank-by-bank data from 2001 to the present without needing dummies for 2008 and 2018. For the linked 90-days overdue measure, analysing from 2001 to the present without using dummies may also be possible. It is straightforward to check empirically whether 2008 and 2018 dummies are relevant. Even for the shorter panels for the two sub-periods, it is likely that the rich variation in bank-specific drivers as set out in Table 8, combined with the aggregate information on debt and house prices to income, mortgage spreads and growth rates of GDP, could aid powerfully in understanding the dynamics of credit impairments and NPLs.

[156] Among model selection methods, the Lasso (least absolute shrinkage and selection operator) method used by Ari *et al.* (2019), see Section 4, struggles with negative correlations and step-wise regression is also problematic, Castle *et al.* (2020). From this perspective, the Autometrics method offers many advantages.

Moreover, from 2008, it is possible to run bank panel models for a variety of disaggregations of default ratios[157] for different loan classes and sectors, such as secured/unsecured and for households and corporates, with the latter divided further into industrial sectors. For these subdivisions, different types of aggregate data are likely to prove relevant. Havrylchyk (2010), for example, uses the terms of trade to model credit impairments for the mining sector in 2001-2007. Building up an aggregate picture from granular analysis at the bank level would complement both aggregate time-series models and the bank-specific stress tests currently being run.

We also explore implications for forecasting models for credit risk indicators, or an early warning system (EWS) approach, which could assist the stress testing exercise by helping to design alterative macro-scenarios. Forecasting credit risk indicators such as NPLs is mainly of interest for macro-prudential policy, for instance, in the stress testing of banks and the different scenarios that might be run, and the NPL position might directly feed into the consideration of options for macroprudential scenarios. The nature of the lags found in the empirical model of the credit impairments ratio in Section 5.2, suggests that practical forecasts can be made for this measure and associated NPL ratios at least one year ahead, and probably two years. For example, it was found that credit conditions in the previous three years and economic growth in the previous two years were strongly significant in explaining the current credit impairments ratio. Moreover, debt and house price-to-income ratios, also highly significant, are very persistent over time. Thus, current ratios should be usefully informative for forecasting one or two years ahead.

It is useful to contrast an early warning forecast model of the NPL to gross outstanding loans ratio with the approach of GDP at risk (Adrian *et al.*, 2019; Prasad *et al.*, 2019). The latter requires a quantile regression technique, which places a bigger weight on falls in GDP than rises. In normal circumstances, NPLs are low and steady, and when there is a growth boost, NPLs are particularly low. However, NPLs have a non-linear reaction relative to the business cycle, and rocket in bad phases of the cycle. This makes NPL models plausibly a convenient proxy for the more complex GDP-at-risk approach, see the discussion at the end of Section 4.

[157] For banks following the IRB approach, consistent 90-day overdue data are available for a wide range of asset classes from 2008, and available from 2012 for banks following the standardised approach.

We also examine other avenues for improving the data towards better modelling of NPLs and other credit risk indicators in South Africa. For South Africa, the focus of the discussion so far has been mainly on data generated by the SARB. We now consider two other potential sources, the National Credit Regulator (NCR) and credit bureaus. Since 2007:Q3, the NCR has been producing data on household mortgage and non-mortgage credit, including performance data. The NCR classify loans by type of loan into 'current' and overdue status broken up into 30 days, 30-61, 61-90, 91-120 and over 120 days. In principle, the number of loans entering the shorter overdue durations could be used to forecast more serious later non-performance. While the coverage of the NCR data includes loans generated by insurance companies and other sources outside the banking system, the dominant share of loans does come from banks.

Turning to credit reference bureaus as a potential data source, Illana Melzer and her group at the Centre for Affordable Housing Finance in Africa (CAHF) have extensive experience analysing mortgage markets in South Africa. They have analysed loan performance in some detail, comparing conventional mortgage loans with those to lower income households, under the Financial Sector Charter (FSC) that came into being in January 2004. Their data refer to the largest four banks and come from one of the credit reference bureaus. From their 2012 presentation, they show how non-performance is strongly affected by the vintage of origination in the years 2004 to 2008, see Figure 10.

In their 2017 presentation (Melzer, 2017), they provide a similar chart for mortgages originated since 2009, see Figure 11. These show that while the 2009 vintage of loans had substantially worse performance than later vintages, there is little difference between any of the vintages between 2010 and 2014, though FSC loans have systematically higher overdue ratios than conventional loans.

Exploiting the vintage dimension in an econometric analysis has significant potential for improving the understanding and forecasting of overdue mortgages. The vintage captures lending conditions at the time mortgages were issued. The evidence we presented in Section 5.2 suggests that mortgage spreads in the vintage year would be an excellent proxy for lending conditions at the time of origination, but it would be worth considering information from loan-to-value ratios. The effect of age, or time expired between origination and the time that performance is being evaluated, is likely to depend on economic conditions in this period, measured by variables such as interest rates, house price to income ratios, income and employment growth and inflation.

Finally, we suggest that a micro-simulation approach developed by analysts at the Centre for Affordable Housing Finance in Africa to analyse the profitability of mortgages could be adapted by the SARB to analyse mortgage defaults and the scale of potential losses. Melzer and Hayworth (2018) have developed a granular simulation model of mortgage profitability, which has considerable potential for being adapted to focus on loan performance for different types of mortgages. Among other things, the model uses a two-state (normal vs 'problematic', where the latter includes missing a payment or early settlement) Markov matrix of transitions between states estimated from historical data to reproduce the payment behaviour for different market segments. The model can be used to simulated different scenarios such as an increase in the probability of a missed payment, a change in the recovery rate, given default, or a change in the interest rate. Since the household sector appears to account for a good deal of the NPLs in the banking system, we suggest there is scope for future interaction with the CAHF group drawing on their extensive experience studying mortgage markets in South Africa.

7 Conclusions

Elevated levels of NPLs are a recurrent characteristic of banking crises. Banking crises are typically preceded by poor quality of lending, excessive credit growth and high levels of leverage. A period of easy credit conditions, resulting in lax lending criteria, tends to create financial vulnerability among borrowers and potentially among lenders, particularly if followed by an economic downturn. The value of non-performing loans, often stable in boom periods, can rise sharply when the crisis breaks. Rising NPLs raise funding costs for banks, damaging their efficiency and profitability. Should credit costs then rise and banks apply tougher lending criteria for firms and households, a credit crunch may follow with falling or stagnant economic growth. The solvency of banks and borrowers may be threatened, with damaging feedbacks onto bank and firm share prices with liquidations, and onto house prices with repossessions. Further negative economic feedbacks could ensue from the spending constraints of indebted households and firms. There is thus *a two-way connection between credit conditions* *and NPLs.* The financial sector interconnectedness in the economy might be large enough to cause systemic risk.

This paper has surveyed international literature on measuring and modelling NPLs. As Bholat *et al.* (2018) and other sources make clear, there can be major problems in consistency across countries and jurisdictions, within countries between different institutions and across time. These problems make it harder to draw firm conclusions from empirical studies, whether from country panels, time series for individual countries or bank-specific panels. Nevertheless, many empirical studies point to the relevance of rates of economic growth in reducing NPLs, and interest rates and the unemployment rate in raising NPLs, among macroeconomic drivers in the recent past. Results are more mixed for inflation and the exchange rate. Higher price and wage inflation can reduce the burden of nominal debt (lowering NPLs), though if living standards fall because of price inflation, the coefficient sign is liable to reverse. The effects of changes in the exchange rate are also ambiguous. High levels of foreign debt make domestic borrowers vulnerable when the exchange rate depreciates, and a depreciation can also be associated with capital outflows or reduced optimism about growth. For countries less dependent on capital flows, a depreciation of the real exchange rate can improve the competitive positions and stimulate growth, with the opposite effect on NPLs. We also considered a host of bank-specific and banking sector and non-financial corporate sector determinants that might influence NPLs, in particular for their potential usefulness in banking panel studies in South Africa. Model selection methods, using reductions of general models to specific parsimonious models, were considered. Sometimes such methods can result in the selection of more than one parsimonious model. In a forecasting context, the literature suggests that pooling or averaging of forecasts from several models often improves robustness, see Hendry and Clements (2004) and, in a bank stress testing context, Gross and Población (2015).

We have noted the exclusion of some relevant drivers, and also the exclusion of possible asymmetries and non-linearities in most studies. Only a small minority of studies focus on real estate-related measures such as house price and debt-to-income ratios, highly relevant for South Africa. Notably, there is a relationship between modelling NPLs and the recent literature on growth at risk using quantile regression methods that emphasises such factors, as well as credit gaps or other measures of excessive credit growth, though this connection is rarely if ever made in studies of NPLs.

We have studied the availability of data on NPLs in South Africa, finding consistent data for some measures in the period 2001 to 2007, and from 2008, though in many cases affected by the 2018 switch in the accounting treatment of impairments. There is only one published time series back to the 1990s but that is not for an NPL concept, rather for ‘credit impairments’, a loan loss provision. We have developed a new empirical model for this series, in itself a useful credit risk indicator. We find that ratios of mortgage debt and house prices to income, credit conditions in the previous three years as measured by credit spreads, and the growth rate for GDP in the previous two years, are major drivers of the credit impairments to gross loans and advances ratio. However, it is very important to allow for breaks in the data from definitional changes in 2001, 2008 and 2018.

Our empirical findings underline for South Africa the two-way connection between credit conditions and credit risk. A model of this kind for aggregate credit impairments or NPL ratios should be an important part of the Core model, since credit risk indicators are likely to affect credit pricing and credit extension by banks, thus improving linkages in the model between the financial sector and the real economy. This new model for South Africa has useful implications for prospective NPL models. We anticipate that the key drivers in the long-run solution would be similar for South African NPL measures but the relative weights on the different drivers and the short-run dynamics could differ.

We suggest several possibilities for how pre- and post-2008 data on different NPL concepts might be joined to permit an analysis of data back to 2001 on reasonably consistent concepts. For modelling purposes, it is important to include data for the volatile period from 2001 up to 2007, for drawing robust economic insights. Aggregate data post-2008 may not give a long enough time span for robustness in time series analyses, though panel data analyses for individual banks and loan classes may be informative from a modelling perspective, compensating for the lack of historical information. Data classified by sector and loan type are likely to be particularly informative as the drivers of credit risk typically differ by sector and loan type.

We conclude that the default ratio measure available since 2008 is the most likely of the three NPL measures considered to be taken back on a consistent basis to 2001[158]. On the 90-day overdue NPL concept, we judge that breaks in the data in 2008 and later, as banks successively switch from the standardised to the IRB method for measuring credit quality, are likely to make it somewhat more difficult to join up data for a consistent series back to 2001. However, we recommend that the attempt be made to link the available data on a bank-by-bank basis to provide more than one NPL concept for modelling. The third measure, impaired loans, published since 2008 is the least satisfactory candidate as a potential NPL measure. It is strongly affected by the accounting switch in 2018, and a further switch will have occurred in 2008 in linking to pre-2008 on credit risk buckets.

We strongly urge the SARB to publish time series data at least for two of the NPL measures we have discussed for the 2001-2007 period and from 2008 onwards, with clear documentation. This would require a special background paper to give transparent methodological detail on joining, using a bank-by-bank basis, the different time segments for the various NPL concepts. For the default ratio, the pre-2008 data require joining with the post-2008 data, for all banks regardless of whether they use the standardised or IRB approaches. For the 90-day overdue ratio, for banks using the standardised approach, the pre-2012 data require joining with the post-2012 data; whereas for IRB banks, data from the last month of their using the standardised approach needs to be joined to the first month that they use the IRB approach. The aggregate time series data, with appropriate qualifications and explanations of methodology should then be routinely published. We also noted that, in addition to data on stocks of NPLs, flows into and out of the non-performing classification can provide additional information on credit risk and propose this as a topic for further investigation. Similar issues concerning regulatory shifts in definitions will apply.

The Financial Stability Department and the Macro-Models team have obviously a *different set of incentives* to the Prudential Authority department. The former two are both backward- and forward-looking, aiming to find long and consistent time series that

[158] Our referees from the Prudential Authority support our suggestion that a potential approach would be to use default for the internal ratings-based approach and the sum of “sub-standard”, “Doubtful” and “Loss” for the standardised approach for data for the period post 2008. Both these measures have day count and qualitative measures. Pre-2008 data (i.e. 2001 to 2007) would, presumably, follow the same approach (using the concepts of substandard, doubtful and loss to obtain a 90 day overdue measure).

will improve the modelling and forecasting at the central bank to strengthen the financial system and the understanding of monetary transmission (see Aron and Muellbauer, 2022a and 2022b), as is commonly the focus at most major central banks. The Prudential Authority probably mainly considers the current state of affairs and its development into the future, for sound supervisory purposes. In our view, before institutional memory is lost, it would be important have dedicated money and time set aside for the cooperation from outstanding, knowledgeable and committed people in the Prudential Authority, to help get consistent NPL series together to this end.

References

Adrian, T., N. Boyarchenko, and D. Giannone. 2019. "Vulnerable Growth" *American Economic Review* 109 (4): 1263-89.

Allen, F. and D. Gale. 2000. "Financial Contagion." *Journal of Political Economy* 108(1): 1–33.

Ari, A., Chen, S. and L. Ratnovski. 2019. "The dynamics of non-performing loans during banking crises: A new database." IMF Working Paper, WP/19/272, International Monetary Fund, December.

Ari, A., Chen, S. and L. Ratnovski. 2021. "The Dynamics of Non-Performing Loans during Banking Crises: A New Database with Post-COVID-19 Implications." *Journal of Banking and Finance*, DOI:10.1016/j.jbankfin.2021.106140 (in press).

Aron, J. and J. Muellbauer. 2011. "Modelling and Forecasting with County Court Data: Regional Mortgage Possession Claims and Orders in England and Wales." Spatial Economic Research Centre (SERC), London School of Economics, Discussion Paper SERCDP No. 0070.

Aron, J. and J. Muellbauer. 2013. "Wealth, credit conditions and consumption: evidence from South Africa." *Review of Income and Wealth* 59 (S1): S161–S196, October.

Aron, J. and J. Muellbauer. 2016. "Modelling and Forecasting UK Mortgage Arrears and Possessions." *Journal of Urban Economics* 94: 32–53, July 2016.

Aron, J. and J. Muellbauer. 2022a. "New Econometric Models for South African House Prices, Mortgage Debt and Residential Investment." Mimeo. Nuffield College, University of Oxford, February 2022.

Aron, J. and J. Muellbauer. 2022b. "Enhancing Financial Stability and Monetary Analysis in the Core Model of the SARB." Mimeo. Nuffield College, University of Oxford, February 2022.

Aron, Janine, Greg Farrell and John Muellbauer. 2020. "Financial Stability and the Housing Market in South Africa." Mimeo. Nuffield College, University of Oxford, February 2020.

Balgova, M., Alexander P., and M. Skrzypinska, 2017, "Reducing non-performing loans: Stylized facts and economic impact." ms., EBRD and Oxford, December.

Barisitz, S. 2011. "Nonperforming Loans in CESEE – What Do They Comprise?" Focus on European Economic Integration, Oesterreichische Nationalbank (Austrian Central Bank), 4: 46-68.

Barisitz, S. 2013a. "Nonperforming Loans in Western Europe – A Selective Comparison of Countries and National Definitions." Focus on European Economic Integration, Oesterreichische Nationalbank (Austrian Central Bank), 1: 28-47.

Barisitz, S. 2013b. "Nonperforming Loans in CESEE – An Even Deeper Definitional Comparison." Focus on European Economic Integration, Oesterreichische Nationalbank (Austrian Central Bank), 3: 67-84.

Basel Committee on Banking Supervision (BCBS). 2017. “Prudential treatment of problem assets—definitions of nonperforming exposures and forbearance, Guidelines.” 4 April 2017.

Basel Committee on Banking Supervision (BCBS). 2004. “International Convergence of Capital Measurement and Capital Standards. Revised Framework.” June 2004.

Basel Committee on Banking Supervision (BCBS). 2001a. “The Internal Ratings-Based Approach.” Supporting Document to the New Basel Capital Accord”. Consultative Document. Issued for comment by 31 May 2001.

Basel Committee on Banking Supervision (BCBS). 2001b. “The Standardised Approach to Credit Risk.” Supporting Document to the New Basel Capital Accord”. Consultative Document. Issued for comment by 31 May 2001.

Baudino, P., J. Orlandi and R. Zamil. 2018. “The identification and measurement of nonperforming assets: a crosscountry comparison.” FSI Insights on policy implementation No 7 (April), Financial Stability Institute, BIS.

Beck, R., Jakubik, P. and A. Piloiu. 2013. “Non-performing loans: What matters in addition to the economic cycle?” European Central Bank, Working Paper Series, No 1515, February 2013.

Belloni, A., and V. Chernozhukov. 2013. “Least Squares After Model Selection in High-Dimensional Sparse Models.” *Bernoulli* 19(2): 521–547.

Belloni, A., D. Chen, V. Chernozhukov and C. Hansen. 2012. “Sparse Models and Methods for Optimal Instruments with An Application to Eminent Domain.” *Econometrica* 80(6): 2369-2429.

Bernanke, B. S., and M. Gertler. 1989. “Agency costs: Net worth and business fluctuations.” *American Economic Review* 79, 14–31.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1996. “The Financial Accelerator and the Flight to Quality.” *Review of Economics and Statistics* 78 (1): 1-15.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. “The Financial Accelerator in a Quantitative Business Cycle Framework.” in J. B. Taylor, and M. Woodford, eds., *Handbook of Macroeconomics*, Amsterdam: North-Holland, 1341-1393.

Bholat, D., Lastra, R. M., Markose, S. M., Miglionico, A., & Sen, K. 2018. Non-performing loans at the dawn of IFRS 9: Regulatory and accounting treatment of asset quality. Journal of Banking Regulation, 19, 33–54.

Bholat, D., Lastra, R. M., Markose, S. M., Miglionico, A., & Sen, K. 2016. “Non-performing loans: regulatory and accounting treatments of assets.” Working Paper no. 594, Bank of England.

Castle, J.L., J.A. Doornik and D.F. Hendry, "Robust Discovery of Regression Models", ms., Economics Department, Oxford, 2020.

Chauvin, V. and J. Muellbauer. 2018. "Consumption, household portfolios and the housing market in France." *Economie et Statistique/Economics and Statistics* 500-501-502: 157-178, plus online complement.

de Jager, S., R. Ehlers, K. Mojapelo and P. Pienaar. 2021. "Short-term impacts and interaction of macroprudential policy tools." Working Paper Series WP/21/20, South African Reserve Bank.

Diamond, D.W. and P.H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy,* 91(3): 401–419.

Elmer, P.J. and S.A. Seelig. 1998. "Insolvency, Trigger Events, and Consumer Risk Posture in the Theory of Single-Family Mortgage Default." FDIC Working Paper 98-3.

European Banking Authority (EBA). 2014. EBA Final Draft Implementing Technical Standards: On Supervisory Reporting on Forbearance and Non-performing Exposures Under Article 99 (4) of Regulation (EU) No 575/2013. Eur Account Rev EBA/- ITS/20:1–65.

European Banking Authority (EBA). 2016. *Report on the dynamics and drivers of nonperforming exposures in the EU banking sector*. European Banking Authority, 22 July 2016.

European Central Bank (ECB). 2017. "Guidance to banks on non-performing loans." Banking Supervision, ECB, March 2017.

Ferrari, S., P. Van Roy and C. Vespro. 2021. "Sensitivity of credit risk stress test results: Modelling issues with an application to Belgium." *Journal of Financial Stability* 52. https://doi.org/10.1016/j.jfs.2020.100805.

Financial Stability Review. Various years. Bi-annual Reports, South African Reserve Bank.

Fitch Ratings. 2021. "What Investors Want to Know: Loan Moratoria in EMEA Emerging Markets." Special Report, March 2021.

Ghosh, A. 2017. "Sector-specific analysis of non-performing loans in the US banking system and their macroeconomic impact." *Journal of Economics and Business* 93: 29-45.

Ghosh, A. 2015. "Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states." *Journal of Financial Stability* 20: 93-104.

Gross, M., and J. Población. 2015. "A false sense of security in applying handpicked equations for stress test purposes." ECB working paper 1845, European Central Bank.

Havrylchyk, O. 2010. "A macroeconomic credit risk model for stress testing the South African banking sector." MPRA Paper No. 21639, and South African Reserve Bank Working Paper 3(10) (March).

Hendry, D. F., and M. P. Clements. 2004. "Pooling of forecasts." *The Econometrics Journal* 7(1), 1–31.

Hendry, D.F. and J.A. Doornik. 2014. "Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics." The MIT Press.

Ho, P. H., Huang, C. W., Lin, C. Y., and J. F. Yen. 2016. "CEO overconfidence and financial crisis: Evidence from bank lending and leverage." *Journal of Financial Economics* 120: 194–209.

International Monetary Fund (IMF). 2005. *The treatment of nonperforming loans.* Washington, DC: International Monetary Fund.

Jiménez, G., and J. Saurina. 2005. "Credit cycles, credit risk, and prudential regulation." *International Journal of Central Banking* 2: 1–38.

Kasinger, J., J.P. Krahnen, S. Ongena, L. Pelizzon, M. Schmeling and M. Wahrenburg. 2021. "Non-performing Loans - New risks and policies?", SAFE White Paper 84, commissioned by the European Parliament upon request of the Economic and Monetary Committee (ECON).

Kau, J.B., Keenan, D.C., Muller, W.J., and J.F. Epperson. 1992. "A Generalized Valuation Model for Fixed-Rate Residential Mortgages." *Journal of Money, Credit and Banking* 24: 279-99.

Keeley, M. C. 1990. "Deposit insurance, risk, and market power in banking." *The American Economic Review* 80: 1183–1200.

Keeton, W., and C. Morris. 1987. "Why do banks' loss differ?" *Economic Review* 72: 3–21.

Kiyotaki, N. and J. Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211-48.

Klein, N. 2013. "Non-performing loans in CESEE: Determinants and impact on macroeconomic performance." IMF Working Paper No. 13/72, International Monetary Fund, Washington.

Krueger, R. 2002. *International Standards for Impairment and Provisions and their Implications for Financial Soundness Indicators (FSIs).* International Monetary Fund, Washington.

Laeven, L. and F. Valencia. 2013. "Systemic banking crises database." IMF Economic Review, 61 (2): 225-270.

Laeven, L. and F. Valencia. 2018. "Systemic Banking Crises Revisited," IMF Working Paper, 18/206, International Monetary Fund, Washington.

Louzis, D. P., Vouldis, A. T., and V. L. Metaxas. 2012. “Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios.” *Journal of Banking and Finance* 36: 1012–1027.

Macháček, M, A. Melecký and M. Šulganová. 2018. "Macroeconomic Drivers of Non-Performing Loans: A Meta-Regression Analysis." *Prague Economic Papers*, Prague University of Economics and Business, 2018(3): 351-374.

Makri, V., Tsagkanos, A., and A. Bellas. 2014. “Determinants of nonperforming loans: The case of Eurozone.” *Panoeconomicus* 61: 193–206.

Melzer, I. 2012. “FSC and Affordable Mortgage Loan Performance Assessment: with REAL Data. Updated analysis.” Presentation. Centre for Affordable Housing Finance in Africa. https://housingfinanceafrica.org/authors/illana-melzer/.

Melzer, I. 2017. “Access, performance and profitability – a mortgage market overview.” Tanzania Mortgage Seminar. Centre for Affordable Housing Finance in Africa, February. https://housingfinanceafrica.org/authors/illana-melzer/.

Melzer, I. and C. Hayworth. 2018. “Bringing life to mortgage markets in South Africa.” Report for Centre for Affordable Housing Finance in Africa, August.

Naili, M. and Y. Lahrichi. 2020. “The determinants of banks' credit risk: Review of the literature and future research agenda.” *International Journal of Finance and Economics*, 2020: 1–27. Wiley.

Ozili, P. K. 2019. “Non-performing loans and financial development: New evidence.” *The Journal of Risk Finance* 20: 59–81.

Ozili, P. K. and E. Outa. 2017. “Bank loan loss provisions research: A review.” *Borsa Istanbul Review* 17-3.

Prasad, A. S. Elekdag, P. Jeasakul, R. Lafarguette, A. Alter, A. Xiaochen Feng, and C. Wang. 2019. “Growth at Risk: Concept and Application in IMF Country Surveillance.” IMF Working Paper, WP/19/36, International Monetary Fund.

Prudential Authority (formerly Bank Supervision Department. Various years. Annual Reports, South African Reserve Bank.

Quarterly Bulletin. Various years. Quarterly Reports, South African Reserve Bank.

Smal, D., C. Pretorius, and N. Ehlers. 2007. “The Core Forecasting Model of the South African Reserve Bank,” SARB Working Paper WP/07/02, June.

S&P Global Market Intelligence. 2021. South African banks set for low returns, but may see impairment reversals., April 2021.

United Nations System of National Accounts. 2008. Statistical Commission 2009. System of National Accounts, 2008 (2008 SNA). New York.

Vandell, K. D. 1995. “How Ruthless is Mortgage Default? A Review and Synthesis of the Evidence.” *Journal of Housing Research* 6: 245-264.

Wang, A. T. 2018. “A reexamination on the effect of bank competition on bank non-performing loans.” *Applied Economics* 50, 6165–6173.

Appendix: Classification categories of non-performing loans prior to 2008, and from 2008 for banks using the standardised approach: 'Special Mention, Substandard, Doubtful and Loss'.

(I) Special mention

Included in the category of special mention are credit exposures in respect of which the obligor is experiencing difficulties that may threaten the bank's position. Ultimate loss is not expected, but may occur if adverse conditions persist.

As a minimum, credit exposure that exhibits one or more of the characteristics specified below shall be included in the category of special mention:

(A) Early signs of liquidity problems exist, such as delay in the servicing of loans.

(B) Loan information is inadequate or incomplete. For example, the reporting bank is unable to obtain from the relevant obligor annual audited financial statements or such statements are not available.

(C) The condition of and control over collateral is questionable.

(D) The bank fails to obtain proper documentation from or co-operation by the obligor or finds it difficult to maintain contact with the obligor.

(E) There is a slowdown in business activity or an adverse trend in the obligor's operations that signals a potential weakness in the financial strength of the obligor, but which may not necessarily have reached a point that threatens the ongoing servicing of the relevant exposure.

(F) Volatility in economic or market conditions is likely to negatively affect the obligor in the future.

(G) Poor performance persists in the industry in which the obligor conducts business.

(H) The relevant obligor, or, in the case of a corporate borrower, a key executive, is in ill health.

(I) The obligor is subject to litigation that is likely to have a significant impact on the financial position of the said obligor.

(J) The obligor is experiencing difficulty with the servicing of other loans from either the reporting bank or other banks.

Provided that any relevant credit exposure amount that is overdue for more than 60 days shall as a minimum be classified as special mention.

(ii) Substandard

Any credit exposure that reflects an underlying, well defined weakness that may lead to probable loss if not corrected should be included in the category of substandard. The risk that such credit exposure may become an impaired asset is probable, and the bank is relying, to a large extent, on available security.

The primary sources of repayment are insufficient to service the remaining contractual principal and interest amounts, and the bank has to rely on secondary sources for repayment, which secondary sources may include collateral, the sale of a fixed asset, refinancing and further capital.

Credit exposures classified as substandard are likely to exhibit one or more of the characteristics specified below:

(A) Repayment of the principal amount and/or accrued interest has been overdue for more than 90 days, and the net realisable value of security is insufficient to cover the payment of the principal amount and accrued interest.

(B) The principal amount and accrued interest are fully secured, but the repayment of the principal amount and/or accrued interest has been overdue for more than 12 months.

(C) Significant deficiencies exist that threaten the obligor's business, cash flow or payment capability, which deficiencies may include the items specified below:

 (i) The credit history or performance record of the obligor is not satisfactory.

 (ii) Labour disputes or unresolved management problems may affect the business, production or profitability of the obligor.

(iii) Increased borrowings are not in proportion with the obligor's business.

(iv) The obligor is experiencing difficulty with the repayment of obligations to other creditors.

(v) Construction delays or other unplanned adverse events resulting in cost overruns are likely to require loan restructuring.

(vi) The obligor is unemployed.

(iii) Doubtful

Credit exposure in the category of doubtful is considered to be impaired, but is not yet considered final loss due to some pending factors, such as a merger, new financing or capital injection, which factors may strengthen the quality of the relevant exposure.

Doubtful credit exposures exhibit not only all the weaknesses inherent in credit exposures classified as substandard but also have the added characteristics that the said exposures are not duly secured. The said weaknesses make collection in full, on the basis of currently existing facts, conditions and values, highly questionable and improbable. The possibility of loss is high, but due to certain important and reasonably specific factors that may strengthen the asset, the classification of the asset as an estimated loss is deferred until a more exact status may be determined.

Credit exposures classified as doubtful exhibit one or more of the characteristics specified below:

(A) Repayment of the principal amount and/or accrued interest has been overdue for more than 180 days, and the net realisable value of security is insufficient to cover the payment of the principal amount and accrued interest.

(B) In the case of unsecured or partially secured credit exposures that have been overdue for less than 180 days, other serious deficiencies, such as default, death, bankruptcy or liquidation of the obligor, are detected or the obligor's whereabouts are unknown.

Credit exposures that have been overdue for 180 days and longer are usually classified as doubtful unless the said exposures are well secured, legal action has actually commenced, and timely realisation of the collateral or enforcement of guarantees obtained will result in the repayment of the relevant principal and interest amounts due, including payments in respect of amounts overdue.

When an account is classified as doubtful, unless particular circumstances pertaining to the relevant obligor dictate otherwise, interest shall no longer be accrued or accrued interest shall be impaired.

(iv) Loss

Credit exposures classified as loss are considered to be uncollectable once collection efforts, such as realisation of collateral and institution of legal proceedings, have been unsuccessful. The relevant exposures are considered of such little value that the said exposures should no longer be included in the net assets of the bank.

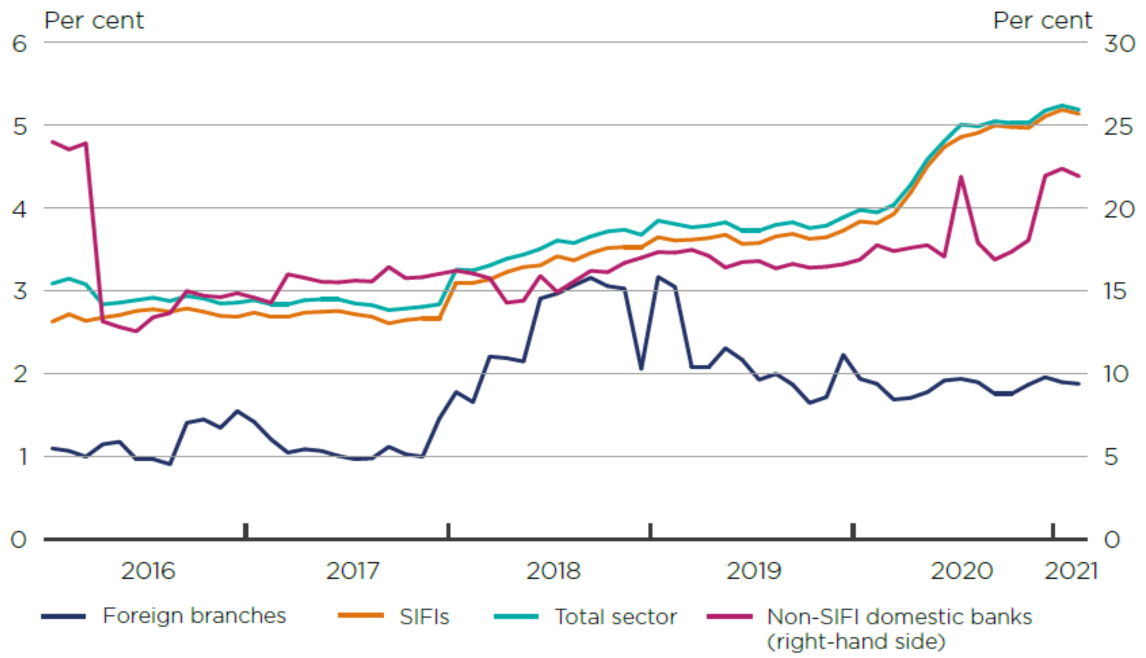
This classification does not necessarily mean that the asset has no recovery value. Instead, it is neither practical nor desirable to defer writing off this basically worthless asset even though partial recovery may be effected in the future, that is, banks should not retain exposures on their books while attempting long-term recoveries.

Non-performing credit exposures that have been overdue for at least one year shall be classified as loss unless such exposures are well secured, legal action has actually commenced, and timely realisation of the collateral or enforcement of guarantees obtained will result in the repayment of the relevant principal and interest amounts due, including payment in respect of amounts overdue.

When an account is classified as loss, unless particular circumstances pertaining to the relevant obligor dictate otherwise, interest shall no longer be accrued or accrued interest shall be impaired.

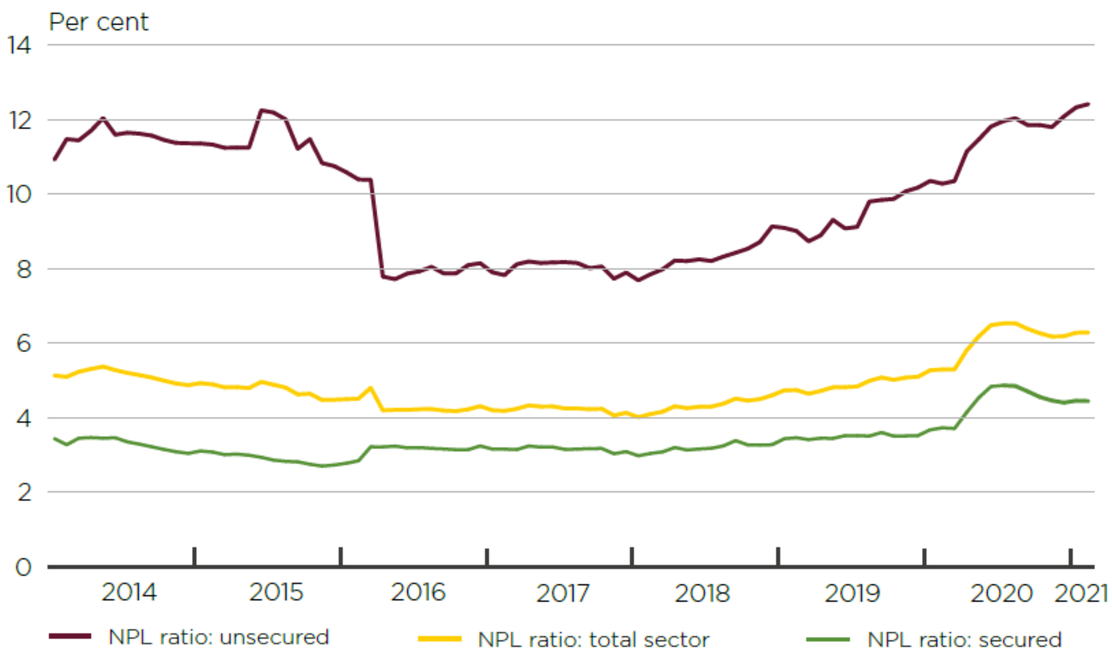
Source: Prudential Authority, SARB.

Figure 1: NPLs by the 'impaired advances' measure - by category of bank



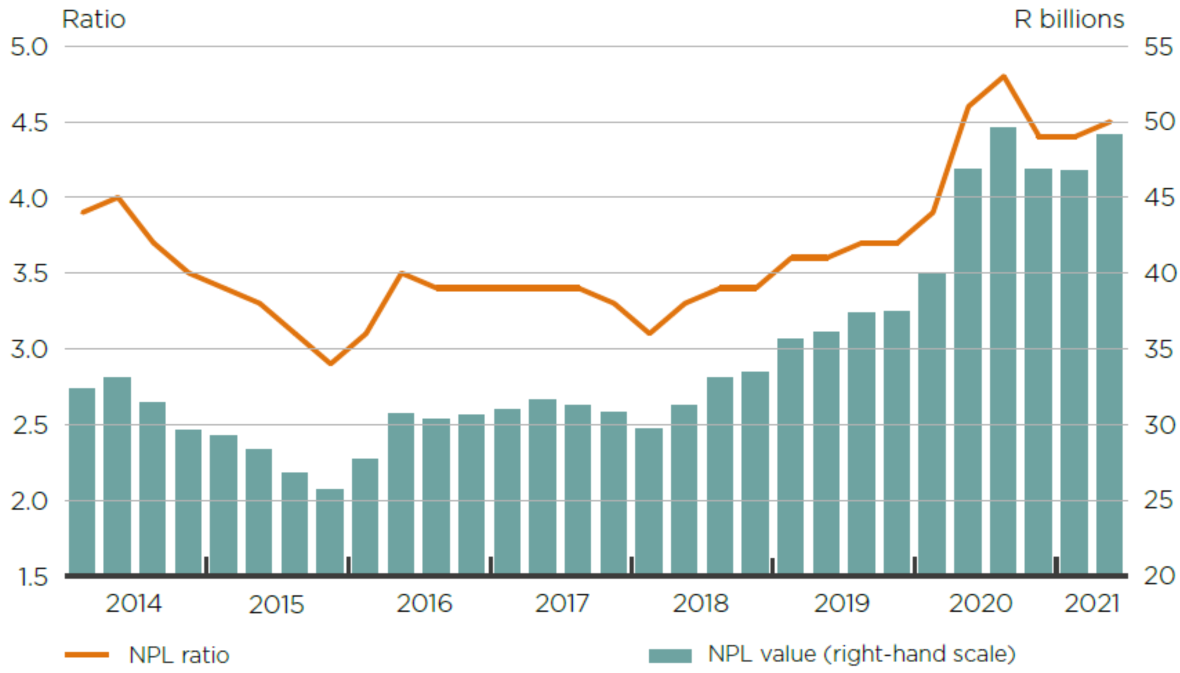
Source: Financial Stability Review, First Edition, 2021; Prudential Authority and SARB.
Notes: The NPL measure is the ratio of impaired advances to gross loans and advances.

Figure 2a: NPLs by the '90-day overdue' measure - household NPLs



Source: Financial Stability Review, First Edition, 2021; Prudential Authority.
Notes: The NPL measure is the ratio of the value of household NPLs to total outstanding household loans. NPLs are defined as loans for which debt-service payments are 90 days or more overdue.

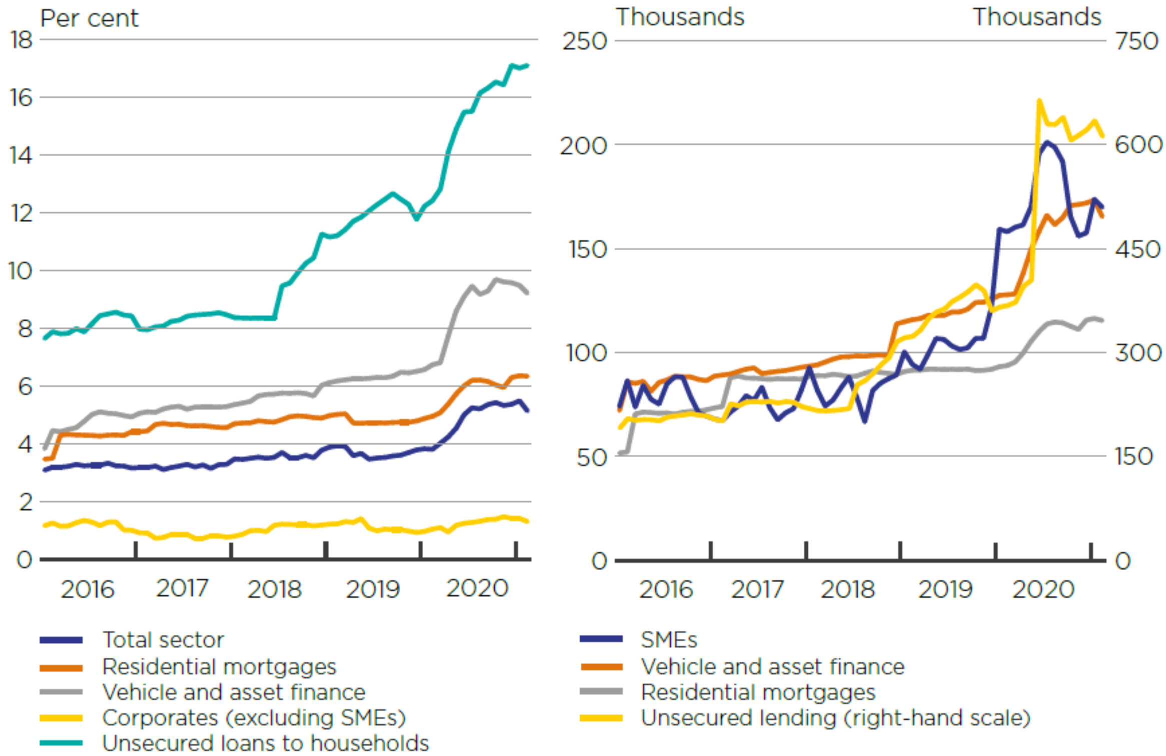
Figure 2b: NPLs by ’90-day overdue’ measure - residential mortgage NPLs



Source: Financial Stability Review, Second Edition, 2021; Prudential Authority.

Notes: The NPL ratio measures the value of mortgage NPLs relative to total mortgage loans and advances. NPLs are defined as loans for which debt-service payments are 90 days or more overdue.

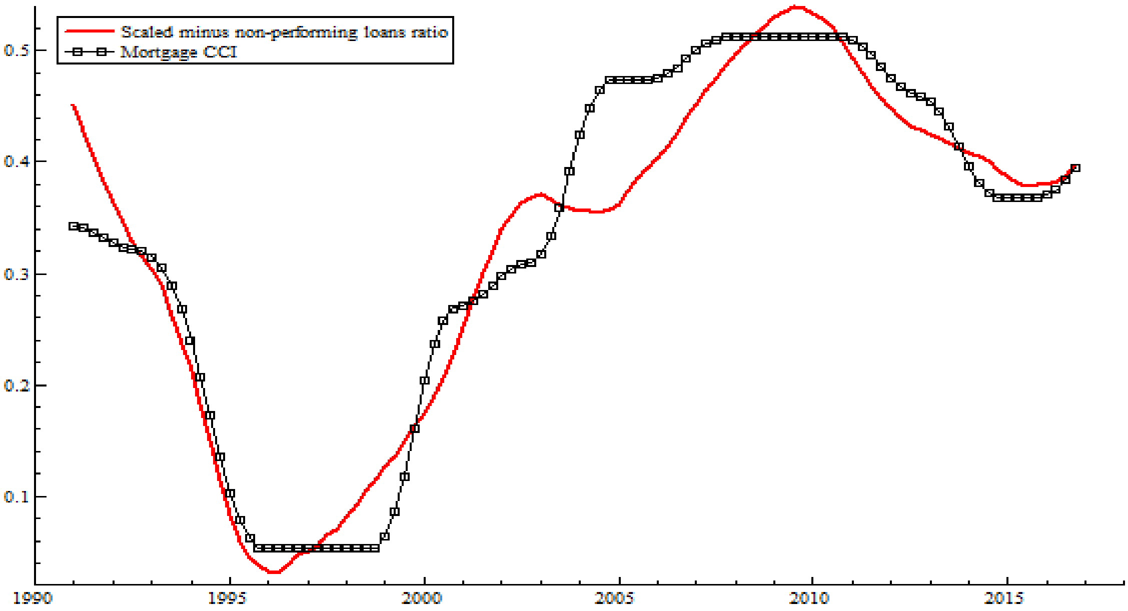
Figure 3: NPLs by the 'default ratio' measure - for selected banking sector portfolios



Source: Financial Stability Review, First Edition, 2021; Prudential Authority.

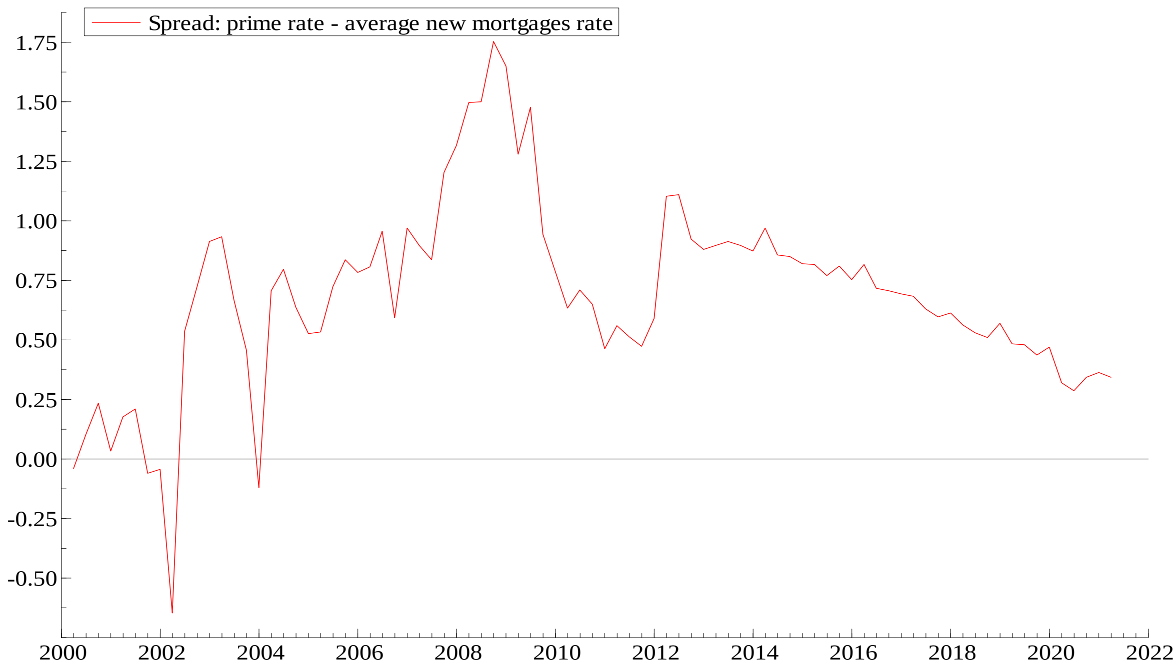
Notes: (i) Default ratios (left) and number of counterparties in default (right). (ii) The data refer to reporting IRB banks only. (iii) The NPL as default ratio is calculated as defaulted exposures as a percentage of the exposures at default, with a higher ratio indicative of increased defaulted exposures in the loan portfolio.

Figure 4: Estimated mortgage credit conditions index and (minus) NPL ratio for France



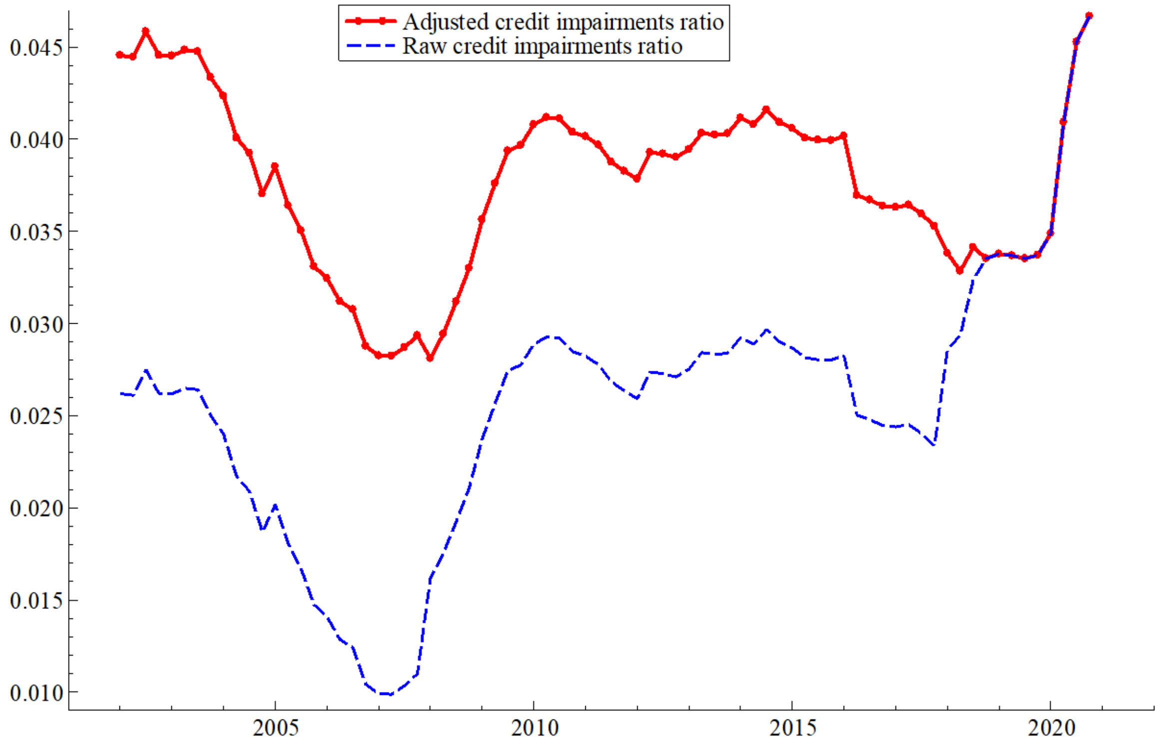
Source: INSEE, Banque de France; authors' calculations.

Figure 5: The mortgage rate spread: prime rate of interest minus average rate on new mortgages



Source: The average mortgage interest rate is BAT9612M, SARB, from 2001Q1. Data for 2000 interpolated by the authors.

Figure 6: Raw credit impairments ratio and adjusted for definitional changes (2008 and 2018)



Source: See Table 9, and for the adjusted ratio, Section 5.2.

Figure 7: Decomposition of the adjusted CIR into three long-run drivers

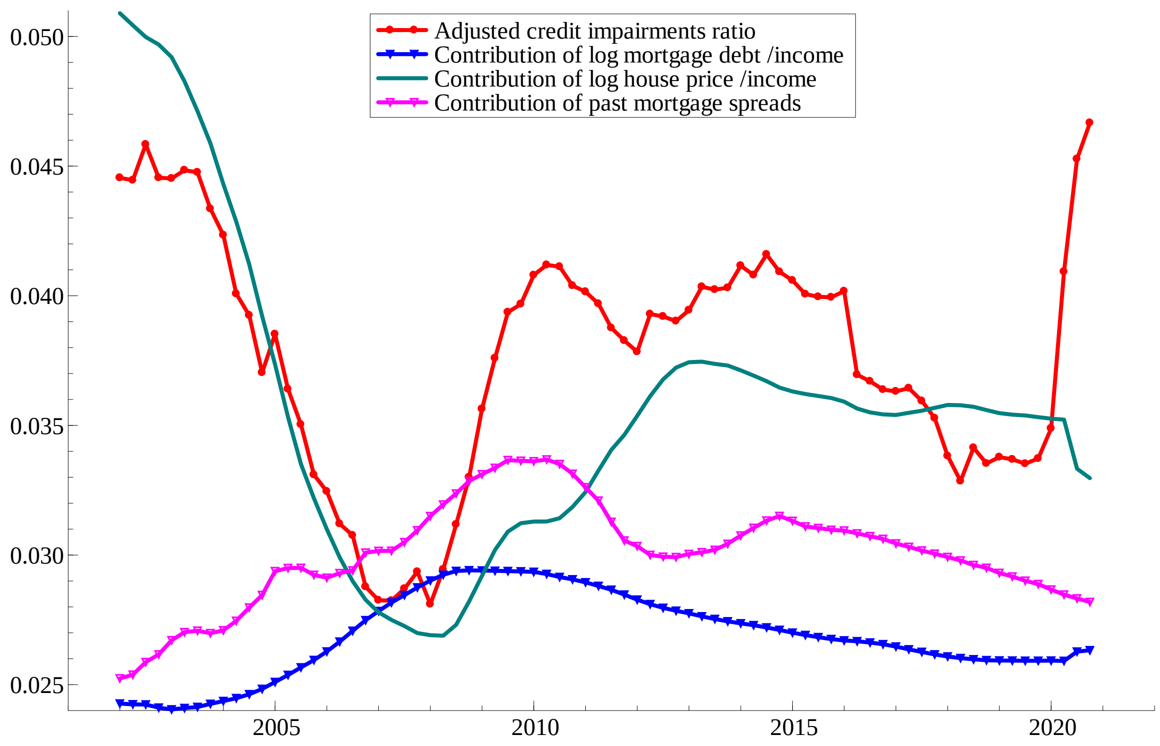


Figure 8: Adjusted CIR and the contemporaneous mortgage interest spread

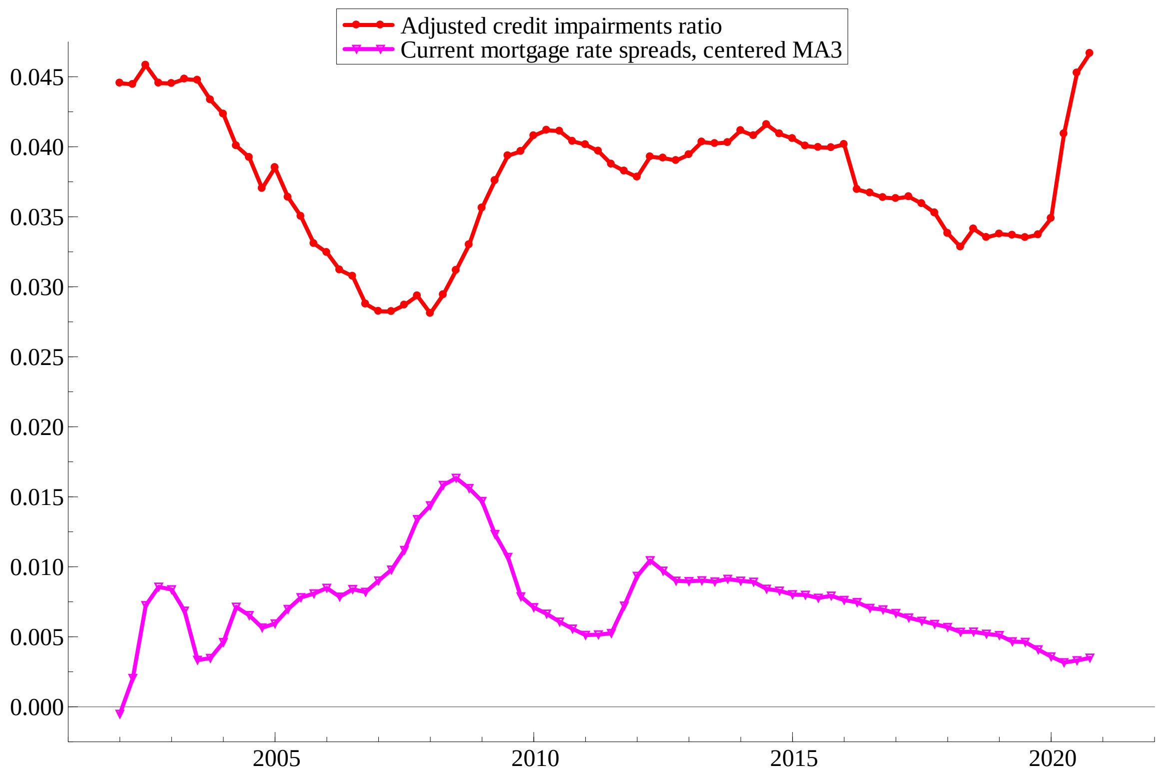


Figure 9: Adjusted CIR and average loan-to-value ratio

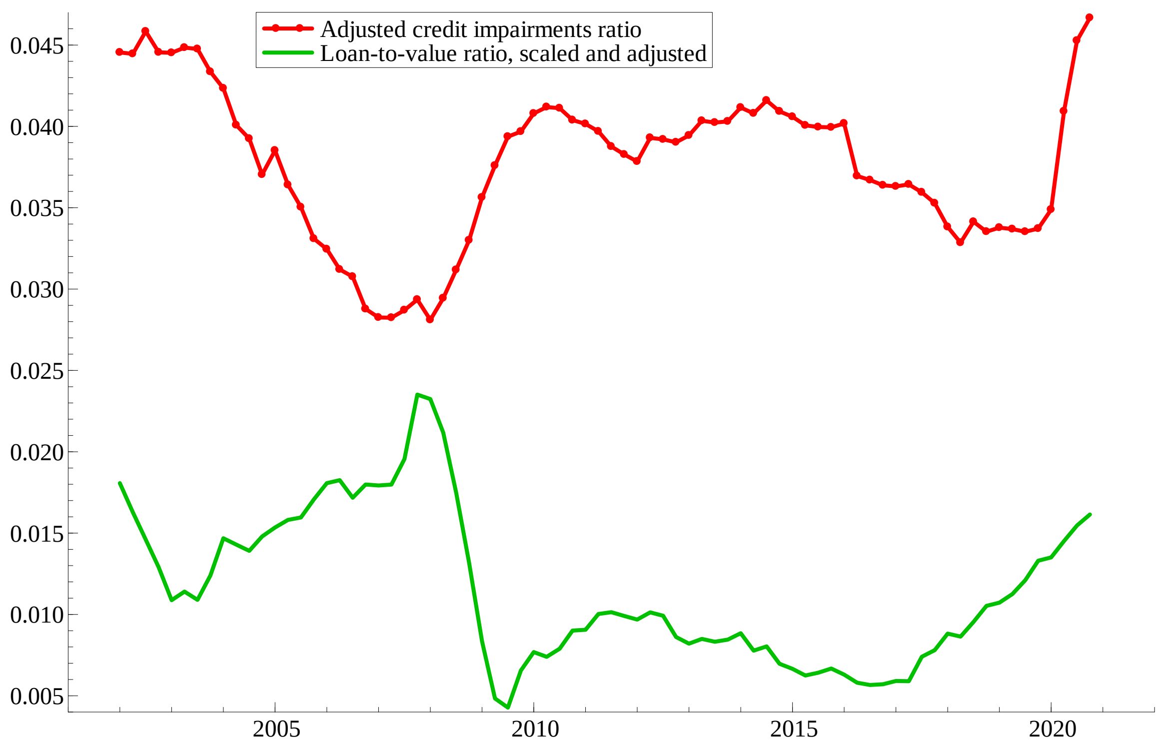
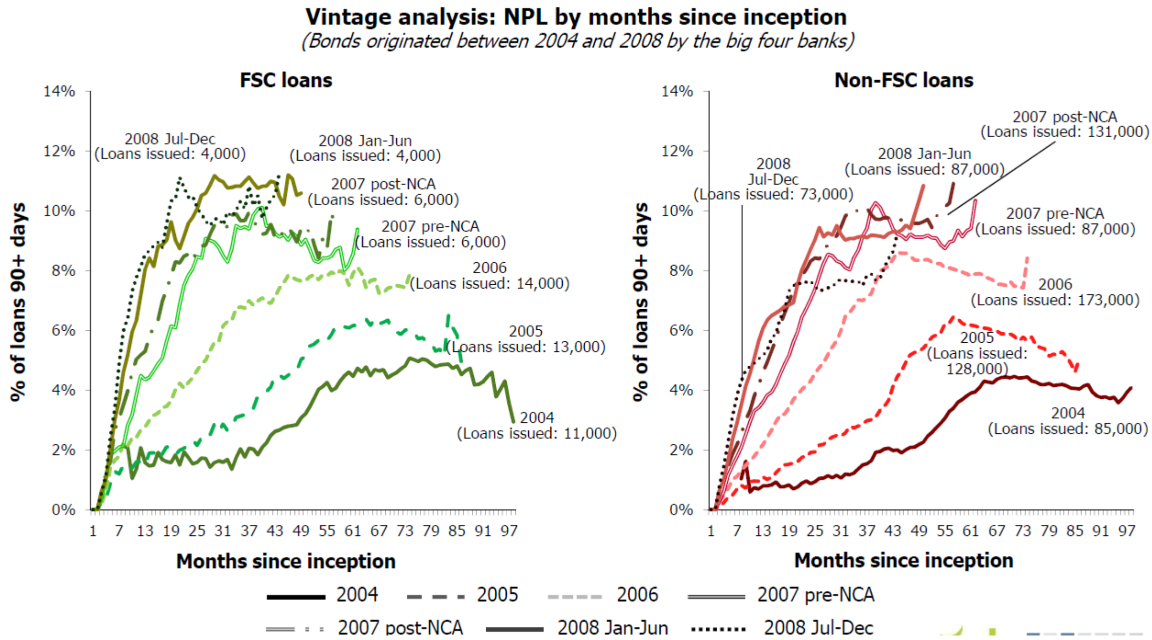
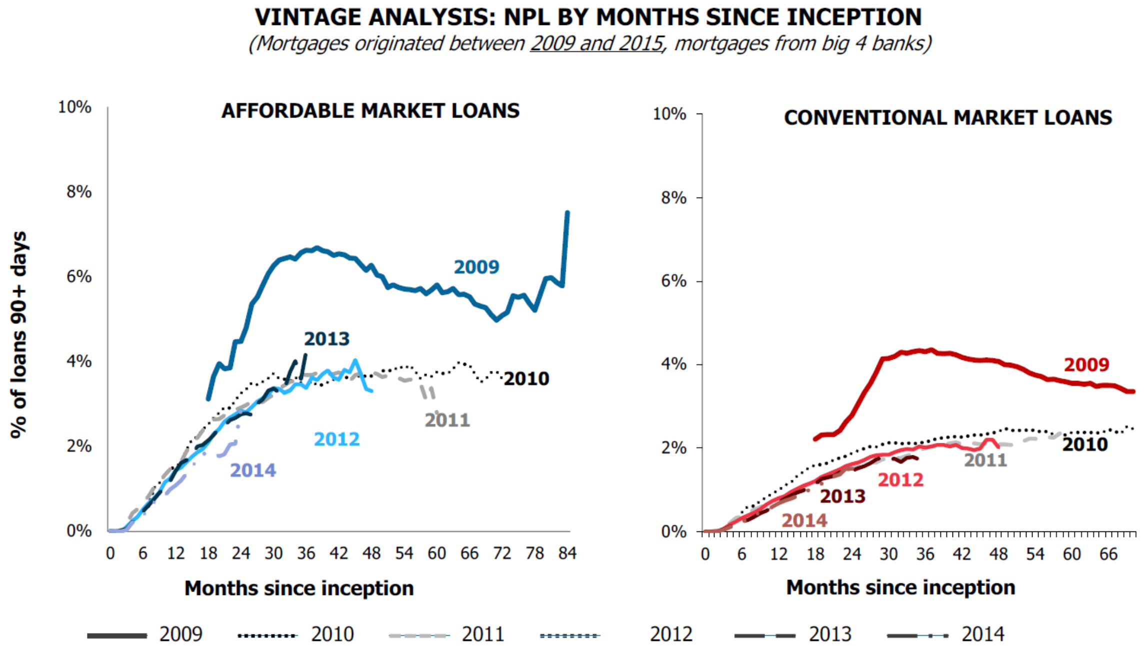


Figure 10: Vintage analysis: NPL by months since inception (origination between 2004 and 2008)



Source: Melzer (2012), Deeds office data sourced from the Affordable Land & Housing Data Centre, HSRC (ALHDC), bureau data from the XDS bureau.

Figure 11: Vintage analysis: NPL by months since inception (origination between 2009 and 2015)



Source: Melzer (2017).

Table 1: Recent global trends in annual NPLs

	2008	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*
Australia	1.36	2.02	2.15	1.97	1.70	1.37	1.01	0.89	0.95	0.86	0.90	0.96	1.11
Botswana					2.62	3.61	3.55	3.73	4.85	5.28	5.43	4.79	4.32
Brazil	3.11	4.21	3.11	3.47	3.45	2.86	2.85	3.31	3.92	3.59	3.05	3.11	2.24
Chile	0.98	2.93	2.69	2.35	2.16	2.11	2.06	1.87	1.83	1.92	1.87	2.06	1.55
Colombia	3.93	4.01	2.86	2.50	2.76	2.77	2.92	2.85	3.12	4.18	4.40	4.17	4.80
Czech Rep.	2.81	4.58	5.39	5.22	5.24	5.20	5.61	5.48	4.59	3.74	3.14	2.70	2.96
Eswatini		5.87	7.81	3.53	10.67	6.79	6.87	6.63	9.62	7.92	9.32	9.45	
France	2.82	4.02	3.76	4.29	4.29	4.50	4.16	4.05	3.70	3.12	2.75	2.47	2.71
Germany	2.85	3.31	3.20	3.03	2.86	2.70	2.34	1.97	1.71	1.50	1.24	1.05	
Greece	4.67	6.95	9.12	14.43	23.27	31.90	33.78	36.65	36.30	45.57	41.99	36.45	26.98
India	2.45			2.67	3.37	4.03	4.35	5.88	9.19	9.98	9.46	9.23	7.94
Indonesia	3.19	3.29	2.53	2.14	1.77	1.69	2.07	2.43	2.90	2.56	2.29	2.43	2.75
Ireland	1.92	9.80	13.05	16.12	24.99	25.71	20.65	14.93	13.61	11.46	5.73	3.36	3.54
Italy	6.28	9.45	10.03	11.74	13.75	16.54	18.03	18.06	17.12	14.38	8.39	6.75	4.36
Lesotho		3.02	3.03	2.10	2.46	3.79	4.23	4.04	3.69	4.42	3.66	3.30	4.20
Mexico	2.97	2.81	2.04	2.12	2.44	3.24	3.04	2.52	2.09	2.09	2.05	2.09	2.43
Mozambique						2.66	3.24	4.31	5.73	12.64	11.12	10.16	
Namibia			1.96	1.49	1.34	1.29	1.45	1.55	1.54	2.59	3.58	4.56	6.39
Netherlands	1.68	3.20	2.83	2.71	3.10	3.23	2.98	2.71	2.54	2.31	1.96	1.86	1.89
Peru			3.03	2.89	3.23	3.50	3.95	3.93	4.29	4.70	3.27	3.37	4.13
Poland	2.82	4.29	4.91	4.66	5.20	4.98	4.82	4.34	4.05	3.94	3.85	3.80	3.71
Portugal	3.60	5.13	5.31	7.47	9.74	10.62	11.91	17.48	17.18	13.27	9.43	6.18	4.86
Romania	2.75	7.89	11.85	14.33	18.24	21.87	13.94	13.51	9.62	6.41	4.96	4.09	3.83
South Africa	3.92	5.94	5.79	4.68	4.04	3.64	3.24	3.12	2.86	2.84	3.73	3.89	5.18
Spain	2.81	4.12	4.67	6.01	7.48	9.38	8.45	6.16	5.64	4.46	3.69	3.15	2.85
United Kingdom	1.56	3.51	3.95	3.96	3.59	3.11	1.65	1.01	0.94	0.73	1.07	1.08	1.22
United States		4.96	4.39	3.78	3.32	2.45	1.85	1.47	1.32	1.13	0.91	0.86	1.07

Source: IMF Financial Soundness Indicators, the ratio of non-performing loans as a percentage of total gross loans. Note, the table draws on figures using different methodologies and definitions across countries, and these may also change over time within countries.

Table 2a: Comparing provisions for impaired exposures under IAS 39 and IFRS 9

IAS 39	**Unimpaired loans** **Impairments:** minimal		**Impaired loans** **Impairments:** lifetime incurred and expected loss
IFRS 9	**Stage 1** Performing loans **Impairment:** 12-month expected loss	**Stage 2** Underperforming loans **Impairment:** lifetime expected loss	**Stage 3** Non-performing loans **Impairments:** lifetime incurred and expected loss

Table 2b: Mapping regulatory frameworks for NPLs with accounting concept of 'impaired'

Number of countries	Framework	Risk buckets	Pass / Normal	Special Mention or Watch	Substandard	Doubtful	Loss
9 (6 Asia, 3 LAC)	Regulatory	5	1	2	3	4	5
2 (1 Asia, US)[31]		5	1	2	3	4	5
3 (2 Asia, 1 LAC)		4	1		2	3	4
1 (Asia)		6	1	2	3	4–5	6
1 (LAC)		6	1	2	3	4	5–6
1 (LAC)		6	1–3		4	5	6
1 (LAC)		8	1–2	3–4	5–7		8
1 (LAC)		9	1–2	3–4	5–7		8–9
1 (LAC)		9	1–7			8	9
1 (LAC)[32]		16	1–6	7–10	11–16		
2 (EU-SSM,[33] 1 Asia[34])		2	Performing		Non-performing		
	Accounting (IFRS)	3	Stage 1 and 2 (unimpaired)		Stage 3 (impaired)		

Source: Baudino *et al.* (2018), see this paper for footnotes 137 to 140 indicated in Table 2b.

Table 3: Institute of International Finance (IIF) loan classification scheme

Category of loan	Definition
Standard	Credit is sound and all principal and interest payments are current. Repayment difficulties are not foreseen under current circumstances and full repayment is expected.
Watch	Asset subject to conditions that, if left uncorrected, could raise concerns about full repayment. These require more than normal attention by credit officers
Substandard	Full repayment is in doubt due to inadequate protection (e.g., obligor net worth or collateral) and/or interest or principal **or both are more than 90 days** overdue. These assets show underlying, well-defined weaknesses that could lead to probable loss if not corrected and risk becoming impaired assets.
Doubtful	Assets for which collection/liquidation in full is determined by bank management to be improbable due to current conditions and/or interest or principal **or both are overdue more than 180 days**. Assets in this category are considered impaired but are not yet considered total losses because some pending factors may strengthen the asset's quality (merger, new financing, or capital injection).
Loss	An asset is downgraded to loss when management considers the facility to be virtually uncollectible and/or principal or interest or both are overdue more than one year.

Source: Krueger (2002).

Table 4: Changing definitions for South African NPL and related data

	Impaired Advances	*Specific Credit Impairment*	*Default Ratio*	*60/90/180 Days Overdue*
Source	PA	PA and QB: KBP1123M	PA	Banking Supervision Department/PA
Type of Concept	Accounting NPL concept.	Accounting concept of loan loss provisions.	Prudential NPL concept.	Prudential NPL concept.
Definition	Advances in respect of which a bank has raised a specific impairment and includes any advance or restructured credit exposures subject to amended terms, conditions or concessions that are not formalised in writing. Expressed in R billions or as a percentage of on-balance-sheet loans and advances.	Provisioning (credit impairments) is an accounting concept, which is an allowance made against losses on loans identified as bad or doubtful, including provisions made against groups of loans based on their age. The coverage ratio is the ratio of specific credit impairments as a percentage of impaired advances. Expressed in R billions, or as ratio to impaired advances, or ratio to gross loans and advances.	The ratio of defaulted exposure as a percentage of total exposure at default.	The *90 days overdue ratio* is defined as all exposures overdue for more than 90 days as a percentage of on-balance sheet exposures. The *past due ratio* is calculated as exposures greater than 60 days but less than or equal to 90 days, as a percentage of total on-balance-sheet exposure.
Frequency	Monthly	Monthly	Monthly	Monthly
Published range	Published from 2008 onwards.	Published from 1991 onwards.	Published as an aggregate for all banks from 2008 onwards.	Published from at least 1994 (since the first online PA/Banking Supervision reports date from 1994) until 2000. For 1994 to 1997, total overdues are published without indicating the relevant age categories. From 1998 to 2000: age analysis is given: <30 days; <90 days; >90 days. From 2001, the DI500 form (last dated December-2007) gives credit quality categories: standard/current; special mention; sub-standard; doubtful; loss. While these are quantitatively related to different days overdue, other qualitative risk factors are incorporated too. From 2012, data exist for 90-days overdue both for banks following the IRB method and the standardised method.

	Impaired Advances	*Specific Credit Impairment*	*Default Ratio*	*60/90/180 Days Overdue*
Potential longer time series? ***See also Section 3.4***	*It is unclear if comparable data exist before 2008 to make a continuous series – even though credit impairments against impaired loans (series KBP1123M, next column) dates back to 1991.*	*Given that the 'impaired advances' are defined as those against which a 'specific provision' has been made, ideally the 'general provisions' – see below under Definitional Changes – should be separated out from 2008 from this KBP1123M credit impairment series for improved consistency in modelling.*	*Default ratios for banks using the standardised approach are defined by adding the three credit risk categories: 'sub-standard; doubtful; loss'. These data can be taken back to 2001.* *Banks reporting on the IRB approach also report default ratios.* *On a bank-by-bank approach, data on the earlier standardised approach can be linked with data on the IRB approach by comparing the last month on the former with the first month of the latter.*	*It is unclear whether the detailed overdue per days data collected after regulatory changes in 2012– see below under Definitional Changes – were also collected from 2008 to 2011. For 2001 to 2007, a very approximate translation of the five risk buckets into days overdue may be possible.*
Disaggregation of the data	Cannot split into asset classes. Can be split by banks: e.g., SIFI versus non-SIFI; Individual banks.	Cannot split into asset classes.	From 2008 for all banks, totals for individual banks, foreign branches, SIFIs. From 2008 for IRB banks: total, retail, corporate and SMEs for IRB (also available for standardised banks from at least 2012). From at least 2015 (but possibly 2008), for the IRB banks, corporate sub-sectors (construction, manufacturing, wholesale trade, retail trade, real estate, electricity-gas-water). Household sector: total, secured and unsecured lending; residential mortgages, revolving credit facilities, and vehicle and asset finance categories. For 2001- 2007, defaults ratios for mortgages, instalment finance and other loans and advances could in principle be separately constructed.	Household sector (90 days overdue) secured and unsecured lending published after adjustment of the bank returns form from 2012. Credit products (mortgages, leases, instalments, other) from 1994 to 2007. From the NCR: multiple days overdue measures are available from 2007Q3 (for the entire financial sector i.e., banks and non-bank credit providers), but only for households (secured and unsecured).
Definitional Changes	The implementation of IFRS 9 in 2018 *caused a jump in these data.*	From January 2001, with new Banking Regulations, only loans	In 2008, 4 banks were given permission to use (advanced or foundation) IRB models for credit rating. The other banks used the	Before 1998, it is unclear how 'overdue loans' as shown in Annual PA reports were defined.

	Impaired Advances	*Specific Credit Impairment*	*Default Ratio*	*60/90/180 Days Overdue*
		classified as "doubtful", and "loss" were regarded as "overdue". *There is a significant fall in the data.* The pre-2008 data are defined as 'specific provisions', while from January 2008 they include 'specific and general provisions. *There is a significant jump (rise) in the data.* From January 2018, 'the measurement of impairments according to the expected credit loss model replaced the incurred loss model' following implementation of IFRS9. *There is a significant jump (rise) in the data.*	standardised approach. A further bank has since migrated to IRB methods. There are likely to be discontinuities in default ratios in the month where the switch to the IRB approach came into force. These can be minimized in the bank-by-bank linkage method discussed above.	From 1998 to 2000, they were defined as 90 or more days overdue. From 2001 to 2007, after Amended Regulations in 2000 relating to banks were implemented on 1 January 2001, only loans classified as "doubtful", and "loss" were regarded as "overdue" for purposes of Annual PA reports. These do not precisely match a days-overdue category. For banks following the IRB method, 90-day overdue data begin in 2008; for the other banks, they begin in 2012. Regulatory changes in 2012, see 12-Dec-2012, Government Notice no 1029, required data to be collected showing *what proportion* of each category in the standardised approach comprised loans on a per days concept, such as 'greater than 90 days'.

Source: Constructed by the authors.

Table 5: Variables used in the Ari *et al.* (2019) study of NPL determinants

<table>
<tr><td colspan="2">DEPENDENT VARIABLES</td></tr>
<tr><td>1</td><td>Likelihood of elevated NPLs, with elevated NPLs defined as larger than 7% of total loans.
(For alternative definition: larger than 5%).</td></tr>
<tr><td>2</td><td>Peak NPLs as a percentage of total loans (note, these are elevated NPLs).
(For alternative definition: relative to NPL ratio at crisis date).</td></tr>
<tr><td>3</td><td>Time to peak from start of crisis.
(For alternative definition: relative to first year when NPL ratio was greater than 7%).</td></tr>
<tr><td>4</td><td>Time to resolve, defined as time for elevated NPLs since start of crisis to fall below 7% of total loans
(For alternative definition: defined relative to the first year when NPLs exceeded 7% of total loans).</td></tr>
<tr><td>5</td><td>Likelihood of timely resolution, defined as whether NPLs decline to under 7% within 7 years from the start of crisis.
(For alternative definition: defined as whether NPLs falling below 25% of peak NPL).</td></tr>
<tr><td colspan="2">PREDICTOR SETS</td></tr>
<tr><td>I</td><td>Covers pre-crisis domestic macroeconomic and external conditions</td></tr>
<tr><td colspan="2">GDP growth, domestic credit to the private sector, unemployment and inflation rates.
Government-debt-to-GDP ratio.
Change in the bilateral nominal exchange rate against the U.S. dollar, and two dummy variables for an exchange rate peg and whether the peg was broken, measured in the 5-year period prior to the crisis.
A country’s GDP per capita as the proxy for institutional strength.</td></tr>
<tr><td>III</td><td>Set I <u>plus</u> predictors reflecting pre-crisis banking sector conditions</td></tr>
<tr><td colspan="2">Bank return on assets and equity, net interest margins, operating-expense-to-net-interest-income ratio, and noninterest-income-to-total-income ratio.
(Bank capitalization is excluded due to the lack of available data).
Measures of bank concentration.
Rule of law index from the World Bank Worldwide Governance Indicators.</td></tr>
<tr><td>IIII</td><td>Set II <u>plus</u> predictors reflecting pre-crisis corporate conditions</td></tr>
<tr><td colspan="2">Non-financial corporate debt-to-assets ratio.
Earnings before interest and taxes (EBIT) to total interest expense ratio.
Share of short-term debt in total debt, and the current-asset-to-liability ratio.
Share of foreign assets in total assets.</td></tr>
</table>

Source: Compiled from Ari *et al.* (2019).

Notes: These authors follow the crisis start dating in Laeven and Valencia (2013, 2018). Predictor variables are measured as averages or cumulative changes over the five years prior to the crisis and all dependent variables on or after the crisis date.

Table 6: Typology of NPL determinants and expected signs

Determinants	Definition	*Meta-study* **Macháček et al. (2017): 37 studies**	*Review* **Naili and Lahrichi (2020): 69 studies**	*Cross-sectional study* **Ari et al. (2019) ***
Macroeconomic variables				
GDP growth	Annual percentage growth rate of GDP	Majority find negative, where significant	Majority reviewed find slower growth raises NPLs	*In Set 1.* Dep3. Time to peak: - Dep5. *p*(Resolution): +
Unemployment	Unemployment rate in year t/change in	Majority find positive, where significant	Majority reviewed find positive relationship between unemployment and NPLs	*In Set 1.* Dep3. Time to peak: - Dep5. *p*(Resolution): +
Inflation	Annual average inflation rate	Ambiguous	Ambiguous outcomes	*In Set 1. Not significant.*
Interest rate	Lending interest rate	Majority find positive, where significant	Generally, higher interest rates raise bank lending rates and NPLs. Fixed versus floating regimes yield differences.	*In Set 1. Not significant.*
Real estate or house prices	Changes in the house price index.**	NA	NA	NA
Exchange rate	Nominal exchange rate depreciation; peg; regime change	Ambiguous outcome.	Ambiguous outcome. NPLs worsen with high proportions of private sector debts dominated in foreign currencies.	*In Set 1.* Dep2. Peak NPL: - Dep3. Time to peak: - Dep5. *p*(Resolution): +
Exchange rate	The change in real exchange rate	NA	With high export volumes and low currency mismatches, depreciation reduces NPLs.	NA
Domestic credit to the private sector	Growth in domestic credit.		NA, but see banking sector results on loan growth below.	*In Set 1.* Dep1. High *p*(NPL): + Dep3. Time to peak: +

Determinants	Definition	*Meta-study* Macháček et al. (2017): 37 studies	*Review* Naili and Lahrichi (2020): 69 studies	*Cross-sectional study* Ari et al. (2019) *
				Dep4. Time to resolve: + Dep5. *p*(Resolution): -
Public debt	Gross government debt as % of GDP or change in.		Generally, there is a positive association between worsening public debt and NPLs.	*In Set 1.* Dep3. Time to peak: + Dep4. Time to resolve: + Dep5. *p*(Resolution): -
Institutional environment /strength	Institutional strength proxied by GDP per capita		NA	*In Set 1.* Dep1. High p(NPL): - Dep3. Time to peak: +
	Rule of law index: The World Bank's *Worldwide Governance Indicators*		NA	*In Set 2. Not significant.*
	Corruption Perception Index		Various studies find a positive association between corruption and NPLs.	NA
Bank-specific variables				
Bank capitalization (CAR)	(Tier 1 Capital+Tier 2 Capital)/Risk Weighted assets Equity/Total Assets	NA	Opposing results. Many studies suggest a high CAR reduces NPLs. Banks with more capital at risk are more likely to engage in prudent lending with adequate loan screening. Prominent studies find the opposite result.	Would have included in *Set 2* if there were data.
Bank size	Natural log of total assets		No clear-cut evidence. "Too big to fail" banks may take excessive risk. Otherwise, size may reduce NPLs as large-sized banks have modern risk management systems and procedures are in a better position to conduct proper loan screening.	NA

Determinants	Definition	*Meta-study* **Macháček et al. (2017): 37 studies**	*Review* **Naili and Lahrichi (2020): 69 studies**	*Cross-sectional study* **Ari et al. (2019) ***
Bank efficiency	Operating expenses/Operating income *(Ari uses operating-expense to net-interest-income ratio)*		Evidence from several studies finds cost inefficiency is linked with poor management and hence worse NPLs. "Skimping" on good loan quality with apparent cost efficiency, may only be reflected in growing future NPLs, however. Shocks can cause NPLs and at the same time decrease apparent efficiency through the extra costly operations required.	*In Set 2.* Dep4. Time to resolve: -
Bank performance	ROE= Net income/Total equity		Profitability is linked with NPLs, but the sign varies. Several studies find that higher profitability reduces NPLs, as there is less inclination for excessive risk-taking with resultant higher quality loan portfolios. But banks could game the market to conceal bad loans by a liberal credit policy with subsequent high NPLs.	*In Set 2. Not significant.*
	ROA= Net income/Total assets			*In Set 2.* Dep2. Peak NPL: -
Bank profitability	Net interest margins - the difference between interest received and interest paid		NA	*In Set 2. Not significant.*
Loan growth	Percentage growth of total loans between two consecutive years		The main finding is that rapid credit growth leads to riskier lending behaviours through adverse selection, inappropriate managerial incentives and limited risk-screening capacity or reduced screening standards in boom periods, worsening credit quality.	NA, but see private sector credit growth above.
Bank diversification	Non-interest income/Total income		Ambiguous results. Limited studies tend to find NPLs are worsened with more diversification through increasing risk from inexperience, lack of comparative advantage and business failure.	*In Set 2.* Dep5. *p*(Resolution): +
Managerial factors	CEO compensation - salary, bonus, long-term incentive plan, other annual compensation, value of option grants, value of restricted stocks grants, value change of existing option holdings, value change of existing restricted stocks, and value change of direct equity holdings.		Scant evidence on a direct link with NPL levels, but varied views that higher executive compensation increases managerial risk-taking in risky banks, while option-based compensation increases CEOs' risk aversion.	NA

Determinants	Definition	*Meta-study* **Macháček et al. (2017): 37 studies**	*Review* **Naili and Lahrichi (2020): 69 studies**	*Cross-sectional study* **Ari et al. (2019) ***
	Banks' overconfidence measured by cash-based or stock option-based incentives		Overconfidence may reduce managerial risk aversion, which can be detrimental, though possibly temporarily rewarding in a boom period. US evidence using the stock options proxy found overconfident managers linked with high levels of NPLs.	NA
Corporate social responsibility (CSR)	CSR Index of a bank (e.g., FTSE4 Good Global Index, EIRIS)		Scant empirical studies on direct CSR-NPL link but tend to suggest CSR-banks incur lower NPLs.	NA
Ownership concentration	Total shares held by stake insiders/Total shares outstanding		Ambiguous results. Dispersed ownership has been linked with poor incentives for proper monitoring of banks, but also more risk taking by controlling interests prioritizing personal interest. Some evidence for concentrated ownership reducing bank risk exposures, and significantly reducing NPLs. Other studies find poorer performance, lower cost efficiency, higher risk-taking and worse loan quality.	NA
	Concentration based on ownership levels 10, 25 or 50%.			NA
Ownership identity	The identity of major shareholder: State or Institutional		Few studies directly on ownership structure and NPLs. State ownership linked to high risk-taking and poor performance. Institutional ownership reduces NPLs.	NA
Concentration in the banking industry	Lerner Index. Boone indicator: elasticity of profits to marginal costs. Concentration ratio: sum of squared market share of largest banks.		Ambiguous. Some evidence supports the competition–fragility hypothesis, that more competition increases NPLs.	*In Set 2. Not significant.* (not sure which measures)
Non-financial corporate sector variables				
Corporate leverage	Debt-to-assets ratio	NA	NA	*In Set 3.* Dep1. High p(NPL): + Dep5. *p*(Resolution): +
Corporate debt service capacity	Earnings before interest and taxes (EBIT) to total interest expense ratio		NA	*In Set 3. Not significant.*

Determinants	Definition	*Meta-study* Macháček et al. (2017): 37 studies	*Review* Naili and Lahrichi (2020): 69 studies	*Cross-sectional study* Ari et al. (2019) *
Corporate debt service capacity	Share of short-term debt in total debt		NA	*In Set 3.* Dep2. Peak NPL: + Dep3. Time to peak: -
Maturity profile of debt and the rollover risk	Current-asset-to-liability ratio		NA	*In Set 3* Dep5. *p*(Resolution): -
International competitiveness	Share of foreign assets in total assets		NA	*In Set 3. Not significant.*

Source: Compiled by the authors.

Notes: *The five dependent variables of Ari *et al*. (2019), see Table 5, are elevated NPLs (*Dep 1*), the peak NPLs as a percentage of total loans (*Dep 2*), the time to reach the NPL peak (*Dep 3*), the time to resolve NPLs (*Dep 4*), and the likelihood of resolution (*Dep 5*). Three sets of variables are used, the first in *Set 1* comprises macro-variables, then in *Set 2* appended by banking variables, which are in *Set 3* appended by non-financial firm/industry variables. The most complete general-to-specific results therefore would be when *Set 3* is included.

** See discussion in the text as none of the above three studies included this variable.

Table 7: Potential NPL dependent variables in SA from bank-issued loans

Type of non-performing loan	*Denominator*	*NPL definition and frequency*	*Source/code and span*
Total			
Value of total NPLs from the banking sector	Gross outstanding loans.	*Monthly* impaired advances. *Monthly* default ratios (Basel II). *Monthly* 90 days overdue.	For impaired loans from 2008 (*and possibly approximately back to 2001, see Section 3.4)*. For default ratios from 2008 (*and close approximation back to 2001, see Section 3.4*). For 90-day overdue (*very approximately from 1998, see Section 3.4)*
Household			
Value of total household NPLs	Total outstanding household loans.	*Monthly* default ratios. *Monthly* 90 days overdue.	For default ratios and 90-days overdue, from 2008 for banks using the IRB approach. For default ratios and 90-days overdue, from 2012 for banks using the standardised approach.
Value of total secured household NPLs (mortgages & other secured debt)	Total outstanding household secured loans.		
Value of total unsecured household NPLs	Total outstanding household unsecured loans.		
Value of total mortgage NPLs (largely comprising residential)	Total outstanding 'Residential' mortgage loans.	*Monthly* default ratios. *Monthly* 90 days overdue.	For default ratios from 2008 (*and close approximation back to 2001, see Section 3.4*). For 90-day overdue (*very approximately from 1998, see Section 3.4)* (Before 2008 only, mortgage data include corporate mortgages).
Non-financial corporate			
Value of commercial & other mortgage NPLs	Total outstanding 'Commercial & other' mortgage loans.	*Monthly* default ratios.	For default ratios from 2008 for banks using the IRB method.
Total non-financial corporates	Total outstanding NFC bank debt.	*Monthly* default ratios.	For default ratios from 2008 for banks using the IRB method. Aggregates only for standardised banks from 2012.
By sector: real estate, mining and quarrying, manufacturing, construction, trade, business services, electricity+gas+water, personal services, transport, and communication	Total outstanding NFC bank debt, by sector.		
Bank-level and for individual banks			
Value of total NPLs, all banks or single banks	Gross outstanding loans for all banks or single banks	*Monthly* impaired advances. *Monthly* default ratios. *Monthly* 90 days overdue.	For impaired loans from 2008 (*and possibly approximately back to 2001, see Section 3.4)*. For default ratios from 2008 (*and to a close approximation for all banks or groups of banks back to 2001, see Section 3.4*). For 90-day overdue (very *approximately from 1998, for all banks or groups of banks see Section 3.4)*
Value of total NPLs, TOP 4 to 8 banks	Gross outstanding loans for TOP 4 to 8 banks		
Value of total NPLs, SIFIs or foreign.	Gross outstanding loans for *residual* smaller banks		

Source: Compiled by the authors.
Notes: This table should be read in conjunction with Table 4, giving the institutional definitional changes over time for the NPL measures. This table does not include the credit risk measure: 'credit impairments', where Havrylchyk (2010) modelled disaggregated data. This table does not refer to NCR data for households, which begin in 2007Q3.

Table 8: Potential NPL driver variables by category in South Africa

Driver	***Frequency***	***Source***
Macro-determinants		
GDP growth	Quarterly	Quarterly Bulletin
Domestic credit extended to the private sector	Monthly	Quarterly Bulletin
Unemployment rate	Monthly	Quarterly Bulletin
Inflation rate	Monthly	Quarterly Bulletin
Lending interest rate	Monthly	Quarterly Bulletin
Government-debt-to-GDP ratio	Quarterly	Quarterly Bulletin
Nominal and real exchange rates	Monthly	Quarterly Bulletin
GDP per capita (*used as a governance indicator*)	Quarterly	Quarterly Bulletin
Macro-determinants seldom included but relevant in SA		
Real estate prices relative to income	Quarterly	Quarterly Bulletin
Household debt relative to household income	Quarterly	Quarterly Bulletin
Corporate debt relative to GDP	Quarterly	Quarterly Bulletin
Real personal income growth	Quarterly	Quarterly Bulletin
Credit conditions proxies e.g., LTVs and interest spreads	Monthly	Quarterly Bulletin, SARB
Banking determinants		
Bank size: Natural log of total assets (total assets in R billions)	Monthly	FSR Appendix/PA
Bank profitability, return on assets and equity: ROE= Net income/Total equity return on equity (smoothed) ROA= Net income/Total assets return on assets (smoothed)	Monthly	FSR Appendix/PA
Bank efficiency: Operating expenses to gross income (smoothed)*	Monthly	FSR Appendix
Bank profitability: Interest margin to gross income (smoothed)	Monthly	FSR Appendix
Bank capitalization (CAR): Total capital adequacy ratio; Tier 1 capital adequacy ratio; and Common equity Tier 1 capital adequacy ratio. Banks' share prices (percentage change).	Monthly	FSR Appendix/PA
Loan growth: Total loans and advances (R billions)	Monthly	FSR Appendix/PA
Lending specialisation: Total loans/total assets ratio	Monthly underlying data should allow construction of ratio	FSR Appendix
Bank diversification: Non-interest income/gross income [*Not clear if there is a separate measure of non-interest income*]		FSR Appendix
Measures of bank concentration: Herfindahl–Hirschman Index (H-index); Gini concentration index; and Market share in terms of assets (top 5 banks)	Monthly	FSR Appendix
	Monthly	FSR Appendix
	Monthly	FSR Appendix
Credit rating: Percentage change in credit rating composition of on-balance-sheet loans	Monthly	FSR, e.g., 2019 FSR I, Table 2, p.11.
Governance indicators: e.g., Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption	Annual from 1996; World Bank reports	World Bank, *Worldwide Governance Indicators*

Driver	***Frequency***	***Source***
Non-financial corporate sector variables		
Corporate debt service capacity: corporate sector interest coverage ratio (ICR) disaggregated by industry (measured as earnings before interest and taxes (EBIT) to total interest expenses).	Quarterly	FSR, e.g., 2019 FSR I, Figure 21, p.24.
Bank credit extended to the corporate sector: instalment sale and leasing finance, mortgage advances, overdrafts, credit card debtors, and other loans and advances [used as a proxy for corporate debt]	Quarterly	FSR, e.g., 2019 FSR I, Table 8, p.23.
Corporate profitability: net operating surpluses of corporations used as proxy for corporate profits		FSR, e.g., 2019 FSR I, Table 5, p.23.
Corporate debt service capacity: Share of short-term debt in total debt [see above for proxy of total debt] [*Not clear if there is a measure of short-term debt*]	Quarterly underlying data should allow construction of ratio	FSR, e.g., 2019 FSR I, Table 8, p.23.
Corporate leverage: corporate debt-to-assets ratio [see above for proxy of total debt] [*Not clear if there is a measure of assets*]		FSR, e.g., 2019 FSR I, Table 8, p.23.
Maturity profile of debt and rollover risk: current-asset-to-liability ratio [see above for proxy of total debt] [*Not clear if there is a measure of assets*]		FSR, e.g., 2019 FSR I, Table 8, p.23.
International exposure: foreign currency composition of non-financial corporate debt	Quarterly	FSR, e.g., 2019 FSR I, Figure 20, p.24.

Source: Compiled by the authors.

*Notes: *Ghosh (2017) measures operational efficiency by non-interest expenses divided by total assets (i.e., banks' overhead costs-to-assets).*

Table 9: Data definitions and sources for the Credit Impairments Ratio (CIR) model

Variable	Definition	Mean	Std. deviation	Minimum	Maximum	Data source
DEPENDENT VARIABLE						
Δ (Credit Impairments Ratio)	Quarterly change in the ratio of credit impairments (from the monthly series KBP1123M) divided by total gross loans and advances issued to the private sector by banks.	0.0162	0.146	-0.322	0.523	Quarterly Bulletin
INDEPENDENT VARIABLES						
Credit Impairments Ratio	The ratio of credit impairments (from the monthly series KBP1123M) divided by total gross loans and advances issued to the private sector by banks.	2.47	0.630	0.989	3.49	Quarterly Bulletin
log (mortgage debt relative to income (ma4))	The log of the 4-quarter moving average of mortgage debt (in current prices) divided by income (in current prices).	-0.944	0.191	-1.28	-0.655	Quarterly Bulletin
log (house price to income ratio (ma4))	The log of the 4-quarter moving average of the house price index divided by per capita nominal household disposable income.	0.732	0.141	0.362	0.946	Quarterly Bulletin and SARB
Δ_4 log (REER)	The yearly change in the log of the real effective exchange rate. A rise is a Rand appreciation.	0.00433	0.104	-0.204	0.262	Quarterly Bulletin
Electric Power Outages (ma3)	The deviation from the lagged 8-quarter moving average of log electricity output; then take the 3-quarter moving average of this deviation.	0.0115	0.0269	-0.0552	0.0716	Quarterly Bulletin
Mortgage rate spread (ma4)	The prime rate minus the average interest rate on new mortgage loans (all expressed as a 4-quarter moving average).	0.761	0.317	-0.135	1.60	SARB
Δ_8 log (per capita real GDP)	Two-year change in the log of GDP (in constant prices) divided by population.	0.0227	0.0302	-0.0224	0.0829	Quarterly Bulletin
Split Dummy 2007Q4	1 up to 2007Q4, zero thereafter					Constructed
Split Dummy 2017Q4	1 up to 2017Q4, zero thereafter					Constructed
Split Dummy 2017Q4 (ma4)	The 4-quarter moving average of the above					Constructed
Split Dummy 2016Q1	1 up to 2016Q1 inclusive, zero thereafter.					Constructed
Δ Dummy 2004Q4	The quarterly difference in the impulse dummy for 2004Q4.					Constructed

Table 10: A new Credit Impairments Ratio (CIR) model for South Africa

Dependent variable: Δ (Credit Impairments Ratio) $_t$	*2002:1 to 2020:1*		*2002:1 to 2020:1*		*2008:2 to 2020:1*	
	Eq. 1		*Eq. 2*		*Eq. 3*	
	Coefficient	*t-Statistic*	*Coefficient*	*t-Statistic*	*Coefficient*	*t-Statistic*
Constant	4.81	13.94	4.45	13.92	5.72	6.21
Credit Impairments Ratio $_{t-1}$	-0.709	-14.27	-0.683	-13.64	-0.819	-7.00
Seasonal Q1 $_t$	0.034	2.29	0.037	2.41	0.036	1.99
Seasonal Q3 $_t$	0.036	2.58	0.036	2.50	0.021	1.17
Split Dummy 2007Q4 $_t$ 1 to 2007Q4, zero thereafter	-0.456	-10.46	-0.459	-10.45	0.000	-
Split Dummy 2017Q4 $_t$ 1 to 2017Q4, zero thereafter	-0.350	-5.77	-0.328	-5.29	-0.270	-3.54
Split Dummy 2017Q4 (ma4) $_t$	-0.494	-5.15	-0.480	-4.87	-0.694	-3.66
Δ Dummy 2004Q4 $_t$	-0.137	-3.82	-0.130	-3.52	0.000	-
Dummy 2016Q1 $_t$	0.202	4.84	0.235	5.75	0.227	5.04
log (mortgage debt relative to income (ma4)) $_{t-2}$	0.605	5.39	0.532	4.93	0.782	4.07
log (house price to income ratio (ma4)) $_{t-2}$	-2.86	-9.62	-2.54	-9.24	-3.40	-4.93
Δ_4 log (REER) $_{t-1}$	0.180	1.83		-		-
Electric Power Outages (ma3) $_{t-1}$	-1.20	-2.42		-		-
Mortgage rate spread (ma4) $_{t-1}$	0.146	3.38	0.168	4.06	0.173	2.19
Mortgage rate spread (ma4) $_{t-5}$	0.146	2.56	0.093	1.72	0.182	2.24
Mortgage rate spread (ma4) $_{t-9}$	0.202	3.82	0.175	3.44	0.213	2.33
Δ_8 log (per capita real GDP) $_{t-1}$	-1.83	-2.13	-3.07	-4.35	-3.16	-3.81
Equation standard error	0.0482		0.0498		0.0498	
Adjusted R-squared	0.891		0.884		0.845	
Durbin-Watson	1.87		1.85		1.86	
Breusch/Godfrey LM: AR/MA4	*p* = [.246]		*p* = [.618]		*p* = [0.728]	
Chow test	*p* = [.579]		*p* = [.709]		*p* = [.652]	
Breusch-Pagan het. Test	*p* = [.618]		*p* = [.814]		*p* = [.632]	

Notes: Estimation performed in TSP 5.0 of Hall and Cumm