Regulatory arbitrage, shadow banking and monetary policy in China

Vo Phuong Mai Le, Kent Matthews, David Meenagh, Patrick Minford, Zhiguo Xiao

aCardiff Business School, Cardiff University, Wales, UK
bFudan School of Management, Fudan University, China

(Revised) July 2022

Abstract

Regulatory arbitrage is a persuasive explanation for the rapid growth in shadow bank credit. In China, the distortions caused by government support to state-owned enterprises (SOEs) and preferential lending by state-owned banks have created an environment for the development of shadow banks that lend to small and medium enterprises (SMEs). The imposition of a loan-to-deposit ratio (LDR) cap of 75% in 2009-2015 gave an additional boost to the growth of shadow bank credit by providing an incentive for conventional banks to bypass regulation and lend to SMEs via the shadow banks. The result is that shadow bank credit varied contra-cyclically to regular commercial bank credit in response to monetary policy shocks, dampening the effectiveness of monetary policy during the period of the LDR cap. This paper presents a model of the Chinese economy using a DSGE framework that accommodates a banking sector which isolates the effects of lending to SMEs by shadow banks. The model which is estimated by the method of indirect inference, allows for bank and shadow bank lending to affect the credit premium on private investment. We show that in general regular bank credit and shadow bank credit varies pro-cyclically with monetary policy but varies contra-cyclically when a LDR cap is imposed. The findings have implications for the policy of de-leveraging followed by China.

Keywords: DSGE model; China; Indirect Inference; Shadow Banking

JEL clasificación: E3; E44; E52; C1;

Corresponding Author: Kent Matthews,
matthewsk@cardiff.ac.uk, Kent.Matthews@nottingham.edu.cn

We acknowledge support from ESRC-Newton Grant ES/P004199/1 and the National Natural Science Foundation of China (Grant #71661137005 and 11571081). We are grateful to an anonymous referee, the editor Jonathan Batten, and the discussant and participants of the 2021 Vietnam Symposium of Banking & Finance 28-30 October (VSBF2021) for helpful comments.
1. Introduction

There is a growing consensus in the literature that the existence of shadow banking dilutes the effectiveness of monetary policy. The argument goes that a tightening of monetary policy reduces bank credit creation in the regular banking sector but a leakage into the shadow banking sector means that total credit is less affected. The background to this thinking is the regulatory response to the Global Financial Crisis (GFC) of 2008 which resulted in a raft of macroprudential policies aimed at reducing the vulnerability of the economy to financial shocks. Prominent among these policies is the increase in risk capital that commercial banks must hold and the countercyclical capital requirements of Basel 3. Increasing capital requirements along with other quantitative controls have engendered credit leakage towards unregulated financial institutions. Opinions differ as to whether this credit leakage to the shadow banking sector increases financial fragility (Meeks et al., 2017), or improves the efficiency of financial intermediation (Fève et al., 2019). However, that the existence of unregulated shadow banks impinges on the efficacy of monetary policy is unquestionable.

What is yet unclear is whether the operation of monetary policy is merely dampened by the leakages caused by shadow banking or fully neutralised so as to render monetary policy ineffective. Ultimately, the issue is empirical, but understanding the monetary transmission mechanism in an economy with the existence of shadow banks requires an empirically testable theoretical framework which this paper, among others, address.

The main factor driving the growth of shadow banking in China is the same as in the West, namely regulatory arbitrage (Acharya et al., 2013; Gorton and Metrick, 2010; Pozsa, et al., 2010). However, that is where the similarity ends. Shadow banking in China differs from that of the developed economies in three distinct ways. First, it is much less complex than their western counterparts. Shadow banks in China supply plain vanilla loans that would otherwise be supplied by the regular banks (Elliot et al., 2015). Second, it is much more interconnected with the Chinese commercial banking system (Elliott and Yan, 2013; Chen, et al., 2018). Third, it was implicitly endorsed by the Chinese government (Wang et al., 2019). Together these distinctive features underpinned the rapid growth of shadow bank lending.

In 2000, shadow banks accounted for less than 10 percent of China’s economy. In response to the GFC, China launched an unprecedented multitrillion RMB stimulus package in 2008. It soon switched to a contractionary monetary policy in 2009-2015, which has been argued by Chen et al. (2018) gave further impetus to the expansion of the shadow banking sector. Shadow
banks reached a peak in 2016 at over 80 percent of gross domestic product (GDP). It has fallen back since then in response to a tightening of regulation and clamp down on various shadow banks such as the P2P platforms (Hsu et al., 2020).

The thinking in official circles is that shadow banks have added to the financial fragility of the Chinese economy, and that the effectiveness of monetary policy has been weakened by its existence. This has given rise to a raft of regulations that have stamped down on shadow bank activity. We show in this paper that the contra-cyclical movements of regular bank credit and shadow bank credit is episodic and a special case of regulatory imposition. Using this experience to design regulatory policy can be dangerous and harmful to the Chinese economy.

In this paper we study the Chinese macroeconomy within a DSGE framework that includes a banking sector and recognise the role of shadow banks. In this endeavour, we are not new. There have been several studies that have taken this route and we tread in their footsteps. In Le et al. (2021a) we estimate and test a DSGE model for China. The model accommodates a banking sector that isolates the effects of lending to the private sector including shadow bank lending. This model can explain the dynamic behaviour of Chinese macroeconomic variables for the period 1992-2016, and regular bank credit responds to monetary policy instruments in the expected way with shadow bank credit moving pro-cyclically. Our contribution here is to modify the Le et al. (2021a) model by imposing a loan-to-deposit ratio (LDR) cap on the bank’s balance sheet for the period 2009-2015. This additional constraint results in shadow bank credit moving contra-cyclically to regular bank credit in response to monetary policy shocks. In contrast to the mainstream, and as an additional contribution, the model is estimated with the method of indirect inference making the model data consistent. We show that the model is data consistent, but the imposition of the LDR cap is rejected by the data for the full period. However, the re-estimated model with the cap for the sub-period 2009-2015 is not rejected, adding empirical weight to our argument.

The rest of this paper is organised in the following way. In the next section, we set out the context and literature review. In section 3, we describe in brief the model framework, and the methodology. In section 4, we set out the empirical results for the model and show that regular and shadow bank credit move together in response to tightened monetary policy but the imposition of the LDR cap results in a contra-cyclical response of regular bank and shadow bank credit to monetary tightening. Our final section concludes, with some reflections on the generality of alternative models that have been used to analyse China’s monetary policy.
2. Shadow banking in China

The Financial Stability Board define shadow banking as credit intermediation involving entities and activities outside the regular banking system (FSB, 2017). To this definition the Peoples Bank of China adds ‘that serves to provide liquidity and credit transformation which could potentially be a source of systemic risk or regulatory arbitrage’, (*PBoC 2013 China Financial Stability Report*). The PBoC definition is pragmatic and works backwards from the perceived danger scenario. In doing so, it recognises that unlike in the West where shadow bank funding is from the market, in China it is strongly connected to the regular banks, earning it the sobriquet ‘the bank’s shadow’ (Sun, 2019).

Several good expositions of the Chinese shadow banking system exist (Elliott et al., 2015; Ehlers et al., 2018) and we refer to these studies for a detailed analysis of the structure, development, and associated risks. What we know about shadow banking in general is that the regulatory gap between the conventional commercial banks and shadow banks is the motivation for the growth in the latter activity. Regulators face a trade-off between tighter regulation of the conventional banks and the credit leakage into the shadow banking system. This structure has been used to analyse the effectiveness of reserve requirement regulations in China within a DSGE framework. Such studies highlight the implications of quantitative restrictions and regulations operating in an environment of exiting distortions caused by government support for state-owned enterprises through soft-loan conditions and implicit government guarantees.

What we know about shadow bank activity in China is that it encompasses a variety of bank-based functions including the issuance of wealth management products, asset management products, entrusted loans, trust loans, finance loans, P2P lending, and other informal lending (Allen and Gu, 2020). Most of the shadow bank credit comes in the form of loans from trust companies (trust loans) and company-to-company loans (entrusted loans) intermediated by the banks (Allen et al., 2019). Because of the support provided by the government to the state-owned sector (Cong et al., 2019), the banks prefer to lend to this relatively low-risk sector, starving SMEs that then depend on the shadow bank for financing (Lu et al., 2015; Tsai, 2017).

---

1 Gebauer and Mazelis (2020) apply a DSGE framework to model the credit leakage to the shadow banking system in the eurozone in response to macroprudential policy by the ECB.

2 Chang et al. (2019) show that raising reserve requirements results in banks shifting credit to the shadow banking sector and lending to the more productive private sector. Reserve requirements can act as a second-best policy to offset the distortions created by government guarantees to less productive state-owned enterprises. Wang et al. (2019) show that tightening monetary policy through reserve requirements dampens the effectiveness of monetary policy through the shift to shadow bank credit.
The four-trillion stimulus package of the Chinese government following the GFC was financed through bank loans by the provincial governments. Chen et al. (2020) show that rollover pressure and credit tightening have led to the provincial governments to refinance using the shadow bank channels to sustain long term capital projects.

The growth of the Chinese shadow banking sector following the GFC has prompted several Chinese scholars to examine its role in dampening the effect of monetary policy. Qiu and Zhou (2014) undertake a systematic investigation into the role of shadow banking in the monetary transmission mechanism using a DSGE framework. They argue that Chinese shadow banking is counter cyclical and reduces the effectiveness of monetary policy. Others such as, Liu et al. (2014), Lin et al. (2016) and Wang et al. (2020) also study Chinese shadow banking using the DSGE framework and claim that shadow banking is pro cyclical. Funke et al. (2015) analyse the impacts of interest rate liberalization on monetary policy transmission. Gao et al. (2018) find a counter cyclical pattern of Chinese shadow banking using a structural vector autoregressive (VAR) model and contend that the effects of monetary policy are dampened by borrowers substituting between regular and shadow bank credit.

These model-based results conform with the analysis of Chen et al. (2018) who use a panel VAR methodology to model the distortionary effect of shadow banking on monetary policy. They argue that the quantitative cap on loans in the form of a 75% maximum loan-to-deposit ratio (LDR) incentivised the banks to invest in risky non-loan assets through the shadow banking channel. Gabriel et al. (2018) also conduct a VAR analysis to show that shadow bank credit amplifies increases in the money supply but weakens the effect of tightening monetary policy on bank credit. Cong (2019) uses similar methodology and finds that tightened monetary policy compresses both regular bank and shadow bank lending, but that shadow bank credit is more affected than regular bank credit.

The LDR cap is the key to understanding the episode of the coincidence of money supply tightening and shadow banks growth. The stimulus package saw bank credit soar in 2008-9 and broad money (M2) rise to nearly 30% in Q4 2009 – see Figure 1. The imposition of the LDR\(^3\) in 2009 saw broad money growth fall to 19% in 2010 and shadow bank credit soar from 19% of GDP in 2008 to nearly 38% of GDP in 2010 (Figure 2). Money growth continued to decline and stayed below 15% through to 2015 but shadow bank credit soared to reach 78% of GDP\(^4\).

\(^3\) The maximum LDR existed since 1994 but was not strictly enforced until 2009.
\(^4\) Estimates from JP Morgan-Chase.
In late 2015, the LDR cap was lifted. During this period entrusted lending rose to fill the credit gap created by contractionary monetary policy (Chen et al., 2018). Other restrictive policies through ‘window guidance’ which placed informal quotas on lending to specific sectors saw a widening funding gap emerge in the real estate sector (Allen and Gu, 2020; Allen et al. 2018). Since 2016, the size of the shadow bank sector has shrunk relative to GDP following the active policy of deleveraging and additional regulations on shadow banks such as the P2P platforms (Hsu et al., 2020).

A recent paper by Yang et al. (2019) sets out a New-Keynesian DSGE model of the Chinese economy that includes, a regular bank sector and a shadow bank sector. In addition, it includes the LDR constraint and loan quota constraints. In keeping with the mainstream literature, the model is estimated by the Bayesian method for the period 2006Q1-2017Q4.

**Figure 1: M2 growth 2000q1-2016q4 (year-on-year) %**

![M2 growth chart](image)

Source: FRED database

---

5 This is not dissimilar to the practice of ‘moral suasion’ practiced by the Bank of England during the 1950s and 1960s as a means of directing or stifling commercial bank credit to specific sectors.
3. Model Framework and Methodology

3.1 The model in brief

The DSGE framework is commonplace as a modelling tool for the investigation of macroeconomic shocks. A strong stream of research has followed the lead of Smets and Wouters (2007) analysis of US business cycles and most recently applied to a variety of monetary and fiscal shocks. Le et al. (2021a) estimate and test a DSGE model for China. The model accommodates a banking sector that isolates the effects of lending to the private sector including shadow bank lending. This model can explain the dynamic behaviour of Chinese macroeconomic variables for the period 1992-2016. It finds that while financial shocks were significant, it is real shocks, specifically productivity shocks, that dominate the variance of output. Looking through the impulse responses analysis, one finding in their paper is that shadow bank lending is procyclical. That is, if the economy is hit with a negative monetary policy shock, the aggregate lending, and the loans to both SOE and SME sectors fall. However, the SMEs investment and output would decline by more than the SOEs because of the financial accelerator mechanism in the SME sector. This result however contrasts with the consensus view that the effect of monetary policy on output is dampened as firms switch from conventional bank credit to shadow bank credit (Chen et al., 2018). This consensus finding is related to the period of credit repression that the PBoC imposed in 2009-2015 in trying to curb

---

6 See for example, Bratsiotis and Theodoridis (2022); Polyzos et al. (2021); and Le et al. (2016b, 2020).
the growth of shadow banking credit, in fear of its possible adverse effects on the real estate sector and the real economy.

In Le et al. (2021a) it is assumed that SOEs can borrow from financial intermediaries at the risk-free rate because of the implicit state guarantee, but SMEs are risky, i.e., they can go bankrupt and there is a costly verification process in recovering their worthiness. Therefore, SMEs must borrow at a premium over the risk-free rate. In this set up, a contractionary monetary policy raises the cost of borrowing for both sectors leading to a decrease in demand for credit and investments and lowering the price of capital. In the SME sector, this results in a lower net worth of the firms and a higher credit premium, further lowering the demand for credit and investment in this sector (Bernanke, Gertler, and Gilchrist, 1999 – BGG). The model did not include any micro-prudential type of quantitative restriction on the lending process, which, however, is one of the instruments used by the Chinese government to reduce the credit growth in the economy. In this paper we impose a loan-deposit ratio cap as a micro-prudential measure.

The Le et al. (2021a) model is modified in the following way. The representative household lend their savings to two types of financial intermediaries: a conventional bank who in turn would lend to SOEs; and a financial intermediary (shadow bank) that lends to a riskier SME\(^7\). When the conventional bank lends to SOEs, it is faced with a loan-to-deposit constraint \(\tau\), the shadow bank’s lending to SMEs, however, are not constrained by this condition. Effectively, the regulation would make SOEs lending relatively more expensive than that to SMEs. In the log-linearised form, the model is expressed as follows.

3.1.1 Consumers:

There is an intertemporal Euler equation that determines optimal consumption. The equation represents a weighted average of current, past, and expected future consumption and labour as a function of the real interest \(r_t - E_t\pi_{t+1}\) and the intertemporal shock to preferences \(\varepsilon_t\).\(^8\)

\(^7\) In reality the deposit taking banks uses some of the funds to lend to Shadow banks through off-balance sheet vehicles (Allen and Gu, 2020). For modelling purposes, we have households lending directly to conventional banks and shadow banks.

\(^8\) Note that consumption and leisure are assumed non-separable in the utility function as in Smets and Wouters (2007). This formulation was also used in Le et al. (2011, 2016a), where the whole model was estimated and tested against the unfiltered US data. To minimise the changes to the theoretical model, we keep this utility function specification.
\[ c_t = \frac{\lambda}{1+\gamma} c_{t-1} + \frac{1}{1+\gamma} E_t c_{t+1} + \left(1 - \frac{1}{(1+\gamma)}\right) \left(l_t - E_t l_{t+1}\right) - \left(\frac{1-\lambda}{(1+\gamma)}\right) \left(r_t - E_t \pi_{t+1}\right) + \varepsilon_t \]  

where \( \sigma_c \) denotes the intertemporal elasticity of substitution, \( \lambda \) the degree of habit formation, \( \beta \) the household's discount factor, \( \gamma \) the trend growth rate of technology, \( \phi \) the cost of adjusting the rate of investment, and \( \frac{W^*L^*}{C^*} \) the steady state ratio of labour income to consumption.

### 3.1.2 Firms

There are two intermediate goods sectors, SMEs and SOEs. In each sector, intermediate-goods firms produce using labour which moves freely across the two sectors so that the labour composite is \( l_t = \frac{N^{SOE}}{N} n_{t}^{SOE} + \frac{N^{SME}}{N} n_{t}^{SME} \), where in each sector labour is a combination of labour hired from imperfect and perfect labour markets, and capital is bought from the sector specific capital producer. These firms produce intermediate goods under perfect competition assumptions, according to the following Cobb-Douglas production functions:

\[
y_t^{SOE} = \phi [\alpha^{SOE} k_{t-1}^{SOE} + (1 - \alpha^{SOE}) n_{t}^{SOE} + \varepsilon_t^{SOE}] \]  
\[
y_t^{SME} = \phi [\alpha^{SME} k_{t-1}^{SME} + (1 - \alpha^{SME}) n_{t}^{SME} + \varepsilon_t^{SME}] \]

where \( \alpha^i \) with \( i = SOE, SME \) denotes the share of capital in production for sector \( i \) and \( \phi \) equals one plus the share of fixed costs in production. The demand for capital and labour in the two sectors are:

\[
mpk_t^{SOE} = y_t^{SOE} - k_{t-1}^{SOE} + pW_t^{SOE} \]  
\[
mpk_t^{SME} = y_t^{SME} - k_{t-1}^{SME} + pW_t^{SME} \]

and

\[
n_t^{SOE} = y_t^{SOE} - w_t + pW_t^{SOE} \]  
\[
n_t^{SME} = y_t^{SME} - w_t + pW_t^{SME} \]

where \( pW_t^i \) is the relative price of wholesale output.

In each sector the price of new capital depends positively on the expected future marginal product of capital and the expected future value of capital, and negatively on the real cost of borrowing as shown in (8) and (9) below:

\[
q_k^{SOE} = \frac{1-\delta}{1-\delta+R^{K,SOE}} E_t q_k^{SOE} + \frac{R^{K,SOE}}{1-\delta+R^{K,SOE}} (y_t^{SOE} - k_{t-1}^{SOE} + pW_t^{SOE}) - (r_t^{SOE} - E_t \pi_{t+1}) \]  
\[ (8) \]
\[ q_t^{SME} = \frac{1-\delta}{1-\delta+R_t^{KSME}} E_t q_t^{SME} + \frac{R_t^{KSME}}{1-\delta+R_t^{KSME}} (y_t^{SME} - k_t^{SME} + pW_t^{SME}) - (r_t^{FI} - E_t \pi_{t+1} + s_t) \] (9)

where \( R_t^{Ki} \) are the steady-state values of the return on capital, and \( \delta \) is the rate of capital depreciation, \( r_t^{SOE} \) is the nominal lending rate to SOEs and \( r_t^{FI} \) is the nominal lending rate to financial intermediaries who bridge between lenders and SMEs.

### 3.1.3 Capital producers

The capital producers are sector specific. The capital accumulation functions are as in (10) and (11).

\[
\begin{align*}
    k_t^{SOE} &= \left(\frac{1-\delta}{y}\right) k_{t-1}^{SOE} + \left(1 - \frac{1-\delta}{y}\right) inv_t^{SOE} + \left(1 - \frac{1-\delta}{y}\right) (1 + \beta y^{1-\sigma_c}) y^2 \varphi \varepsilon_t^{SOE inv} \quad (10) \\
    k_t^{SME} &= \left(\frac{1-\delta}{y}\right) k_{t-1}^{SME} + \left(1 - \frac{1-\delta}{y}\right) inv_t^{SME} + \left(1 - \frac{1-\delta}{y}\right) (1 + \beta y^{1-\sigma_c}) y^2 \varphi \varepsilon_t^{SME inv} \quad (11)
\end{align*}
\]

where \( \delta \) is the depreciation rate and assumed to be the same between the two sectors.

The Euler equations for investment in SOE and SME sectors specify the optimal investment plan. The weighted average of past, current, and future investment in each sector depends on the price of new capital, \( q_t \).

\[
\begin{align*}
    inv_t^{SOE} &= \frac{1}{1+\beta y^{(1-\sigma_c)}} inv_{t-1}^{SOE} + \frac{\beta y^{(1-\sigma_c)}}{1+\beta y^{(1-\sigma_c)}} E_t inv_{t+1}^{SOE} + \frac{1}{(1+\beta y^{(1-\sigma_c)}) y^2 \varphi} q_t^{SOE} + \varepsilon_t^{SOE inv} \quad (12) \\
    inv_t^{SME} &= \frac{1}{1+\beta y^{(1-\sigma_c)}} inv_{t-1}^{SME} + \frac{\beta y^{(1-\sigma_c)}}{1+\beta y^{(1-\sigma_c)}} E_t inv_{t+1}^{SME} + \frac{1}{(1+\beta y^{(1-\sigma_c)}) y^2 \varphi} q_t^{SME} + \varepsilon_t^{SME inv} \quad (13)
\end{align*}
\]

Aggregate investment is a combination of two sectors investment

\[
inv_t = \frac{INV^{SOE}}{INV} inv_t^{SOE} + \frac{INV^{SME}}{INV} inv_t^{SME} \quad (14)
\]

### 3.1.4 Banks

The costs of borrowing for SOEs and SMEs are determined as follows. Households make deposits at perfectly competitive conventional banks who then lend to SOEs directly and also make deposits at shadow banks who then lend to SMEs so that the aggregate bank sector balance sheet is \( D_t = L_t^c + L_t^S \). Sector profit \( R_t D_t = R_t^c L_t^c + R_t^S L_t^S \) is maximised subject to the
balance sheet implies that households make deposits to conventional bank and financial intermediaries at the deposit rate, \( R_t^d = R_t^c = R_t^{S} \). The former would pass on the risk-free rate to SOEs, while the latter would charge SMEs a premium over the risk-free rate.

However, the loan deposit cap assumption implies that loans to SOEs is less than in the previous unconstrained case, so that now \( L_t^{SOE} = \omega L_t^c \). The conventional bank obtains funds at the rate \( R_t \) and charges SOEs at the rate \( R_t^{SOE} \), where \( R_t^{SOE} = \frac{R_t}{\omega} > R_t \) to satisfy the zero profit condition \( R_t L_t^c = R_t^{SOE} \omega L_t^c \). The remainder \( (1 - \omega) L_t^c \) is available funding for the shadow banks, so that the zero profit problem becomes \( R_t D_t = R_t^{SOE} \omega L_t^c + R_t^S (1 - \omega) L_t^c + L_t^S \). This equation can be rewritten as \( R_t = \tau R_t + R_t^{SOE} \omega L_t^c + R_t^S \left( (1 - \omega) L_t^c + L_t^S \right) \). The remainder \( (1 - \omega) L_t^c \) is available funding for the shadow banks, and with more funding available for the shadow banks, the cost of funds to the shadow banks is \( R_t^S = \frac{(1 - \tau)}{(1 - \omega) \tau + (1 - \tau)} R_t < R_t \). Therefore, the loan to deposit ratio makes the SOEs loans relatively more expensive than loans to SMEs.

The interest rate that the SOE must pay for the loans is:

\[
\tau_t^{SOE} = \frac{1}{\tau} \tau_t, \quad (15)
\]

where \( \tau \) equals to 1 if there is no loan-to-deposit regulation, otherwise it equals to 0.75. The SME would pay:

\[
\tau_t^S = \frac{(1 - \tau)}{(1 - \omega) \tau + (1 - \tau)} R_t \quad (15')
\]

plus an external credit premium of

\[
s_t = E_t c y_{t+1} - (r_t - E_t \pi_{t+1}) = \chi (q q_t^{SME} + k_t^{SME} - n w_t^{SME}) - \psi_1 m_t + \varepsilon_t^S \quad (16)
\]

The risk premium is reduced with a higher cash collateral \((m_t)\) and a higher net worth \((n w_t)\) relative to the gross value of capital \(-(q q_t^{SME} + k_t^{SME} - n w_t^{SME})\), and rises with more exogenous shocks \((\varepsilon_t^{PR})\). We assume that in every period a fixed death rate \((1 - \theta)\) happens so that the stock of SME firms is kept constant by an equal birth rate of new firms, and their net worth remains below the demand for capital. This means that the SMEs net worth is the past

\[\text{The balance sheet of the firm is fixed capital and cash as assets and loans and equity (net worth) as liabilities. The loans and equity are used to purchase capital goods. The higher the equity the lower the dependence on bank debt and the lower the riskiness of the firm and consequently the lower the risk premium. See Le et al. (2016a) for a theoretical underpinning.}\]
net worth of surviving firms plus their total return on capital minus the expected return (which is paid out in borrowing costs) on the externally financed part of their capital stock

\[ nw_t^{SME} = \theta nw_{t-1}^{SME} + \frac{K^{SME}}{NW^{SME}}(cy_t - E_{t-1}cy_t) + E_{t-1}cy_t + \zeta_{2t} + \epsilon_t^{NW} \]  

(17)

where \( \frac{K^{SME}}{NW^{SME}} \) is the steady state ratio of SMEs’ capital expenditures to SMEs net worth, \( cy_t \) is the real external financing rate, \( \zeta_{2t} \) is a regulatory specific shock to net worth, and \( \epsilon_t^{NW} \) represents all other net worth shocks. The regulations are the instruments that the central bank can also use to influence the credit premium and indirectly the net worth of the firms. However, in the absence of observable data on micro-prudential measures, for simplicity we include these in the errors \( \epsilon_t^x \) and \( \epsilon_t^{NW} \).

While the BGG net worth channel has been used extensively in the literature and is well known, the cash collateral variant warrants further elaboration. The idea of costly state verification is that net worth is all invested in plant, machinery and other capital and thus cannot be recovered at original value and has less value when the firm goes bankrupt because it has become specialised to the firm’s activities. It is normal for banks also to request an amount of collateral (Kiyotaki and Moore, 1997). If this collateral was in terms of cash, i.e., a firm holds some cash on its balance sheet, this can be recovered directly without loss of value and no verification cost. The elimination of the collateral cost helps to lower the credit premium for given net worth and it allows firms to increase leverage and so raise their expected returns. It assumes that banks and SME firms have a mutual interest in firms holding as much cash as can be acquired for collateral. The process of cash being used as collateral is as follows. The central bank issues cash through open market operations to households in exchange for government bonds; households deposit cash with banks; firms want to acquire as much of this cash as possible for their collateral needs (they invest their net worth in cash to the maximum available with the rest going into other collateral and capital). In practice the firms’ profits are continuously paid out as dividends to the banks which lend to them, so they have nothing with which to acquire these assets if they do not collaborate with banks. They achieve this balance sheet outcome by agreeing with the banks that, as a minimum counterpart to the credit advanced, they will hold the maximum cash collateral available, which is M0. Thus, all M0 at once finds its way to firm’s balance sheet, where it is securely pledged to the banks in the event
of bankruptcy.\textsuperscript{10}

3.1.5 Final Goods Producer and Labour Supplier

The final goods producer would gather these intermediate goods with a CES production function into final goods and pay intermediate firms $pW_t^i$ with $i = SOE, SME$. The final output together is

$$y_t = \eta y_t^{SOE} + (1 - \eta)y_t^{SME}$$

(18)

where $\eta$ is the share of SOE output in total output.

Cost minimisation implies the following demand for intermediate goods:

$$y_t^{SOE} = -\varepsilon pW_t^{SOE} + y_t$$

(19)

$$y_t^{SME} = -\varepsilon pW_t^{SME} + y_t$$

(20)

where $\varepsilon$ is the elasticity of substitution between the two intermediate goods.

Following Le et al. (2011), we assume that the final goods producers then sell a part of final goods in the competitive market at marginal cost, and it differentiates the rest and then marks up for sale in the market characterised by the nominal rigidities. Therefore, the model introduces monopolistic power and nominal price rigidities at the retail level. For simplicity, we solve for prices under the competitive market assumption and then under the imperfect competition assumption take the weighted average between the two as the solution for the price. Labour supply also works in the same way.\textsuperscript{11} Households supply labour to a regulated labour market and to a competitive labour market, so that the aggregate wage index is a weighted average of the perfectly competitive and imperfectly competitive wage levels. The imperfectly competitive market set-up gives the New Keynesian (NK) Phillip curve where a weighted average of current, past, and expected future inflation depends on the price mark-up and an exogenous cost-push shock to prices, $\varepsilon_P^C$:

$$\pi_t = \frac{\beta(1 - \sigma)}{1 + \beta(1 - \sigma)\lambda_p}E_t \pi_{t+1} + \frac{\lambda_p}{1 + \beta(1 - \sigma)\lambda_p} \pi_{t-1} +$$

\textsuperscript{10} There are many ways that money can be brought into a model like this. The way we have done it is in the spirit of the credit channel where cash is pledged as collateral and serves to reduce the risk premium. A real-world feature is the availability of liquidity to the financial system which reduces interest rates and spreads.

\textsuperscript{11} As we are unaware of a long time series of wage for SOEs and SMEs, this modelling convenience enables us to work with a single wage series.
\[
\frac{1}{1+\beta (1-\sigma_c)} \left( \frac{1-\beta (1-\sigma_c) c_p}{1-\beta (1-\sigma_c) c_{p+1}} \right) (\eta p W_t^{SOE} + (1-\eta) p W_t^{SME}) + \epsilon_t^p
\]  
(21)

and the weighted average of current, past, and expected future wages depends on the wage mark up, inflation and a cost push shock to wages, \( \epsilon_t^W \):

\[
w_t = \frac{\beta}{1+\beta (1-\sigma_c)} (E_t w_{t+1} + E_t \pi_{t+1}) + \frac{1}{1+\beta (1-\sigma_c)} w_{t-1} - \frac{1+\beta (1-\sigma_c) \tau}{1+\beta (1-\sigma_c)} \pi_t
\]

\[
+ \frac{\tau}{1+\beta (1-\sigma_c)} \pi_{t-1}
\]

\[
- \left( \frac{1}{1+\beta (1-\sigma_c)} \right) \left( \frac{1-\beta (1-\sigma_c) c_w}{1-\beta (1-\sigma_c) c_{w+1}} \right) \left( w_t - \sigma l_t - \left( \frac{1}{1-\gamma} \right) \left( c_t - \frac{\lambda}{\gamma} c_{t-1} \right) \right) + \epsilon_t^W
\]  
(22)

The perfectly competitive market set-up also produces the labour supply that reacts to expected inflation

\[
w_t = \sigma l_t + \left( \frac{1}{1-\gamma} \right) \left( c_t - \frac{\lambda}{\gamma} c_{t-1} \right) - (\pi_t - E_{t-1} \pi_t)
\]  
(23)

and the natural log of real marginal costs for the final goods producer must be equal to zero

\[
\epsilon p W_t^{SOE} + (1-\epsilon) p W_t^{SME} = 0
\]  
(24)

3.1.6 Resource constraint

The resource constraint states that aggregate output \( y_t \) depends on consumption, investment and an exogenous (government spending and net trade) shock, \( \epsilon_t^G \) and it ignores the contribution from the entrepreneurs’ consumption as it is negligible:

\[
y_t = \frac{c}{\gamma} c_t + \frac{l}{\gamma} inv_t + \epsilon_t^G
\]  
(25)

3.1.7 Monetary authority

To close the model, we allow the short-term rate on official lending to the banks to be set by the PBoC in accordance with a Taylor Rule. We assume that the PBoC enforces this rule via open market operations. That is, households hold part of their savings in government bonds and the rest in bank sector deposits, which pay the short-term interest rate also obtainable on short term government bonds. To control the short-term rate, the PBoC would sell/buy long-term government bonds to buy/sell short-term government bonds to influence the prices of these assets and thus their rates. Cash is issued in this model, but is only held by firms, as
households have no use for it and deposit it in the banking sector where it is lent to firms to hold as collateral and affect the credit premium. Along with the micro-prudential regulations, the monetary authorities have two further instruments, $M_0$ and $r$:

$$r_t = \rho r_{t-1} + (1 - \rho) \left( \rho_\pi \pi_t + \rho_y y_t + \rho_{\Delta y} (y_t - y_{t-1}) \right) + \varepsilon_t^m$$

$$\Delta m_t = \psi_2 \Delta M_t + \varepsilon_t^{m2}$$ (26) (27)

Where $\psi_2$ is positive and $M_t$ is the supply of money, which is defined as the sum of deposits and base money. Using the firms’ balance sheet, the money supply is expressed as a function of base money, capital, and net worth, and $v$, $\mu$, $c$ are respectively the ratios of net worth, $M_0$ and collateral to money. So:

$$M_t = (1 + v - \mu) k_t + \mu m_t - v n_t$$ (27)

This now gives our monetary authorities three instruments: base money, the interest rate, and micro-prudential policy.

3.2 The method of indirect inference

Using the method of Indirect Inference to evaluate structural macroeconomic models was first proposed by Minford et al. (2009) with further refinements in Le et al. (2011) who evaluated the power of the test using Monte Carlo experiments. Indirect Inference provides a classical statistical inferential framework for judging whether a model could have generated the behaviour found in the data. It does this by comparing an auxiliary model based on the data with the same auxiliary model based on the model simulations. The auxiliary model is independent of the theoretical model and is employed to form a criterion function in the indirect inference test. As with the auxiliary model we choose the parameters of a VAR with the main macroeconomic variables (output, inflation, and interest rate) as the endogenous variables. The VAR parameters are used as they are a good descriptor of the data.

After simulating the model, we calculate the following Wald statistic

$$W = (\beta^a - \bar{\beta}^z)' \Omega^{-1} (\beta^a - \bar{\beta}^z)$$ (28)

where $\beta^a$ is a vector of the auxiliary model parameter estimates from the actual data, $\bar{\beta}^z$ is the mean of the auxiliary model parameter estimates from the simulated data, and $\Omega$ is the

---

12 A complete elaboration of the methodology can be found in Le et al. (2016a)
The variance-covariance matrix of the distribution of the simulated estimate $\beta^s$. This Wald statistic measures the distance between the VAR coefficients from simulated and actual data. If the structural model is correct, then the simulated data (and the VAR estimates based on these) will not be significantly different from the actual data. For the model not to be rejected this Wald statistic should lie within the 95% bounds from the simulated Wald statistics. Model estimation involves the search for the set of parameters of the theoretical model that lie within the 95% bounds and minimise the Wald statistic.

4. Empirical results
Using the Indirect Inference method to estimate and test a macroeconomic model for China, Le et al. (2021a) show that this estimated model without the LDR condition, is data consistent with a p-value of 0.060, for the full sample period, which means it cannot be rejected as the true model that can generate the observed data. In this paper we conduct an experiment to see whether it is possible to find a model that includes the LDR constraint that is consistent with the full sample period using the same methodology. Our attempt was unsuccessful, the search did not find any set of parameters to accept the model including the LDR constraint that is consistent with the data. The reason could have been that LDR was not a strict constraint for most of the full period.$^{13}$

To evaluate the effect of the LDR imposition, we proceed in two stages. First, we impose the LDR constraint on the model estimated for the full period and examine the response of regular bank credit and shadow bank credit to monetary tightening. This is compared with the same response in the unconstrained case. Next, we re-estimate the model for the period 2009-2015 as this was the period when the LDR constrained operated. The indirect inference test gives a p-value of 0.098. Using the standard significance level of 5% the model with the LDR is therefore not rejected and is consistent with data for this subsample. That means that the estimated model can generate the data for this subperiod 2009-2015 and is reliable for policy analysis. We then examine the response of regular bank credit and shadow bank credit to monetary tightening in the re-estimated model. Table 1 presents the results and the p-values of Wald statistics for the full sample and sub-sample period.

---

$^{13}$ There were other restrictions such as controls on the bank deposit rate and loan rate ceilings and others as in Funke et al. (2015)
The parameters of the model for the two periods differ. Most of the big changes relate to policy derived parameters and parameters reflecting changes to the stochastic environment. We find that for the LDR period the new-Keynesian (NK) weight on wages has decreased from 0.998 to 0.508, so the wage determination environment has gone from being nearly completely NK to half NK and half neo-classical\textsuperscript{14}. Wage indexation has also decreased, and the probability of

\textsuperscript{14} The NK weight on inflation has increased from .09 to .664. This could be due to the decrease in the mean of inflation and decrease in the standard deviation of inflation during the LDR period, so more firms will be NK.
not changing wages decreased significantly from 0.812 to 0.065. The main effect of these last two changes is to make NK wages unresponsive to the marginal product of labour.

However, the parameters of the elasticity of consumption and the elasticity of labour supply differ strongly in the sub-period from the full period. There is no suggestion that the behavioural parameters obtained from this method are the deep structural ‘Lucas’ parameters and are therefore invariant to the sample size. As in Le et al. (2017) and, also Iskrev (2010), we follow the normal practice of defining the estimated ‘parameters’ as operative parameters, which themselves combine deeper parameters that cannot be retrieved. This issue lies at the heart of the problem of weak identification in DSGE models which is addressed in Le et al. (2017) using Monte-Carlo methods\textsuperscript{15}. In addition, indirect inference produces parameter estimates with smaller bias than maximum likelihood (Meenagh et al., 2019). This property comes from the high power of the test in rejecting false parameters. However, the power of the test declines with a smaller sample size, but clearly II does not eliminate small sample bias which may be an extra issue for the estimates from the sub-sample period. Hence a change in the data sample could change the estimated parameters because the new parameters provide a better fit.

In the following exercises we examine the behaviour of the model with and without the LDR constraint, according to the estimated parameters in Table 1. Figure 3 shows that the response of the model to a tightening of monetary policy in the unconstrained version of the model estimated for the full period\textsuperscript{16}. The frequency of the data is quarterly, so the x-axis of the IRFs is in quarters and the y-axis is in percentage terms. The results are conventional. Investment by both SOEs and SMEs respond negatively, the latter more strongly because monetary tightening raises the premium. Output in both sectors decline as does borrowing. Regular bank lending to the SOEs fall initially but shadow bank lending to the SMEs falls sharply as the rise in the premium hits home.

Figure 4 compares the effect of the same shock with the LDR constraint imposed alongside the unconstrained version of Figure 3. The responses are very similar, but the degree varies by sector. The negative response to SOE investment is greater under the LDR regulation than the

\textsuperscript{15} For a full discussion of the identification problem and the indirect inference solution see also Meenagh et al. (2019).

\textsuperscript{16} We do not report standard error bounds as the method of indirect inference does not produce parameter standard errors.
negative response of investment to SMEs. The LDR cap is binding on lending by the regular banks to the SOEs but not on the unregulated shadow banks’ lending to the SMEs. A tightening of monetary policy under the LDR cap incentivises the regular banks to leak funding to the shadow banks; effectively creating a substitution in lending away from the SOEs towards the SMEs. This is seen more clearly in the contracyclical response of borrowing as shadow bank credit to the SMEs rise and regular bank credit to the SOEs decline. Looking at the response of total output and that of the SOEs, and SMEs, we see that the initial effect of a contraction in response to tighter monetary policy is lower under the LDR cap than in the unregulated case. While this constraint is not consistent with the data for the full sample period, it provides support for the finding by others that shadow bank activity weakens the response of the economy to monetary policy. However, this result is a special case to a specific policy in a specific period.

We now turn to the examination of the model properties estimated for the sub-period 2009-2015. In Figure 5 we use the model estimated for this sub-period which imposes the LDR constraint to evaluate the effect of monetary tightening. In the presence of the LDR regulation the estimated model predicts a substitution effect between loans to the SME and the SOE sectors. That is, as the interest rate increases, the cost of borrowing rises and given the extra LDR constraint levied on lending to SOEs, the funds get channelled towards SMEs. Therefore, our analysis shows that conventional bank credit and shadow bank credit vary pro-cyclically with monetary policy when there is no LDR condition but varies countercyclically when the LDR constraint is imposed.

---

17 High LDR banks have an incentive to issue principal-floating wealth management products which is a form of shadow banking through their off-balance sheet activity. See Li et al. (2022).
Figure 3: Contractionary monetary shock of 1% under unconstrained model whole period data (1992-2016).

Figure 4: Contractionary monetary shock of 1% under the model without the LDR for the whole period (1992-2016) and with the LDR imposed.
This exercise shows that micro-prudential financial regulations imposed during the period 2009-2015 explain the accepted finding that conventional and shadow bank credit responded in opposite directions following a monetary tightening. Thus, contractionary monetary policy may have had muted effects on the real sector although the IRFs from a 1% monetary shock suggests otherwise. The strict LDR constraint was lifted in 2016 but monetary tightening continues to be used in support of the Chinese government’s policy of deleveraging. The arguments made in Le et al. (2021a, 2021b) that deleveraging through monetary tightening creates more damage to the SME sector threatening the long-term growth prospects for the Chinese economy.

5. Conclusion and discussion
This paper has explored the proposition that the existence of shadow banking in China has dampened the effectiveness of monetary policy by providing a channel for credit leakage at times of monetary tightening. The accepted view expounded by Chen et al. (2018) is that contractionary monetary policy in the period 2009-2015 caused the rise in shadow bank lending. While contractionary monetary policy reduces bank loans in the conventional way, the existence of the LDR encourages banks to undertake risky lending via off-balance sheet vehicles that channel funds to the shadow banking system. We have no quarrel with this interpretation, but we argue in this paper that this channel is a special case that arises...
specifically out of the LDR constraint on banks’ balance sheets. The imposition of a loan-to-deposit cap raises the cost of credit to SOEs and provides an incentive for substitution of credit to the SMEs via the shadow banks. The policy implication of Chen et al. (2018) that loosening the constraints on China’s financial system weakens the authority’s ability to undertake effective monetary policy, implies that a gradual approach to financial reform be adopted. While this may be a valid policy prescription, it is not one that is founded on the role of shadow banks and the efficacy of monetary policy.

Micro-prudential regulations of the LDR type yield unnecessary distortions to the financial system and, as has correctly been argued, weaken the effectiveness of monetary policy. Shadow banks have a valid role to play in the process of financial intermediation and contribute only marginally to the financial fragility of the economy. We have argued elsewhere (Le et al., 2021b), that financial fragility is minimised through the application of sound monetary policy and not through the imposition of distortionary micro-prudential policies. The conclusion that monetary policy is weakened through the existence of shadow banks is not general. In the absence of LDR, a tightening of monetary policy reduces both conventional and shadow bank credit alike, affecting the latter more strongly. The policy of deleveraging, which had been carried out until recently, may generate an unfriendly financial environment for the growth-creating SME sector, endangering the long-term growth potential of the Chinese economy.
References


Financial Stability Board, (2017), *Assessment of shadow banking activities, risks and the adequacy of post-crisis policy tools to address financial stability concerns*, 3 July 2017,


