Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

____________________ ____________________

Aliaksandr Zaretski Date

Essays on banking crises, sovereign default, and macroprudential policy

By

Aliaksandr Zaretski
Doctor of Philosophy

Economics

Juan Rubio-Ramírez, Ph.D.
Advisor

Zhanwei (Vivian) Yue, Ph.D.
Advisor

Kaiji Chen, Ph.D.
Committee Member

Daniel Waggoner, Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D., MPH
Dean of the James T. Laney School of Graduate Studies

Date

Essays on banking crises, sovereign default, and macroprudential policy

By

Aliaksandr Zaretski
M.Sc., Center for Monetary and Financial Studies, 2016
B.Sc., Belarusian State University, 2012

Advisor: Juan Rubio-Ramírez, Ph.D.
Advisor: Zhanwei (Vivian) Yue, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2022

Abstract

Essays on banking crises, sovereign default, and macroprudential policy
By Aliaksandr Zaretski

Chapter 1 studies the optimal regulation of a banking system in a quantitative general equilibrium environment with endogenously incomplete markets. Financial intermediaries issue deposits to households under limited liability and with limited enforcement, invest in the real economy with state-contingent returns, and face survival risk. Pecuniary externalities affect the forward-looking bank value and the value of default through asset prices and asset returns, justifying system-wide regulation, e.g., balance sheet taxation or minimum capital requirements. An alternative way to improve welfare is through "preemptive bailouts": expected future transfers relax the current financial constraints, mitigating the enforcement friction and decreasing the probability of financial crises.

Chapter 2 explores the normative implications of the transmission of sovereign risk to the banking sector. Both banks and foreign lenders invest in risky sovereign debt. The sovereign's financial standing is a two-state Markov chain calibrated to match the observed sovereign default and exclusion events. The resulting Markov-switching DSGE model is solved using global methods. Subsidizing lending to the real sector in both good and bad times improves welfare and reduces the probability of banking crises. A bank net worth subsidy in good times combined with a tax in bad times achieves qualitatively similar but quantitatively smaller effects.

Chapter 3 characterizes optimal government policy in a sticky-price economy with different types of consumers and endogenous financial constraints in the banking and entrepreneurial sectors. The competitive equilibrium allocation is constrained inefficient due to pecuniary externalities and other externalities arising from consumer type heterogeneity. These externalities can be corrected with appropriate fiscal instruments. Independently of the availability of such instruments, optimal monetary policy aims to achieve price stability in the long run, as in the conventional New Keynesian environment. Compared to the competitive equilibrium, the constrained efficient allocation significantly improves between-agent risk sharing, approaching the unconstrained Pareto optimum and leading to sizable welfare gains. Such an allocation has lower leverage in the banking and entrepreneurial sectors and is less prone to the boom-bust financial crises and zero-lower-bound episodes observed occasionally in the decentralized economy.

Essays on banking crises, sovereign default, and macroprudential policy

By

Aliaksandr Zaretski
M.Sc., Center for Monetary and Financial Studies, 2016
B.Sc., Belarusian State University, 2012

Advisor: Juan Rubio-Ramírez, Ph.D.
Advisor: Zhanwei (Vivian) Yue, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2022

Acknowledgments

I believe that this dissertation is an outcome of a journey that started decades ago. Some comments seem appropriate here.

First, I would like to thank my parents for their endless love and patience and my sister for multiple letters and postcards that have reached me worldwide.

Second, I would like to thank my math teacher Aksana Leanidaŭna for her early mentorship, introducing me into the world of research, believing in me, and encouraging to pursue a career in science. We had some successes, and I wish we had achieved more during those years.

Third, I would like to thank my undergraduate mentor Dzima—thanks to whom I had not lost interest in economics when I was going to—and an informal mentor Katya, who convinced me of the value of getting a PhD abroad and who supported me despite initial failures. I am grateful to my former colleagues at the IPM Research Center, who helped me mature as an economist working on real-world policy problems. Sadly, my home country has become a place where such expertise is no longer sought.

Fourth, I would like to thank my master's thesis advisor Carlos, who provided many important early suggestions on doing rigorous academic research. I am equally indebted to an informal mentor Javier, who continued giving me valuable feedback after graduation.

Fifth, I would like to thank my current mentors Juan, Vivian, Federico, Kaiji, Dan, and Toni, who all spent many hours attempting to make me a good economist. I wish it had been more successful, but I hope I have not failed you. I am grateful to Daniel for the math camp and informal mentoring early in the program and to Alessandro for countless useful suggestions during the later stages of my dissertation work.

I have met many good people during this journey. I am especially grateful to my classmate and friend Carla for her tenacious kindness during the PhD years.

Contents

Figures

Tables

Chapter 1

Optimal macroprudential policy with preemptive bailouts

1.1 Introduction

The global financial crisis and the Great Recession of the late 2000s raised several challenging normative questions. Is there a need for macroprudential regulation, and if so, which policy instruments are effective? Should regulatory requirements be conditioned on individual-specific characteristics? Is "too big to fail" a problem? Are bailouts justified? In the recent decade, considerable progress has been made in understanding the rationales for macroprudential regulation in small open economies that borrow in the international financial market at exogenously determined interest rates. At the same time, our knowledge of the optimal regulation of banks in quantitative general equilibrium remains limited. Banking crises were at the heart of the global financial crisis of 2007–2009, including in the US and the UK, and many developed and developing economies have a bank-based financial system. Therefore, it is crucial to understand how to regulate banks optimally over the business cycle.

This paper considers a quantitative general equilibrium environment with endogenous financial frictions in the banking system. In this environment, multiple externalities arise, justifying system-wide regulation, e.g., bank balance sheet taxation or minimum capital requirements. Without regulation, occasional large drops in net worth lead to financial crises when endogenous financial constraints switch from being slack to binding. An alternative way to decrease the probability of financial crises and improve welfare is through "preemptive bailouts." Expected future transfers relax the current financial constraints guaranteeing bank solvency and alleviating the limited enforcement friction in the first place. Such transfers are systemic, not being a source of moral hazard. Addressing pecuniary externalities and mitigating the enforcement friction generally constitutes a trade-off between limiting excessive borrowing and lending by banks ex ante and relaxing their financial constraints ex post. Quantitatively, unregulated banks overborrow and overlend compared to the Markov perfect equilibrium outcome but underborrow and underlend compared to the Ramsey outcome.

The economy I consider is a real business cycle model with a banking sector (Gertler and Kiyotaki, 2010; Gertler and Karadi, 2011) and a nonlinear investment technology (Lucas and Prescott, 1971). The banking system consists of heterogeneous banks that issue deposits to households, invest in the real economy with state-contingent returns, and face survival risk. Financial intermediation is imperfect due to the limited enforcement of deposit contracts and the resulting enforcement constraint faced by individual banks. The enforcement constraint posits that the forward-looking bank value must be at least as great as the value of default—running away with a fraction of assets—at all possible contingencies. The constraint is thus similar to that studied by Kehoe and Perri (2002) in an international real business cycle model with endogenously incomplete markets. In addition, banks can become insolvent under limited liability, with the deposit insurance agency financed lump sum guaranteeing that deposits are risk free.

I highlight two distinct sources of the inefficiency of the competitive equilibrium allocation. First, there are pecuniary externalities: individual bankers do not internalize that their private portfolio decisions influence the price of claims on firm profits and the future return on bank assets, affecting both the forward-looking bank value and the value of default of all banks. The equilibrium asset price at t is a function of the aggregate capital stock chosen at $t-1$ and t, depending negatively on the former and positively on the latter. Consequently, in the partial derivative sense, greater bank lending to the real sector at t is linked to a greater demand for capital goods and asset price at t, increasing the value of default; furthermore, there is a lower asset price and marginal product of capital at $t+1$, decreasing the ex-post return on bank assets at $t+1$ and the bank value at t. Both effects tighten enforcement constraints of all banks at t. On the other hand, by decreasing the asset price at $t+1$, greater bank lending at t decreases the value of default at $t+1$, relaxing future enforcement constraints. Hence, while the former effects are consistent with excessive borrowing and lending in the competitive equilibrium, the latter effect contributes to insufficient borrowing and lending, and in both cases, the effect on bank borrowing is due to the balance sheet identity. There are, moreover, additional externalities, many of which depend on the extent of commitment by the policymaker. A planner that can choose a policy plan at the beginning of time once and for all internalizes the effect of allocations at t on $t-1$ expectations. As a result, the planner with commitment internalizes that greater bank lending at t increases the bank value and relaxes the enforcement constraint at $t-1$, which is due to a positive effect on the asset price and thus on the ex-post asset return at t. This effect is absent without commitment when the planner must consider how current decisions affect the endogenous state and optimal decisions in the future.

The second type of inefficiencies is the very nature of the limited enforcement friction that restricts bank borrowing and lending compared to the economy without such a friction. Intuitively, if the enforcement constraint is binding at t, one can achieve a strict welfare improvement by promising a greater future bank value conditional on survival to $t+1$, relaxing the financial constraint at t. Formally, this goal can be achieved by manipulating the future bank value distribution between entrants and survivors. This strategy has a limitation in that an implementable distribution must be constrained to a half-open unit interval: entrants must have a positive bank value to operate. Since the feasible space is not compact, it is not guaranteed that the maximum can be attained: indeed, to relax the enforcement constraint in some contingencies, the planner might want to choose a distribution that is infinitely close to the feasible boundary. To avoid this caveat, I conduct the normative analysis either for a given distribution or under a constraint specifying that the distribution must be in a certain sense consistent with that endogenously arising in the competitive equilibrium. Under the assumption of commitment, the planner internalizes how allocations affect the future distribution. E.g., a greater future bank debt decreases the future bank value of survivors and tightens the current enforcement constraint, thus contributing to potential overborrowing and overlending in the competitive equilibrium taking the enforcement friction as given.

Formally, I characterize the constrained efficient allocation, which results from a planning problem of a benevolent policymaker that maximizes household welfare subject to the competitive equilibrium implementability constraints except for the optimality conditions of bankers: that is, the policymaker makes portfolio decisions on behalf of the banking system. I study this problem both under the assumption of commitment and no commitment on the planner's side. The no-commitment case corresponds to a Markov perfect equilibrium of a non-cooperative game between successive policymakers (Klein et al., 2008). As explained in the previous paragraphs, both types of constrained efficient allocations highlight similar distortions in the competitive equilibrium. There are, however, two key differences. First, the competitive equilibrium deposit supply is efficient in the intertemporal sense when compared to the commitment allocation for a given bank value distribution. There is no wedge between the agent's and the planner's bank debt Euler equations. At the same time, due to the balance sheet constraint, the quantity of deposits need not be efficient if bank loans are not. Second, by construction of the Markov allocation, the time-consistent planner cannot affect the future bank value distribution except by affecting the future endogenous state. Therefore, the distribution is always taken as given in the time-consistent analysis. Although the argument about the potential welfare benefit of the survivors-biased bank value distribution generally applies, there are crucial quantitative differences from the case of commitment.

I show how to implement both types of constrained efficient allocations in a regulated competitive equilibrium with two types of policy instruments that address the two types of inefficiencies. The externalities can be corrected either by linear taxes on deposits and loans balanced in the aggregate or by one of these types of taxes rebated lump sum, by targeting the bank capital ratio, or—under certain assumptions—with minimum state-contingent capital requirements. The problem with the latter is in its asymmetric nature, which does not allow closing the wedges in the contingencies in which the optimal credit spread is too low. The given bank value distribution can be achieved with entrants/survivors-specific transfers—preemptive bailouts—that either add up to zero or match the aggregate lump-sum transfer that rebates the proceeds from the linear taxes. For most computations, I use global projection methods to fully account for precautionary savings effects when the occasionally binding enforcement constraint is about to switch from the slack to the binding regime.

Quantitatively, in both the Markov perfect and Ramsey equilibria, the enforcement constraint is binding by an order of magnitude less often than in the competitive equilibrium. Both normative arrangements generate sizable consumption-equivalent welfare gains: from a 0.57% state-space median at the Markov allocation to a 0.75% ergodic distribution average at the commitment allocation.

At the same time, there are crucial differences in the nature of the optimal bank value distribution and the relative magnitude of bank assets and debt. In the Markov perfect equilibrium, the optimal distribution is more entrants biased than in the competitive equilibrium, and banks generally borrow and lend less than in the competitive equilibrium. Conversely, in the Ramsey equilibrium, an opposite situation occurs: the optimal distribution is more survivors biased, and there is more borrowing and lending than in the competitive equilibrium. At the same time, the Ramsey allocation has less borrowing and lending than in the unconstrained competitive equilibrium—that is, the competitive equilibrium in the environment without the enforcement friction. These differences reflect a trade-off that the planner faces. Without commitment, limiting excessive borrowing and lending ex ante to address the pecuniary externalities is the key pursuit, as the planner has limited ability to affect the future bank value distribution. A more entrants-biased distribution helps to achieve this goal. However, preemptive bailouts have more power with commitment, and the planner leans toward a more survivors-biased distribution that generates greater equilibrium borrowing and lending. At the same time, it is not optimal to choose an extremely biased distribution, and under both arrangements, transfers to survivors increase around crises. We can thus identify a two-sided objective: on the one hand, to prevent banks from becoming too large ex ante; on the other hand, to provide preemptive support to older and larger banks when financial constraints bind ex post.

Related literature This paper is related to the literature on financial crises and pecuniary externalities arising from endogenous financial constraints. Lorenzoni (2008) is the first to highlight overborrowing in the competitive equilibrium due to a pecuniary externality in the price of capital in a three-period model with two-sided limited commitment and direct finance; Dávila and Korinek (2018) provide a comprehensive theoretical analysis of pecuniary externalities. Bianchi (2011) considers a quantitative endowment (small open) economy with two goods and a flow collateral constraint that depends on the relative price of nontradable goods. He shows that overborrowing due to the pecuniary externality can be corrected with a state-contingent debt tax. In the same model, Benigno et al. (2016) show that policies that relax financial constraints ex post achieve greater welfare than the optimal debt tax since the former can implement the unconstrained first-best allocation. The competitive equilibrium features underborrowing compared to that allocation.

Moreover, Schmitt-Grohé and Uribe (2021) emphasize that there exist reasonable parameterizations under which multiple equilibria arise in that model, including a self-fulfilling equilibrium that features underborrowing. Benigno et al. (2013) find that underborrowing arises in a related production economy and highlight the importance of ex-post policies. In small open and endowment economies with stock collateral constraints, Bianchi and Mendoza (2018) and Jeanne and Korinek (2019) identify overborrowing in the competitive equilibrium compared to the Markov perfect equilibrium and characterize the optimal time-consistent debt tax. The current paper focuses on the implications of pecuniary externalities due to asset prices and asset returns that affect the forward-looking enforcement constraint faced by financial intermediaries. My quantitative findings correlate with the findings in the small open and endowment economy contexts. I identify overborrowing by banks in the competitive equilibrium compared to the Markov perfect equilibrium outcome but underborrowing compared to the Ramsey outcome.

This paper is also related to the literature on financial crises, bailouts, and optimal financial regulation. Allen and Gale (2004) generalize the model of Diamond and Dybvig (1983) and show that with complete markets and limited market participation, the competitive equilibrium is either incentive efficient or constrained efficient, defaults and financial crises occur in equilibrium with incomplete contracts, and no regulation is warranted. However, with incomplete markets, there is room for liquidity regulation.[1] In the current paper, endogenous market incompleteness gives rise to pecuniary externalities and inefficient financial intermediation. Farhi and Tirole (2012) demonstrate that imperfectly targeted time-consistent accommodating interest-rate policies lead to multiple equilibria, increase correlation in risk-taking behavior by financial intermediaries and sow the seeds of future crises. Regulation in the form of a cap on short-term debt reduces the set of equilibria to a singleton that corresponds to the commitment benchmark. The current paper focuses on the symmetric equilibrium in the banking system to facilitate aggregation and permit tractable positive and normative analysis. Even though preemptive bailouts are imperfectly targeted within entrants and survivors, the symmetric equilibrium is not subject to collective moral hazard. Bianchi (2016) studies the implications of a pecuniary externality in an equity constraint that depends on the market wage rate and emphasizes the benefits of a systemic debt relief policy—a proportional

[1] Farhi et al. (2009) show that the competitive equilibrium in the model of Diamond and Dybvig (1983) is constrained inefficient even with complete markets if agents can engage in hidden trades.

reduction in debt repayments—that helps relax equity constraints during crises. The objective of relaxing financial constraints is similar to the objective of preemptive bailouts in the current paper, but the latter constitute a somewhat different policy—group-dependent lump-sum transfers provided to banks at $t+1$ to relax financial constraints at t. Chari and Kehoe (2016) develop a model where costly firm bankruptcies occur in the competitive equilibrium, which is both ex-ante and ex-post efficient compared to the commitment benchmark. Without commitment, inefficient bailouts will arise, and regulation in the form of a limit on the debt-to-value ratio and the tax on firm size is desirable to achieve a sustainably efficient outcome. In the current paper, the competitive equilibrium is constrained inefficient compared to the commitment benchmark, while preemptive (not actual) bailouts help mitigate the source of endogenous market incompleteness.

As part of smaller quantitative literature, Boissay et al. (2016) develop a real business cycle model with a banking sector that features an interbank market. High-skilled banks borrow from low-skilled banks and households to lend to firms and may decide to divert borrowed funds to invest in the storage technology subject to diversion costs. A relevant incentive compatibility constraint eliminates the former possibility. The authors briefly discuss the constrained inefficiency of the competitive equilibrium and compute welfare losses. A crucial difference from the current paper is that the bank's problem is static, and the incentive constraint is always binding in equilibrium; therefore, the sources of the inefficiency of the competitive equilibrium are utterly different from those in the current paper. Indeed, in the current paper, the competitive equilibrium is constrained efficient if the enforcement constraint is always binding, and the bank value distribution externality arises because the enforcement constraint is forward looking. Collard et al. (2017) study locally Ramsey-optimal bank capital requirements and monetary policy. In their model, sufficiently high capital requirements help eliminate risky lending in equilibrium. On the contrary, in the current paper, capital requirements do not generally constitute an effective policy instrument, and their role is different—to force individual banks to internalize pecuniary externalities due to the enforcement constraint. In a continuous-time environment, Di Tella (2019) demonstrates how the possibility of hidden trades in physical capital by intermediaries inflates the asset price and risk exposure of other intermediaries. The constrained efficient allocation can be implemented with a tax on assets, while bank capital requirements are ineffective. Van der Ghote (2021) develops a continuous-time model with nominal rigidities and a banking sector similar to that in the current paper but with the capital requirement constraint imposed as part of the environment. The author restricts attention to Markov equilibria and acknowledges the presence of pecuniary externalities, discussing them intuitively and computing the optimal capital requirement numerically. The current paper instead characterizes constrained efficient allocations that do not depend on the presence of specific policy instruments. Indeed, as mentioned above, capital requirements might not be effective for correcting the externalities. Moreover, the current paper identifies the bank value distribution externality and conducts the normative analysis both with and without commitment on the planner's side.

Several papers studied the welfare implications of specific policies in related environments, assuming that the enforcement constraint is always binding so that one can use smooth local approximations (Gertler and Kiyotaki, 2010; Gertler and Karadi, 2011; Gertler et al., 2012; De Paoli and Paustian, 2017). In the current paper, the enforcement constraint is occasionally binding, and the model is solved using global or quasi-global methods. Moreover, as mentioned above, the competitive equilibrium is constrained efficient if the enforcement constraint is always binding, so in that case, regulation might not be desirable. Gertler et al. (2020b) and Akinci and Queralto (forthcoming) also use global methods, but they do not study efficiency, restricting the analysis to specific policy rules.

The rest of this paper proceeds as follows. Section 1.2 describes the theoretical model and defines the sequential and recursive competitive equilibria. Section 1.3 conducts the normative analysis. Section 1.4 presents quantitative results. Section 1.5 concludes. Appendix contains proofs of theoretical results.

1.2 Model

The model is an extension of Gertler and Kiyotaki (2010) and Gertler and Karadi (2011). An economy is populated by households, firms, and the government. Each household has a unit measure of members: $f \in (0, 1)$ bankers and $1 - f$ workers at each point in time. Bankers and workers switch occupations, but their measures remain constant. A head of household makes family consumption, saving, and labor supply decisions. Bankers manage banks that combine their net worth and deposits issued to other households to invest in final-good producing firms. Final goods are produced using capital and labor with constant returns to scale. The purchase of capital goods requires external financing. Capital goods are produced from final goods subject to external adjustment costs. The government provides deposit insurance through lump-sum taxation of households.

As will be clear momentarily, there are two sources of aggregate uncertainty in our economy; therefore, all allocations and prices are history-contingent functions. We will denote a specific history as $z^t \in Z^t$ and the time-0 probability or density of this history as $\pi(z^t)$. We will only make the history dependence explicit when it is essential for clarity of the argument.

1.2.1 Households

On behalf of a family, the head of the household decides how much to consume C_t, save in one-period risk-free deposits $\frac{D_{t+1}}{R_t}$ with the gross return R_t, and how much labor L_t to supply at the wage rate W_t. The budget constraint of the household is

$$C_t + \frac{D_{t+1}}{R_t} \leq W_t L_t + D_t + \Pi_t - T_t,$$

where Π_t denotes net transfers from the ownership of banks and firms, and T_t is a lump-sum tax.

The household's preferences are represented by $\mathbb{E}_0[\sum_{t=0}^{\infty} \beta^t U(C_t, L_t)]$, where $\beta \in (0,1)$, $U : \mathbb{R}_+^2 \to \mathbb{R}$ is twice continuously differentiable and strictly concave with $U_C > 0$, $U_L < 0$, and $\lim_{C\to 0} U_C(C, L) = \infty$ for all $L \geq 0$. Let $\Lambda_{t,s} \equiv \beta^{s-t} \frac{U_{C,s}}{U_{C,t}}$ denote a stochastic discount factor (SDF) for periods $s \geq t \geq 0$. The necessary conditions for optimality include the budget constraint holding as equality, the labor supply condition that links the wage to the marginal rate of substitution of consumption for leisure (1.1), and the Euler equation that prices deposits (1.2).

$$W_t = -\frac{U_{L,t}}{U_{C,t}}, \tag{1.1}$$

$$1 = R_t \mathbb{E}_t(\Lambda_{t,t+1}). \tag{1.2}$$

Combined with initial and transversality conditions on $\{D_t\}$, the above equations are also sufficient to determine the household's optimal plan given prices and government policies.

1.2.2 Bankers

A banker manages a bank that invests net worth n_t and deposits $\frac{d_{t+1}}{R_t}$ into firms' equity s_t at a price Q_t. The banker may also decide to pay dividends x_t to their household. The bank's balance sheet constraint is then

$$Q_t s_t \leq n_t - x_t + \frac{d_{t+1}}{R_t}.$$

Bankers are assumed to stay in the banking business for a finite expected time. Specifically, a banker in the period t remains to be a banker in the period $t+1$ with the probability $\sigma \in [0,1)$ and becomes a worker with the probability $1-\sigma$. Hence, the expected lifetime of the banking business is $\frac{1}{1-\sigma}$. A banker that exits transfers the accumulated net worth to their household. Accordingly, $(1-\sigma)f$ workers start a banking business each period, being endowed with a startup net worth $n_t^0 > 0$ by their households. The future net worth is the difference between the ex-post returns on assets and liabilities:

$$n_{t+1} = R_{t+1}^K Q_t s_t - d_{t+1},$$

where R_t^K is the gross return on firms' equity. We will make the following assumption to ensure that the aggregate net worth of the banking system does not explode over time.

Assumption 1.1. *$\sigma < \beta$.*

Unlike in the standard framework, we will explicitly allow for the possibility of becoming insolvent. If a banker has survived to period $t > 1$ but $n_t \leq 0$, the banker cannot operate and remains inactive until becoming a worker. If a banker has become a worker in t and $n_t \leq 0$, the banker cannot transfer anything to their household. Hence, there is limited liability on the banker's side. In the baseline analysis, we will assume that there is a deposit insurance agency that guarantees households a risk-free return on deposits and is funded lump sum.

Let $\eta_t \equiv \mathbf{1}_{\mathbb{R}_{++}}(n_t)$. The bank value v_t satisfies a stochastic difference equation:

$$\begin{aligned} v_t &= x_t + \mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[(1-\sigma)n_{t+1} + \sigma v_{t+1}]\} \\ &= x_t + \mathbb{E}_t\left\{\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\eta_{t+j}\right)\Lambda_{t,t+i}\sigma^{i-1}[(1-\sigma)n_{t+i} + \sigma x_{t+i}]\right\}. \end{aligned}$$

Since households are identical, the same stochastic discount factor applies to future payoffs of all bankers.

We assume that at the end of a period t, a banker could divert a fraction $\theta \in [0, 1]$ of assets to their household. In that case, the bank would default, while other households could recover only the remaining fraction $1 - \theta$ of assets. Consequently, other households will be willing to lend to the bank only if the following incentive compatibility constraint holds:

$$v_t \geq \theta Q_t s_t.$$

This constraint is an enforcement constraint (EC) of the type studied in Kehoe and Perri (2002). Since v_t, and thus the banker's budget set, depends on infinitely many future controls, a recursive representation of the banker's problem need not exist (Marcet and Marimon, 2019). The standard approach in the literature is to guess that v_t is linear in individual net worth and reformulate the problem in terms of the state-contingent marginal value of net worth common to all bankers. Although this approach identifies a solution to the banker's problem, other—nonlinear—solutions could exist. To investigate this possibility, we will solve the general banker's problem, not taking ex-ante assumptions on the form of v_t.[2]

The banker's problem is

$$\max_{\{d_{t+1}, s_t, v_t, x_t\}} v_0$$

subject to the non-negativity, balance sheet, bank value, and ECs. Let $\nu_t(z^t)$, $\gamma_t(z^t)$, and $\lambda_t(z^t)$ denote the Lagrange multipliers—normalized by $\beta^t \pi(z^t)$—associated with the latter three constraints and define the scaled multipliers

$$\bar{\nu}_t \equiv \frac{\frac{\nu_t}{\gamma_t}}{1 - \frac{\lambda_t}{\gamma_t}}, \qquad \bar{\lambda}_t \equiv \frac{\frac{\lambda_t}{\gamma_t}}{1 - \frac{\lambda_t}{\gamma_t}}.$$

Consider the following assumption.

Assumption 1.2. *At an optimal plan, for all $t \geq 0$, $z^t \in Z^t$, and all continuations of z^t, a sequence $n \mapsto \beta^n \prod_{i=1}^{n} \eta_{t+i}(z^{t+i})(1 + \bar{\lambda}_{t+i-1}(z^{t+i-1}))$ is bounded.*

Note that $\lim_{n\to\infty} \beta^n \prod_{i=1}^{n} \eta_{t+i} = 0$ for all (t, z^t) and all continuations of z^t since $\beta < 1$ and

[2] As shown by Marcet and Marimon (2019), the Lagrangian in these types of problems admits a recursive representation on the expanded state space. The solution characterized below can be equivalently derived using the reformulated Lagrangian.

$\prod_{i=1}^{n} \eta_{t+i} \in \{0,1\}$ for all $n \in \mathbb{N}$. Assumption 1.2 thus requires that $\prod_{i=1}^{n}(1+\bar{\lambda}_{t+i-1})$ does not grow too fast in those histories $z^\infty \in Z^\infty \mid z^t$ in which bankers are solvent forever. Conversely, a sufficient condition for assumption 1.2 is that for any history $z^\infty \in Z^\infty$ and any $t \geq 0$, there exists $i > 0$, such that $\eta_{t+i}(z^{t+i}) = 0$. In other words, independently of the current history, the banker will become insolvent at some point in the future with probability one. Under the latter stronger assumption, all sequences in assumption 1.2 are not only bounded but converge to zero.

The next proposition characterizes the solution to the banker's problem.

Proposition 1.1. *Under assumption 1.2, the unique bounded solution to the banker's problem has $v_t = \bar{\nu}_t n_t$. At the optimal plan, $x_t = 0$ for all (t, z^t) and the following Karush—Kuhn—Tucker (KKT) conditions hold:*

$$\bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})]R_t, \tag{1.3}$$

$$\theta\bar{\lambda}_t + \bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})R^K_{t+1}], \tag{1.4}$$

$$0 = \bar{\lambda}_t(v_t - \theta Q_t s_t), \qquad \bar{\lambda}_t \geq 0.$$

We have shown that the linear solution to the banker's problem is indeed the unique solution, so the conventional approach in the literature is without loss of generality. The risk-neutrality of bankers is critical for this result. According to the expression for the value function, the marginal value of net worth equals the scaled multiplier $\bar{\nu}_t$. The intuition is that net worth is more valuable when the balance sheet constraint is tighter: the greater the original multiplier on the balance sheet constraint ν_t, the greater the marginal value of net worth $\bar{\nu}_t$.

As shown in the proof of proposition 1.1, at the optimal plan, $\gamma_t = 1+\sum_{j=0}^{t}\lambda_j$. Remember that γ_t affects the scaled multipliers $\bar{\nu}_t$ and $\bar{\lambda}_t$; therefore, similar to Kehoe and Perri (2002) and Marcet and Marimon (2019), the solution to the banker's problem depends on the history of Lagrange multipliers associated with the EC. At the same time, the scaled multipliers $\bar{\nu}_t$ and $\bar{\lambda}_t$ are sufficient statistics for the characterization of the optimal plan. For this reason, the banker's problem admits a recursive representation, as we will see in a later subsection.

The KKT conditions (1.3) and (1.4) imply that the risk-adjusted credit spread $\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})(R^K_{t+1}-R_t)]$ is entirely determined by the scaled multiplier $\bar{\lambda}_t$. The tighter the EC, the greater the $\bar{\lambda}_t$, and the greater the spread. Due to limited liability, the greater the probability of future insolvency, the lower the marginal cost of issuing deposits and the marginal benefit of extending credit, which is a standard source of moral hazard.

Since the marginal value of net worth is a function of the aggregate state only, the aggregate bank value V_t is a multiple of the aggregate net worth N_t, that is, $V_t = \bar{\nu}_t N_t$. By aggregating the individual ECs, one then obtains the aggregate EC:

$$\bar{\nu}_t N_t \geq \theta Q_t S_t^K, \tag{1.5}$$

where S_t^K denotes the aggregate holdings of firms' equity. The constraint (1.5) is binding if and only if all individual ECs are binding, if $\bar{\lambda}_t > 0$. Similarly, (1.5) is not binding if and only if at least one individual EC is not binding, only if $\bar{\lambda}_t = 0$. Hence,

$$0 = \bar{\lambda}_t(\bar{\nu}_t N_t - \theta Q_t S_t^K), \qquad \bar{\lambda}_t \geq 0. \tag{1.6}$$

Aggregating the individual balance sheet constraints—all of which bind since $\bar{\nu}_t > 0$—the banking sector balance sheet is

$$Q_t S_t^K = N_t + \frac{D_{t+1}}{R_t}. \tag{1.7}$$

The aggregate net worth N_t at $t > 0$ is a sum of the aggregate net worth of survivors and entrants. To derive the aggregate net worth of survivors, define the aggregate net payoff on assets $\widetilde{N}_t \equiv R_t^K Q_{t-1} S_{t-1}^K - D_t$. In a symmetric equilibrium, $\widetilde{N}_t > 0$ if and only if $n_t > 0$ for all survivors; hence, $\eta_t = \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_t)$. Since the exit probability is the same across bankers, the aggregate net worth of survivors is $\sigma \max(\widetilde{N}_t, 0)$. I assume that the aggregate endowment of entrants satisfies $N_t^0 = \bar{N} + \omega Q_t S_{t-1}^K$, where $(\bar{N}, \omega) \in \mathbb{R}_+^2$. Hence, $N_t = \sigma \max(\widetilde{N}_t, 0) + N_t^0$, and thus

$$N_t = \bar{N} + \sigma \max(R_t^K Q_{t-1} S_{t-1}^K - D_t, 0) + \omega Q_t S_{t-1}^K. \tag{1.8}$$

Combining (1.7) and (1.8), one can derive the law of motion for the aggregate net worth

$$N_{t+1} = \bar{N} + \sigma \max[(R_{t+1}^K - R_t) Q_t S_t^K + \sigma R_t N_t, 0] + \omega Q_{t+1} S_t^K.$$

Based on this equation, a necessary condition for the existence of a deterministic steady state with $\widetilde{N} > 0$ is $\sigma R < 1$, which ensures that the initial net worth $N_0 > 0$—determined by the initial conditions $D_0 \geq 0$ and $K_0 > 0$—vanishes as $t \to \infty$. Taking into account (1.2), we see that the former condition is equivalent to that stated in assumption 1.1. Quantitatively, assumption 1.1 is also necessary for the existence of an ergodic distribution. Intuitively, $\sigma R_t < 1$ must hold "on average" to have $\lim_{t\to\infty} \mathbb{E}(N_t) \in \mathbb{R}_{++}$.

1.2.3 Firms

The economy is populated by firms that produce final and capital goods.

Final good producers

Firms that produce the final good demand labor L_t and purchase machines and equipment K_t from capital good producers. The technology is represented by a production function $F : \mathbb{R}_+^2 \to \mathbb{R}_+$, which is twice continuously differentiable, satisfies Inada conditions, and exhibits constant returns to scale. Firms rely on external financing from banks to purchase capital goods by offering state-contingent securities, which correspond to the quantity of capital goods demanded. By no-arbitrage, the price of both securities and capital goods equals Q_t.

The representative firm's profit is $\Pi_t^F \equiv A_t F(\xi_t K_t, L_t) - W_t L_t - R_t^K Q_{t-1} K_t + Q_t(1-\delta)\xi_t K_t$, where A_t is the total factor productivity (TFP), ξ_t represents capital quality, and $\delta \in [0, 1]$ is the depreciation rate. Both $\{A_t\}$ and $\{\xi_t\}$ are exogenous stochastic processes. Profit maximization implies a labor demand condition

$$W_t = A_t F_{L,t}. \tag{1.9}$$

The financial contract between banks and firms posits that the revenue net of labor costs is paid as a dividend; hence,

$$R_t^K = \frac{A_t F_{K,t} + Q_t(1-\delta)}{Q_{t-1}} \xi_t. \tag{1.10}$$

Capital good producers

New capital goods are produced using final goods as inputs. A stock of capital K_t and inputs I_t allow to produce $\Phi\left(\frac{I_t}{K_t}\right) K_t$, where $\Phi : \mathbb{R}_+ \to \mathbb{R}$ satisfies $\Phi' > 0$, $\Phi'' \leq 0$, $\lim_{x\to 0} \Phi'(x) = \infty$, and $\lim_{x\to\infty}\Phi'(x)=0$, similar to Lucas and Prescott (1971). The firm chooses $\{I_t\}$ to maximize $\mathbb{E}_0\left\{\sum_{t=0}^{\infty}\Lambda_{0,t}\left[Q_t\Phi\left(\frac{I_t}{K_t}\right)K_t-I_t\right]\right\}$. The supply curve for new capital goods is described by

$$Q_t=\left[\Phi'\left(\frac{I_t}{K_t}\right)\right]^{-1}. \tag{1.11}$$

1.2.4 Government

The government provides deposit insurance financed lump sum, implying the budget constraint

$$T_t=-\min(\widetilde{N}_t,0). \tag{1.12}$$

1.2.5 Market clearing

The capital, securities, and final good markets clear as follows:

$$K_{t+1}=(1-\delta)\xi_tK_t+\Phi\left(\frac{I_t}{K_t}\right)K_t, \tag{1.13}$$

$$S_t^K=K_{t+1}, \tag{1.14}$$

$$A_tF(\xi_tK_t,L_t)=C_t+I_t. \tag{1.15}$$

1.2.6 Competitive equilibrium

We will now define the sequential and recursive competitive equilibria (CE) in the unregulated economy.

A sequential CE (SCE) can be defined as follows.

Definition 1.1. *Given initial conditions D_0, K_0, transversality conditions, and exogenous stochastic processes $\{A_t,\xi_t\}$, an SCE is represented by the following measurable functions that map Z^t to $\mathbb{R}$ for all $t\geq 0$:*

- *allocations C_t, D_{t+1}, I_t, K_{t+1}, L_t, N_t, S_t^K;*
- *prices Q_t, R_t, R_t^K, W_t;*
- *transformed Lagrange multipliers $\bar{\lambda}_t$, $\bar{\nu}_t$;*

- *deposit insurance tax T_t.*

The functions are consistent with (1.1)–(1.15) *for all $t \geq 0$.*

The linearity of the banker's problem allows the construction of a set of allocations of individual bankers consistent with the SCE.

We need to introduce additional notation to define a recursive CE (RCE). Denote as $\mathcal{X} \subseteq \mathbb{R}^2$ and $\mathcal{Z} \subseteq \mathbb{R}^2$ the spaces of endogenous and exogenous state variables $X \in \mathcal{X}$ and $z \in \mathcal{Z}$. We have $X = (D, K)$ and $z = (A, \xi)$. Let $\mathcal{S} \equiv \mathcal{X} \times \mathcal{Z}$ with $S \in \mathcal{S}$. To simplify notation, we will often use the subscripts S and S' to denote the values of functions evaluated at those states. The problems of all agents except bankers could be identically set up recursively, yielding the recursive analogs of the corresponding optimality conditions. As discussed in the description of the banker's problem, the EC generally depends on future control variables and thus does not allow setting up a recursive problem directly. At the same time, proposition 1.1 showed that the banker's value function in the sequential problem is linear in net worth. Therefore, the banker's problem admits a recursive representation, as stated in the following lemma.

Lemma 1.1. *The banker's problem can be represented by the Bellman equation*

$$v(n, S) = \max_{(d,s) \in \Gamma(n,S)} \mathbb{E}_z\{\eta' \Lambda_{S,S'}[(1-\sigma)n' + \sigma v(n', S')]\},$$

where $\eta' \equiv \mathbf{1}_{\mathbb{R}_{++}}(n')$ and the correspondence $\Gamma : \mathbb{R}_+ \times \mathcal{S} \to \mathcal{P}(\mathbb{R}^2_+)$ is defined by the following constraints

$$\begin{aligned} \nu: \quad & 0 \leq n + \frac{d}{R_S} - Q_S s, \\ \lambda: \quad & 0 \leq \mathbb{E}_z\{\eta' \Lambda_{S,S'}[(1-\sigma)n' + \sigma v(n', S')]\} - \theta Q_S s, \\ & n' = R^K_{S'} Q_S s - d. \end{aligned}$$

1. *In a symmetric equilibrium, $\eta \in \{0, 1\}$, $\lambda \geq 0$, and $\nu \geq 1$ are independent of n.*

2. The KKT conditions, in addition to constraints, are

$$\nu_S = (1+\lambda_S) R_S \mathbb{E}_z[\eta_{S'} \Lambda_{S,S'} (1-\sigma+\sigma\nu_{S'})], \tag{1.16}$$

$$\theta\lambda_S + \nu_S = (1+\lambda_S) \mathbb{E}_z[\eta_{S'} \Lambda_{S,S'} (1-\sigma+\sigma\nu_{S'}) R^K_{S'}], \tag{1.17}$$

$$0 = \lambda_S(\nu_S n - \theta Q_S s), \qquad \lambda_S \geq 0.$$

3. The solution to the Bellman equation is $v(n,S) = \nu_S n$.

The Euler equations (1.16) and (1.17) are equivalent to their sequential counterparts (1.3) and (1.4). So are the expressions for the value functions. The aggregate bank value is $V_S = \nu_S N_S$, and the aggregate EC is

$$\nu_S N_S \geq \theta Q_S S^K_S. \tag{1.18}$$

Similar to the discussion after proposition 1.1, the aggregate complementary slackness conditions are

$$0 = \lambda_S(\nu_S N_S - \theta Q_S S^K_S), \qquad \lambda_S \geq 0. \tag{1.19}$$

We are now ready to define an RCE.

Definition 1.2. *Given the exogenous Markov processes $\{A, \xi\}$, an RCE is represented by the following measurable functions that map $\mathcal{S}$ to $\mathbb{R}$:*

- *allocations C, D', I, K', L, N, S^K;*
- *prices Q, R, R^K, W;*
- *Lagrange multipliers λ, ν;*
- *deposit insurance tax T.*

The functions are consistent with (1.16)–(1.19) and the recursive versions of (1.1), (1.2), (1.7)–(1.15). The aggregate law of motion $S \mapsto S'$ is generated by D', K', and Markov transitions $z \to z'$.

1.3 Normative analysis

This section studies the problem of a benevolent social planner who will maximize household welfare, internalizing the determination of market prices and making the optimal portfolio decisions on behalf of the banking system subject to the aggregate EC. We will characterize the constrained efficient allocation under commitment (CEA) and the Markov-perfect constrained efficient allocation (MCEA). We will show how to implement the CEA and MCEA in the regulated CE with either affine taxes on bank assets and liabilities or state-contingent capital requirements to address the pecuniary externalities and bank entrants/survivors-specific transfers to achieve the targeted bank value distribution.

1.3.1 Sources of inefficiency

To proceed with the formal characterization of the CEA and MCEA, we must derive the aggregate EC of the banking system. Doing so will also clarify the nature of distortions in the CE on an intuitive level.

Let us index the existing bankers with $i \in [0, f]$. We can assume without loss of generality that survivors are always in the $[0, \sigma f]$ interval, and entrants are in the $(\sigma f, f]$ interval. Hence, the indices of $(1-\sigma)\sigma f$ current survivors that will exit the next period will be filled by $\sigma(1-\sigma)f$ current entrants that will survive to the next period. Let $v^1_{i,t+1}(z^{t+1})$ denote the bank value of the banker i conditional on survival from z^t to z^{t+1}. Let $\sigma^1_t \equiv \frac{V^1_t}{V_t}$, where $V^1_t \equiv \int_0^{\sigma f} v_{i,t}\,\mathrm{d}i$ is the aggregate bank value of survivors. It follows that the aggregate bank value of the banking system satisfies

$$
\begin{aligned}
V_t &\equiv \int_0^f v_{i,t}\,\mathrm{d}i \\
&= \mathbb{E}_t\left\{\eta_{t+1}\Lambda_{t,t+1}\left[(1-\sigma)\int_0^f n_{i,t+1}\,\mathrm{d}i + \sigma\int_0^f v^1_{i,t+1}\,\mathrm{d}i\right]\right\} \\
&= \mathbb{E}_t\left\{\eta_{t+1}\Lambda_{t,t+1}\left[(1-\sigma)\widetilde{N}_{t+1} + \int_0^{\sigma f} v_{i,t+1}\,\mathrm{d}i\right]\right\} \\
&= \mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[(1-\sigma)\widetilde{N}_{t+1} + \sigma^1_{t+1}V_{t+1}]\},
\end{aligned}
\tag{1.20}
$$

where the second equality follows from the definition of the individual bank value of the banker

$i \in [0, f]$, the third equality follows from the fact that each banker i has the same probability of survival from z^t to z^{t+1}, and the fourth equality follows from definitions of V^1 and σ^1.

Using (1.14), the aggregate EC is

$$V_t \geq \theta Q_t K_{t+1}. \tag{1.21}$$

Further using the definition of $\widetilde{N}_t$ and (1.10) and substituting (1.20) in (1.21), we obtain

$$\mathbb{E}_t\Big\{\eta_{t+1}\Lambda_{t,t+1}\Big[(1-\sigma)\{[A_{t+1}F_K(\xi_{t+1}K_{t+1}, L_{t+1}) + Q(K_{t+1}, K_{t+2}, \xi_{t+1})(1-\delta)]\xi_{t+1}K_{t+1} - D_{t+1}\} \\ + \sigma^1_{t+1}V_{t+1}\Big]\Big\} \geq \theta Q(K_t, K_{t+1}, \xi_t)K_{t+1},$$

where the asset price function Q is defined by (1.11) and (1.13) as

$$Q(\underset{-}{K_t}, \underset{+}{K_{t+1}}, \xi_t) = \left[\Phi'\left(\Phi^{-1}\left(\frac{K_{t+1}}{K_t} - (1-\delta)\xi_t\right)\right)\right]^{-1}.$$

The function Q is decreasing in the first argument and increasing in the second argument, which follows from Φ being strictly increasing and concave.

There are two broad sources of potential distortions in the CE allocation. The first, highlighted in red, arises because individual bankers do not internalize how their asset allocations affect the current asset price and the future asset returns. The second, highlighted in green, reflects that the future continuation value of the banking system conditional on survival might be inefficiently low.

The first type of distortions reflects pecuniary externalities working through the asset price Q and the asset return R^K, affecting both the bank value and the value of default—running away with a fraction of assets. First, private bankers do not internalize that higher investment in the real sector—higher K_{t+1} in the aggregate—decreases the future asset returns by decreasing both the future marginal product of capital and the future asset price, which, in turn, decreases the current bank value and makes the ECs of all banks more likely to be binding at t. Second, individual bankers do not internalize that greater K_{t+1} increases the current asset price Q_t, making the default option more attractive and further increasing the probability that ECs of all bankers are binding at t. Third, since greater K_{t+1} decreases the future asset price, it has a negative effect on the future value of default, relaxing the future ECs. Fourth, from the perspective of the planner that has commitment, a higher K_{t+1} increases the $t-1$ expectation of the current asset return, thus relaxing the EC at $t-1$. Fifth, from the perspective of the planner that does not have commitment and limits its policies to Markovian ones, the changes in D_{t+1} and K_{t+1} are the changes in the endogenous state variables of the "future" planner, having multiple additional effects through the future policy functions. Therefore, the private portfolio decisions might be distorted through multiple channels working in opposite directions, some of which depend on the assumption of commitment from the planner's side. We will study these channels in more detail in the following subsections.

The nature of the second type of potential inefficiencies is linked to how the future bank value conditional on survival affects the current value of the banking system. From the perspective of an individual banker, the continuation value is a product of the constant survival probability σ and the future bank value v_{t+1}. From the planner's perspective, the aggregate continuation value equals the aggregate bank value of the survived banks V^1_{t+1}, which is a state-contingent share σ^1_{t+1} of the aggregate future bank value V_{t+1}. If the planner could choose $\{\sigma^1_t\}$, it would generally be optimal to increase it in all contingencies to relax the aggregate EC and thus expand the feasible set, potentially leading to welfare gains.

We are now ready to proceed with the formal characterization of the constrained efficient allocations, both with and without commitment.

1.3.2 Constrained efficient allocation under commitment

Consider the sequential planning problem of optimizing the household welfare by choosing infinite sequences of history-contingent allocations at $t = 0$ subject to relevant infinite sequences of history-contingent CE implementability constraints. By definition 1.1, the complete set of CE implementability conditions is (1.1)–(1.15). Since we let the planner optimize on behalf of the banking system, the constraints (1.3), (1.4), and (1.6) are not applicable. Consequently, we replace (1.5) with the definition of the aggregate bank value (1.20) and the aggregate EC (1.21). We can use (1.2), (1.8), (1.9), (1.10), (1.11), (1.12), (1.13), and (1.14) to solve for R_t, N_t, W_t, R^K_t, Q_t, T_t, I_t, and S^K_t, respectively. It is also convenient to define the investment and asset price functions I and Q. (The latter has already been defined in the previous subsection.) The investment function I is defined based on (1.13) as

$$I(\underset{-^*}{K_t}, \underset{+}{K_{t+1}}, \xi_t) = \Phi^{-1}\left(\frac{K_{t+1}}{K_t} - (1-\delta)\xi_t\right) K_t,$$

where * in $-^*$ indicates a numerical statement, although it is true under any reasonable calibration.

Before describing the planning problem, we must decide how to handle σ^1_{t+1} appearing in (1.20). By definition, we must have $\sigma^1_t(z^t) \in [0, 1)$ for all $t \geq 0$ and $z^t \in Z^t$. To see that the right bound is not included, note that otherwise, we would have $v_{i,t} = 0$ for all entering bankers $i \in (\sigma f, f]$. By the individual EC, we would then have $Q_t s_{i,t} = 0$ for all such i, implying that all entrants could not operate. Note that $\sigma^1_t = 0$ is possible since survivors can become insolvent. Suppose the planner considers $\{\sigma^1_t\}$ as a control variable. Since the latter affects the continuation value in the EC only, it may be optimal to set $\sigma^1_{t+1}(z^{t+1}) \to 1$ in the histories $z^{t+1} \in Z^{t+1} \mid z^t$ in which $\eta_{t+1}(z^{t+1}) = 1$ if the EC is binding at z^t. In such a case, the maximum cannot be attained. To avoid this problem, let us, first, define the CE-consistent bank value distribution $\{\widehat{\sigma}^1_t\}$, where

$$\widehat{\sigma}^1_t \equiv \eta_t \frac{\sigma \widetilde{N}_t}{N_t}.$$

We will then conduct the analysis under the assumption that $\{\sigma^1_t\}$ is either given or satisfies $\{\sigma^1_t\} = \{\widehat{\sigma}^1_t\}$. Since the feasible set for $\{\sigma^1_t\}$ is a space of sequences of functions that map to an open unit interval, we can explore the implications of alternative distributions $\{\sigma^1_t\}$ quantitatively in a straightforward manner.

The sequential planning problem is, therefore,

$$\max_{\{C_t, D_{t+1}, K_{t+1}, L_t, V_t\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t U(C_t, L_t)\right]$$

subject to

$$0 = N_t - Q(K_t, K_{t+1}, \xi_t)K_{t+1} + \beta \mathbb{E}_t \left(\frac{U_C(C_{t+1}, L_{t+1})}{U_C(C_t, L_t)}\right) D_{t+1},$$

$$0 = \beta \mathbb{E}_t \left\{\eta_{t+1} \frac{U_C(C_{t+1}, L_{t+1})}{U_C(C_t, L_t)} [(1-\sigma)\widetilde{N}_{t+1} + \sigma^1_{t+1} V_{t+1}]\right\} - V_t,$$

$$0 \leq V_t - \theta Q(K_t, K_{t+1}, \xi_t)K_{t+1},$$

$$0 = U_C(C_t, L_t)A_t F_L(\xi_t K_t, L_t) + U_L(C_t, L_t),$$

$$0 = A_t F(\xi_t K_t, L_t) - C_t - I(K_t, K_{t+1}, \xi_t),$$

where

$$\widetilde{N}_t \equiv [A_t F_K(\xi_t K_t, L_t) + Q(K_t, K_{t+1}, \xi_t)(1-\delta)]\xi_t K_t - D_t, \qquad \eta_t \equiv \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_t),$$

$$N_t \equiv \bar{N} + \sigma\eta_t\widetilde{N}_t + \omega Q(K_t, K_{t+1}, \xi_t)K_t,$$

and $\{\sigma_t^1\}$ is either given or satisfies

$$\sigma_t^1 = \sigma^1(\underset{-}{D_t}, \underset{-^*}{K_t}, \underset{+^*}{K_{t+1}}, \underset{+}{L_t}, A_t, \xi_t) = \widehat{\sigma}_t^1,$$

where we again use the notation $-^*$ and $+^*$ to indicate numerical statements.

Let us denote the Lagrange multipliers on the planner's constraints—normalized by $\beta^t\pi(z^t)$—as $\widetilde{\nu}_t$, $\widetilde{\gamma}_t$, $\widetilde{\lambda}_t$, λ_t^L, and λ_t^Y, respectively. Define $\nu_t \equiv \frac{\widetilde{\nu}_t}{U_{C,t}}$, $\lambda_t \equiv \frac{\widetilde{\lambda}_t}{U_{C,t}}$, and $\gamma_t \equiv \frac{\widetilde{\gamma}_t}{U_{C,t}}$. As in the CE, define $\widehat{x}_t \equiv \frac{x_t}{\gamma_t}$ and $\bar{x}_t \equiv \frac{\widehat{x}_t}{1-\widehat{\lambda}_t}$ for $x \in \{\nu, \lambda, \lambda^L, \lambda^Y\}$.

As discussed in the previous subsection, there are multiple potential sources of the inefficiency of the CE allocation. The next proposition provides a formal validation.

Proposition 1.2. *The CE (SCE) allocation is generally inefficient compared to the CEA. The CEA analogs of the Euler equations* (1.3) *and* (1.4) *are*

$$\bar{\nu}_t = (1+\bar{\lambda}_t)R_t\mathbb{E}_t\left[\eta_{t+1}\Lambda_{t,t+1}\left(1-\sigma+\sigma_{t+1}^1\sigma\bar{\nu}_{t+1} \quad \underbrace{-\frac{\partial\sigma_{t+1}^1}{\partial D_{t+1}}V_{t+1}}_{\textit{future distribution }(+)}\right)\right],$$

$$\theta\bar{\lambda}_t + \bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma_{t+1}^1\sigma\bar{\nu}_{t+1})R_{t+1}^K] + \Psi_t^K,$$

where Ψ_t^K *satisfies*

$$
\begin{aligned}
Q_t\Psi_t^K \equiv\ & \underbrace{\bar{\nu}_t Q_{2,t}\{[\sigma\eta_t(1-\delta)\xi_t+\omega]K_t - K_{t+1}\}}_{\textit{balance sheet } (-^*)} + \underbrace{(1+\bar{\lambda}_t)\mathbb{E}_t\left(\eta_{t+1}\Lambda_{t,t+1}\frac{\partial\sigma_{t+1}^1}{\partial K_{t+1}}V_{t+1}\right)}_{\textit{future distribution } (-^*)} \\
& \underbrace{-\bar{\lambda}_t\theta Q_{2,t}K_{t+1}}_{\textit{value of default } (-)} \quad \underbrace{-\frac{\bar{\lambda}_t^Y}{U_{C,t}}I_{2,t}}_{\textit{consumption } (-^*)} + \underbrace{\frac{\mathbf{1}_{\mathbb{N}}(t)\eta_t}{\sigma_t^1}\Bigg[\underbrace{(1-\sigma)Q_{2,t}(1-\delta)\xi_t K_t}_{\textit{asset return } (+)} + \underbrace{\frac{\partial\sigma_t^1}{\partial K_{t+1}}V_t}_{\textit{distribution } (+^*)}\Bigg]}_{t-1 \textit{ expectations}} \\
& + \underbrace{(1+\bar{\lambda}_t)\mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma_{t+1}^1\sigma\bar{\nu}_{t+1})[A_{t+1}F_{KK,t+1}\xi_{t+1}+Q_{1,t+1}(1-\delta)]\xi_{t+1}K_{t+1}\}}_{\textit{future asset return } (-)} \\
& + (1+\bar{\lambda}_t)\mathbb{E}_t\Bigg(\eta_{t+1}\Lambda_{t,t+1}\sigma_{t+1}^1\Bigg\{\underbrace{\bar{\nu}_{t+1}[\omega(Q_{1,t+1}K_{t+1}+Q_{t+1})-Q_{1,t+1}K_{t+2}]}_{\textit{future balance sheet } (+^*)} \\
& \underbrace{-\bar{\lambda}_{t+1}\theta Q_{1,t+1}K_{t+2}}_{\textit{future value of default } (+)} + \underbrace{\bar{\lambda}_{t+1}^L A_{t+1}F_{KL,t+1}\xi_{t+1}}_{\textit{future wage } (+^*)} + \underbrace{\frac{\bar{\lambda}_{t+1}^Y}{U_{C,t+1}}(A_{t+1}F_{K,t+1}\xi_{t+1}-I_{1,t+1})}_{\textit{future consumption } (+^*)}\Bigg\}\Bigg).
\end{aligned}
$$

Moreover, the following holds.

1. *If* (1.21) *at the CEA is either binding almost surely (a.s.) or slack a.s. for all* $t \geq 0$*, then the CEA is time consistent. Otherwise, it is generally time inconsistent.*

2. *If* $\{\sigma_t^1\} = \{\widehat{\sigma}_t^1\}$ *and* (1.21) *at the CEA is binding a.s. for all* $t \geq 0$*, and the CEA satisfies* $\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}f_{t+1}(R_{t+1}^K - R_t)] \geq 0$ *for all* $\{f_t\}_{t\geq 1}$ *with* $f_t : Z^t \to \mathbb{R}_{++}$*, then the CEA equals the CE allocation, that is, the latter is constrained efficient.*

3. *Given* $\{\sigma_t^1\}$*, for all* $t \geq 0$*, let* $\bar{Z}^t \subseteq Z^t$ *be the set of histories at which* (1.21) *is strictly binding—in the sense that the corresponding Lagrange multiplier is positive. If* $\bar{Z}^t$ *is of positive measure at least for some* $t \geq 0$*, then there exists* $\{\widetilde{\sigma}_t^1\}$ *with* $\widetilde{\sigma}_t^1(z^t) \in [\sigma_t^1(z^t), 1)$ *for all* (t, z^t) *such that* $\{\widetilde{\sigma}_t^1\}$ *is strictly preferred to* $\{\sigma_t^1\}$.

Our first observation is that if the planner takes the distribution as given, there is no distortion in the choice of deposits, consistent with our intuitive analysis in the previous subsection.[3] If

[3] One might notice a slight difference in the Euler equations: instead of $\sigma\bar{\nu}_{t+1}$ in proposition 1.1, we have $\sigma_{t+1}^1\sigma\bar{\nu}_{t+1}$ in proposition 1.2—this difference is solely due to how the Lagrange multiplier on the bank value constraint is related to the multiplier on the EC. All original multipliers are stationary in the CEA, unlike in the CE. If one writes the deposit Euler equation in terms of the original multipliers, it will be symbolically equivalent to that in the CE, as one can verify in the proofs of propositions 1.1 and 1.2.

the planner internalizes the determination of the distribution, a wedge between the deposit Euler equations does appear: the social marginal cost of deposits is greater than the private marginal cost because the planner understands that greater borrowing at t has a negative effect on the future net worth of survived banks and, therefore, on their relative bank value. At the same time, the presence of the wedge in the Euler equation should not necessarily lead to overborrowing in the CE because the deposit Euler equation is essentially a fixed-point equation in the transformed multiplier $\{\bar{\nu}_t\}$ conditional on other variables, and the multipliers in the CE and CEA are generally different.

The wedge between the asset Euler equations Ψ_t^K consists of multiple terms with opposing effects on the wedge sign. If $\{\sigma_t^1\} = \{\widehat{\sigma}_t^1\}$, there are two terms (highlighted in green) capturing the effect of the choice of capital on the bank value distribution. On the one hand, greater capital negatively affects the future distribution through the negative effect on the future asset price and asset return. On the other hand, greater capital positively affects the $t-1$ expectation of the distribution at t through the positive effect on the current asset price. Both effects rely on the nature of commitment.

Consider the remaining terms that do not depend on the ability to affect the distribution. Greater capital affects the current asset price positively, increasing the ex-post asset return and net worth of both survivors and entrants (the liability side) while also directly increasing the value of bank assets. These two balance sheet margins typically have a negative net effect on the planner's marginal value of capital. By increasing bank assets, greater capital immediately increases the value of default, thus negatively affecting the marginal benefit of capital. Moreover, it increases investment and lowers consumption, generating an additional negative partial effect. The final negative effect is due to the negative impact of greater capital on the future asset price and asset return.

There are several positive effects. With commitment, greater capital and a greater asset price at t affect the $t-1$ expectation of the current asset return positively, increasing the bank value and relaxing the EC at $t-1$. Furthermore, the balance sheet, value of default, and consumption channels described in the previous paragraph have their future counterparts since the asset price function depends on both the beginning-of-the-period and end-of-the-period capital stock. Greater capital has a negative effect on the future asset price; therefore, the future balance sheet, value of default, and consumption effects have positive signs. Furthermore, greater capital increases the future marginal product of labor and the wage rate, having an additional positive effect.

The inefficiency of the CE allocation relative to the CEA and the fact that the planner chooses allocations that must be consistent with the forward-looking household Euler equation (1.2) and the definition of the forward-looking aggregate bank value (1.20) imply that the CEA is generally time inconsistent. There is a special case when the CEA is time consistent: it happens if the EC is either always binding or always slack at the CEA. In such a case, the implementability constraints completely determine the CEA. These constraints can be formulated recursively as a system of functional equations on the state space (D, K, z); therefore, the CEA must be time consistent. If this situation occurs with $\{\sigma_t^1\} = \{\widehat{\sigma}_t^1\}$, the CEA implementability constraints are necessary for the CE. They are, moreover, sufficient if the expected credit spread discounted with the pricing kernel $\eta_{t+1}\Lambda_{t,t+1}f_{t+1}$ for positive-valued f is nonnegative in the CEA. (This condition guarantees that the CE Lagrange multiplier $\bar{\lambda}_t$ is nonnegative.) The described situation does not arise quantitatively: the EC is only occasionally binding. Nevertheless, this result has an implication for computing macro-banking models similar to that in this paper. If we computed such a model ignoring the occasionally binding constraint, assuming that it is always binding, we would not be able to identify the externalities and would wrongly conclude that the CE allocation is efficient. Note that the issue here is not in the order of local approximation—the allocations would seem identical independently of the order—but in accounting for the occasionally binding constraint properly.

The final part of proposition 1.2 states that for a given distribution $\{\sigma_t^1\}$, we can generally find an alternative distribution $\{\widetilde{\sigma}_t^1\}$ which is at least weakly preferred to $\{\sigma_t^1\}$ as long as the CEA at the original distribution has contingencies in which the EC is binding. The alternative distribution increases the future bank value of survivors, which automatically increases the current bank value both at z^t and the preceding contingencies, relaxing the EC at those contingencies and expanding the planner's feasible set. Again, this argument relies on the nature of commitment: the planner relaxes the EC at t by promising more survivors-biased distribution at $t + 1$, bearing similarity with forward guidance for monetary policy. Note that ex post, the planner is indifferent between honoring such promises or not because $\{\sigma_t^1\}$ affects the planner's constraints only through the continuation value in the forward-looking bank value. This channel is thus *not* a source of time inconsistency.

1.3.3 Markov perfect equilibrium

Since the CEA is generally time inconsistent, a thorough and complete investigation of the constrained efficiency of the economy considered in this paper requires exploring the implications of the lack of commitment by the policymaker. To do so, we will study a Markov perfect equilibrium (MPE) of a non-cooperative game between sequential—"current" and "future"—social planners (Klein et al., 2008). We will focus on the concept of MPE due to its quantitative tractability, following Bianchi (2016) and Bianchi and Mendoza (2018), who applied this approach in the analysis of optimal macroprudential policy in small open economies. Other concepts of time-consistent policies exist, such as sustainable policies (Chari and Kehoe, 1990), and Markov policies are generally inferior to history-contingent sustainable policies. It is, however, harder to compute the latter policies in our environment.

Denote the future planner's value and policy functions as $\bar{V}^h$, $\bar{C}$, $\bar{K}'$, $\bar{L}$, $\bar{V}^1$, where all functions map $\mathcal{S} \to \mathbb{R}$. Since the current planner can affect the future bank value of survivors $\bar{V}^1$ only indirectly by affecting $S' = (D', K', z')$, we can use (1.20) to solve for V, removing it from the set of implementability conditions. The current planner's best response to the future planner's decisions is represented by

$$V^h(S) = \max_{(C,D',K',L)\in\mathcal{G}(S)} U(C,L) + \beta\mathbb{E}_z(\bar{V}^h(S')),$$

where $\mathcal{G} : \mathcal{S} \to \mathcal{P}(\mathbb{R}^4_+)$ is defined by the constraints

$$\begin{aligned}
0 &= \bar{N} + \sigma\eta_S\widetilde{N}_S + Q(K,K',\xi)(\omega K - K') + \beta\mathbb{E}_z\left(\frac{U_C(\bar{C}_{S'},\bar{L}_{S'})}{U_C(C,L)}\right)D',\\
0 &\leq \beta\mathbb{E}_z\left\{\eta_{S'}\frac{U_C(\bar{C}_{S'},\bar{L}_{S'})}{U_C(C,L)}[(1-\sigma)\widetilde{N}_{S'} + \bar{V}^1_{S'}]\right\} - \theta Q(K,K',\xi)K',\\
0 &= U_C(C,L)AF_L(\xi K,L) + U_L(C,L),\\
0 &= AF(\xi K,L) - C - I(K,K',\xi),
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{N}_S &\equiv [AF_K(\xi K,L) + Q(K,K',\xi)(1-\delta)]\xi K - D, \qquad \eta_S \equiv \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_S),\\
\widetilde{N}_{S'} &\equiv [A'F_K(\xi' K',\bar{L}_{S'}) + Q(K',\bar{K}'_{S'},\xi')(1-\delta)]\xi' K' - D', \qquad \eta_{S'} \equiv \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_{S'}).
\end{aligned}$$

In an MPE for a given distribution $\sigma^1 : \mathcal{S} \to [0,1)$, $V^h \equiv \bar{V}^h$ solves the Bellman equation, and policy functions of the current and future planners coincide. In particular, V^1 satisfies

$$V_S^1 = \sigma_S^1 \mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\widetilde{N}_{S'} + V_{S'}^1]\}.$$

Consistent with the notation used so far, let us denote the Lagrange multipliers on the planner's constraints as $\widetilde{\nu}$, $\widetilde{\lambda}$, λ^L, and λ^Y, respectively. Define $\nu_S \equiv \frac{\widetilde{\nu}_S}{U_{C,S}}$ and $\lambda_S \equiv \frac{\widetilde{\lambda}_S}{U_{C,S}}$. The next proposition parallels proposition 1.2 in the context of the MPE.

Proposition 1.3. *The CE (RCE) allocation is generally inefficient compared to the MCEA. Under the assumption of differentiability of the policy functions, the MCEA generalized Euler equations associated with D' and K'—corresponding to (1.16) and (1.17)—can be represented as*

$$\nu_S = R_S \mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\lambda_S + \sigma\nu_{S'}]\} \underbrace{-\frac{R_S \Xi_S^D}{U_{C,S}}}_{\textit{future policy}},$$

$$\theta\lambda_S + \nu_S = \mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\lambda_S + \sigma\nu_{S'}]R_{S'}^K\} + \Omega_S^K + \underbrace{\frac{\Xi_S^K}{Q_S U_{C,S}}}_{\textit{future policy}},$$

where for $X \in \{D, K\}$,

$$\Xi_S^X \equiv \beta\nu_S \mathbb{E}_z(\underbrace{U_{CC,S'}\bar{C}_{X,S'} + U_{CL,S'}\bar{L}_{X,S'}}_{\textit{SDF in deposit rate } (-*)})D'_S + \beta\lambda_S \mathbb{E}_z\Big(\eta_{S'}\Big[(\underbrace{U_{CC,S'}\bar{C}_{X,S'} + U_{CL,S'}\bar{L}_{X,S'}}_{\textit{SDF in aggregate bank value } (-*)})$$
$$\times [(1-\sigma)\widetilde{N}_{S'} + \bar{V}_{S'}^1] + U_{C,S'}\{\underbrace{(1-\sigma)[A'F_{KL,S'}\bar{L}_{X,S'} + Q_{2,S'}\bar{K}'_{X,S'}(1-\delta)]\xi' K'_S + \bar{V}_{X,S'}^1}_{\textit{future asset return and bank value of survivors}}\}\Big]\Big)$$

is the combined marginal effect of X' on the current planner's Lagrangian through the policy func-

tions of the future planner $\bar{C}$, $\bar{L}$, $\bar{K}'$, and $\bar{V}^1$. The capital wedge satisfies

$$
\begin{aligned}
Q_S\Omega_S^K \equiv\ & \underbrace{\nu_S Q_{2,S}\{[\sigma\eta_S(1-\delta)\xi+\omega]K-K'_S\}}_{\textit{balance sheet } (-^*)} \ \underbrace{-\lambda_S\theta Q_{2,S}K'_S}_{\textit{value of default } (-)} \ \underbrace{-\frac{\lambda_S^Y}{U_{C,S}}I_{2,S}}_{\textit{consumption } (-^*)} \\
& + \underbrace{\mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\lambda_S+\sigma\nu_{S'}][A'F_{KK,S'}\xi'+Q_{1,S'}(1-\delta)]\xi' K'_S\}}_{\textit{future asset return } (-)} \\
& + \mathbb{E}_z\Bigg(\Lambda_{S,S'}\Bigg\{\underbrace{\nu_{S'}[\omega(Q_{1,S'}K'_S+Q_{S'})-Q_{1,S'}K'_{S'}]}_{\textit{future balance sheet } (+^*)} \ \underbrace{-\lambda_{S'}\theta Q_{1,S'}K'_{S'}}_{\textit{future value of default } (+)} \ + \underbrace{\lambda_{S'}^L A'F_{KL,S'}\xi'}_{\textit{future wage } (+^*)} \\
& + \underbrace{\frac{\lambda_{S'}^Y}{U_{C,S'}}(A'F_{K,S'}\xi'-I_{1,S'})}_{\textit{future consumption } (+^*)}\Bigg\}\Bigg).
\end{aligned}
$$

First, as in the case of commitment, we must be aware that the planner's (transformed) Lagrange multipliers are generally different from those in the CE. Moreover, the direct quantity effects in the planner's Euler equations (right-hand sides without the wedges) are symbolically different from those in (1.16) and (1.17): the planner's $(1-\sigma)\lambda_S+\sigma\nu_{S'}$ corresponds to the individual banker's $(1+\lambda_S)(1-\sigma+\sigma\nu_{S'})$, which both reflect the direct effects on the future net worth and the (relevant) continuation value. In the individual banker's problem, the bank value appears in the EC and the objective function—hence, the multiplication by $1+\lambda_S$. Moreover, the shadow value of net worth ν is linked to the derivative of the banker's value function v. In the planner's problem, the objective is the household welfare, so the bank value appears once in the EC (multiplication by λ_S only). Moreover, the shadow value of net worth is linked to the derivatives of the household value function V^h, not being related to the EC—therefore, there is no multiplication by $1+\lambda_S$.

Now consider the wedges. Without commitment, the current planner must consider how its current decisions affect the future endogenous state and the decisions of the future planner, which introduces the Ξ_S^D and Ξ_S^K terms reflecting those effects. These objects have a symmetric structure, capturing three main transmission mechanisms. First, D' and K' affect the future consumption $\bar{C}$ and labor $\bar{L}$ decisions and thus the future marginal utility of consumption and the SDF, which affects the deposit rate according to the household Euler equation (1.2). Second, there is a similar effect on the SDF implicit in the forward-looking bank value (1.20). Third, D' and K' affect the future net worth at exit and the future bank value of survivors $\bar{V}^1$ conditional on survival, where the former is generated by the impact on both the future marginal product of capital through $\bar{L}$ and the future asset price through $\bar{K}$. Intuitively, we can expect that the derivatives of the policy functions with respect to K are generally nonnegative since greater K is associated with both greater output and a greater bank net worth. On the contrary, a greater bank debt D has a negative effect on net worth, investment, and the household value function, so we can expect that the derivatives of the policy functions are generally nonpositive. The combined effects and the signs of Ξ_S^D and Ξ_S^K remain ambiguous.[4]

The additional capital wedge Ω_S^K corresponds to a similar term arising under commitment. Contrary to the latter, the time-consistent planner cannot affect the $t-1$ expectations of the asset return and the bank value distribution at t. Likewise, without commitment, the planner cannot affect the future distribution except for the indirect impact through the future endogenous states. For this reason, we did not make the distribution explicit in the continuation value of survivors $\bar{V}^1$. The remaining effects—the negative balance sheet, value of default, and consumption channels, the corresponding positive future effects, and the negative impact on the future asset return—are identical to the case of commitment. A quantitative exploration is generally required to assess which effects dominate. Indeed, as we will see, the combined effect is typically not uniformly positive or negative but state-contingent, allowing for both excessive and insufficient borrowing and lending in the CE.

Unlike in the case of commitment, we do not have a formal statement on the welfare ranking of Markov perfect outcomes corresponding to different σ^1. A shift in σ^1 directly affects the fixed point as we iterate on $\bar{V}^1$, so the welfare effects may have different signs in different regions of the state space. We can, however, state with certainty that a uniform positive shift in σ^1 must increase welfare in the steady state in which the EC is binding.

1.3.4 Implementation with taxes, transfers, and capital requirements

The presence of two broad sources of inefficiencies—various pecuniary externalities and a potentially suboptimal bank value distribution—generally requires two types of policy instruments to implement the CEA (MCEA) in a regulated CE. A given distribution σ^1 can naturally be achieved with

[4] Our quantitative approach is to find a fixed point in the Bellman equation and the policy functions directly instead of solving the KKT conditions, so we will not be assuming that the policy functions are differentiable.

entrants/survivors-specific transfers within the banking system. The wedges in the Euler equations can be addressed with proportional taxes on bank deposits and assets or, under some assumptions, with state-contingent capital requirements. The next proposition formalizes the alternative ways of implementing the CEA (MCEA) in a regulated CE. We will use the sequential (CEA) notation where $\{x_t\}$ denotes a sequence of functions $x_t : Z^t \to \mathbb{R}$, while the implicit recursive (MCEA) analog is a single function $x : \mathcal{S} \to \mathbb{R}$.

Proposition 1.4. *Consider a regulated CE that differs from those in definitions 1.1 and 1.2 in that the banker $i \in [0, f]$ now has the budget constraint*

$$(1 + \tau_t^K) Q_t s_{i,t} \leq n_{i,t} + (1 - \tau_t^D) \frac{d_{i,t+1}}{R_t} + \tau_{i,t},$$

faces an additional regulatory constraint

$$n_{i,t} \geq \kappa_t Q_t s_{i,t},$$

where $\kappa_t \leq 1$, and there is a budget constraint $\tau_t^D \frac{D_{t+1}}{R_t} + \tau_t^K Q_t K_{t+1} = \int_0^f \tau_{i,t} \, \mathrm{d}i$ of the macroprudential authority.

The CEA (MCEA) can be implemented in a regulated CE above as follows. If $\{\sigma_t^1\} = \{\widehat{\sigma}_t^1\}$, we can set $\tau_{i,t} = 0$ for all (i, t, z^t). Otherwise, $\{\tau_{i,t}\}$ can be set to achieve the targeted distribution $\{\sigma_t^1\}$. The following instruments can be used to account for the wedges.

1. *If $\int_0^f \tau_{i,t} \, \mathrm{d}i = 0$, we can use $\{\tau_t^D, \tau_t^K\}$ and set $\kappa_t = -\infty$ for all (t, z^t).*

2. *If $\int_0^f \tau_{i,t} \, \mathrm{d}i \neq 0$, we can use $\{\tau_t^D\}$ ($\{\tau_t^K\}$) and set $\tau_t^K = 0$ ($\tau_t^D = 0$) and $\kappa_t = -\infty$ for all (t, z^t).*

3. *Independently of $\{\tau_{i,t}\}$, if the CEA (MCEA) satisfies $\mathbb{E}_t[\eta_{t+1} \Lambda_{t,t+1} f_{t+1} (R_{t+1}^K - R_t)] \geq 0$ for all (t, z^t) and all $\{f_t\}_{t \geq 1}$ with $f_t : Z^t \to \mathbb{R}_{++}$, then we can use $\{\kappa_t\}$ and set $\tau_t^D = \tau_t^K = 0$ for all (t, z^t). Without loss of generality, we can set $\kappa_t \equiv \frac{N_t}{Q_t K_{t+1}}$, where the right-hand side is evaluated at the CEA (MCEA), in which case the regulatory constraint is always binding in the regulated CE.*

The CEA (MCEA) and the policy that implements it constitute a Ramsey (Markov perfect) equilibrium.

As explained in the proof of proposition 1.4, we construct all the policies using the primal approach. For example, in all variants of the implementation with taxes, the optimal tax rate τ_t^D satisfies

$$\tau_t^D = 1 - \frac{\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})]R_t}{\bar{\nu}_t},$$

conditional on the optimal allocations and the regulated CE multiplier $\bar{\nu}_t$, which is a function of the optimal allocations. If $\tau_{i,t} = 0$ for all (i, t, z^t), the individual banker's value function is still linear in the individual net worth, so we can immediately solve for $\bar{\nu}_t = \frac{V_t}{N_t}$, where the right-hand side is evaluated at the CEA (MCEA). In general, $\{\bar{\nu}_t\}$ solves a fixed-point equation, which differs based on whether we allow for aggregate lump-sum transfers ($\int_0^f \tau_{i,t}\,\mathrm{d}i \neq 0$). In the latter case, we need only one proportional tax.

Instead of linear taxes, we can also implement the optimal allocation by introducing a regulatory capital requirement constraint. Capital requirements alone are sufficient to account for the wedges if and only if a measure of a discounted credit spread stays nonnegative in the CEA (MCEA). A sufficient condition is that $\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}f_{t+1}(R_{t+1}^K - R_t)] \geq 0$ for all positive-valued f_{t+1}. A necessary and sufficient condition is that it holds for $f_{t+1} = 1 - \sigma + \sigma(\bar{\nu}_{t+1} + \bar{\xi}_{t+1})$, where $\bar{\xi}_t$ is a transformation of the Lagrange multiplier on the regulatory constraint. A difficulty is that $\{\bar{\nu}_t, \bar{\xi}_t\}$ solve a system of two stochastic difference equations (the banker's deposit and asset Euler equations) conditional on the optimal allocation. Quantitatively, the required assumption is not always valid: the planner can optimally choose to have a negative discounted credit spread in some contingencies. In this case, capital requirements alone fail to be effective, although they would still be effective if augmented, for example, with a linear deposit subsidy.

Define $\bar{N}_t^1 \equiv (\bar{\nu}_t + \bar{\xi}_t)[\sigma_t^1(N_t + T_t^b) - \eta_t\sigma\widetilde{N}_t]$, where $\bar{\xi}_t$ is a transformation of the Lagrange multiplier on the regulatory constraint and $T_t^b \equiv \int_0^f \tau_{i,t}\,\mathrm{d}i$ is the aggregate lump-sum transfer. In the case of the implementation with taxes, the regulatory constraint is irrelevant, so $\bar{\xi}_t = 0$. If we do not allow the aggregate lump-sum transfer, $T_t^b = 0$. Note that $\bar{\nu}_t + \bar{\xi}_t$ is the total shadow value of wealth for bankers, and $\sigma_t^1(N_t + T_t^b) - \eta_t\sigma\widetilde{N}_t$ is survivors' targeted net worth gain from more survivors-biased bank value distribution. As shown in the proof of proposition 1.4, the aggregate transfer to survivors $\tau_t^1 \equiv \int_0^{\sigma f} \tau_{i,t}\, \mathrm{d}i$ can be expressed as

$$\tau_t^1 = \frac{1}{\bar{\nu}_t}\left\{\bar{N}_t^1 + (\sigma_t^1 - \sigma)\mathbb{E}_t\left[\eta_{t+1}\sum_{i=0}^{\infty}\left(\prod_{j=0}^{i-1}\sigma_{t+1+j}^1\eta_{t+2+j}\right)^{\mathbf{1}_{\mathbb{N}}(i)}\Lambda_{t,t+1+i}\bar{N}_{t+1+i}^1\right]\right\}.$$

If $\sigma_t^1 = \sigma$, the transfer is simply proportional to the targeted gain in net worth. Otherwise, there is an additional dynamic component—an expected discounted sum of future net worth gains that correspond to the targeted distribution $\{\sigma_t^1\}$.

The final part of proposition 1.4 is about the equivalence between the Ramsey problem conditional on the corresponding set of policy instruments and the CEA (MCEA) planning problem. Hence, each policy from proposition 1.4 is Ramsey optimal.[5] This equivalence is a consequence of applying the primal approach to construct a policy that implements the CEA (MCEA).

1.4 Quantitative results

This section describes the model calibration and conducts a multifaceted comparison of the CE, MCEA, and CEA allocations. We will investigate the efficiency of borrowing and lending by the banking system, explore the properties of optimal policies, analyze welfare gains, compare the economic dynamics around financial crises, and study the implications of alternative bailout policies.

1.4.1 Calibration and computation

I assume separable preferences for households: $U(C, L) = \lim_{x\to\gamma}\frac{C^{1-x}-1}{1-x} - \chi\frac{L^{1+\phi}}{1+\phi}$ with $(\gamma, \phi, \chi) \in \mathbb{R}_+^3$. The final and capital good production technologies are $F(\xi K, L) = (\xi K)^\alpha L^{1-\alpha}$ with $\alpha \in (0, 1)$, and $\Phi(x) = \zeta + \kappa_1 x^\psi$ with $\zeta \in \mathbb{R}$, $\kappa_1 > 0$, and $\psi \in (0, 1]$. The logs of exogenous stochastic processes $\{A_t\}$ and $\{\xi_t\}$ are AR(1) with autocorrelations (ρ_a, ρ_ξ) and standard deviations (σ_a, σ_ξ).

Table 1.1 reports the parameter values that are mostly set to reflect long-run facts about the US economy in 1990–2019. The Cobb—Douglas elasticity α targets the average labor share in

[5] Traditionally, the term “Ramsey” applies to a sequential problem where the planner chooses a policy plan at $t = 0$. In the context of an MPE, by “Ramsey,” we mean a planning problem similar to the MCEA problem; that is, a planner without commitment sets the policy optimally, taking into account the impact on the decisions of the future planner.

Table 1.1. Parameter values

Parameter	Value	Target
Preferences and technology		
α	0.404	labor share $\approx$ 59.6%
β	0.995	annualized real interest rate = 2%
γ	1	log preferences from consumption
δ	0.02	annual depreciation rate $\approx$ 7.6%
ζ	-0.007	$\frac{I}{K} = \delta$
κ_1	0.499	$Q = 1$
ϕ	0.625	microfounded aggregate Frisch elasticity = 1.6
χ	0.86	$L = 1$
ψ	0.75	panel data estimates in the literature
Banking		
$\bar{N} = 0$	0	linear endowment rule
σ	0.976	bank exit probability $\approx$ 0.091
θ	0.216	$N/(QK) = 0.125$, annualized credit spread = 0.5%
ω	0.001	
Exogenous stochastic processes		
ρ_a	0.935	$\text{corr}(\widehat{Y}_t, \widehat{Y}_{t-1}) \approx 0.886$, $\text{corr}(\widehat{I}_t, \widehat{I}_{t-1}) \approx 0.894$, $\text{sd}(\widehat{Y}_t) \approx 0.013$, $\text{sd}(\widehat{I}_t) \approx 0.045$
ρ_ξ	0.956	
σ_a	0.006	
σ_ξ	0.002	

Note. $\widehat{X}_t$ denotes the cyclical component of $\ln(X_t)$ extracted using the HP filter with $\lambda = 1600$.

the nonfarm business sector based on the data from the US Bureau of Labor Statistics. The discount factor β corresponds to the annualized real interest rate of 2%. The risk-aversion γ is set to unity, implying log preferences from consumption, as common in the literature. The depreciation rate δ proxies the average depreciation rate of the current-cost net stock of private fixed assets and consumer durables in the Bureau of Economic Analysis data. The capital production technology parameters (ζ, κ_1) are set to have $\frac{I}{K} = \delta$ and normalize $Q = 1$ in the deterministic steady state, while ψ is set as in Gertler et al. (2020a) to match panel data estimates. The labor disutility elasticity ϕ—an inverse of the Frisch elasticity of labor supply— targets the average of the microfounded estimates of the aggregate Frisch elasticity for males (Erosa et al., 2016) and females (Attanasio et al., 2018). The labor disutility scale χ corresponds to a normalization $L = 1$ in the steady state.

It is computationally convenient to set $\bar{N} = 0$ so that the aggregate endowment of entrants is linear in the assets of exiting bankers. I set the survival probability σ based on the average establishment exit rate in finance, insurance, and real estate according to the Business Dynamics Statistics data. The remaining banking parameters (θ, ω) target the average capital ratio of 12.5%—consistent with the evidence in Begenau et al. (2020) that for most banks, regulatory constraints are not binding—together with the annualized credit spread of 0.5% so that the EC binds in the CE less than half of the time.

The AR(1) parameters—autocorrelations (ρ_a, ρ_ξ) and standard deviations (σ_a, σ_ξ)—target the autocorrelations and standard deviations of output and investment, using the National Income and Product Accounts data. Each variable is logged and detrended using the HP filter with $\lambda = 1600$, a standard value for quarterly data.

To compute the CE and MCEA, I use global projection methods (Judd, 1998) to fully address the nonlinearities due to the occasionally binding EC and limited liability. Specifically, I approximate the CE and MCEA unknown functions with linear 2D splines for each $z \in \widehat{\mathcal{Z}} \subset \mathcal{Z}$. (Accordingly, I approximate the exogenous stochastic process $\{A_t, \xi_t\}$ by a finite-state Markov chain $z \mapsto z'$.) In the case of the CEA, I employ both the global projection method—linear 4D splines—and the local piecewise linear perturbation method (Guerrieri and Iacoviello, 2015) that respects occasionally binding constraints but not precautionary savings. The latter method serves as the baseline, but I verify some results with the global method on a coarse grid. Since Lagrange multipliers γ_{t-1} and ν_{t-1} must be treated as state variables, the complexity of the Ramsey problem combined with the course of dimensionality makes fully global approximation challenging.

Instead of the natural endogenous state (D, K), I work with a rotated state space based on $(\log(D), \log(K))$. This way, we can account for the strong positive correlation between $\log(D)$ and $\log(K)$, which is illustrated in figure 1.1 in the CE case.

1.4.2 Bank solvency and EC regimes

The model has two main nonlinearities. First, banks can become insolvent, in which case they must default under limited liability. Second, the EC is occasionally binding. When the constraint binds, banks are indifferent between continuing the business and running away with a fraction of assets. As illustrated in figure 1.2, these two binary events divide the underlying endogenous state space into three regions: banks are solvent and unconstrained (highlighted in yellow), solvent but

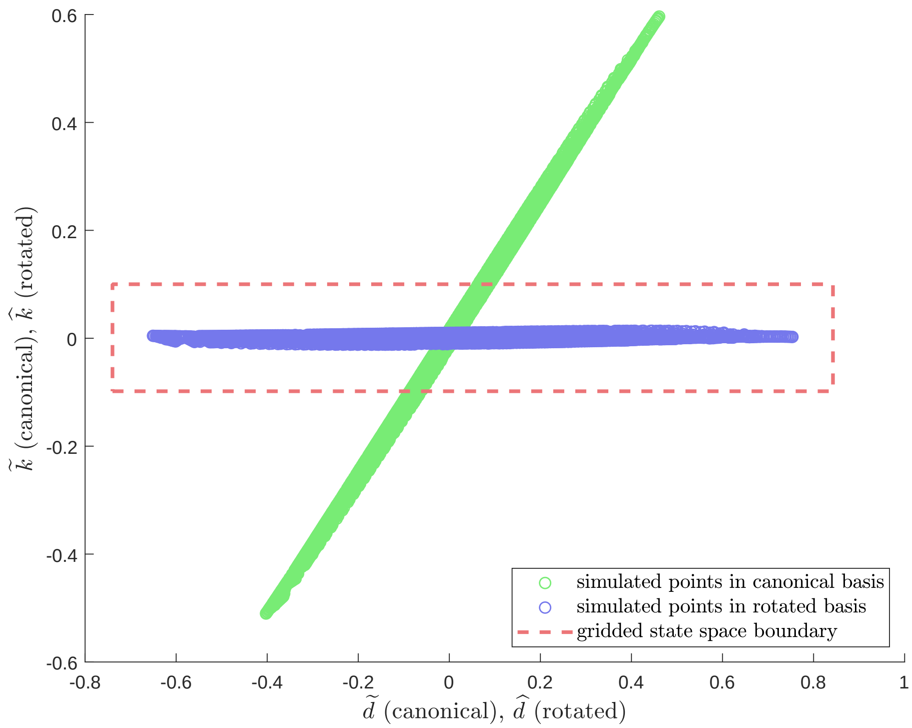

Figure 1.1. Endogenous state space, CE ergodic distribution. For $X \in \{D, K\}$, $\widetilde{x} \equiv \log(X) - \widehat{\mathbb{E}}(\log(X))$, and $(\widehat{d}, \widehat{k})$ are obtained by rotating $(\widetilde{d}, \widetilde{k})$ clockwise at the angle $\arctan\left(\frac{\widehat{\operatorname{cov}}(\widetilde{d},\widetilde{k})}{\widehat{\operatorname{var}}(\widetilde{d})}\right)$.

constrained (light green), and insolvent and constrained (dark green). Banks cannot be insolvent and unconstrained in the CE simultaneously. If the survived banks are insolvent, their bank value is zero, so the EC must be binding for them. Since the Lagrange multipliers depend only on the aggregate state, the entering banks must also be constrained.

According to figure 1.2, banks are solvent and unconstrained when the initial capital stock K is sufficiently large compared to bank debt D. There generally exist thresholds $\bar{K}(D, z)$ and $\widehat{K}(D, z)$, such that banks are solvent when $K > \bar{K}(D, z)$ and are, moreover, unconstrained when $K > \widehat{K}(D, z) \geq \bar{K}(D, z)$. Based on the figure, we can conjecture that both $\bar{K}$ and $\widehat{K}$ are decreasing in z in the sense that $\bar{K}(D, z_2) \leq \bar{K}(D, z_1)$ when $A_2 > A_1$ and $\xi_2 > \xi_1$. The thresholds are also generally increasing in D. Analytic characterization of $\bar{K}$ and $\widehat{K}$ does not seem possible, but the conjectured properties are intuitive.

Although the area of the insolvency region might seem significant, the model does not typically visit those states. According to figure 1.1, the ergodic set is a thin ellipse inside the gridded state space (the dashed parallelogram in figure 1.2). Insolvency is more likely in the worst exogenous

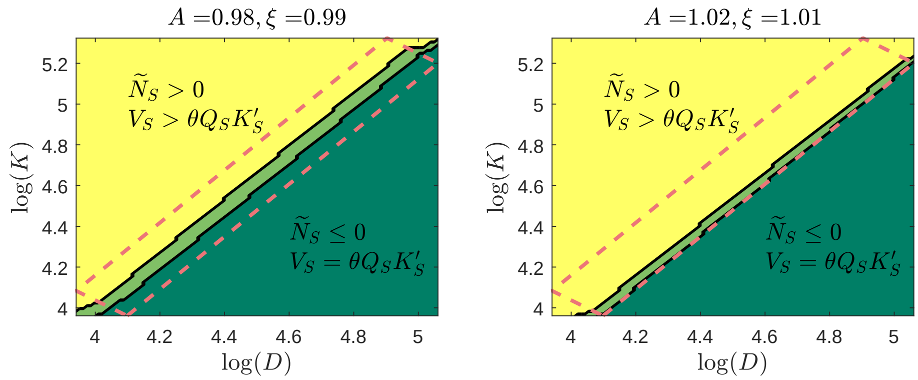

Figure 1.2. Bank solvency and EC regimes in the worst and best exogenous states in the CE. In the yellow region, banks are solvent, and the EC is slack. In the light green region, banks are solvent, but the EC is binding. In the dark green region, banks are insolvent, and the EC is binding. The dashed parallelogram (not a rectangle due to scaling) is the boundary of the endogenous state space represented in the canonical basis.

state, but it does not typically occur even in that case. On the other hand, the model stays in the binding EC regime approximately 40% of the time in the CE.

Figure 1.2 confirms the potential welfare benefits from preemptive bailouts. By keeping banks away from the solvent-but-constrained buffer zone, the policymaker escapes the potentially harmful effects of being in the constrained regime and decreases the probability of ending up in the insolvency region, at which point the banking system would collapse.

Figure 1.3 further explores how the magnitude of the distance between the aggregate bank value and the value of default $V_S - \theta Q_S K'_S$ varies in the state space. We now focus on the gridded

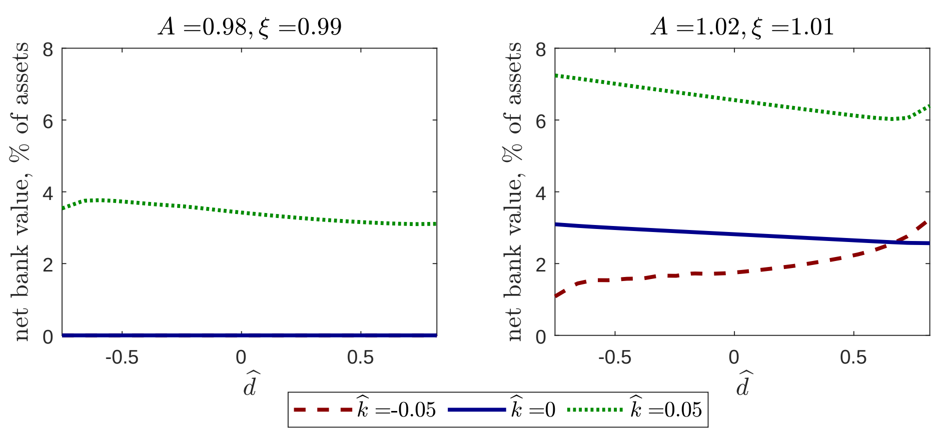

Figure 1.3. Net bank value in the CE. Slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid. The y-axis is $V_S - \theta Q_S K'_S$ in % of $Q_S K'_S$.

state space where the model is solved. Figure 1.3 displays the variation along the $\widehat{d}$ dimension, that is, moving from the southwest to northeast inside the dashed parallelogram in figure 1.2 at the quartiles of the $\widehat{k}$ grid. In the lowest exogenous state, the EC is typically binding at higher leverage ratios, e.g., at or below the median of $\widehat{k}$. The constraint is slack when banks are more capitalized (higher $\widehat{k}$), and the slack in proportion to bank assets slightly decreases as the balance sheet expands (larger $\widehat{d}$). In the highest exogenous state, the expected asset returns are greater and financial constraints are mostly slack, especially at the higher capital ratios. As with the lowest state, the relative slack generally decreases as the balance sheet expands at higher capital ratios; however, there is an opposite relationship when banks are more leveraged. These regularities indicate that a way to improve over the CE is to relax the binding ECs when exogenous conditions are worse.

1.4.3 Financial crises in the unregulated economy

When the banker's EC binds, a banker is indifferent between continuing to run the banking business and defaulting on liabilities and running away with a fraction of assets. Our convention is that bankers continue to operate at the point of indifference. The instances where the aggregate EC is about to switch from being slack to binding and the ensuing spells in the binding regime with the associated deleveraging can naturally correspond to the build-up of systemic risk and financial crises. The risk is systemic because our banks make symmetric decisions: when the EC binds for one bank, it binds for all. This subsection explores the economic dynamics around such episodes.

Define a financial crisis that starts at t as an event that satisfies two conditions on the behavior of the aggregate EC: it is slack for at least five years before the crisis ($[t-20, t-1]$) and then binding for at least one year ($[t, t+3]$). Figure 1.4 illustrates the typical dynamics around such crises. The figure is obtained by simulating the CE for 1,000,000 periods (quarters), selecting crisis episodes as defined above, and averaging the simulated paths. There are 8,106 such crises, which corresponds to approximately 3.2 financial crises per century, consistent with the findings in the related literature (Mendoza, 2010).

Financial crises have a boom-bust pattern. Ahead of a crisis, output, consumption, and investment are increasing, and the balance sheet of the banking system is expanding. A leading indicator of the crisis is the gradually falling forward-looking asset price. The aggregate EC binds when a bad exogenous state occurs, typically due to a decrease in capital quality. The asset price and the

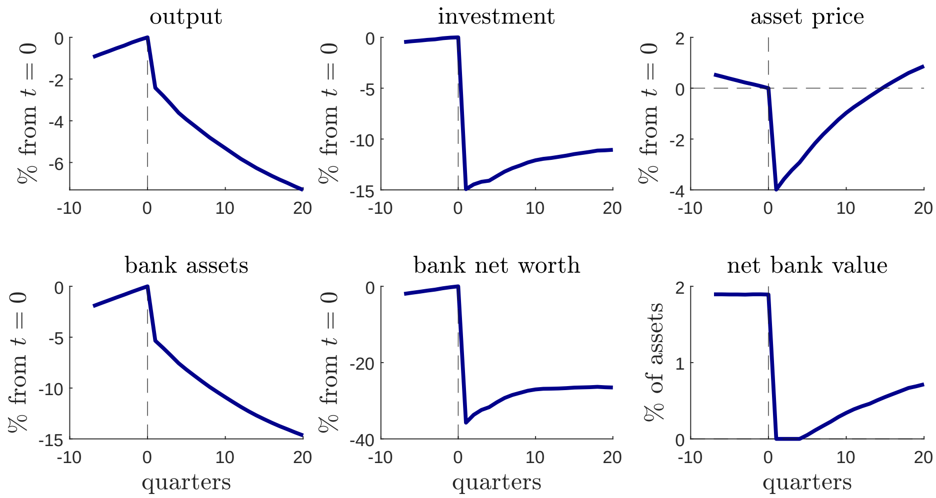

Figure 1.4. Financial crises in the CE. Averages over a 1,000,000-period simulation.

realized return on bank assets drop, which triggers a sharp fall in bank net worth—the bust starts. As banks deleverage, balance sheets shrink, firms cut investment, and an economic recession starts. There is a slight rise in consumption on impact due to the fall in the deposit rate and the increase in labor supply, but the effect is short-term, as consumption starts to fall next period. Meanwhile, the forward-looking asset price starts to recover, and so does bank net worth and the aggregate investment. As bank deleveraging continues, the EC switches to being slack again, and the bank value slowly begins to recover. The fall in output gradually slows down, but the recession and financial deleveraging persist.

1.4.4 Markov perfect equilibrium

This subsection explores the optimal time-consistent allocation. The CE and MCEA have the same underlying state space, so we can directly compare the policy functions. Specifically, we will focus on differences in bank deposits and loans, welfare gains, and optimal policies. We will also compare the economic dynamics around events identified as financial crises in the CE.

We will focus on the MCEA computed conditional on the distribution $\sigma^1 : \mathcal{S} \to \mathbb{R}$, the optimal linear transformation of the CE distribution. Define the CE distribution $\bar{\sigma}^1_S \equiv \eta_S \frac{\sigma \widetilde{N}_S}{N_S}$ for all $S \in \mathcal{S}$, where the right-hand side is evaluated at the CE allocation. Then $\sigma^1 = \lambda^* \bar{\sigma}^1$, where

$$\lambda^* = \underset{\lambda \in [0,\bar{\lambda})}{\arg\sup}\, \mathbb{E}(V^h_S \mid \sigma^1 = \lambda \bar{\sigma}^1), \qquad \bar{\lambda} = \sup\{\lambda \mid \sup_{S \in \mathcal{S}} (\lambda \bar{\sigma}^1_S) \le 1\}.$$

Numerically, the upper bound for λ is $\bar{\lambda} \approx 1.004$, while $\lambda^* \approx 0.995$—that is, the CE bank value distribution must be scaled down to maximize the unconditional welfare over the MCEA ergodic distribution.

The planner finds it optimal to scale down the distribution because it complements the planner's efforts to address the pecuniary externalities in the CE. In turn, correcting the pecuniary externalities helps to relax the EC. When the constraint is mostly slack, it might be inferior to increase λ further up due to the general nonconcavity of the value function. Later we will explore the implications of different values of λ.

Bank borrowing and lending

Proposition 1.3 identified multiple channels through which the time-consistent planner's marginal cost of bank borrowing and marginal benefit of bank lending differ from those in the CE. The channels have opposing signs, and the net effect is theoretically ambiguous. Let us now resolve the ambiguity numerically.

Figure 1.5 displays the histograms of bank deposits $\frac{D_{t+1}}{R_t}$ and bank loans $Q_t K_{t+1}$ from a 1,000,000-period simulation of the CE and MCEA with the same sequence of exogenous state variables $\{A_t, \xi_t\}$ and initial conditions (D_0, K_0). The histograms demonstrate that the CE allo-

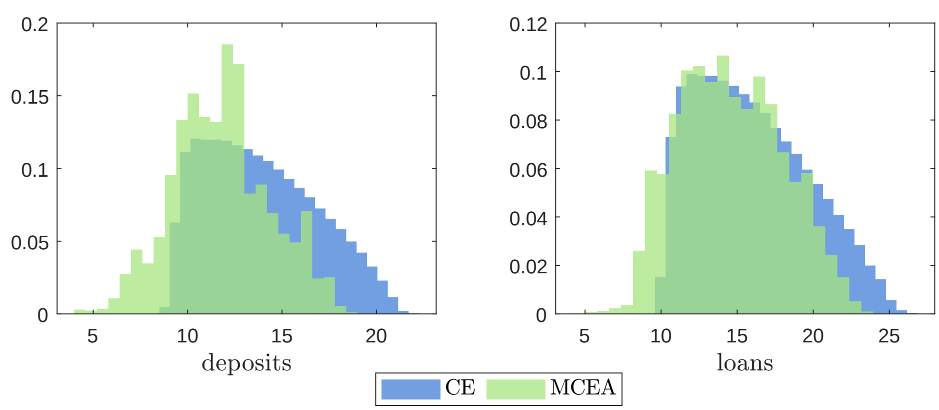

Figure 1.5. Bank borrowing and lending in the CE and MCEA. Histograms based on the 1,000,000-period simulation with the same sequence of exogenous shocks and initial conditions. Variables are normalized by the average CE output; the y-axis has the estimated probability density function (pdf) normalization.

cation has both overborrowing and overlending by the banking system compared to the MCEA. The efficient amount of borrowing and lending is characterized by a lower mean, variance, and skewness. Excessive borrowing and lending in the CE are mainly reflected in the longer right tail of the distributions. Specifically, the constraint is in the binding regime at about 40% of the time in the CE but less than 5% in the MCEA. The MCEA planner internalizes how asset prices affect the bank value and the value of default and optimally chooses a buffer to insure away from the constrained regime, so the distribution of deposits and loans is less skewed.

Figure 1.6 illustrates the % difference in the quantity of deposits $\frac{D'_S}{R_S}$ in the MCEA relative to the CE along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid. For convenience, the bottom part of the figure contains histograms of $\{\widehat{d}_t\}$ conditional on the corresponding exogenous states. Remember that an increase in $\widehat{d}$ corresponds to an increase in both bank debt D and capital stock

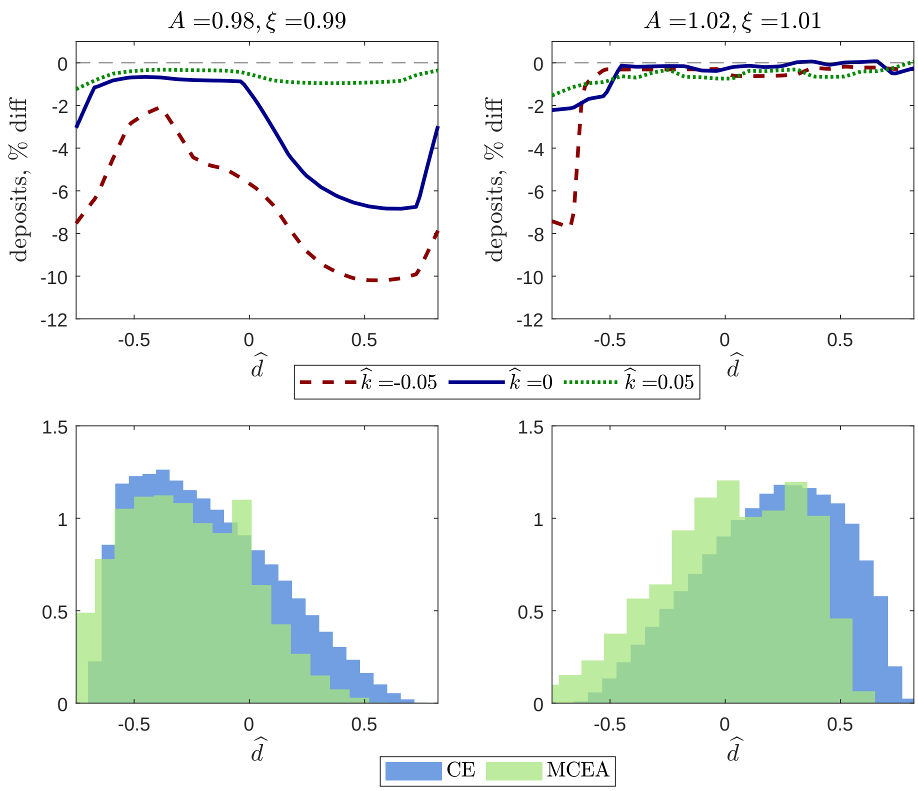

Figure 1.6. Overborrowing. In the top row, the % difference in $\frac{D'_S}{R_S}$ between MCEA and CE based on the slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid. In the bottom row, histograms of $\{\widehat{d}_t\}$ conditional on the corresponding exogenous state (1,000,000-period simulation, pdf normalization on the y-axis).

K linked to bank assets. An increase in $\widehat{k}$ corresponds to a *decrease* in D and an increase in K, which approximately corresponds to a decrease in the leverage ratio (an increase in the bank capital ratio).

The majority of the state space is characterized by overborrowing by the banking system in the CE relative to the MCEA. The extent of overborrowing is not uniform, and there are indeed some states where we observe slight *underborrowing* instead. Overborrowing is smaller when banks are well-capitalized. Overborrowing is generally severe when banks are highly leveraged at the low quantities of debt. In figure 1.3, we see that in such states, the EC is either binding or close to being so in the CE. On the contrary, as illustrated in figure 1.7, the constraint is slack in the MCEA (in the lowest exogenous state—only slightly).

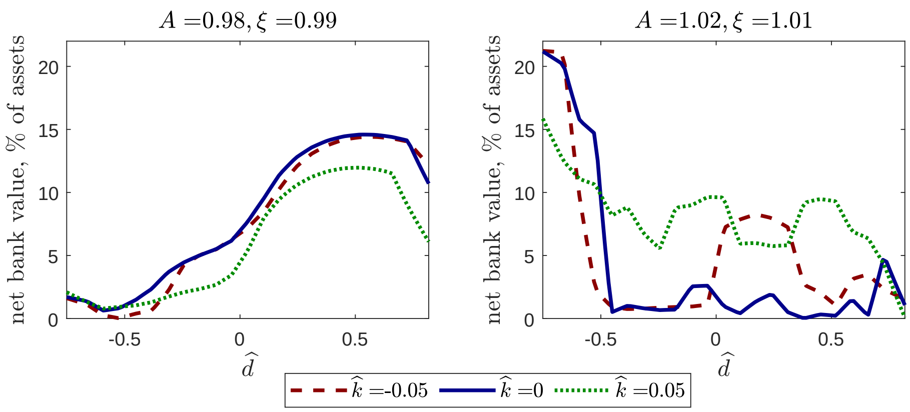

Figure 1.7. Net bank value in the MCEA. Slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid. The y-axis is $V_S - \theta Q_S K'_S$ in % of $Q_S K'_S$.

Conditional on the lowest exogenous state, the magnitude of overborrowing has an inverted S-like shape. The global minimum of overborrowing is around the first quartile of $\widehat{d}$—in that region, the MCEA EC is close to being binding, which indicates that the planner is constrained in their ability to improve over the CE allocation. As bank debt increases, the financial constraint becomes slack in the MCEA, and the magnitude of both the net bank value and overborrowing is at their maximum near the third quartile of the $\widehat{d}$ grid. This regularity is particularly striking when $\widehat{k}$ is lower and banks are more leveraged. Indeed, in the latter states, the EC stays binding in the CE, while the time-consistent planner moves away from the binding region quite significantly, realizing the harmful effects of entering the debt-deflation spiral at larger debt values. When the balance sheet size is closer to the upper bound of the grid, the relative net bank value slightly decreases in the MCEA, so the extent of overborrowing in the CE also decreases. Looking at the bottom row of figure 1.6, we must observe that the region of the state space between the first and second quartiles of the $\widehat{d}$ grid is more likely to occur since it is problematic to expand the balance sheet significantly conditional on the lowest exogenous state (when asset prices are low).

Conditional on the highest exogenous state, the situation is quite different. We still have significant overborrowing when banks are more leveraged at low values of debt, as the relative slack of the planner's EC is large there, but as the balance sheet expands, the magnitude of overborrowing is close to zero, and there are some states where we observe slight underborrowing. The reason is that the EC is already slack in the CE, so externalities are less pronounced—mathematically, many terms in the wedges in proposition 1.3 vanish. In this case, the planner is not building substantial buffers to insure away from the binding regime, as the consequences of the latter are less severe when exogenous conditions are good.

Figure 1.8 parallels the top row of figure 1.6, illustrating the differences in bank lending. The

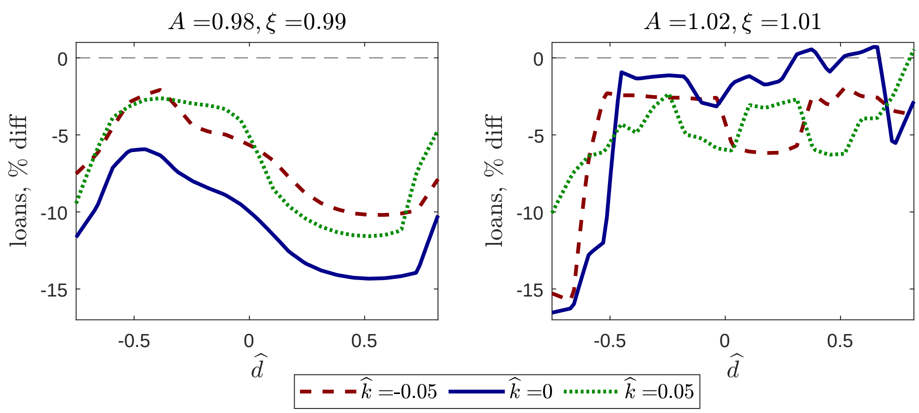

Figure 1.8. Overlending. Slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid. % difference in $Q_S K_S'$ (MCEA relative to CE) on the y-axis.

patterns are qualitatively very similar to those in the case of deposits, which is not surprising due to the bank balance sheet constraint. The magnitude of overlending is generally more significant than that of overborrowing since pecuniary externalities directly impact bank asset allocation, while the effect on deposits is indirect through the bank balance sheet. Related to the latter, the magnitude of overlending tracks more closely the magnitude of the planner's net bank value: overlending in the CE is greater in those states where the planner's EC is slacker.

Optimal policies

Let us now turn attention to optimal policies that implement the MCEA. Figure 1.9 shows the policy functions for the optimal deposit tax rebated lump sum in the aggregate and the corresponding transfer to survived banks τ_S^1 that supports the optimal distribution σ^1. (We will refer to this policy as optimal affine taxation.) The policy function for the tax parallels the policy functions for

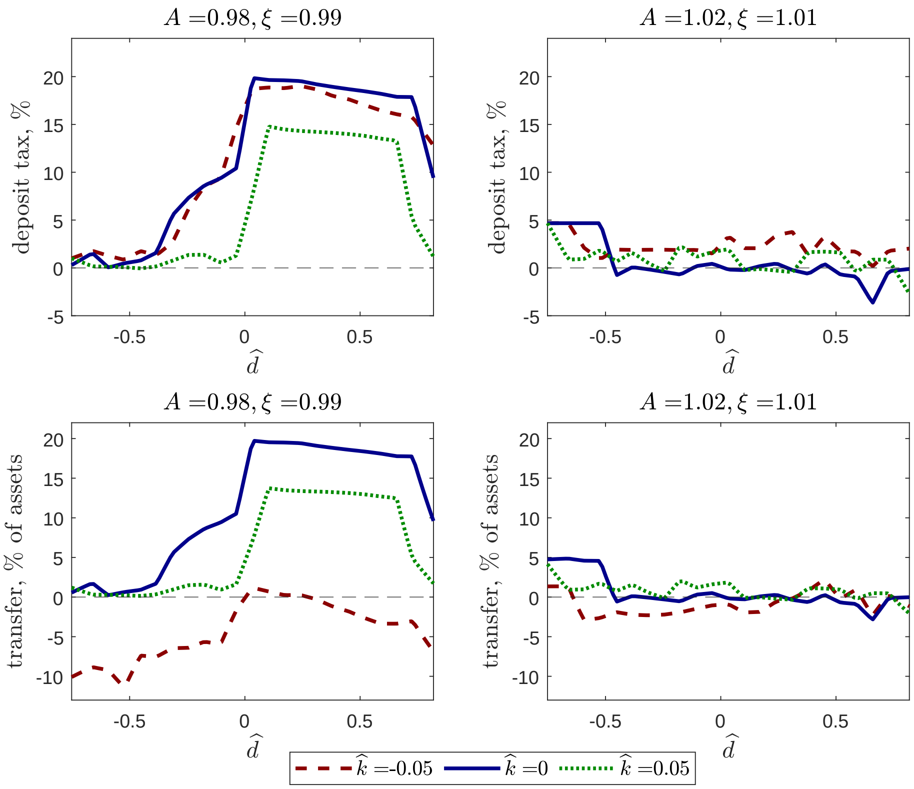

Figure 1.9. Optimal deposit tax rebated lump sum and the transfer τ^1 supporting the optimal distribution σ^1. Slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid.

the net bank value, overborrowing, and overlending. Inefficiencies are manifested to the greatest extent when exogenous conditions are bad. The policies are the most active in those states, and the deposit tax is primarily positive. The tax is greater when the planner's EC is slacker and banks have more debt. On the contrary, the tax varies about zero in the good state and can be negative, reflecting underborrowing by banks.

The variations in the optimal transfer to survived banks (as a percentage of bank assets) over the state space are qualitatively similar to those in the optimal deposit tax due to the government budget constraint $\tau_S^D \frac{D_S'}{R_S} = \tau_S^1 + \int_{\sigma^f}^{f} \tau_{i,S}\,\mathrm{d}i$. In the worst exogenous state, when banks are the most leveraged and the planner's constraint is almost binding, the transfer is mostly negative, encouraging deleveraging. When the agent's constraint is binding but the planner's constraint is slack, the transfer increases since it helps to relax the EC. In the best state, financial constraints are mostly slack in both the CE and MCEA, so the transfer is close to zero.

Consider now the policy scheme in which the aggregate lump-sum transfers are forbidden, that is, $\int_0^f \tau_{i,S}\,\mathrm{d}i = 0$. In this case, the planner must balance the budget with a linear tax on bank assets. The planner can still distribute entrants/survivors-specific lump-sum transfers that must vanish in the aggregate. Figure 1.10 displays the optimal policy in the described situation. (We will refer to this policy as optimal linear taxation.) Qualitatively, the deposit tax has similar trends to those in

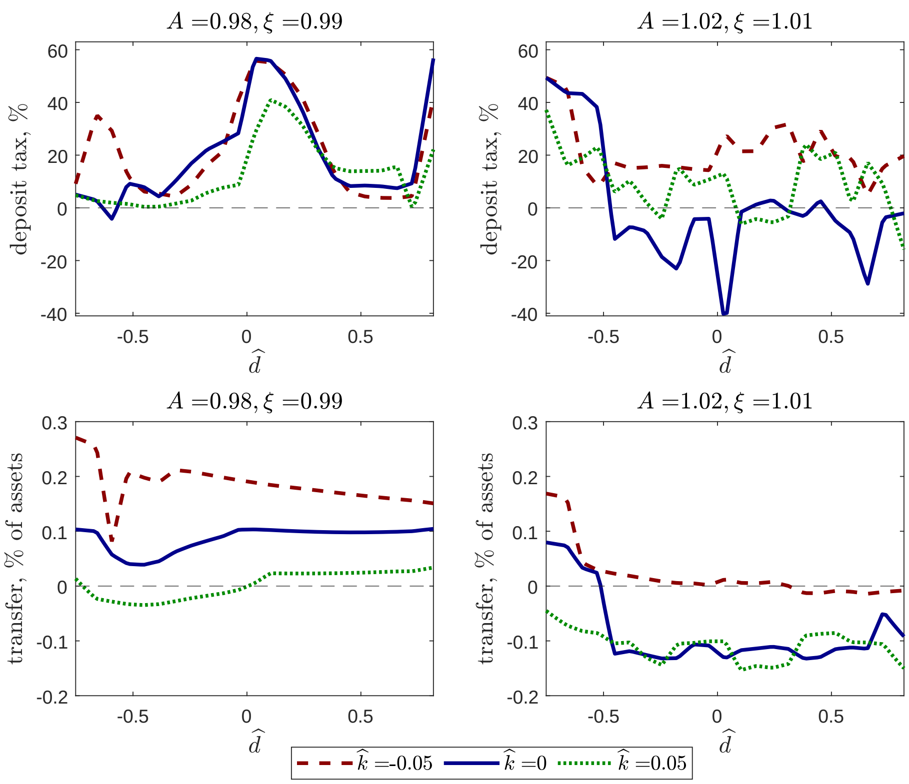

Figure 1.10. Optimal linear deposit tax and the transfer τ^1 supporting the optimal distribution σ^1 when $\int_0^f \tau_{i,S}\,\mathrm{d}i = 0$. Slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid.

the affine scheme, but the tax dispersion in the state space increases. In the bad exogenous state, the optimal tax is mostly positive and now reaches up to 60% at some debt values. The tax varies about zero in the good exogenous state, but the fluctuations are more pronounced than before, ranging from a 40% subsidy to a 50% tax. However, such extreme values are not typically observed in the ergodic distribution, which has a mode around the first quartile of $\widehat{d}$ in the worst state and around the third quartile in the best state.

The optimal transfer to survivors is now fully funded by a tax on entrants. The magnitude of the transfer is significantly less than that in the case of affine taxation since the transfer is now not directly related to system-wide proportional taxes. In the bad state, the transfer to survived banks is positive at greater leverage ratios, helping to relax the EC in the CE, where it is binding. The transfer is also generally more significant in the regions where the planner's constraint is slacker—at both shallow and large values of bank debt. The EC is mostly slack in the CE in the good state, except at the very low debt values when the transfer is greater. Otherwise, the transfer is either close to zero or negative.

Numerical results indicate that a possible implementation with minimum capital requirements is not always possible: the optimal credit spread and the implied Lagrange multiplier could be negative in some states. In other words, conditional on the same state, the optimal bank capital ratio in the MCEA can be smaller than in the CE. On the other hand, the ergodic state space is different in the MCEA due to the general overborrowing and overlending in the CE. If we compare the empirical distributions of the bank capital ratio, the MCEA distribution is shifted to the right compared to the CE distribution. This fact is illustrated in figure 1.11 together with the implied optimal transfer and all other policy schemes.

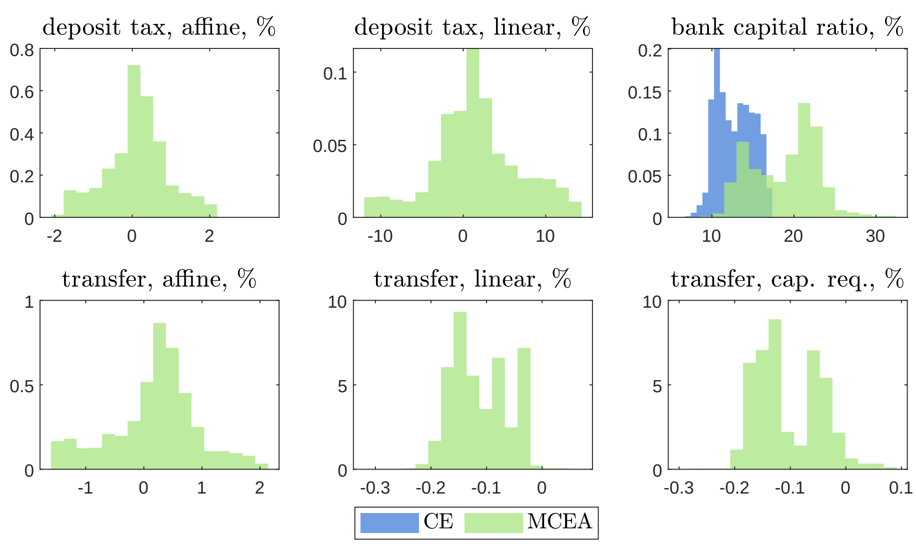

Figure 1.11. Optimal policies in the ergodic distribution. Each column corresponds to an alternative decentralization scheme (1,000,000-period simulation, pdf normalization on the y-axis). Outliers are removed. Transfers (τ_t^1) are in % of bank assets. The transfer in the last column is meaningful only when the implied Lagrange multiplier on the regulatory constraint is nonnegative, which does not always hold.

In the ergodic distribution, the optimal taxes have more mass in the positive region, reflecting overborrowing and overlending in the CE. The deposit tax and the optimal transfer in the affine scheme have long right tails not shown for the sake of clear illustration. As discussed above, the magnitude of optimal deposit taxes in the linear scheme is generally greater, while the magnitude of optimal transfers is smaller and primarily negative. The optimal bank capital ratio is typically greater than in the CE, with a mean of about 19%. The transfer corresponding to a possible implementation with capital requirements is only relevant when the Lagrange multiplier on the regulatory constraint is nonnegative, and the latter is not always the case. Generally, capital requirements alone are insufficient to implement the optimal allocation and must be augmented with other instruments.

Welfare gains and the role of optimal transfers

We have seen that the extent of overborrowing and overlending and, correspondingly, the magnitude of optimal taxes and transfers can be quite significant. Figure 1.12 illustrates how welfare gains from the baseline MCEA relative to the CE allocation vary in the state space. Welfare gains are

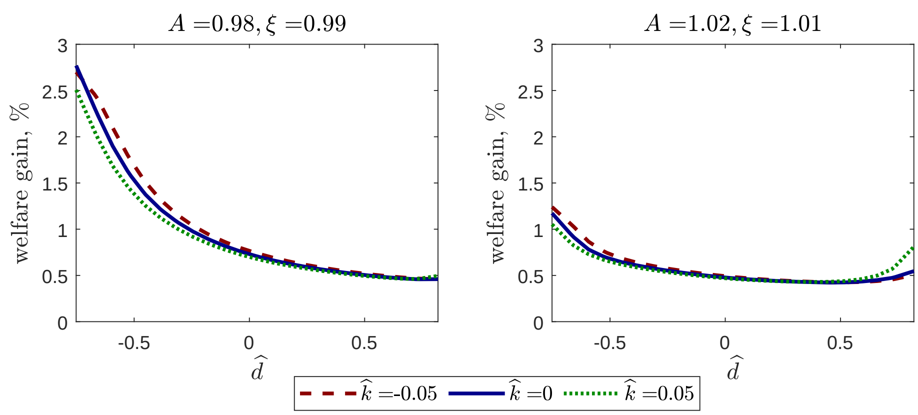

Figure 1.12. Welfare gains. Slices of the underlying surfaces along the $\widehat{d}$ dimension at the quartiles of the $\widehat{k}$ grid. Consumption-equivalent welfare gain of the MCEA relative to CE in % on the y-axis.

generally greater when exogenous conditions are bad, banks are more leveraged, and the CE EC is more binding—at the lower values of assets and debt. The mean welfare gain over the whole state space is 0.75% of consumption, and the median welfare gain is 0.57% of consumption. These numbers are about twice as large as those found in the open economy international finance context by Bianchi and Mendoza (2018).

Figure 1.13 illustrates how the average welfare gain from the MCEA varies as a function of

λ—the scale of the CE bank value distribution. At $\lambda < 0.9$ (approximately), the MPE does not

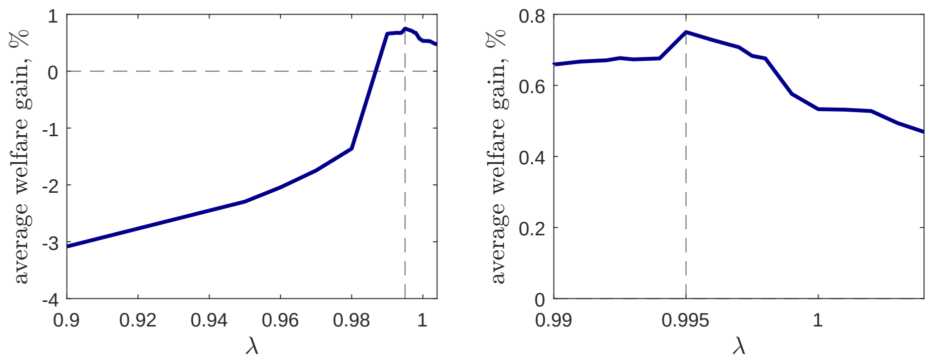

Figure 1.13. Welfare gains from the MCEA conditional on alternative distributions $\sigma^1 = \lambda\bar{\sigma}^1$. The y-axis is the average welfare gain over the state space in % of consumption. The right plot zooms in on the right edge of the left plot. The vertical dashed line corresponds to $\lambda = \lambda^*$.

exist since the EC cannot be satisfied in some states. At $\lambda \in (0.9, 0.98]$, the EC is binding almost everywhere in the state space, and as we increase λ, welfare losses steadily decrease as constraints are relaxed. When we move to $\lambda = 0.99$, λ is large enough for the EC to be mostly slack, and average welfare jumps to the welfare gain region. As we increase λ from 0.99 to 0.994, the measure of the state space where the constraint is slack continues to increase slightly, after which there is a significant jump when we move to $\lambda^* = 0.995$. A further increase in λ does not lead to an increase in the measure of the slack region—on the contrary, it decreases slightly, and the decrease is more significant as we move to $\lambda = 0.999$ and further above.

The reason welfare starts to decrease at $\lambda > \lambda^*$ is that the value function is generally not globally concave. A typical situation is that there are two local maxima: one where the EC is binding, another where the EC is slack, and the latter is typically quite distant from the binding region. When the constraint is mostly slack, but we keep increasing λ, the feasible set expands, and some of the global maxima switch from the slack to the binding region. The switching affects the continuation value at other states, and when the value function converges, we can observe a decrease in welfare. Despite the decrease, welfare gains remain sizable.

Let us note here that we also considered an alternative situation where the distribution σ^1 is restricted to satisfy $\sigma^1_S = \eta_S \frac{\sigma \widetilde{N}_S}{N_S}$ for all $S \in \mathcal{S}$. Note that we should not confuse this case with the case of $\lambda = 1$. In the latter case, σ^1 is fixed ex ante at the CE distribution $\bar{\sigma}^1$; that is, the values σ^1_S are given for all $S \in \mathcal{S}$. In the former case, the distribution object σ^1 is part of the

MPE—σ^1 is updated at each iteration to satisfy the proportionality constraint. The corresponding MPE outcome does not require any transfers for implementation since the equilibrium distribution is consistent with the linearity of the individual bank value in the individual net worth. In this case, the average welfare gain is 0.67%, and the median welfare gain is 0.5%, lower than under the optimal linear transformation that serves as the baseline in our analysis.

It is worth emphasizing that one can construct infinitely many distributions that are not limited to linear transformations and dominate the baseline distribution. It is interesting to investigate some of those possibilities, although the investigation is hindered by the computationally intensive reality of finding the MPE in alternative cases.

Financial crises

We will conclude the analysis of the MPE by exploring the economic dynamics around periods identified as financial crises in the CE. Figure 1.14 compares the dynamics around such events in the CE and MCEA. The most striking difference is that the EC remains slack during the whole

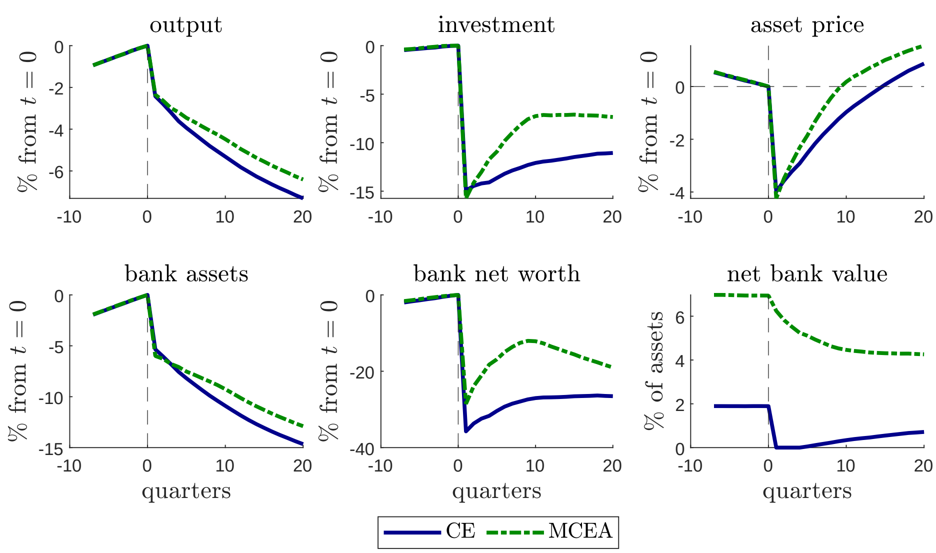

Figure 1.14. Financial crises, CE and MCEA. Averages over a 1,000,000-period simulation.

crisis window in the MCEA. As bad shocks hit, the bank value decreases gradually, but the buffer over the value of default is safely sufficient to evade transitioning to the binding regime. Bank net worth falls on impact by about 28.5% compared to 35.7% in the CE and starts to recover much faster. There is a slightly greater fall in bank assets and the asset price on impact, but both rebound faster than in the CE. There is a greater relative fall in bank deposits during the first 2.5 years, which contributes to the slackness of the EC. We also notice that banks are much less leveraged before the crisis, as the bank capital ratio is at about 20.3% at $t = 0$ in the MCEA compared to 14.3% in the CE. A faster recovery in bank assets and the asset price is reflected in a much faster recovery in investment, which crowds out consumption to a certain extent. In sum, the cumulated fall in output is lower in the MCEA compared to the CE.

Figure 1.15 further explores how potential MCEA decentralization policies behave around crises. During the two years ahead of the average crisis, the optimal deposit tax in the affine scheme

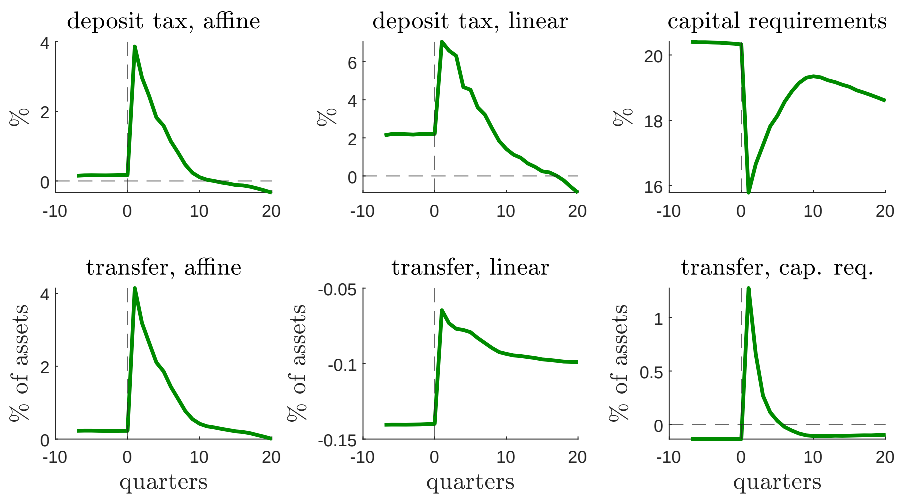

Figure 1.15. MCEA decentralization policies around financial crises. Averages over a 1,000,000-period simulation. Each column corresponds to an alternative implementation mechanism. Transfers refer to τ^1.

increases from about 15 to 17 basis points. Accordingly, the optimal transfer to survivors stays modest. The optimal tax is much larger in the linear taxation scheme, growing from 213 to 221 basis points before the crisis. The magnitude of the transfer, on the other hand, is much smaller in the linear scheme, as it is not linked to the tax in the government budget constraint. Moreover, the transfer is negative, reflecting that the optimal distribution σ^1 is uniformly lower than the CE distribution—therefore, the planner supports entrants in good times. As for capital requirements, they are much larger than the CE capital ratio before the crisis, while the transfer is negative similar to the case of linear taxation.

When the bad shock hits and the EC binds in the CE, the optimal deposit tax—in both the affine and linear schemes—increases significantly, contributing to faster deleveraging and evading the binding constraint entirely in the MPE. The optimal transfer to survivors rises substantially in both taxation mechanisms (in the linear scheme, the negative transfer decreases), compensating for the rise in the deposit tax. This rise reflects the preemptive bailout: the rise in the transfer supports the value of the bank so that it stays above the value of default, and the EC remains slack. The optimal capital requirements—which, by proposition 1.4, correspond to the optimal capital ratio—fall significantly when the bad shock hits but stay well above the capital ratio in the CE allocation. Hence, the optimal capital requirements have a macroprudential nature. Note that, unlike in the whole state space, the implementation with capital requirements is effective around the potential crises. In general, however, they must be augmented by taxes to be effective.

1.4.5 Ramsey equilibrium

In this subsection, we will explore the implication of the Ramsey equilibrium, relating them to the findings discussed so far. The Ramsey allocation is not recursive but history-dependent; therefore, we cannot directly compare policy functions with those in the CE.[6] We will thus focus on comparing the empirical distributions and the economic dynamics around financial crises in the CE.

As in section 1.4.4, the baseline analysis is conditional on the optimal bank value distribution among a certain class of distributions. The baseline computation of the CEA relies on piecewise linear perturbation about the steady state.[7] For this reason, it is more convenient to focus on constant distributions $\sigma_t^1(z^t) = \sigma^1$ for all (t, z^t). The optimal distribution in the class of constant distributions is $\sigma^1 \approx 0.9985$, the smallest value at which the aggregate EC is slack in the steady state. Note that the steady-state value of the CE distribution is $\bar{\sigma}^1 \approx 0.9911$.

Unlike in the MPE, where the CE distribution has to be optimally scaled down, it is optimal to scale it up in the Ramsey equilibrium. The Ramsey planner finds it optimal to promise sufficiently large transfers to the banks, such that the EC becomes just slack in the steady state—that is, banks are exactly at the boundary of the constrained and unconstrained regions. Both with and without commitment, there is a similar rationale to provide just enough transfers to have financial

[6] Since the Ramsey equilibrium is recursive on the state space augmented with Lagrange multipliers, it is possible to compare policy functions conditional on specific values of Lagrange multipliers.

[7] For consistency, in this subsection, we use the same method to compute the CE. The CE simulation will thus differ from the baseline simulation from the previous analyses. The piecewise linear perturbation accounts for the occasionally binding constraint but does not account for precautionary savings. The computational burden of simulating the model using this approach is significant, so the simulation length is reduced from 1,000,000 to 100,000.

constraints relaxed, but pecuniary externalities present an opposing force that prevents the planner from providing excessive preemptive bailouts. Commitment matters for the location of the optimal transfer boundary. The boundary is further above with commitment, so the Ramsey planner supports more bank debt and credit, as we will see momentarily.

Later we will explore the welfare implications of alternative σ^1, as we did with the MPE analysis.

Bank borrowing, lending, and optimal policies

We begin by looking at the empirical distributions of bank deposits and loans. Figure 1.16 shows the corresponding histograms, where in addition to the CE and CEA, we have histograms from the frictionless (unconstrained) CE (UE), in which $\theta = 0$ and all other parameters are identical to those in the baseline CE. By construction, in the UE, the EC is always slack since the aggregate net worth and bank value are strictly positive.

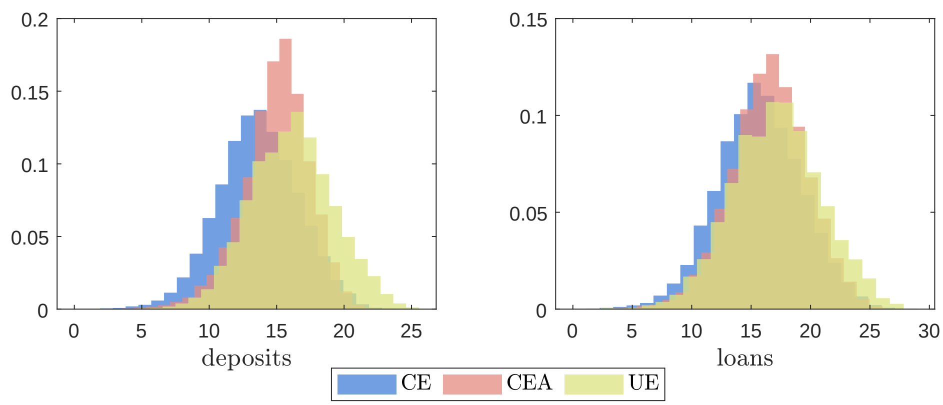

Figure 1.16. Bank borrowing and lending in the CE, CEA, and UE. The latter refers to a frictionless (unconstrained) CE with $\theta = 0$. Histograms based on the 100,000-period simulation with the same sequence of exogenous shocks. Variables are normalized by the average CE output; the y-axis has the pdf normalization.

An immediate implication of the optimal preemptive bailout policy that supports the relative bank value of survived banks at a greater value than in the CE is the expansion of bank balance sheets. With commitment, we observe *underborrowing* and *underlending* by the banking sector in the CE compared to the CEA. Bank deposits and loans have a greater mean and variance in the Ramsey equilibrium than in the CE. The CEA histograms are more skewed to the left, so the median deposits and loans are even greater. However, the optimal balance sheets are smaller than in the UE, where the EC is always slack and the limited enforcement friction is shut down.

Hence, the Ramsey planner alleviates the friction with preemptive bailouts but does not eliminate it, reflecting a trade-off between preventing excessive borrowing and lending ex ante and relaxing financial constraints ex post.

Figure 1.17 displays the empirical distributions of the alternative CEA implementation policies. Although there is greater bank borrowing and lending in the CEA than in the CE, the optimal

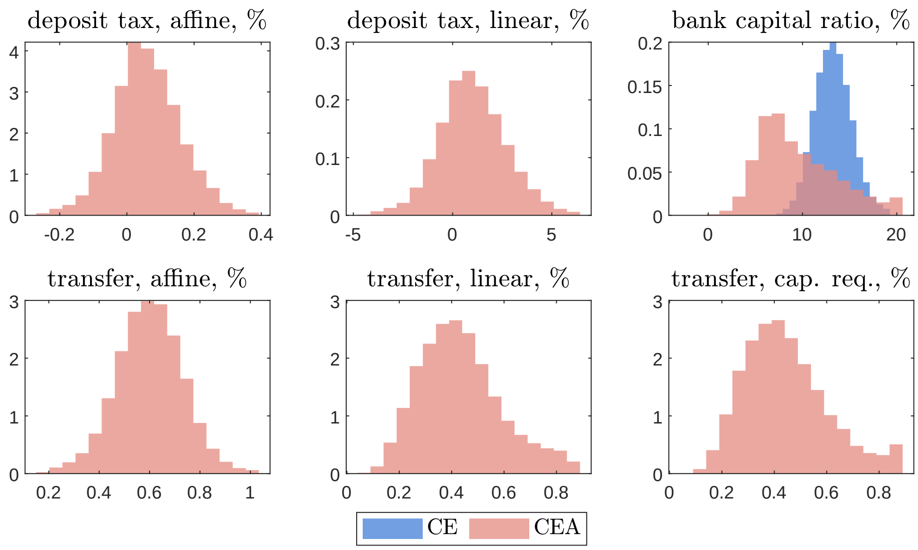

Figure 1.17. Optimal policies under commitment. Histograms based on the 100,000-period simulation with the same sequence of exogenous shocks. Each column corresponds to an alternative CEA decentralization scheme (pdf normalization on the y-axis). Outliers are removed. Transfers (τ_t^1) are in % of bank assets. The transfer in the last column is meaningful only when the implied Lagrange multiplier on the regulatory constraint is nonnegative, which does not always hold.

deposit taxes have more mass in the positive region, similar to what we found in the MPE. The Ramsey planner uses taxes to correct the pecuniary externalities, which prevents borrowing and lending from being excessively large, even though it is larger than in the CE due to optimal transfers. Similar to the MCEA, when aggregate transfers are forbidden (the linear implementation scheme), the magnitude of the taxes is generally greater.

The optimal bank capital ratio has a lower mean and median than in the CE but a greater variance and a much greater skewness to the right. Hence, under commitment, the optimal capital ratios are generally lower than in the CE, reflecting the increased borrowing and lending, but there is a nontrivial measure of contingencies in which the planner finds it optimal for banks to be sufficiently more capitalized than in the CE.

The optimal transfers to survived banks are uniformly positive independently of the CEA implementation mechanism. This fact contrasts with the MCEA, where transfers were primarily negative since it was optimal to scale down the CE bank value distribution. The optimal transfers have a comparable magnitude across implementation schemes with a mean of about 0.5% of bank assets. As with the MCEA, we must note that the decentralization with capital requirements alone does not always succeed; therefore, the optimal transfers are only valid conditional on having the Lagrange multiplier on the regulatory constraint nonnegative in the relevant contingencies.

Welfare gains and the role of optimal transfers

Since the CEA is not recursive, instead of exploring how welfare gains vary in the state space, we will focus on welfare gains based on the ergodic mean of the value function. Figure 1.18 illustrates how the ergodic welfare gain varies as a function of σ^1. The ergodic welfare gain from the CEA

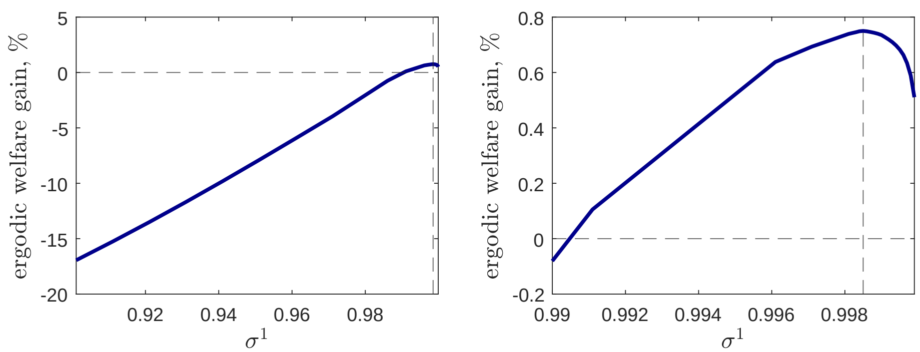

Figure 1.18. Welfare gains from the CEA conditional on alternative distributions σ^1. The y-axis is the welfare gain based on the ergodic means of the value functions in % of consumption. The right plot zooms in on the right edge of the left plot. The vertical dashed line corresponds to $\sigma^1 = 0.9985$.

conditional on the optimal distribution σ^1 is about 0.75% of consumption.[8] When $\sigma^1 \in [0.9, 0.99]$, we are in the ergodic welfare loss region, and the losses dramatically decrease as we increase σ^1, and the planner's EC is relaxed in more and more contingencies. When we move to the CE distribution with $\bar{\sigma}^1 \approx 0.9911$, we finally get a welfare gain of 0.11%, and the EC is slack about 68.6% of the time compared to 50.7% in the CE. When we go up to the optimal distribution $\sigma^1 \approx 0.9985$, the constraint is slack 95.4% of the time and is now slack in the steady state. As we increase σ^1 further up to 0.9999, the constraint becomes slack 99.6% of the time, but the trade-off between the

[8] Although this number may seem identical to the baseline welfare gain from the MCEA, note again the difference in welfare gain concepts. In the MCEA, we would detect an ergodic welfare loss due to the extent of overborrowing and overlending in the CE.

excessive borrowing and lending ex ante and the slackness of the EC ex post swings to the former, so welfare gains decrease down to 0.51% of consumption, which is still significant.

The moral of the story is that it is optimal to relax the EC in most contingencies but not necessarily in all possible contingencies: the optimal transfers should be large enough but not excessively large. The Ramsey planner commits to providing enough help to older and larger banks when financial constraints bind ex post while discouraging banks from growing too large ex ante. In other words, "too big to fail" is a problem that must be addressed ex post, but it is better to evade it ex ante.

Financial crises

Finally, we will look at the economic dynamics around financial crises. Remember that in this subsection, we use a different approach to compute the CE for consistency with the computation of the CEA. Our identification of financial crises changes slightly: instead of requiring the EC to be slack for twenty quarters before the crisis, we look for at least ten quarters, which allows obtaining a similar frequency of financial crises of about 3.1 crises per century.

Figure 1.19 illustrates the dynamics around crises. The general trends are quite similar to

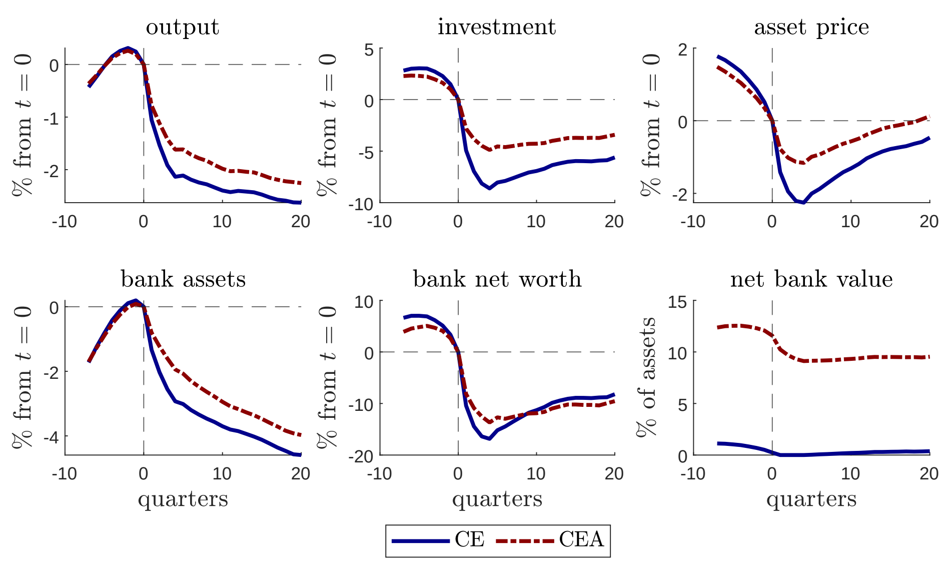

Figure 1.19. Financial crises, CE and CEA. Averages over a 100,000-period simulation.

those in figure 1.14, which, in particular, confirms that the alternative computational approach is adequate. The most striking difference from the behavior of the optimal time-consistent allocation is that the optimal bank capital ratio is now uniformly lower during the crises than in the CE. These dynamics reflect the comparison of empirical distributions in figure 1.17. Since the planner finds it optimal to provide sufficient support to survivors through transfers, they generally borrow more and become more leveraged on average. A lower capital ratio is not a problem since the very purpose of those transfers—or preemptive bailouts—is to prevent the EC from switching to the binding regime, which is achieved successfully—the net bank value in the CEA generally remains slack around crises and to a greater extent than in the MCEA.

The boom-bust dynamics in the CEA are generally less pronounced than in the CE, as both real and financial variables are less volatile in such episodes and recover faster after the bad shock hits. In particular, we observe a faster recovery in the asset price and bank assets and liabilities, which does not allow investment to drop as severely as in the CE. Consumption also varies less, and output rebounds faster.

Figure 1.20 focuses on the dynamics of the CEA decentralization policies. Some curves are not

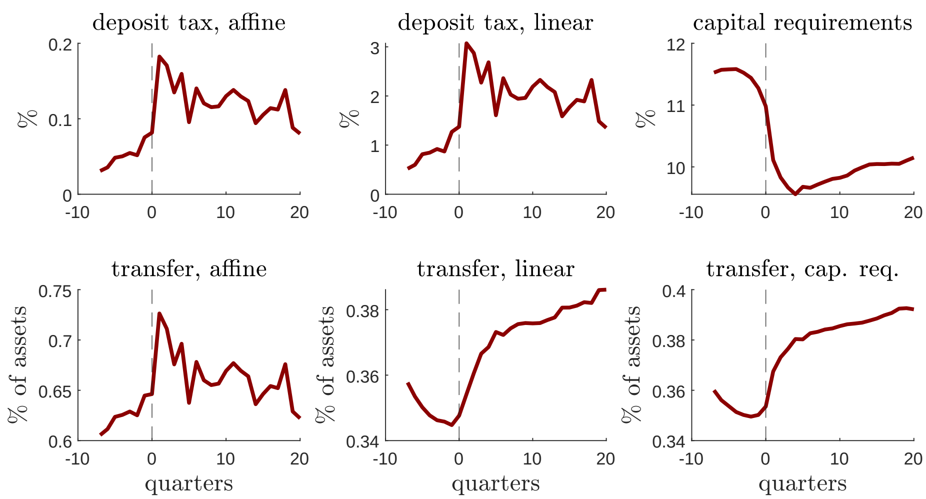

Figure 1.20. CEA decentralization policies around financial crises. Averages over a 100,000-period simulation. Each column corresponds to an alternative implementation mechanism. Transfers refer to τ^1.

as smooth as in figure 1.15 due to a lower simulation length, but the trends are clear. As in the time-consistent case, the optimal deposit taxes are increasing ahead of a crisis, albeit with a greater magnitude, and jump when the bad shock arrives to encourage faster deleveraging and keep the banking sector in the unconstrained regime. The increase is followed by a gradual decline as both exogenous and endogenous conditions improve.

By construction, the optimal transfer in the affine scheme tracks the dynamics of the deposit tax to a great extent. In the linear scheme, the optimal transfer is falling slightly ahead of a crisis, which is an additional way to encourage deleveraging ex ante. When the shock arrives, the trend is reversed, and the transfer increases to relax the EC. We observe very similar behavior in the optimal transfer conditional on the implementation with capital requirements. Unlike in the MPE, the optimal transfers in all implementation schemes are generally positive around crises.

As in the MPE, the implementation with capital requirements is effective around crises since the implied Lagrange multiplier on the regulatory constraint stays positive. At the same time, the optimal capital ratio is generally lower than in the CE. It might seem unintuitive, but remember that the optimal constant bank value distribution scales the CE analog up in the Ramsey equilibrium, so the corresponding transfers would decrease the CE capital ratio to even lower values in the environment without additional regulation. Therefore, with optimal capital requirements and preemptive bailouts, the regulatory constraint would bind in the regulated CE.

1.5 Conclusion

This paper has characterized the optimal regulation of a banking system in a quantitative general equilibrium environment. We have found that a benevolent policymaker generally faces a trade-off between limiting excessive borrowing and lending by banks ex ante in normal times and supporting the banking system ex post in bad times. The optimal policy requires a combination of system-wide deposit taxes or state-contingent capital requirements—that address pecuniary externalities implicit in the banking system enforcement constraint—and bank entrants/survivors-specific transfers that achieve the optimal bank value distribution. We have referred to the optimal transfers as preemptive bailouts, as their goal is to prevent financial constraints from becoming binding, guaranteeing bank solvency.

We have studied the optimal policy in the Markov perfect equilibrium and the Ramsey equilibrium, which differ in whether the policymaker can commit. Independently of the latter, the optimal transfer policy generally ensures that the enforcement constraint is slack in most but not all states/contingencies, and it is just slack in the long run. The presence of commitment has, however, striking quantitative implications. We generally observe overborrowing and overlending by banks in the competitive equilibrium compared to the Markov perfect equilibrium outcome, and the optimal transfers are generally negative. There is, however, mostly underborrowing and underlending in competitive markets compared to the Ramsey outcome, and the optimal transfers are generally positive in this case. On the other hand, the behavior of optimal policies around financial crises is quite similar: optimal taxes are mostly procyclical, while optimal transfers and bank capital requirements are countercyclical.

The present analysis can be extended in various ways. We could consider alternative environments in which banks can self-insure with endogenous equity issuance or can invest in other types of assets, such as government debt, which will potentially introduce additional externalities. It is also interesting to generalize the model and explore the implications for optimal monetary policy.

Appendix

1.A Proofs

1.A.1 Proposition 1.1

Consider the problem of a banker that enters the banking business at $t = 0$. Define $\eta^t \equiv \prod_{j=0}^{t} \eta_t$, where by construction $\eta_0 = 1$. Let γ_t, λ_t, and ν_t denote the normalized Lagrange multipliers on the bank value, enforcement, and balance sheet constraints, respectively. Also, let ϕ_t denote the multiplier on the constraint $x_t \geq 0$. The banker's Lagrangian is then

$$\mathcal{L} = v_0 + \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \eta^t \sigma^t \Lambda_{0,t} \left\{ \gamma_t [x_t + \mathbb{E}_t \{ \eta_{t+1} \Lambda_{t,t+1} [(1-\sigma) n_{t+1} + \sigma v_{t+1}] \} - v_t] \right. \right.$$
$$\left. \left. + \lambda_t (v_t - \theta Q_t s_t) + \nu_t \left(n_t - x_t + \frac{d_{t+1}}{R_t} - Q_t s_t \right) + \phi_t x_t \right\} \right],$$

where

$$n_t \equiv \begin{cases} R_t^K Q_{t-1} s_{t-1} - d_t & \text{survivor} \\ n_t^0 & \text{entrant} \end{cases},$$

and n_t^0 depends on the aggregate state only. The first-order conditions (FOCs) with respect to v_t, x_t, d_{t+1}, and s_t are

$$\begin{aligned}
\gamma_t &= \mathbf{1}_{\{0\}}(t) + \mathbf{1}_{\mathbb{N}}(t)\gamma_{t-1} + \lambda_t, \\
\phi_t &= \nu_t - \gamma_t, \\
\nu_t &\geq R_t \mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[(1-\sigma)\gamma_t + \sigma\nu_{t+1}]\}, \qquad \text{equality if } d_{t+1} < \bar{D}, \\
\lambda_t\theta + \nu_t &\geq \mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[(1-\sigma)\gamma_t + \sigma\nu_{t+1}]R_{t+1}^K\}, \qquad \text{equality if } s_t > 0.
\end{aligned}$$

The complementary slackness conditions are

$$\begin{aligned}
0 &= \phi_t x_t, \qquad \phi_t \geq 0, \qquad x_t \geq 0, \\
0 &= \lambda_t(v_t - \theta Q_t s_t), \qquad \lambda_t \geq 0, \qquad v_t \geq \theta Q_t s_t.
\end{aligned}$$

Optimal dividend

The FOC with respect to v_t implies $\gamma_t = 1 + \sum_{j=0}^{t}\lambda_j$. Since $\lambda_t \geq 0$, we have $\gamma_t \geq 1$. Since $\phi_t \geq 0$, the FOC with respect to x_t implies $\nu_t \geq \gamma_t \geq 1$. Moreover, $\phi_t > 0$ if and only if $\nu_t > \gamma_t$. The deposit Euler equation implies

$$\begin{aligned}
\frac{\nu_t}{\gamma_t} &= R_t\mathbb{E}_t\left[\eta_{t+1}\Lambda_{t,t+1}\left(1 - \sigma + \sigma\frac{\nu_{t+1}}{\gamma_{t+1}}\frac{\gamma_{t+1}}{\gamma_t}\right)\right] \\
&\geq R_t\mathbb{E}_t\left[\eta_{t+1}\Lambda_{t,t+1}\left(1 + \sigma\frac{\lambda_{t+1}}{\gamma_t}\right)\right],
\end{aligned}$$

where the second line uses $\nu_{t+1} \geq \gamma_{t+1}$ and $\gamma_{t+1} = \gamma_t + \lambda_{t+1}$. We can get different sufficient conditions for $\nu_t > \gamma_t$ based on these inequalities. For example, this would be the case if $\eta_{t+1}(z^{t+1}) = 1$ almost everywhere and $\{\lambda_{t+1}(z^{t+1}) > 0\}$ is of positive measure. If, otherwise, $\nu_t = \gamma_t$, the FOC with respect to x_t implies $\phi_t = 0$, and the complementary slackness conditions imply that any $x_t \in \left[0, n_t + \frac{d_{t+1}}{R_t}\right]$ is a solution. Hence, we can set $x_t = 0$ without loss of generality.

Dependence of Lagrange multipliers on the aggregate state

The Euler equations can only depend on the individual net worth through the Lagrange multipliers or the indicator function η_t. Due to the linearity of the problem, I focus on a symmetric equilibrium where all existing banks either default or not at the same time. Suppose $\{\lambda_t\}$, and thus $\{\gamma_t\}$, depends on the aggregate state only. By iterating forward the Euler equation for deposits, one can solve for ν_t as a function of $\{\gamma_t, \eta_t\}$, stochastic discount factors, and deposit rates. Hence, $\{\nu_t\}$ depends on the aggregate state only. Conversely, suppose $\{\nu_t\}$ depends on the aggregate state only. The same Euler equation immediately implies that $\{\gamma_t\}$, and thus $\{\lambda_t\}$, depend on the aggregate state only. The Euler equation for securities is consistent with the fact that the Lagrange multipliers do not depend on n_t.

Stationary transformation

For any history $z^\infty \in Z^\infty$, the sequence $\{\gamma_t\}_{t=0}^\infty$ is nondecreasing and generally unbounded. Let $\widehat{\lambda}_t \equiv \frac{\lambda_t}{\gamma_t}$ and $\widehat{\nu}_t \equiv \frac{\nu_t}{\gamma_t}$. By construction, $\widehat{\lambda}_t \in [0,1)$ and $\widehat{\nu}_t \geq 1$. Note that

$$\frac{\gamma_{t+1}}{\gamma_t} = 1 + \frac{\lambda_{t+1}}{\gamma_t} = 1 + \widehat{\lambda}_{t+1}\frac{\gamma_{t+1}}{\gamma_t} = \frac{1}{1-\widehat{\lambda}_{t+1}}.$$

Now define $\bar{\nu}_t \equiv \frac{\widehat{\nu}_t}{1-\widehat{\lambda}_t}$ and $\bar{\lambda}_t \equiv \frac{\widehat{\lambda}_t}{1-\widehat{\lambda}_t}$. Note that $1+\bar{\lambda}_t = \frac{1}{1-\widehat{\lambda}_t}$. It follows that

$$\bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})]R_t,$$
$$\theta\bar{\lambda}_t + \bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})R^K_{t+1}].$$

By construction, $\lambda_t > 0$ if and only if $\bar{\lambda}_t > 0$ and $\lambda_t = 0$ if and only if $\bar{\lambda}_t = 0$. The complementary slackness conditions on the EC can, therefore, be written as

$$0 = \bar{\lambda}_t(v_t - \theta Q_t s_t), \qquad \bar{\lambda}_t \geq 0.$$

Value function

Define $\mu_t \equiv v_t - \bar{\nu}_t n_t$. Guess that $v_t = \bar{\nu}_t n_t$. Substituting in the bank value constraint

$$
\begin{aligned}
(1+\bar{\lambda}_t)v_t &= (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})(R^K_{t+1}Q_t s_t - d_{t+1})] + \sigma(1+\bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}) \\
&= (\theta\bar{\lambda}_t + \bar{\nu}_t)Q_t s_t - \bar{\nu}_t\frac{d_{t+1}}{R_t} + \sigma(1+\bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}) \\
&= \bar{\lambda}_t v_t + \bar{\nu}_t n_t + \sigma(1+\bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}),
\end{aligned}
$$

where the first line uses the definition of net worth, the second line uses the Euler equations, and the third line uses the balance sheet constraint and the complementary slackness conditions. It follows that

$$\mu_t = \sigma(1+\bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}).$$

Trivially, $\{\mu_t\} = \{0\}$ is a solution. Suppose there exists another solution $\{\mu_t\}$ with $\mu_t(z^t) \neq 0$ for some (t, z^t). Conditional on this (t, z^t), define a sequence $x_n(z^t) : \mathbb{N} \to \mathbb{R}$ as

$$x_n = \frac{\sigma^n}{U_{C,t}}\mathbb{E}_t\left\{\beta^n\left[\prod_{i=1}^{n}\eta_{t+i}(1+\bar{\lambda}_{t+i-1})\right]U_{C,t+n}\mu_{t+n}\right\}.$$

Clearly, $x_n(z^t) = \mu_t(z^t)$ for all $n \in \mathbb{N}$. Therefore, $\{x_n\}$ is a constant sequence and $\lim_{n\to\infty} x_n = \mu_t \neq 0$. On the other hand, by assumption, a sequence $n \mapsto \beta^n \prod_{i=1}^n \eta_{t+i}(1+\bar{\lambda}_{t+i-1})$ is bounded conditional on all continuations of z^t, which implies that $\mathbb{E}_t\left\{\beta^n\left[\prod_{i=1}^n \eta_{t+i}(1+\bar{\lambda}_{t+i-1})\right]U_{C,t+n}\mu_{t+n}\right\}$ is bounded at any bounded plan. Since $\sigma < 1$, it must be that $\lim_{n\to\infty} x_n = 0$, which is a contradiction. Therefore, $v_t = \bar{\nu}_t n_t$ is the unique solution to the banker's problem. ■

1.A.2 Lemma 1.1

By proposition 1.1, the value function of an individual banker is linear in net worth: $v_t = \bar{\nu}_t n_t$, where $\bar{\nu}_t$ is common to all bankers. Since v_{t+1} depends on n_{t+1} and the aggregate state, and n_{t+1} depends only on the current control variables and the future states, v_t and the EC do not depend on future controls. It follows that one can, equivalently, study the recursive problem

$$v(n, S) = \max_{(d,s)\in\Gamma(n,S)} \mathbb{E}_z\{\eta'\Lambda_{S,S'}[(1-\sigma)n' + \sigma v(n', S')]\},$$

where $\eta' \equiv \mathbf{1}_{\mathbb{R}_{++}}(n')$ and the correspondence $\Gamma : \mathbb{R}_+ \times \mathcal{S} \to \mathcal{P}(\mathbb{R}^2_+)$ is defined by the following constraints

$$\begin{aligned}
\nu: \quad & 0 \leq n + \frac{d}{R_S} - Q_S s, \\
\lambda: \quad & 0 \leq \mathbb{E}_z\{\eta' \Lambda_{S,S'}[(1-\sigma)n' + \sigma v(n', S')]\} - \theta Q_S s, \\
& n' = R^K_{S'} Q_S s - d.
\end{aligned}$$

The envelope condition is

$$\frac{dv(n,S)}{dn} = \nu.$$

The Euler equations are

$$\nu_S = (1+\lambda_S) R_S \mathbb{E}_z[\eta_{S'} \Lambda_{S,S'}(1 - \sigma + \sigma \nu_{S'})],$$
$$\theta \lambda_S + \nu_S = (1+\lambda_S) \mathbb{E}_z[\eta_{S'} \Lambda_{S,S'}(1 - \sigma + \sigma \nu_{S'}) R^K_{S'}],$$

imposing that ν, λ, and η are functions of the aggregate state only in a symmetric equilibrium we consider. The envelope condition implies $v(n, S) = \mu_S + \nu_S n$, where $\mu : \mathcal{S} \to \mathbb{R}$ is an unknown function. Using the Bellman equation,

$$\mu_S = (1+\lambda_S)\sigma \mathbb{E}_z(\eta_{S'} \Lambda_{S,S'} \mu_{S'}),$$

which has a trivial solution $\mu_S = 0$ for all $S \in \mathcal{S}$. The complementary slackness conditions are

$$0 = \lambda_S(\nu_S n - \theta Q_S s), \qquad \lambda_S \geq 0. \quad \blacksquare$$

1.A.3 Proposition 1.2

The sequential planning problem is

$$\max_{\{C_t, D_{t+1}, K_{t+1}, L_t, V_t\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \right]$$

subject to

$$
\begin{aligned}
\widetilde{\nu}_t: &\quad 0 = N_t - Q(K_t, K_{t+1}, \xi_t)K_{t+1} + \beta\mathbb{E}_t\left(\frac{U_C(C_{t+1}, L_{t+1})}{U_C(C_t, L_t)}\right)D_{t+1},\\
\widetilde{\gamma}_t: &\quad 0 = \beta\mathbb{E}_t\left\{\eta_{t+1}\frac{U_C(C_{t+1}, L_{t+1})}{U_C(C_t, L_t)}[(1-\sigma)\widetilde{N}_{t+1} + \sigma^1_{t+1}V_{t+1}]\right\} - V_t,\\
\widetilde{\lambda}_t: &\quad 0 \le V_t - \theta Q(K_t, K_{t+1}, \xi_t)K_{t+1},\\
\lambda^L_t: &\quad 0 = U_C(C_t, L_t)A_tF_L(\xi_tK_t, L_t) + U_L(C_t, L_t),\\
\lambda^Y_t: &\quad 0 = A_tF(\xi_tK_t, L_t) - C_t - I(K_t, K_{t+1}, \xi_t),
\end{aligned}
$$

where

$$
\begin{aligned}
&\widetilde{N}_t \equiv [A_tF_K(\xi_tK_t, L_t) + Q(K_t, K_{t+1}, \xi_t)(1-\delta)]\xi_tK_t - D_t, \qquad \eta_t \equiv \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_t),\\
&N_t \equiv \bar{N} + \sigma\eta_t\widetilde{N}_t + \omega Q(K_t, K_{t+1}, \xi_t)K_t,
\end{aligned}
$$

and $\{\sigma^1_t\}$ is either given or satisfies $\{\sigma^1_t\} = \{\widehat{\sigma}^1_t\}$. In the latter case, the partial derivatives are

$$
\begin{aligned}
&\frac{\partial\sigma^1_t}{\partial D_t} \equiv -\sigma\frac{N^0_t}{N^2_t} < 0,\\
&\frac{\partial\sigma^1_t}{\partial K_t} \equiv \sigma\frac{N^0_t}{N^2_t}\{[A_tF_{KK,t}\xi_t + Q_{1,t}(1-\delta)]\xi_tK_t + R^K_tQ_{t-1}\} - \sigma\frac{\widetilde{N}_t}{N^2_t}\omega(Q_{1,t}K_t + Q_t),\\
&\frac{\partial\sigma^1_t}{\partial K_{t+1}} \equiv \frac{\sigma Q_{2,t}K_t[N^0_t(1-\delta)\xi_t - \omega\widetilde{N}_t]}{N^2_t}, \qquad \frac{\partial\sigma^1_t}{\partial L_t} \equiv \sigma\frac{N^0_t}{N^2_t}A_tF_{KL,t}\xi_tK_t > 0,
\end{aligned}
$$

where $N^0_t \equiv \bar{N} + \omega Q_tK_t$. Otherwise, all derivatives are zero.

Define $\nu_t \equiv \frac{\widetilde{\nu}_t}{U_{C,t}}$, $\gamma_t \equiv \frac{\widetilde{\gamma}_t}{U_{C,t}}$, and $\lambda_t \equiv \frac{\widetilde{\lambda}_t}{U_{C,t}}$. The FOCs are

$$
\begin{aligned}
C_t: &\quad 0 = U_{C,t} + U_{CC,t}[\nu_t(N_t - Q_tK_{t+1}) - \gamma_tV_t] + \lambda^L_t(U_{CC,t}A_tF_{L,t} + U_{CL,t}) - \lambda^Y_t\\
&\qquad + \mathbf{1}_{\mathbb{N}}(t)U_{CC,t}\{\nu_{t-1}D_t + \gamma_{t-1}\eta_t[(1-\sigma)\widetilde{N}_t + \sigma^1_tV_t]\},\\
D_{t+1}: &\quad \nu_t = R_t\mathbb{E}_t\left\{\eta_{t+1}\Lambda_{t,t+1}\left[\left(1 - \sigma - \frac{\partial\sigma^1_{t+1}}{\partial D_{t+1}}V_{t+1}\right)\gamma_t + \sigma\nu_{t+1}\right]\right\},
\end{aligned}
$$

$$
\begin{aligned}
K_{t+1}: \quad & 0 = \nu_t(Q_{2,t}\{[\sigma\eta_t(1-\delta)\xi_t + \omega]K_t - K_{t+1}\} - Q_t) - \lambda_t\theta(Q_{2,t}K_{t+1} + Q_t) - \frac{\lambda_t^Y}{U_{C,t}} I_{2,t} \\
& + \mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}[(1-\sigma)\gamma_t + \sigma\nu_{t+1}]\{[A_{t+1}F_{KK,t+1}\xi_{t+1} + Q_{1,t+1}(1-\delta)]\xi_{t+1}K_{t+1} \\
& + R_{t+1}^K Q_t\}) + \mathbb{E}_t\Bigg(\Lambda_{t,t+1}\Bigg\{\nu_{t+1}[\omega(Q_{1,t+1}K_{t+1} + Q_{t+1}) - Q_{1,t+1}K_{t+2}] \\
& - \lambda_{t+1}\theta Q_{1,t+1}K_{t+2} + \lambda_{t+1}^L A_{t+1}F_{KL,t+1}\xi_{t+1} + \frac{\lambda_{t+1}^Y}{U_{C,t+1}}(A_{t+1}F_{K,t+1}\xi_{t+1} - I_{1,t+1})\Bigg\}\Bigg) \\
& + \gamma_t\mathbb{E}_t\left(\eta_{t+1}\Lambda_{t,t+1}\frac{\partial\sigma_{t+1}^1}{\partial K_{t+1}}V_{t+1}\right) + \mathbf{1}_{\mathbb{N}}(t)\gamma_{t-1}\eta_t\left[(1-\sigma)Q_{2,t}(1-\delta)\xi_t K_t + \frac{\partial\sigma_t^1}{\partial K_{t+1}}V_t\right], \\
L_t: \quad & 0 = U_{L,t} + \nu_t[U_{C,t}\sigma\eta_t A_t F_{KL,t}\xi_t K_t + U_{CL,t}(N_t - Q_t K_{t+1})] - \gamma_t U_{CL,t}V_t \\
& + \lambda_t^L(U_{CL,t}A_t F_{L,t} + U_{C,t}A_t F_{LL,t} + U_{LL,t}) + \lambda_t^Y A_t F_{L,t} + \mathbf{1}_{\mathbb{N}}(t)\Bigg[\nu_{t-1}U_{CL,t}D_t \\
& + \gamma_{t-1}\eta_t\left\{U_{CL,t}[(1-\sigma)\widetilde{N}_t + \sigma_t^1 V_t] + U_{C,t}\left[(1-\sigma)A_t F_{KL,t}\xi_t K_t + \frac{\partial\sigma_t^1}{\partial L_t}V_t\right]\right\}\Bigg], \\
V_t: \quad & \gamma_t = \mathbf{1}_{\mathbb{N}}(t)\eta_t\sigma_t^1\gamma_{t-1} + \lambda_t.
\end{aligned}
$$

The complementary slackness conditions are

$$
0 = \lambda_t(V_t - \theta Q_t K_{t+1}), \qquad \lambda_t \geq 0.
$$

Conditional on the multipliers $\{\nu_t, \gamma_t\}$, the planner's FOC for D_{t+1} is equivalent to the individual FOC for d_{t+1} if the planner cannot internalize σ^1, that is, when $\frac{\partial\sigma_{t+1}^1}{\partial D_{t+1}} = 0$. However, the original multipliers are generally nonstationary in the CE and stationary in the CEA. We must apply a transformation for closer comparison, as we did with the CE. However, unlike in the CE, we can have $\gamma_t = 0$. To apply the transformation, we will assume that $\gamma_t(z^t) > 0$ for all (t, z^t). We thus consider $t \geq t^*$ such that $\lambda_{t^*} > 0$ and assume that for all $t > t^*$ where $\eta_t = 0$, we have $\lambda_t > 0$. We will not be relying on this assumption in the computation. Define $\widehat{x}_t \equiv \frac{x_t}{\gamma_t}$ and $\bar{x}_t \equiv \frac{\widehat{x}_t}{1-\widehat{\lambda}_t}$ for $x \in \{\nu, \lambda, \lambda^L, \lambda^Y\}$. The FOC for V_t implies $\frac{\gamma_t}{\gamma_{t-1}} = \frac{\mathbf{1}_{\mathbb{N}}(t)\eta_t\sigma_t^1}{1-\widehat{\lambda}_t}$, which implies $\frac{\gamma_{t+1}}{\gamma_t} = \frac{\eta_{t+1}\sigma_{t+1}^1}{1-\widehat{\lambda}_{t+1}}$. We can, therefore, write the FOCs for D_{t+1} and K_{t+1} as

$$
\begin{aligned}
\bar{\nu}_t &= (1+\bar{\lambda}_t)R_t\mathbb{E}_t\left[\eta_{t+1}\Lambda_{t,t+1}\left(1 - \sigma + \sigma_{t+1}^1\sigma\bar{\nu}_{t+1} - \frac{\partial\sigma_{t+1}^1}{\partial D_{t+1}}V_{t+1}\right)\right], \\
\theta\bar{\lambda}_t + \bar{\nu}_t &= (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1 - \sigma + \sigma_{t+1}^1\sigma\bar{\nu}_{t+1})R_{t+1}^K] + \Psi_t^K,
\end{aligned}
$$

where

$$
\begin{aligned}
Q_t\Psi_t^K \equiv\ & \bar{\nu}_t Q_{2,t}\{[\sigma\eta_t(1-\delta)\xi_t+\omega]K_t-K_{t+1}\}+(1+\bar{\lambda}_t)\mathbb{E}_t\left(\eta_{t+1}\Lambda_{t,t+1}\frac{\partial\sigma_{t+1}^1}{\partial K_{t+1}}V_{t+1}\right)-\bar{\lambda}_t\theta Q_{2,t}K_{t+1}\\
& -\frac{\bar{\lambda}_t^Y}{U_{C,t}}I_{2,t}+\frac{\mathbf{1}_{\mathbb{N}}(t)\eta_t}{\sigma_t^1}\left[(1-\sigma)Q_{2,t}(1-\delta)\xi_t K_t+\frac{\partial\sigma_t^1}{\partial K_{t+1}}V_t\right]\\
& +(1+\bar{\lambda}_t)\mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma_{t+1}^1\sigma\bar{\nu}_{t+1})[A_{t+1}F_{KK,t+1}\xi_{t+1}+Q_{1,t+1}(1-\delta)]\xi_{t+1}K_{t+1}\}\\
& +(1+\bar{\lambda}_t)\mathbb{E}_t\Bigg(\eta_{t+1}\Lambda_{t,t+1}\sigma_{t+1}^1\Bigg\{\bar{\nu}_{t+1}[\omega(Q_{1,t+1}K_{t+1}+Q_{t+1})-Q_{1,t+1}K_{t+2}]\\
& -\bar{\lambda}_{t+1}\theta Q_{1,t+1}K_{t+2}+\bar{\lambda}_{t+1}^L A_{t+1}F_{KL,t+1}\xi_{t+1}+\frac{\bar{\lambda}_{t+1}^Y}{U_{C,t+1}}(A_{t+1}F_{K,t+1}\xi_{t+1}-I_{1,t+1})\Bigg\}\Bigg).
\end{aligned}
$$

The arguments for bullet points are as follows.

1. If the EC is always binding at the CEA, the CEA is completely determined by the planner's constraints, and the planner's Euler equations for D_{t+1} and K_{t+1} determine the social Lagrange multipliers $\{\bar{\nu}_t, \bar{\lambda}_t\}$. The implementability constraints can be formulated in recursive form, constituting a system of functional equations on a state-space with endogenous states (D, K) and exogenous states (A, ξ). If the EC is always slack at the CEA, the balance sheet, bank value, and ECs must be redundant. Indeed, the sequence of history-contingent balance sheet constraints determines $\{D_{t+1}\}$, and the sequence of bank value constraints determines $\{V_t\}$ conditional on $\{C_t, K_{t+1}, L_t\}$. The remaining labor market clearing and resource constraints are static. Therefore, the CEA must be recursive. In general, however, due to forward-looking constraints, many KKT conditions are different for $t = 0$ and $t > 0$; hence, the CEA is generally time inconsistent.

2. If the EC is always binding at the CEA, the CEA is determined by the implementability constraints. With $\{\sigma_t^1\} = \{\widehat{\sigma}_t^1\}$, those constraints are necessary for the CE. They are also sufficient, since we can use (1.3) and (1.4) to back out the transformed CE multipliers $\{\bar{\nu}_t, \bar{\lambda}_t\}$, where the condition $\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}f_{t+1}(R_{t+1}^K - R_t)] \geq 0$ applied to $f_t \equiv 1-\sigma+\sigma\bar{\nu}_t$ guarantees that $\bar{\lambda}_t \geq 0$ for all (t, z^t). Therefore, the CEA equals the CE allocation, and the latter is constrained efficient.

3. Given $\{\sigma_t^1\}$, consider any $\bar{Z}^t$ that is of positive measure. The definition of the forward-looking bank value—the aggregate bank value constraint—and the fact that $\sigma_t^1(z^t) < 1$ for

all (t, z^t) imply that for all $z^t \in \bar{Z}^t$, there exists $\{\epsilon_{t+1}(z^{t+1})\}_{z^{t+1}\in Z^{t+1}|z^t}$, where $\epsilon_{t+1}(z^{t+1}) \in (0, 1-\sigma^1_{t+1}(z^{t+1}))$ for all $z^{t+1} \in Z^{t+1} \mid z^t$. Let us define $\{\widetilde{\sigma}^1_t\}$ as follows. For all $z^t \in \bar{Z}^t$, set $\widetilde{\sigma}^1_{t+1}(z^{t+1}) \equiv \sigma^1_{t+1}(z^{t+1}) + \epsilon_{t+1}(z^{t+1})$ for all z^{t+1} that continue from z^t. Set $\widetilde{\sigma}^1_t(z^t) \equiv \sigma^1_t(z^t)$ for all other (t, z^t). The alternative distribution constructed this way would strictly relax ECs at $z^t \in \bar{Z}^t$ and strictly expand the planner's feasible set. Since, conditional on $\{\sigma^1_t\}$, the optimal allocation at $z^t \in \bar{Z}^t$ is strictly at the boundary of the feasible set, the optimal allocation over the expanded feasible set will be strictly outside the boundary. The strict concavity of U and the positive measure of $\bar{Z}^t$ imply that the CEA conditional on $\{\widetilde{\sigma}^1_t\}$ must attain strictly greater welfare than the original CEA. ■

1.A.4 Proposition 1.3

The current planner's best response is

$$V^h(S) = \max_{(C,D',K',L)\in\mathcal{G}(S)} U(C,L) + \beta\mathbb{E}_z(\bar{V}^h(S')),$$

where $\mathcal{G} : \mathcal{S} \to \mathcal{P}(\mathbb{R}^4_+)$ is defined by the constraints

$$\begin{aligned}
\widetilde{\nu}: &\quad 0 = \bar{N} + \sigma\eta_S\widetilde{N}_S + Q(K,K',\xi)(\omega K - K') + \beta\mathbb{E}_z\left(\frac{U_C(\bar{C}_{S'},\bar{L}_{S'})}{U_C(C,L)}\right)D', \\
\widetilde{\lambda}: &\quad 0 \le \beta\mathbb{E}_z\left\{\eta_{S'}\frac{U_C(\bar{C}_{S'},\bar{L}_{S'})}{U_C(C,L)}[(1-\sigma)\widetilde{N}_{S'} + \bar{V}^1_{S'}]\right\} - \theta Q(K,K',\xi)K', \\
\lambda^L: &\quad 0 = U_C(C,L)AF_L(\xi K,L) + U_L(C,L), \\
\lambda^Y: &\quad 0 = AF(\xi K,L) - C - I(K,K',\xi),
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{N}_S &\equiv [AF_K(\xi K,L) + Q(K,K',\xi)(1-\delta)]\xi K - D, \qquad \eta_S \equiv \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_S), \\
\widetilde{N}_{S'} &\equiv [A'F_K(\xi' K',\bar{L}_{S'}) + Q(K',\bar{K}'_{S'},\xi')(1-\delta)]\xi' K' - D', \qquad \eta_{S'} \equiv \mathbf{1}_{\mathbb{R}_{++}}(\widetilde{N}_{S'}).
\end{aligned}$$

Define $\nu_S \equiv \frac{\widetilde{\nu}_S}{U_{C,S}}$ and $\lambda_S \equiv \frac{\widetilde{\lambda}_S}{U_{C,S}}$. The FOCs are

$$C: \quad \lambda^Y_S = U_{C,S} + U_{CC,S}[\nu_S(N_S - Q_S K'_S) - \lambda_S\theta Q_S K'_S] + \lambda^L_S(U_{CC,S}AF_{L,S} + U_{CL,S}),$$

$$D': \quad 0 = \beta \mathbb{E}_z(\bar{V}^h_{D,S'}) + \nu_S U_{C,S} \mathbb{E}_z(\Lambda_{S,S'}) + \lambda_S U_{C,S} \mathbb{E}_z[\eta_{S'} \Lambda_{S,S'} (1-\sigma)(-1)] + \Xi^D_S,$$

$$K': \quad 0 = \beta \mathbb{E}_z(\bar{V}^h_{K,S'}) + \nu_S U_{C,S}[Q_{2,S}\{[\sigma\eta_S(1-\delta)\xi + \omega]K - K'_S\} - Q_S]$$
$$+ \lambda_S U_{C,S}\{\mathbb{E}_z[\eta_{S'}\Lambda_{S,S'}(1-\sigma)\widetilde{N}_{K,S'}] - \theta(Q_{2,S}K'_S + Q_S)\} - \lambda^Y_S I_{2,S} + \Xi^K_S,$$

$$L: \quad 0 = U_{L,S} + \nu_S \left[U_{CL,S}(N_S - Q_S K'_S) + U_{C,S}\sigma\eta_S A F_{KL,S}\xi K\right] - U_{CL,S}\lambda_S\theta Q_S K'_S$$
$$+ \lambda^L_S[A(U_{CL,S}F_{L,S} + U_{C,S}F_{LL,S}) + U_{LL,S}] + \lambda^Y_S A F_{L,S},$$

where $\widetilde{N}_{K,S} \equiv [AF_{KK,S}\xi + Q_{1,S}(1-\delta)]\xi K + [AF_{K,S} + Q_S(1-\delta)]\xi$, and for $X \in \{D, K\}$,

$$\Xi^X_S \equiv \beta\nu_S \mathbb{E}_z(U_{CC,S'}\bar{C}_{X,S'} + U_{CL,S'}\bar{L}_{X,S'})D'_S + \beta\lambda_S\mathbb{E}_z(\eta_{S'}[(U_{CC,S'}\bar{C}_{X,S'} + U_{CL,S'}\bar{L}_{X,S'})$$
$$\times [(1-\sigma)\widetilde{N}_{S'} + \bar{V}^1_{S'}] + U_{C,S'}\{(1-\sigma)[A'F_{KL,S'}\bar{L}_{X,S'} + Q_{2,S'}\bar{K}'_{X,S'}(1-\delta)]\xi' K'_S + \bar{V}^1_{X,S'}\}])$$

is the combined marginal effect of X' on the current planner's Lagrangian through the policy functions of the future planner $\bar{C}$, $\bar{L}$, $\bar{K}'$, and $\bar{V}^1$. The envelope conditions are

$$V^h_{D,S} = -\nu_S U_{C,S}\sigma\eta_S,$$
$$V^h_{K,S} = \nu_S U_{C,S}[\sigma\eta_S\widetilde{N}_{K,S} + \omega(Q_{1,S}K + Q_S) - Q_{1,S}K'_S] - \lambda_S U_{C,S}\theta Q_{1,S}K'_S + \lambda^L_S U_{C,S} A F_{KL,S}\xi$$
$$+ \lambda^Y_S(AF_{K,S}\xi - I_{1,S}).$$

The complementary slackness conditions are

$$0 = \lambda_S[\mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\widetilde{N}_{S'} + \bar{V}^1_{S'}]\} - \theta Q_S K'_S], \qquad \lambda_S \geq 0.$$

Substituting the envelope conditions in the FOCs for D' and K' and rearranging, we can write

$$\nu_S = R_S \mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\lambda_S + \sigma\nu_{S'}]\} - \frac{R_S \Xi^D_S}{U_{C,S}},$$
$$\theta\lambda_S + \nu_S = \mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}[(1-\sigma)\lambda_S + \sigma\nu_{S'}]R^K_{S'}\} + \Omega^K_S + \frac{\Xi^K_S}{Q_S U_{C,S}},$$

where

$$Q_S\Omega_S^K \equiv \nu_S Q_{2,S}\{[\sigma\eta_S(1-\delta)\xi+\omega]K - K_S'\} - \lambda_S\theta Q_{2,S}K_S' - \frac{\lambda_S^Y}{U_{C,S}}I_{2,S} + \mathbb{E}_z\{\eta_{S'}\Lambda_{S,S'}$$
$$\times[(1-\sigma)\lambda_S+\sigma\nu_{S'}][A'F_{KK,S'}\xi' + Q_{1,S'}(1-\delta)]\xi' K_S'\} + \mathbb{E}_z\Bigg(\Lambda_{S,S'}\Bigg\{\nu_{S'}[\omega(Q_{1,S'}K_S' + Q_{S'})$$
$$- Q_{1,S'}K_{S'}'] - \lambda_{S'}\theta Q_{1,S'}K_{S'}' + \lambda_{S'}^L A'F_{KL,S'}\xi' + \frac{\lambda_{S'}^Y}{U_{C,S'}}(A'F_{K,S'}\xi' - I_{1,S'})\Bigg\}\Bigg). \quad \blacksquare$$

1.A.5 Proposition 1.4

We will prove the CEA implementation. The arguments for the MCEA implementation are identical.

Decentralization with linear taxes and transfers

Consider a regulated CE with linear taxes on bank assets τ_t^K and liabilities τ_t^D and a lump-sum transfer $\tau_{i,t}$. The proportional taxes are common to all bankers, the lump-sum transfer is individual-specific, and the taxation is balanced in the aggregate so that (1.7) holds. The balance sheet of an individual bank is

$$(1+\tau_t^K)Q_t s_{i,t} \le n_{i,t} + (1-\tau_t^D)\frac{d_{i,t+1}}{R_t} + \tau_{i,t}.$$

Define the aggregate lump-sum transfer $T_t^b \equiv \int_0^f \tau_{i,t}\,\mathrm{d}i$ that satisfies the government budget constraint $T_t^b = \tau_t^D\frac{D_{t+1}}{R_t} + \tau_t^K Q_t K_{t+1}$. With taxes, the individual Euler equations are

$$(1-\tau_t^D)\bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})]R_t,$$
$$\theta\bar{\lambda}_t + (1+\tau_t^K)\bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})R_{t+1}^K],$$

Many policies can implement the CEA. Consider next several arrangements that address the Euler equation distortions. We will focus on ensuring a given bank value distribution at the end. When we refer to any variables based on allocations, they will be based on the CEA. We must show that there exist policies and Lagrange multipliers that, combined with the CEA, satisfy the banker's Euler equations and the aggregate complementary slackness conditions in the regulated CE. In other words, we use the primal approach to construct the policies.

Suppose $\tau_{i,t} = 0$ for all i and t, so that we must have $\tau_t^K = -\tau_t^D\frac{D_{t+1}}{R_t Q_t K_{t+1}}$. In this case, the individual bank value is still linear in the individual net worth, that is, $v_t = \bar{\nu}_t n_t$, which implies $V_t = \bar{\nu}_t N_t$. We can, therefore, use the CEA to back out $\bar{\nu}_t = \frac{V_t}{N_t}$. Note that by substituting the first Euler equation above into the second one and applying some algebra, we can arrive at the definition of the aggregate bank value, which makes the second Euler equation redundant in our construction—it will be satisfied for any $\{\tau_t^D\}$ given the solution for $\{\bar{\nu}_t\}$ and the tax balance condition. We thus have one equation—the regulated deposit Euler equation—to determine $\{\bar{\lambda}_t, \tau_t^D\}$ consistent with the CEA. If the EC is slack in the CEA, the policy must be such that $\bar{\lambda}_t = 0$, in which case we solve

$$\tau_t^D = 1 - \frac{\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})]R_t}{\bar{\nu}_t}. \tag{1.22}$$

If the EC is binding in the CEA, infinitely many policies are consistent with the latter. Without loss of generality, we can focus on the policy, at which $\bar{\lambda}_t = 0$. Intuitively, this case corresponds to an upper bound for τ_t^D. Therefore, (1.22) can be applied in all contingencies.

Suppose $\tau_{i,t} \neq 0$ for some i, and we have a restriction $\int_0^f \tau_{i,t}\,\mathrm{d}i = 0$. Hence, we must balance the budget with τ_t^K as above. The two regulated Euler equations put a restriction on three unknowns $\{\bar{\lambda}_t, \bar{\nu}_t, \tau_t^D\}$. When the EC is slack, we get a unique solution for current variables conditional on the future ones. If the EC is binding, there is a multiplicity of solutions, and we can again consider the one that implies $\bar{\lambda}_t = 0$. Using (1.22) to determine the deposit tax and substituting in the asset Euler equation, we obtain

$$\bar{\nu}_t = \frac{\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})\widetilde{N}_{t+1}]}{N_t}.$$

As in the previous paragraph, suppose $\tau_{i,t} \neq 0$ for some i, but we can use T_t^b to balance the budget. In this case, we can set $\tau_t^K = 0$. If $\bar{\lambda}_t = 0$ as before, the second Euler equation immediately implies

$$\bar{\nu}_t = \mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}(1-\sigma+\sigma\bar{\nu}_{t+1})R_{t+1}^K].$$

We determine the deposit tax from (1.22).

Finally, if $\tau_{i,t} \neq 0$, the individual bank value is not linear in net worth, and we must ensure that the distribution σ^1 implied by the regulated CE coincides with the distribution taken as given by the CEA planner. Define $\mu_{i,t} \equiv v_{i,t} - \bar{\nu}_t n_{i,t}$. Using the definition of the bank value, the Euler equations, the complementary slackness conditions, and the bank balance sheet constraint, we find that $\{\mu_{i,t}\}$ solves a stochastic difference equation:

$$\mu_{i,t} = \bar{\nu}_t \tau_{i,t} + (1 + \bar{\lambda}_t)\sigma \mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{i,t+1}).$$

Consider entrants/survivors-specific transfers in the sense that all survivors get $\frac{1}{\sigma f}\tau_t^1$ independently of the level of individual net worth, and, similarly, all entrants get $\frac{1}{(1-\sigma)f}\tau_t^0$. By definition, $T_t^b = \tau_t^0 + \tau_t^1$, so that $\tau_t^0 = T_t^b - \tau_t^1$, where T_t^b is already known from the construction above. Since the transfer is identical across entrants or survivors, the value function intercepts are too within those groups. Define $\mu_t^1 \equiv \int_0^{\sigma f} \mu_{i,t}\, \mathrm{d}i$, $\mu_t^0 \equiv \int_{\sigma f}^{f} \mu_{i,t}\, \mathrm{d}i$, and $\mu_t \equiv \mu_t^0 + \mu_t^1$. Remember that we use a convention that survivors are always ordered to be in the $[0, \sigma f]$ interval. We, therefore, have

$$\begin{aligned}
\mu_t^1 &= \bar{\nu}_t \tau_t^1 + (1 + \bar{\lambda}_t)\sigma \mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}^1), \\
\mu_t^0 &= \bar{\nu}_t \tau_t^0 + (1 + \bar{\lambda}_t)(1-\sigma) \mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}^1), \\
\mu_t &= \bar{\nu}_t T_t^b + (1 + \bar{\lambda}_t) \mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}^1).
\end{aligned}$$

Note that conditional on $\{\tau_t^1\}$, we can back out $\{\mu_t^1, \mu_t^0, \mu_t\}$ since other objects are either part of the CEA or follow from the construction above. The distribution consistency requires choosing $\{\tau_t^1\}$ such that $\sigma_t^1 = \frac{V_t^1}{V_t}$ for all t. Aggregating the individual bank values and using the above relationships, we get an implicit solution for τ_t^1:

$$\tau_t^1 = \sigma_t^1(N_t + T_t^b) - \eta_t \sigma \widetilde{N}_t + \frac{(\sigma_t^1 - \sigma)(1 + \bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}^1)}{\bar{\nu}_t}.$$

Substituting it back into the fixed-point equation in $\{\mu_t^1\}$, we obtain

$$\mu_t^1 = \bar{\nu}_t[\sigma_t^1(N_t + T_t^b) - \eta_t \sigma \widetilde{N}_t] + \sigma_t^1(1 + \bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{t+1}^1).$$

As discussed above, we can set $\bar{\lambda}_t = 0$ without loss of generality in all the previous expressions. Let

$\bar{N}_t^1 \equiv \bar{\nu}_t[\sigma_t^1(N_t + T_t^b) - \eta_t\sigma\widetilde{N}_t]$. Solving forward the last equation, we obtain

$$\mu_t^1 = \mathbb{E}_t\left[\sum_{i=0}^{\infty}\left(\prod_{j=0}^{i-1}\sigma_{t+j}^1\eta_{t+1+j}\right)^{\mathbf{1}_{\mathbb{N}}(i)}\Lambda_{t,t+i}\bar{N}_{t+i}^1\right].$$

Therefore,

$$\tau_t^1 = \frac{1}{\bar{\nu}_t}\left\{\bar{N}_t^1 + (\sigma_t^1 - \sigma)\mathbb{E}_t\left[\eta_{t+1}\sum_{i=0}^{\infty}\left(\prod_{j=0}^{i-1}\sigma_{t+1+j}^1\eta_{t+2+j}\right)^{\mathbf{1}_{\mathbb{N}}(i)}\Lambda_{t,t+1+i}\bar{N}_{t+1+i}^1\right]\right\}.$$

Decentralization with capital requirements and transfers

Now suppose that the policymaker can impose state-contingent capital requirements $\kappa_t \in [0,1]$ to address the Euler equation distortions. We still allow for entrants/survivors-specific transfers to implement the targeted distribution. The budget and regulatory constraints for the individual banker are

$$Q_t s_{i,t} \leq n_{i,t} + \frac{d_{i,t+1}}{R_t} + \tau_{i,t},$$
$$n_{i,t} \geq \kappa_t Q_t s_{i,t},$$

with the government budget constraint $\int_0^f \tau_{i,t}\,\mathrm{d}i = 0$. Let ξ_t be the Lagrange multiplier on the regulatory constraint. The Euler equations are now

$$\bar{\nu}_t = (1+\bar{\lambda}_t)R_t\mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[1-\sigma+\sigma(\bar{\nu}_{t+1}+\bar{\xi}_{t+1})]\},$$
$$\kappa_t\bar{\xi}_t + \theta\bar{\lambda}_t + \bar{\nu}_t = (1+\bar{\lambda}_t)\mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[1-\sigma+\sigma(\bar{\nu}_{t+1}+\bar{\xi}_{t+1})]R_{t+1}^K\},$$

where $\bar{\xi}_t \equiv \frac{\xi_t}{\gamma_t(1-\widehat{\lambda}_t)}$, consistent with the notation for other multipliers. Aggregating the individual complementary slackness conditions associated with the regulatory constraint, we have

$$0 = \bar{\xi}_t(N_t - \kappa_t Q_t K_{t+1}), \qquad \bar{\xi}_t \geq 0, \qquad N_t \geq \kappa_t Q_t K_{t+1}.$$

At any contingency, we must have $\kappa_t \leq \frac{N_t}{Q_t K_{t+1}}$, where the right-hand side is evaluated at the CEA.

If $V_t > \theta Q_t K_{t+1}$, we must have $\bar{\lambda}_t = 0$, and conditional on $\{\bar{\nu}_{t+1}, \bar{\xi}_{t+1}\}$,

$$\bar{\nu}_t = R_t \mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[1-\sigma+\sigma(\bar{\nu}_{t+1}+\bar{\xi}_{t+1})]\},$$
$$\bar{\xi}_t = \frac{\mathbb{E}_t\{\eta_{t+1}\Lambda_{t,t+1}[1-\sigma+\sigma(\bar{\nu}_{t+1}+\bar{\xi}_{t+1})]R^K_{t+1}\} - \bar{\nu}_t}{\kappa_t}.$$

If $\bar{\xi}_t > 0$, the capital requirement constraint is binding and $\kappa_t = \frac{N_t}{Q_t K_{t+1}}$. If $\bar{\xi}_t = 0$, any $\kappa_t \leq \frac{N_t}{Q_t K_{t+1}}$ is admissible. If $\bar{\xi}_t < 0$, the complementary slackness condition is violated, and the implementation with capital requirements alone fails. A sufficient condition for $\bar{\xi}_t \geq 0$ is $\mathbb{E}_t[\eta_{t+1}\Lambda_{t,t+1}f_{t+1}(R^K_{t+1} - R_t)] \geq 0$ for all positive-valued f_{t+1}. If $V_t = \theta Q_t K_{t+1}$, we can focus on the case of $\bar{\lambda}_t = 0$ as with proportional taxation, in which case the construction remains unchanged. The actual computation will involve finding a fixed point in the system of stochastic difference equations.

Now consider how the individual bank value changes with the presence of the regulatory constraint. Define $\mu_{i,t} \equiv v_{i,t} - (\bar{\nu}_t + \bar{\xi}_t)n_{i,t}$. Using the definition of the bank value, the Euler equations, the complementary slackness conditions, and the bank balance sheet constraint, we find that $\{\mu_{i,t}\}$ solves the identical stochastic difference equation as in the case of proportional taxation:

$$\mu_{i,t} = \bar{\nu}_t \tau_{i,t} + (1+\bar{\lambda}_t)\sigma \mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu_{i,t+1}).$$

Note that if $\tau_{i,t} = 0$ for all i, the individual bank value is linear in net worth, so we could use the relationship $V_t = (\bar{\nu}_t + \bar{\xi}_t)N_t$ instead of one of the Euler equations to back out the transformed multipliers. Otherwise, we proceed in the same way as with proportional taxation to find the required transfer to survivors. The aggregate bank value intercepts for entrants and survivors take the same expressions. Targeting $\sigma^1_t = \frac{V^1_t}{V_t}$ requires setting

$$\tau^1_t = \left(1 + \frac{\bar{\xi}_t}{\bar{\nu}_t}\right)(\sigma^1_t N_t - \eta_t \sigma \widetilde{N}_t) + \frac{(\sigma^1_t - \sigma)(1+\bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu^1_{t+1})}{\bar{\nu}_t}.$$

Substituting it back into the fixed-point equation in $\{\mu^1_t\}$, we obtain

$$\mu^1_t = (\bar{\nu}_t + \bar{\xi}_t)(\sigma^1_t N_t - \eta_t \sigma \widetilde{N}_t) + \sigma^1_t (1+\bar{\lambda}_t)\mathbb{E}_t(\eta_{t+1}\Lambda_{t,t+1}\mu^1_{t+1}).$$

As discussed above, we can set $\bar{\lambda}_t = 0$ without loss of generality in all the previous expressions.

Ramsey equilibrium

Consider a Ramsey planner that takes the determination of the distribution σ^1 as given and chooses optimally either proportional taxes on bank assets and liabilities rebated lump sum or capital requirements. The planner also sets the entrants/survivors-specific transfers that achieve the targeted distribution.

Suppose the planner has access to τ_t^D and τ_t^K, which must satisfy the government budget constraint $0 = \tau_t^D \frac{D_{t+1}}{R_t} + \tau_t^K Q_t K_{t+1}$. The regulated Euler equations for deposits and securities determine $\{\tau_t^D, \bar{\nu}_t\}$ conditional on other variables. Guess that the aggregate complementary slackness conditions associated with the EC are not binding in the sense that they can be constructed ex post. The constraint set of the Ramsey planner is then equivalent to that in the CEA. If the EC is slack in the CEA, set $\bar{\lambda}_t = 0$. Otherwise, there are multiple choices, and one can set $\bar{\lambda}_t = 0$ without loss of generality. The CE complementary slackness conditions are satisfied. Set the aggregate transfer to survivors τ_t^1 as discussed in the CEA decentralization part to achieve the targeted distribution. Hence, conditional on $\{\sigma_t^1\}$, $\{\tau_t^D, \tau_t^K, \tau_t^1\}$ and the CEA constitute a Ramsey equilibrium.

The argument for other cases—either τ_t^D and T_t^b or capital requirements κ_t—is identical to the above. In the former case, we use the regulated Euler equations to solve for τ_t^D and $\bar{\nu}_t$, and the CE complementary slackness conditions can be constructed ex post. In the latter case, the Euler equations determine $\bar{\nu}_t$ and $\bar{\xi}_t$, and the two sets of complementary slackness conditions can be constructed ex post. Specifically, we can set the regulatory constraint to be always binding without loss of generality. (The regulatory constraint gives us κ_t.) ■

Chapter 2

Sovereign risk, banking crises, and macroprudential policy

2.1 Introduction

Since banks hold government debt, sovereign default risk transmits to the banking sector, increasing the likelihood of banking crises. This issue was documented empirically by Reinhart and Rogoff (2009) and has recently been studied theoretically by Bocola (2016), among others. I extend Bocola (2016)'s model and take a normative perspective: what should be the policy response in such an environment?

The policy I study is macroprudential, affecting the whole banking sector uniformly. I consider various taxes on bank balance sheet components as proxy policy instruments. I found that a proportional tax that subsidizes lending to the real sector in both good times and bad times—when a country has defaulted on its debt or is in exclusion from financial markets—improves welfare and reduces the probability of banking crises. The reason is that subsidizing credit to the real sector instead of investing in sovereign bonds strengthens the economy in good times and helps it recover faster in bad times. A bank net worth subsidy in good times combined with a tax in bad times is also welfare-improving and facilitates the prevention of banking crises. This finding is consistent with the overborrowing story explored recently by Bianchi (2011), Bianchi and Mendoza (2018), and Jeanne and Korinek (2019).

The model I use builds on Bocola (2016), which is, in turn, a version of Gertler and Kiyotaki (2010) and Gertler and Karadi (2011) with government borrowing and exogenous default risk. I extend Bocola (2016)'s model in several dimensions. First, like in the contemporary quantitative sovereign default literature started by Aguiar and Gopinath (2006) and Arellano (2008), I assume that when the government defaults on its debt, it is excluded from financial markets for a random number of periods. I capture this by specifying the exogenous "financial standing" process (in the language of Na et al. (2018)) as an irreducible two-state Markov chain. This specification implies that, unlike in Bocola (2016), the sovereign cannot borrow in the same period it defaults. Second, following the sovereign default literature, I open the economy, allowing for international borrowing. The presence of external debt introduces an important additional margin: while the default is harmful to the economy, it frees up additional resources that can be used domestically. Unlike in basic endogenous default models, the default cost is endogenous in my framework. The source is the same as in Bocola (2016): when the government defaults on its debt, banks are hurt, which transmits to the real economy through the credit channel.

The paper is related to various strands of literature. First, it advances the literature that studies the interaction between banking and sovereign debt in a quantitative framework. As described above, the paper builds on Bocola (2016), extending it and studying government policy. Perez (2018) also builds on Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), significantly simplifying it, and merges the framework with the endogenous sovereign default model in line with Eaton and Gersovitz (1981). He then considers two policy measures. The first measure is a proportional subsidy for banks on sovereign debt holdings—he finds a moderate subsidy to be welfare-enhancing due to a disciplining effect on the government through raising the endogenous default cost. He then studies the effects of a floor on bank sovereign debt holdings and finds this measure to be welfare-reducing. Compared to Perez (2018), I do not have endogenous sovereign default, but I have a richer economy and study other policy instruments. Similar to Perez (2018), Sosa-Padilla (2018) merges a simple banking model with an endogenous default model. Importantly, there is no external debt in the model. He then uses the model to explain the data, abstaining from the normative analysis. Both authors need to impose significant simplifying assumptions to avoid the curse of dimensionality since they need to solve for the Markov perfect equilibrium using value function iteration. On the contrary, Ari (2018) limits his analysis to exogenous sovereign default risk, modeled through a shock that follows logistic distribution as in Bocola (2016). In his model, some bankers only care about non-sovereign-default states, which leads to a possibility of bank default, reflected in limited liability. Apart from a logistic shock, there is no other aggregate risk in the model. The author does need to introduce an exogenous default cost through a lower productivity constant in default states. The author also assumes that the economy is in the steady state after the sovereign default. He then studies several liquidity provision measures, deposit insurance, and a leverage constraint.

The paper is organized as follows. Section 2.2 describes the model. Section 2.3 investigates several quantitative properties of the model economy. Section 2.4 conducts the normative analysis. Appendix contains computational details.

2.2 Model

The model features households (workers and bankers), final and capital good producers, foreign lenders, and the government—the fiscal and macroprudential authorities. Workers supply labor to final good producers and save through risk-free bank deposits. Bankers use deposits and net worth to lend to final good producers and the government by purchasing sovereign bonds. Final good producers need loans to purchase physical capital from capital good producers. Foreign lenders are risk neutral. The government borrows domestically and abroad and implements a tax/transfer policy on workers and bankers to finance government expenditures.

As in Bocola (2016), I use a stationary recursive notation for all the problems, denoting next period objects with a prime. The aggregate state vector is $S \equiv (A_B, B, K, P, \mathcal{I}, z)$, where A_B are domestic bond holdings, B are total bond holdings, K is physical capital, P are deposits to be repaid, z is the log of firm productivity, and $\mathcal{I}$ is defined below.

2.2.1 Sovereign debt market

Define the random variable

$$\mathcal{I} = \begin{cases} 1 & \text{if sovereign debt market is open} \\ 0 & \text{if sovereign debt market is closed} \end{cases} .$$

This notation is the same as in Na et al. (2018), except that $\mathcal{I}$ is endogenous in their case.

I assume that $\mathcal{I}$ is a Markov chain with the transition matrix

$$\Pi \equiv \begin{pmatrix} \Pi_{11} & \Pi_{10} \\ \Pi_{01} & \Pi_{00} \end{pmatrix}.$$

The first row describes the transition from $\mathcal{I} = 1$: the probabilities that the government repays (Π_{11}) or defaults (Π_{10}) on its debt. The second row describes the transition from $\mathcal{I} = 0$. When $\mathcal{I} = 0$, the government cannot issue new debt and is in financial exclusion, which is persistent, following Aguiar and Gopinath (2006) and Arellano (2008). Thus, Π_{01} is the probability of regaining access to financial markets, while Π_{00} is the probability of remaining in exclusion.

Let Q_B denote the sovereign bond price. W.l.o.g. and to simplify notation, assume that $Q_B(S) = 1$ when $\mathcal{I} = 0$. Any non-zero number would work too.

2.2.2 Households

Each household has a unit measure of members: a measure $1 - f \in (0, 1)$ are workers and f are bankers. Workers and bankers switch occupations with constant probabilities independent of previous histories. Workers consume the final good, supply labor to final good producers, and save through bank deposits (in banks managed by bankers from other households), which are one-period risk-free bonds. A household's problem is

$$v^h(b; S) = \max_{c, b', l} \left\{ u(c, l) + \beta \mathbb{E}_S v^h(b'; S') \right\}$$

subject to

$$\begin{aligned} c + \frac{b'}{R(S)} &\leq b + (1 - \tau_L(S)) W(S) l + \Pi(S) - T^h(S), \\ S' &= \Gamma(S), \end{aligned}$$

where c is consumption, b are bank deposits, l is labor, $u(\cdot)$ is strictly concave, strictly increasing in c, strictly decreasing in l, satisfies Inada conditions, $\beta \in (0, 1)$ is the discount factor, R is the gross deposit rate, W is wage, Π are profits of capital good producers and transfers from exiting bankers net of the provision of initial net worth to new bankers, T^h is lump-sum tax from the government, Γ is the law of motion for S.

The consumption Euler equation and labor supply equation are

$$1 = R(S)\mathbb{E}_S\{\Lambda(S', S)\}, \tag{2.1}$$

$$(1 - \tau_L(S))W(S) = -\frac{u_l(C(S), L(S))}{u_c(C(S), L(S))}, \tag{2.2}$$

where C and L are aggregate consumption and labor, while

$$\Lambda(S', S) \equiv \beta \frac{u_c(C(S'), L(S'))}{u_c(C(S), L(S))} \tag{2.3}$$

is the stochastic discount factor, common to all households since they are identical.

2.2.3 Bankers

A banker manages a bank that collects deposits from households (distinct from the banker's household) and invests in government bonds and the real sector. A banker stays in the banking business next period with a probability $\psi \in (0, 1)$, otherwise becoming a worker. In the latter case, the banker transfers the accumulated net worth to her household. Hence, $(1 - \psi)f$ bankers become workers each period, while the same measure of workers become bankers, keeping the composition of households unchanged. A household provides an initial net worth to its new bankers, which is based on the value of assets of bankers that have exited the business.

The balance sheet of a bank is

$$\sum_{j \in \{B,K\}} (1 + \tau_j(S))Q_j(S)a'_j = (1 + \tau_N(S))n + (1 - \tau_P(S))\frac{b'}{R(S)} - \frac{T^b(S)}{f},$$

where Q_B and Q_K are sovereign and corporate bond prices, a'_B and a'_K are the corresponding quantities demanded[1], n is a bank's net worth. Finally, τ_j for $j \in \{B, K, N, P\}$ are proportional tax/subsidy rates, T^b is the lump-sum tax—the *macroprudential instruments*.

[1] One corporate bond corresponds to one unit of physical capital purchased by producers. The price of these bonds is the same as the price of capital by no-arbitrage.

The next period net worth of a continuing banker is defined as

$$n' = \sum_{j \in \{B,K\}} R_j(S', S) Q_j(S) a'_j - b',$$

where R_B and R_K are gross returns to be defined later.

The balance sheet identity and the net worth definition can be combined to yield the evolution of a bank's net worth:

$$n' = \sum_{j \in \{B,K\}} \left[R_j(S', S) - \hat{R}_j(S) \right] Q_j(S) a'_j + \hat{R}(S) n - \frac{\hat{T}(S)}{f},$$

where

$$\hat{R}_j(S) \equiv \frac{1 + \tau_j(S)}{1 - \tau_P(S)} R(S), \qquad \hat{R}(S) \equiv \frac{1 + \tau_N(S)}{1 - \tau_P(S)} R(S), \qquad \hat{T}(S) \equiv \frac{1}{1 - \tau_P(S)} R(S) T^b(S).$$

It is assumed that after borrowing, lending, and paying taxes, a banker may divert a fraction $\lambda \in (0, 1)$ of assets and transfer them to her household. In that case, the bank defaults, and other households can recover the remaining fraction $1 - \lambda$ of the bank's assets. Anticipating such a possibility, households require a banker to satisfy an appropriate incentive compatibility constraint such that default never happens.

If $\mathcal{I} = 1$, the problem is

$$v^b(n; S) = \max_{a'_B, a'_K} \mathbb{E}_S \left\{ \Lambda(S', S) \left[(1 - \psi) n' + \psi v^b(n'; S') \right] \right\}$$

subject to

$$\begin{aligned} n' &= \sum_{j \in \{B,K\}} \left[R_j(S', S) - \hat{R}_j(S) \right] Q_j(S) a'_j + \hat{R}(S) n - \frac{\hat{T}(S)}{f}, \\ v^b(n; S) &\geq \lambda \sum_{j \in \{B,K\}} Q_j(S) a'_j, \\ S' &= \Gamma(S). \end{aligned}$$

If $\mathcal{I} = 0$, a'_B is not chosen ($a'_B = 0$).

The solution features

$$v^b(n;S) = \alpha_1(S)n + \frac{1}{f}\alpha_2(S),$$

where

$$\alpha_1(S) = \frac{\mathbb{E}_S\left\{\hat{\Lambda}(S',S)\right\}\hat{R}(S)}{1-\mu(S)}, \tag{2.4}$$

$$\alpha_2(S) = \frac{\psi\mathbb{E}_S\{\Lambda(S',S)\alpha_2(S')\} - \mathbb{E}_S\left\{\hat{\Lambda}(S',S)\right\}\hat{T}(S)}{1-\mu(S)}, \tag{2.5}$$

μ is the Lagrange multiplier on the incentive-compatibility constraint, common to all bankers due to linearity, and

$$\hat{\Lambda}(S',S) \equiv \Lambda(S',S)(1-\psi+\psi\alpha_1(S')).$$

A proof of this result is straightforward (guess-and-verify). It is crucial that α_1 and α_2 are functions of S only, but not of n or other individual variables. This fact is a consequence of the linearity of the problem and allows easy aggregation. Also, notice that $\alpha_2(S) \equiv 0$ is a solution when $T^b(S) \equiv 0$.

Taking into account the form of the value function, the optimality conditions at the banking sector level are

$$0 = \begin{cases} \mathbb{E}_S\left\{\hat{\Lambda}(S',S)\left[R_B(S',S) - \hat{R}_B(S)\right]\right\} - \lambda\mu(S) & \text{if } \mathcal{I}=1 \\ A'_B(S) & \text{if } \mathcal{I}=0 \end{cases}, \tag{2.6}$$

$$0 = \mathbb{E}_S\left\{\hat{\Lambda}(S',S)\left[R_K(S',S) - \hat{R}_K(S)\right]\right\} - \lambda\mu(S), \tag{2.7}$$

$$0 = \mu(S)\left[\alpha_1(S)N(S) + \alpha_2(S) - \lambda\sum_{j\in\{B,K\}} Q_j(S)A'_j(S)\right], \tag{2.8}$$

where A_B, A_K, and N are the aggregate sovereign and corporate bond holdings and net worth of the banking sector. The first two are Euler equations for sovereign and corporate bonds. The last one is the complementary slackness condition.

The aggregation also implies

$$\frac{P'(S)}{R(S)} = \sum_{j\in\{B,K\}} Q_j(S)A'_j(S) + \Upsilon(S) - N(S), \tag{2.9}$$

$$N(S') = \psi \left[\sum_{j \in \{B,K\}} R_j(S', S) Q_j(S) A'_j(S) - P'(S) \right] + \omega \sum_{j \in \{B,K\}} \mathcal{I}'_j Q_j(S') A'_j(S), \tag{2.10}$$

where P are the aggregate deposits, $\omega \in (0, 1-\psi]$ controls the share of aggregate assets of exiting bankers transferred to new bankers, $\mathcal{I}_B \equiv \mathcal{I}$, $\mathcal{I}_K \equiv 1$, and

$$\Upsilon(S) \equiv \sum_{j \in \{B,K\}} \tau_j(S) Q_j(S) A'_j(S) + \tau_P(S) \frac{P'(S)}{R(S)} + T^b(S) - \tau_N(S) N(S) \tag{2.11}$$

is the budget surplus of the macroprudential authority.

2.2.4 Final good producers

Anticipating market clearing, the aggregate production function is

$$Y(S) = K^\theta (e^z L(S))^{1-\theta}, \tag{2.12}$$

where Y is output, K is capital, L is labor, z is the exogenous productivity, and $\theta \in (0, 1)$.

The labor demand is determined by its marginal product:

$$W(S) = (1-\theta) \frac{Y(S)}{L(S)}. \tag{2.13}$$

Final good producers need to borrow from banks to purchase capital for production next period. The financial contract specifies that the gross revenue net of labor costs, together with the value of the undepreciated capital, is used to repay the loan. Hence, the gross rate of return on bank loans satisfies

$$R_K(S', S) \equiv \frac{\theta \frac{Y(S')}{K'(S)} + Q_K(S')(1-\delta)}{Q_K(S)}. \tag{2.14}$$

2.2.5 Capital good producers

The capital good is produced using the final good as an input, subject to adjustment costs. A producer's problem is

$$v^c(S) = \max_i \left\{ Q_K(S) \Phi \left(\frac{i}{K} \right) K - i \right\},$$

where i is the quantity of the final good used in production, $\Phi(\cdot)$ is strictly increasing and strictly concave. Hence,

$$Q_K(S) = \left[\Phi'\left(\frac{I(S)}{K}\right)\right]^{-1}, \tag{2.15}$$

where $I(S)$ is the aggregate investment.

I follow Bocola (2016) in using the specification of adjustment costs above (Lucas and Prescott, 1971), as opposed to the one in Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), because it allows having one less conditional expectation term, which is important when solving the model globally.

2.2.6 Foreign lenders

Following Leland (1994) and Chatterjee and Eyigungor (2012), debt contracts are long term, and each bond has a probability $\pi \in [0,1]$ to mature in any period, which is independent of the previous history. A bond provides a coupon payment, which I assume is also paid in the period of maturity, as in Aguiar et al. (2016). Hence, if $\mathcal{I} = 1$, a share π of a bond portfolio matures and provides the face value and coupon payment, while a share $1 - \pi$ provides the coupon payment and the continuation value based on the current bond price in the secondary market. The bond prices in the secondary and primary markets are identical because a bond issued in the past and sold in the secondary market today has the same expected payoff as a new bond issued in the current period. If $\mathcal{I} = 0$, no one would buy bonds in the secondary market, the value of Q_B is irrelevant and can be set to 1 w.l.o.g. as assumed.

If $\mathcal{I} = 1$, a foreign lender solves

$$\max_{a_B^{*\prime}} \left\{ -Q_B(S)a_B^* + \frac{1}{R^*(S)}\mathbb{E}_S\left\{\mathcal{I}'[\bar{\pi} + (1-\pi)(\iota + Q_B(S'))]a_B^{*\prime}\right\}\right\},$$

where $a_B^{*\prime}$ is the quantity of bonds demanded, D is the fraction of bonds defaulted on in the case of default, $\bar{\pi} \equiv (1+\iota)\pi$, and r is the required expected return—the opportunity cost that satisfies

$$R^*(S) = 1 + \bar{r} + h\left(\frac{A_B^{*\prime}(S)}{Y(S)}\right), \tag{2.16}$$

where $\bar{r} \geq 0$, A_B^* is the aggregate external debt, $h(\cdot)$ is a strictly increasing function that determines the country-specific risk premium required by foreign lenders. A non-degenerate risk premium is required to achieve a well-defined domestic/external bond allocation. This issue is similar to solving a non-stationarity problem in small open economy models, as summarized by Schmitt-Grohé and Uribe (2003). A similar interest rate specification was used as early as by Murphy (1991).

The return on government bonds is

$$R_B(S', S) \equiv \mathcal{I}' \frac{\bar{\pi} + (1-\pi)(\iota + Q_B(S'))}{Q_B(S)}. \tag{2.17}$$

Hence,

$$Q_B(S) = \mathcal{I} \frac{\mathbb{E}_S \left\{ \mathcal{I}'[\bar{\pi} + (1-\pi)(\iota + Q_B(S'))] \right\}}{R^*(S)} + 1 - \mathcal{I}, \tag{2.18}$$

where $Q_B(S) = 1$ when $\mathcal{I} = 0$ w.l.o.g.

2.2.7 Government

The government budget constraint is

$$gY(S) = \tau_L(S)W(S)L(S) + T^h(S) + \Upsilon(S) + \mathcal{I}\{Q_B(S)B'(S) - [\bar{\pi} + (1-\pi)(\iota + Q_B(S))]B\}, \tag{2.19}$$

where B is the stock of government bonds, $g \geq 0$ is a constant ratio of government expenditures to output.

If $\mathcal{I} = 1$, the government sets $T^h(S)$ according to a fiscal rule described below (as in Bocola (2016)), while the budget constraint determines the required amount of debt. If $\mathcal{I} = 0$, the government cannot borrow and thus cannot set $T^h(S)$ independently. Instead, the budget constraint determines the required $T^h(S)$, which is captured by (2.19) together with

$$0 = \begin{cases} T^h(S) - \gamma_0 - \gamma_1 B & \text{if } \mathcal{I} = 1 \\ {A_B^*}'(S) & \text{if } \mathcal{I} = 0 \end{cases}, \tag{2.20}$$

where $\gamma_0, \gamma_1 \in \mathbb{R}$.

2.2.8 Market clearing

The market clearing in the sovereign and corporate bond, capital, and final good markets is described below.

$$B'(S) = A'_B(S) + {A^*_B}'(S), \tag{2.21}$$

$$K'(S) = A'_K(S), \tag{2.22}$$

$$K'(S) = (1-\delta)K + \Phi\left(\frac{I(S)}{K}\right)K, \tag{2.23}$$

$$(1-g)Y(S) = C(S) + I(S) + \mathcal{I}\left\{[\bar{\pi} + (1-\pi)(\iota + Q_B(S))]A^*_B - Q_B(S){A^*_B}'(S)\right\}. \tag{2.24}$$

Notice that external borrowing introduces net exports in the model: borrowing corresponds to imports, and repayment corresponds to exports. The derivation of (2.24) is presented in appendix 2.A.

2.3 Quantitative analysis

2.3.1 Calibration

I specify u and Φ the same as in Bocola (2016) as $u(c,l) = \ln c - \chi\frac{l^{1+\nu}}{1+\nu}$ with $\nu, \chi \geq 0$ and $\Phi(x) = a_1 x^{1-\xi} + a_2$ with $a_1 > 0$, $a_2 \in \mathbb{R}$, and $\xi \in (0,1)$. The function h is the same as in Schmitt-Grohé and Uribe (2003), that is, $h(x) = \phi[\exp(x-\bar{x}) - 1]$ with $\phi, \bar{x} \geq 0$. It follows that $u_c(c,l) = \frac{1}{c}$, $u_l(c,l) = -\chi l^{\nu}$, and $\Phi'(x) = a_1(1-\xi)x^{-\xi}$.

I assume that productivity evolves according to

$$z' = \rho_z z + \sigma_z \epsilon, \quad \epsilon \sim \mathcal{N}(0,1),$$

where $\rho_z \in [0,1)$, $\sigma_z > 0$.

Table 2.1 summarizes the parameter values I use in the analysis. The unit of time is a quarter. Most parameter values are based on posterior estimates in Bocola (2016) based on the Italian data. Some parameter values were changed to less extreme (but nevertheless standard) values to facilitate computation: β, δ. Since the default process is completely different in my model and

Table 2.1. Parameter values

Parameter	Value	Definition
Structural		
a_1	0.189	Capital production function slope
a_2	-0.004	Capital production function intercept
β	0.99	Discount factor
χ	6.627	Labor disutility scale
D	1	Sovereign default haircut share
δ	0.01	Depreciation rate
g	0.198	Ratio of government expenditures to output
γ_0	2.799	Fiscal rule intercept
γ_1	1	Fiscal rule slope
ι	0.003	Sovereign bond coupon rate
λ	0.205	Bank assets run-away share
ν	0.5	Inverse Frisch elasticity
ω	0.007	New bankers' endowment share in total assets
ϕ	0.001	Foreign lenders' risk premium, slope
π	0.055	Share of maturing bonds
ψ	0.968	Bankers' survival probability
$\bar{r}$	0.01	Foreign lenders' required return, intercept
θ	0.3	Capital share
$\bar{x}$	0.5	Foreign lenders' required return, debt-to-GDP reference
ξ	0.42	Capital production function elasticity
Exogenous stochastic processes		
ρ_z	0.89	Productivity process, persistence
σ_z	0.012	Productivity shock, standard deviation
Π_{11}	0.999	Debt repayment probability
Π_{00}	0.962	Probability to remain in financial exclusion

reflects endogenous default literature, I set $D = 1$. This fact implies that policy functions in the exclusion state do not depend on the level of debt (equal to zero).

The slope of the risk premium function is set the same as in Schmitt-Grohé and Uribe (2003). This parameterization, together with $\bar{x} = 0.5$, implies the share of bank sovereign debt exposure to total assets of 15.5%, the share of domestic debt to total debt of 83.0%, and the share of total debt to output of 92.4%, which is roughly consistent with the evidence in Reinhart and Rogoff (2011a), Reinhart and Rogoff (2011b), and Sosa-Padilla (2018). At the same time, the share of external debt to output is only 15.7%. Note that in the sovereign default literature, significantly lower values of the discount factor are required to achieve higher values of external debt. Finally, the implied capital to output ratio is 2.7, consistent with the data.

I calibrate the "financial standing" transition matrix to match the average default frequency of 2.6% and average exclusion time of 6.5 years, the same as in Na et al. (2018).

The global solution method is summarized in appendix 2.B.

2.3.2 Crisis experiments

Figure 2.1 shows the dynamics of the economy after a negative productivity shock of three standard deviations that hits in the fifth quarter. In this simulation, the economy is always in good standing

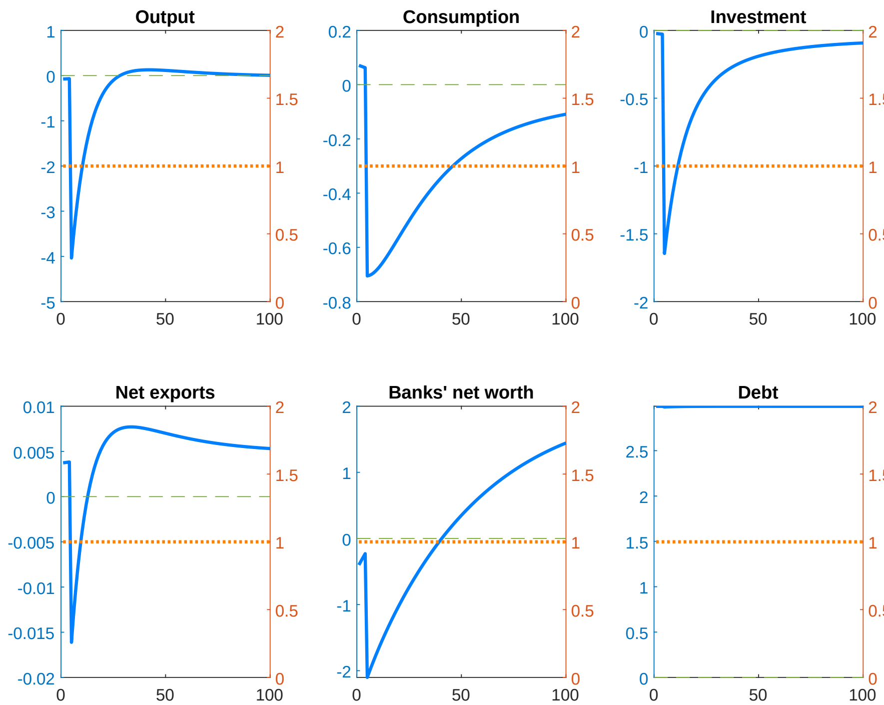

Figure 2.1. Typical recession. Output, consumption, investment, net worth are in percent deviations from the deterministic steady state with $\mathcal{I} = 1$. Net exports and debt are in levels.

($\mathcal{I} = 1$), although agents expect that default may happen in the future.

A negative productivity shock decreases output, consumption, and investment. Net exports decrease on impact but quickly become positive, hinting at a desired countercyclical relationship with output. The banking sector is hurt due to a fall in credit. However, total debt is almost unchanged since a fall in domestic debt on impact is compensated by a rise in external debt.

Figure 2.2 shows the dynamics of the economy after default. Productivity shocks do not hit in this simulation (although agents expect that they may happen). The economy starts in good standing and remains for a year, after which default happens in the fifth quarter. The economy

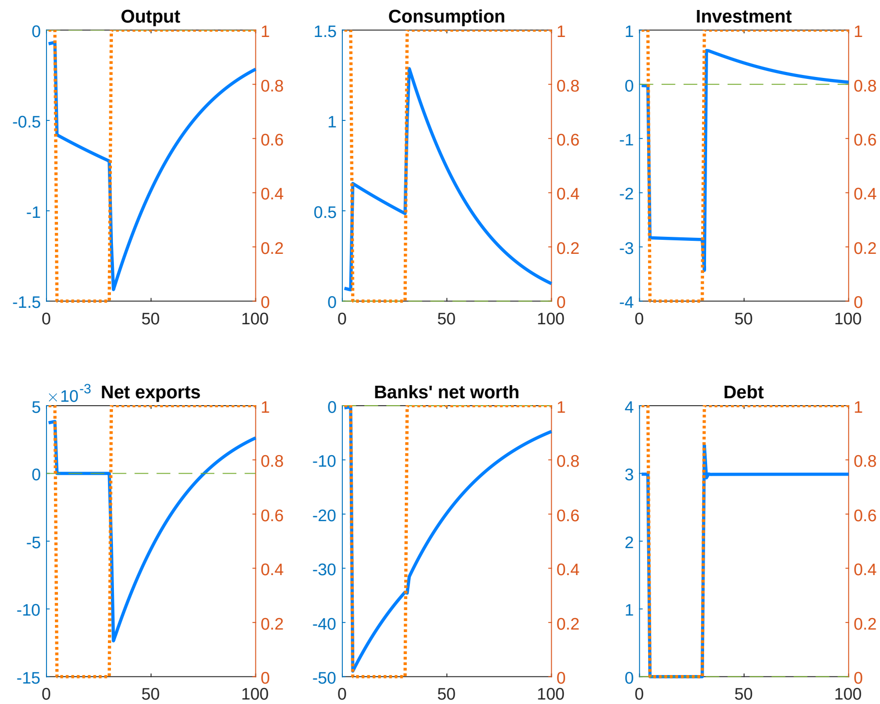

Figure 2.2. Typical default. Output, consumption, investment, net worth are in percent deviations from the deterministic steady state with $\mathcal{I} = 1$. Net exports and debt are in levels.

remains in exclusion for 6.5 years and then regains access to financial markets.

When default happens, the banking sector is hurt severely, credit and investment plummet, and output declines. The reduced demand for deposits from the banking sector makes households save less and substitute for consumption. Being shut out of financial markets is equivalent to being excluded from trade in my model; hence, net exports and debt become zero. When the economy regains access to financial markets, output drops even more on impact due to a rise in borrowing and imports, reflected in the trade deficit. As the banking system recovers, consumption and investment jump and then converge to the stochastic steady state. There is a debt overshooting on impact, but it quickly stabilizes afterward.

Figure 2.3 shows the dynamics of the economy after a negative productivity shock of three standard deviations and default that both hit in the fifth quarter. The economy starts in good standing and remains for a year, after which recession and default happen. The economy remains in exclusion for 6.5 years and then regains access to financial markets.

Both events reinforce each other. The banking sector is hurt even more, which generates a more significant fall in credit and investment. The recession is so severe that consumption also

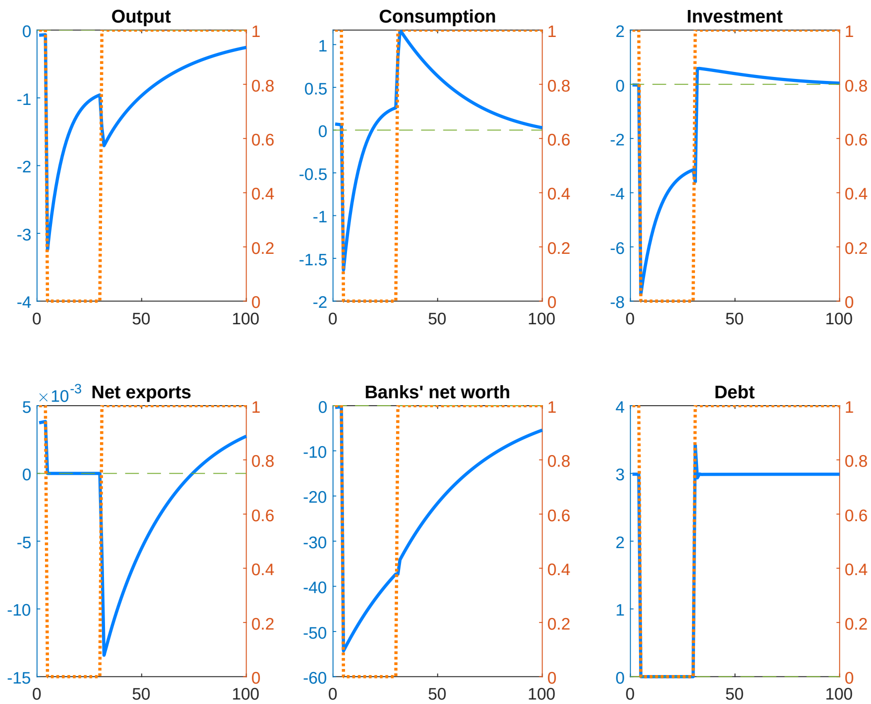

Figure 2.3. Recession and default. Output, consumption, investment, net worth are in percent deviations from the deterministic steady state with $\mathcal{I} = 1$. Net exports and debt are in levels.

falls, though it recovers to the initial value during exclusion.

2.4 Macroprudential policy

Now consider how different government policies affect welfare and the frequency of banking crises, defined as the instances of a binding incentive-compatibility constraint. I define welfare as

$$\mathcal{W} \equiv \mathbb{E}\{v^h(P; S)\}.$$

To compare welfare across policies, I compute the difference in consumption units, which in the case of log utility from consumption takes the following form (in %):

$$\lambda_\omega = \left\{ \exp\left[(1-\beta)\left(\widetilde{\mathcal{W}} - \mathcal{W} \right) \right] - 1 \right\} * 100,$$

where $\widetilde{\mathcal{W}}$ is the welfare under the alternative policy, and $\mathcal{W}$ is the welfare when all taxes are set to zero.

In each exercise below, I simulate the economy for 50,000 periods, achieving an observed frequency of sovereign defaults of 2.6%, consistent with the targeted stationary probability. I use the identical pseudo-random sequences of standard normal and uniform shocks and thus the same productivity and financial standing process sequences in each exercise.

I limit the analysis to constant policies, except that they may be different across financial standing states, which I denote with a 1 or 0 superscript; for example, T^1 and T^0 would denote a constant lump-sum tax or transfer when $\mathcal{I} = 1$ and $\mathcal{I} = 0$, respectively. In each analysis, I set all other policies to zero. To remind, the pool of policies contain the lump-sum tax/transfer T^b and proportional taxes/subsidies τ_B, τ_K, τ_N, τ_P.

Table 2.2 summarizes the effects of different lump-sum tax arrangements when all proportional taxes are set to zero. None of the policies considered could improve welfare, but a positive transfer

Table 2.2. Effects of a lump-sum tax

T^1	T^0	λ_ω, %	Banking crisis frequency, %
-0.005	-0.005	-0.0015	**0.15**
-0.005	0.000	-0.0015	0.20
-0.005	0.005	-0.0015	0.30
0.000	-0.005	-0.0013	0.47
0.000	0.000	**0.0000**	0.61
0.000	0.005	-0.0013	0.87
0.005	-0.005	-0.0012	1.24
0.005	0.000	-0.0012	1.47
0.005	0.005	-0.0012	2.03

in both states significantly reduces the probability of banking crises. Notably, the transfer should be given in good states for the largest effect. Just imposing a transfer in the bad state reduces the probability of crises only slightly. One reason these policies are not welfare-improving is that households eventually finance the “bailout” of the banking sector.

Table 2.3 summarizes the effects of different sovereign bond tax arrangements. I only considered taxes because subsidies of the same magnitude lead to numerical issues, although it would be interesting to compare the predictions with those in Perez (2018). None of the policies considered could improve welfare, but a tax on sovereign bond holdings significantly reduces the probability of banking crises since lower sovereign exposures imply lower losses in the case of sovereign default. Notice that it does not matter whether to impose the tax in both good and bad financial standing

Table 2.3. Effects of a tax on sovereign bond holdings

τ_B^1	τ_B^0	λ_ω, %	Banking crisis frequency, %
0.000	0.000	**0.0000**	0.61
0.000	0.001	0.0000	0.61
0.001	0.000	-1.1788	**0.28**
0.001	0.001	-1.1788	0.28

or only in a good one because no borrowing is allowed in the bad standing.

Table 2.4 summarizes the effects of different corporate bond tax arrangements. Unlike in the previous exercises, some policies are welfare-improving. A subsidy in both states improves welfare and reduces the probability of crises. The intuition for this result is that subsidizing credit to the real sector instead of investing in sovereign bonds strengthens the economy in good times and helps it recover faster in bad times. On the contrary, a tax is welfare-reducing and leads to a higher probability of banking crises.

Table 2.4. Effects of a tax on corporate bond holdings

τ_K^1	τ_K^0	λ_ω, %	Banking crisis frequency, %
-0.001	-0.001	**1.2596**	**0.36**
-0.001	0.000	1.2244	0.63
-0.001	0.001	1.1898	0.99
0.000	-0.001	0.0325	0.44
0.000	0.000	0.0000	0.61
0.000	0.001	-0.0401	0.93
0.001	-0.001	-1.1727	0.55
0.001	0.000	-1.2080	0.67
0.001	0.001	-1.2427	0.95

Table 2.5 summarizes the effects of different net worth subsidy arrangements. Some policies are welfare-improving. The most successful policy subsidizes net worth in good times and taxes it in bad times. In other words, banks are encouraged to borrow less from households in good times and borrow more in bad times. This result is thus consistent with the overborrowing story of Bianchi (2011), Bianchi and Mendoza (2018), and Jeanne and Korinek (2019).

Table 2.6 summarizes the effects of different deposit tax arrangements. Again, only the tax could be considered due to numerical reasons. None of the policies are welfare-improving. The tax in good times reduces the banking crisis probability significantly. There should not be a tax in bad

Table 2.5. Effects of a net worth subsidy

τ_N^1	τ_N^0	λ_ω, %	Banking crisis frequency, %
-0.001	-0.001	-0.0024	1.66
-0.001	0.000	-0.0252	1.40
-0.001	0.001	-0.0484	1.16
0.000	-0.001	0.0174	0.78
0.000	0.000	0.0000	0.61
0.000	0.001	-0.0208	0.51
0.001	-0.001	**0.0459**	0.26
0.001	0.000	0.0231	0.20
0.001	0.001	-0.0001	**0.17**

Table 2.6. Effects of a tax on deposits

τ_P^1	τ_P^0	λ_ω, %	Banking crisis frequency, %
0.000	0.000	**0.0000**	0.61
0.000	0.001	-0.0563	0.80
0.001	0.000	-2.3453	**0.17**
0.001	0.001	-2.3914	0.21

times, consistent with the overborrowing story.

2.5 Conclusion

I compared different macroprudential policy arrangements using the quantitative banking model enriched with sovereign default risk. I found that a proportional tax that subsidizes lending to the real sector in good and bad financial standing improves welfare and reduces the probability of banking crises. The reason is that subsidizing credit to the real sector instead of investing in sovereign bonds strengthens the economy in good times and helps it recover faster in bad times. A bank net worth subsidy in good times combined with a tax in bad times is also welfare-improving and facilitates the prevention of banking crises. This finding is consistent with the overborrowing story explored recently by several authors.

This paper is only the beginning of the research agenda I plan to pursue. The global solution method used in the analysis must be improved. The optimal policy must be derived, studied, and explained.

Appendix

2.A Aggregate budget constraint

Aggregating an individual household's budget constraint, we obtain

$$C(S) + \frac{P'(S)}{R(S)} = P + (1-\tau_L(S))W(S)L(S) + \Pi(S) - T^h(S).$$

Using (2.9) and (2.13),

$$C(S) + \sum_{j\in\{B,K\}} Q_j(S)A'_j(S) + \Upsilon(S) = N(S) + P + (1-\tau_L(S))(1-\theta)Y(S) + \Pi(S) - T^h(S).$$

Note that

$$\begin{aligned}
N(S) + P + \Pi(S) &= \psi\left[\sum_{j\in\{B,K\}} R_j(S,S^-)Q_j(S^-)A_j - P\right] + \omega \sum_{j\in\{B,K\}} \mathcal{I}_j Q_j(S)A_j + P \\
&\quad + (1-\psi)\left[\sum_{j\in\{B,K\}} R_j(S,S^-)Q_j(S^-)A_j - P\right] - \omega \sum_{j\in\{B,K\}} \mathcal{I}_j Q_j(S)A_j \\
&\quad + Q_K(S)\Phi\left(\frac{I(S)}{K}\right)K - I(S) \\
&= \sum_{j\in\{B,K\}} R_j(S,S^-)Q_j(S^-)A_j + Q_K(S)\Phi\left(\frac{I(S)}{K}\right)K - I(S) \\
&= \mathcal{I}[\bar{\pi} + (1-\pi)(\iota + Q_B(S))]A_B + \theta Y(S) + Q_K(S)(1-\delta)K \\
&\quad + Q_K(S)\Phi\left(\frac{I(S)}{K}\right)K - I(S),
\end{aligned}$$

where (2.10), (2.14) and (2.17) were used. Hence, using (2.19), (2.21), (2.22) and (2.23),

$$(1-g)Y(S) = C(S) + I(S) + \mathcal{I}\left\{[\bar{\pi} + (1-\pi)(\iota + Q_B(S))]A_B^* - Q_B(S)A_B^{*\prime}(S)\right\},$$

which is (2.24) above.

2.B Solution method

The aggregate state is $S = (A_B, A_B^*, K, P, z, \mathcal{I})$. The government policy is $\{\tau_B, \tau_K, \tau_N, \tau_P, T^b\}(S)$. The system of functional equations can be reduced to

$$\alpha_1(S) = \frac{\left[1 - \psi + \beta\psi R(S)C(S)\mathbb{E}_S\left\{\frac{\alpha_1(S')}{C(S')}\right\}\right]\frac{1+\tau_N(S)}{1-\tau_P(S)}}{1-\mu(S)}, \tag{2.25}$$

$$\alpha_2(S) = \frac{\beta\psi C(S)\mathbb{E}_S\left\{\frac{\alpha_2(S')}{C(S')}\right\} - \alpha_1(S)(1-\mu(S))\frac{T^b(S)}{1+\tau_N(S)}}{1-\mu(S)}, \tag{2.26}$$

$$\frac{C(S)}{Q_K(S)} = \frac{\lambda\mu(S) + \alpha_1(S)(1-\mu(S))\frac{1+\tau_K(S)}{1+\tau_N(S)}}{\beta\mathbb{E}_S\left\{\frac{\theta\frac{Y(S')}{K'(S)} + Q_K(S')(1-\delta)}{C(S')}(1-\psi+\psi\alpha_1(S'))\right\}}, \tag{2.27}$$

$$0 = \begin{cases} \frac{\lambda\mu(S)+\alpha_1(S)(1-\mu(S))\frac{1+\tau_B(S)}{1+\tau_N(S)}}{\beta\mathbb{E}_S\left\{\frac{\mathcal{I}'[\bar{\pi}+(1-\pi)(\iota+Q_B(S'))]}{C(S')}(1-\psi+\psi\alpha_1(S'))\right\}} - \frac{C(S)}{Q_B(S)} & \text{if } \mathcal{I} = 1 \\ (1-g)Y(S) - C(S) - I(S) & \text{if } \mathcal{I} = 0 \end{cases}, \tag{2.28}$$

$$Q_B(S) = \mathcal{I}\frac{\Pi_{\mathcal{I}1}[\bar{\pi} + (1-\pi)\iota] + (1-\pi)\mathbb{E}_S\{\mathcal{I}'Q_B(S')\}}{R^*(S)} + 1 - \mathcal{I}, \tag{2.29}$$

$$R(S) = \frac{1}{\beta C(S)\mathbb{E}_S\left\{\frac{1}{C(S')}\right\}}. \tag{2.30}$$

The laws of motion for endogenous states are

$$K'(S) = \left[a_1\left(\frac{I(S)}{K}\right)^{1-\xi} + a_2 + 1 - \delta\right]K, \tag{2.31}$$

$$\begin{aligned} A_B'(S) = {} & \mathcal{I}\frac{(1-\tau_P(S))\Big(Y(S) - C(S) - I(S) - T^h(S) + [\bar{\pi} + (1-\pi)(\iota + Q_B(S))]A_B\Big)}{(1+\tau_B(S))Q_B(S)} \\ & + \mathcal{I}\frac{(\tau_N(S)+\tau_P(S))N(S) - (\tau_K(S)+\tau_P(S))Q_K(S)K'(S) - T^b(S)}{(1+\tau_B(S))Q_B(S)}, \end{aligned} \tag{2.32}$$

$$B'(S) = \mathcal{I}\frac{gY(S) - T^h(S) - \Upsilon(S) + [\bar{\pi} + (1-\pi)(\iota + Q_B(S))]B}{Q_B(S)}, \tag{2.33}$$

$$P'(S) = R(S)\left(\sum_{j\in\{B,K\}} Q_j(S)A_j'(S) + \Upsilon(S) - N(S)\right). \tag{2.34}$$

The remaining variables are obtained through a sequential solution as follows:

$$\Lambda(S', S) \equiv \beta\frac{C(S)}{C(S')}, \qquad L(S) = \left[\frac{1-\theta}{\chi}\frac{(1-\tau_L(S))K^\theta e^{(1-\theta)z}}{C(S)}\right]^{\frac{1}{\nu+\theta}}, \qquad Y(S) = K^\theta(e^z L(S))^{1-\theta},$$

$$
\begin{aligned}
I(S) &= [a_1(1-\xi)Q_K(S)]^{\frac{1}{\xi}}K,\\
N(S) &= \psi\left\{\mathcal{I}[\bar{\pi}+(1-\pi)(\iota+Q_B(S))]A_B+\theta Y(S)+Q_K(S)(1-\delta)K-P\right\}\\
&\quad +\omega[\mathcal{I}Q_B(S)A_B+Q_K(S)K], \qquad A_K'(S)=K'(S),\\
\Xi(S) &\equiv \alpha_1(S)N(S)+\alpha_2(S)-\lambda\sum_{j\in\{B,K\}}Q_j(S)A_j'(S),\\
\mu(S) &= \mathbf{1}_{\mathbb{R}_-}(\Xi(S))\left(1-\frac{\left[1-\psi+\beta\psi R(S)C(S)\mathbb{E}_S\left\{\frac{\alpha_1(S')}{C(S')}\right\}\right]\frac{1+\tau_N(S)}{1-\tau_P(S)}N(S)}{\lambda\sum_{j\in\{B,K\}}Q_j(S)A_j'(S)-\alpha_2(S)}\right),\\
\Upsilon(S) &= \sum_{j\in\{B,K\}}\frac{\tau_j(S)+\tau_P(S)}{1-\tau_P(S)}Q_j(S)A_j'(S)+\frac{T^b(S)}{1-\tau_P(S)}-\frac{\tau_N(S)+\tau_P(S)}{1-\tau_P(S)}N(S),\\
T^h(S) &= \begin{cases}\gamma_0+\gamma_1 B & \text{if } \mathcal{I}=1\\ gY(S)-\Upsilon(S) & \text{if } \mathcal{I}=0\end{cases}, \qquad {A_B^*}'(S)=B'(S)-A_B'(S),\\
R^*(S) &= 1+\bar{r}+\phi\left[\exp\left(\frac{{A_B^*}'(S)}{Y(S)}-\bar{x}\right)-1\right], \qquad R_K(S',S)\equiv\frac{\theta\frac{Y(S')}{K'(S)}+Q_K(S')(1-\delta)}{Q_K(S)},\\
R_B(S',S) &\equiv \mathcal{I}'\frac{\bar{\pi}+(1-\pi)(\iota+Q_B(S'))}{Q_B(S)}, \qquad W(S)=\chi C(S)L(S)^{\nu}.
\end{aligned}
$$

We approximate α_1, α_2, C, Q_B, Q_K, R using the *Smolyak algorithm*, following Krueger and Kubler (2004) and Malin et al. (2011). Since $\mathcal{I}$ is a binary variable, it cannot be used directly in the Smolyak algorithm—instead, I index the functions by $\mathcal{I}$. Note that Q_B^0 is known, and Q_K^0 can be obtained recursively; hence, there are ten unknown functions. The initial guess for unknown functions is set to constant functions equal to the values of the corresponding variables in the deterministic steady state with constant $\mathcal{I}$. I solve for unknown coefficients using fixed-point iteration as advocated by Judd (1998). If necessary, I use the homotopy method. In particular, one can first solve for the equilibrium of an economy with $\Pi_{11}=\Pi_{00}=1$ and then gradually transition to a solution of the desired economy.

The system of equations (2.25)–(2.30) needs to be evaluated at each $\mathcal{I}\in\{0,1\}$ for each $S\setminus\{\mathcal{I}\}$ in the Smolyak grid. Each such evaluation requires the following.

1. Given the guesses for unknown functions, use (2.31)–(2.34) to compute the endogenous components of S'.

2. Compute

$$\mathbb{E}_S\left\{\frac{1}{C(S')}\right\}, \qquad \mathbb{E}_S\left\{\frac{\alpha_1(S')}{C(S')}\right\}, \qquad \mathbb{E}_S\left\{\frac{\alpha_2(S')}{C(S')}\right\},$$
$$\mathbb{E}_S\left\{\frac{\theta\frac{Y(S')}{K'(S)} + Q_K(S')(1-\delta)}{C(S')}(1-\psi+\psi\alpha_1(S'))\right\},$$
$$\mathbb{E}_S\left\{\frac{\mathcal{I}'[\bar{\pi}+(1-\pi)(\iota+Q_B(S'))]}{C(S')}(1-\psi+\psi\alpha_1(S'))\right\}, \qquad \mathbb{E}_S\{\mathcal{I}'Q_B(S')\}.$$

The last two conditional expectations are needed only if $\mathcal{I}=1$. See below for computational details.

3. Evaluate the system.

For an arbitrary function $h(S', S)$, we have

$$\begin{aligned}
&\mathbb{E}_S\{h(S',S)\} \\
&= \sum_{\mathcal{I}'\in\{0,1\}} \Pi_{\mathcal{I}\mathcal{I}'} \int_{\mathbb{R}} h\Big(\{A_B'(S), B'(S), K'(S), P'(S), \rho_z z + \sigma_z \epsilon, \mathcal{I}'\}, S\Big) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\epsilon^2}{2}\right) \mathrm{d}\epsilon \\
&= \sum_{\mathcal{I}'\in\{0,1\}} \Pi_{\mathcal{I}\mathcal{I}'} \int_{\mathbb{R}} h\Big(\{A_B'(S), B'(S), K'(S), P'(S), \rho_z z + \sigma_z \sqrt{2}x, \mathcal{I}'\}, S\Big) \frac{1}{\sqrt{\pi}} \exp\left(-x^2\right) \mathrm{d}x \\
&\approx \sum_{\mathcal{I}'\in\{0,1\}} \Pi_{\mathcal{I}\mathcal{I}'} \sum_{j=1}^{n} h\Big(\{A_B'(S), B'(S), K'(S), P'(S), \rho_z z + \sigma_z \sqrt{2}x_j, \mathcal{I}'\}, S\Big) \frac{1}{\sqrt{\pi}} \omega_j,
\end{aligned}$$

where $\{x_j, \omega_j\}_{j=1}^{n}$ is the set of Gauss-Hermite nodes and weights.

Chapter 3

Financial constraints, risk sharing, and optimal monetary policy

3.1 Introduction

In the past decade, there has been a surge in research on externalities stemming from financial constraints.[1] This paper studies the implications of such externalities for optimal monetary policy in an economy with a banking sector and different types of consumers. This economy has a conventional pecuniary externality working through the collateral asset price and other externalities arising from consumer type heterogeneity. To identify the externalities, I characterize a constrained efficient allocation (CEA) chosen by a benevolent social planner who faces the same constraints as private agents but internalizes the determination of market prices. The wedges between the competitive equilibrium (CE) and CEA arise in both the real and the financial sectors of the economy. The real wedges represent the inefficient demand for labor and capital. The financial wedges reflect the inefficient supply of deposits by the banking sector and demand for loans by the entrepreneurial sector, typically resulting in both overborrowing and overlending by banks. These wedges can be addressed with the appropriate fiscal instruments. A key finding of the paper is that the ability to correct the wedges with fiscal instruments does not impact the fundamental nature of Ramsey-optimal monetary policy. The latter prescribes price stability in the long run and approximate price stability in the short run, as in the basic New Keynesian environment.

[1] Dávila and Korinek (2018) present a unifying treatment of such externalities.

The object of the analysis is a New Keynesian economy with different types of consumers—workers, bankers, and entrepreneurs—and a financial sector. Workers are savers who are not directly subject to financial frictions. Bankers manage banks that issue deposits to workers and extend loans to wholesale firms subject to a leverage constraint. Entrepreneurs are the managers of wholesale firms and raise external financing subject to a collateral constraint. The entrepreneur's capital stock serves as collateral and is produced by competitive firms with a nonlinear technology. A monopolistically competitive retail sector is subject to nominal rigidities: the opportunity to adjust prices arrives stochastically according to the Calvo-pricing mechanism.

The normative analysis proceeds step-by-step, starting from a special case of a flexible-price economy with perfectly competitive markets. In this setting, I define a flexible-price competitive equilibrium (FCE) and characterize the flexible-price constrained efficient allocation (FCEA). Due to consumer type heterogeneity, the price externalities are not limited to a conventional pecuniary externality working through the collateral asset price. The social planner is subject to a consolidated budget constraint of bankers and entrepreneurs, which depends on the asset price and the wage rate. Moreover, the planner must respect the private complementary slackness conditions associated with the bank leverage constraint. As a result, the FCE has multiple wedges relative to the FCEA that arise in both the real and the financial sectors of the economy. The real wedges are in the entrepreneur's demand for labor and capital—the factors of production. The labor wedge constitutes the only intratemporal distortion and arises from consumer type heterogeneity, particularly the wage externality. The capital wedge stems from an externality due to the entrepreneur's impatience, both first-order and second-order externalities arising through the capital good production technology, and a pecuniary externality in the collateral constraint. The financial wedges are in the banker's supply of deposits and the entrepreneur's demand for loans, and they mainly result from the differences in patience, reflecting consumer type heterogeneity.

In a special case where the worker's preferences are separable in consumption and leisure and logarithmic in consumption, and the technology is such that capital good producers earn zero profits in the steady state, the FCEA has quantitatively perfect consumption risk sharing between all types of consumers, approaching the unconstrained first-best allocation. The FCEA can be decentralized in a regulated FCE with state-contingent linear taxes on the banker's supply of deposits and the entrepreneur's demand for loans, labor, and capital. I also consider a situation when the complete set of taxes is not available to the policymaker but the leverage limits—bank capital requirement and loan-to-value (LTV) ratio—can be set optimally. The resulting Ramsey allocation has the potential to enhance risk sharing but is typically inferior to the FCEA.

The analysis then moves to the benchmark sticky-price economy. Under an assumption that the social planner takes monetary policy as given, the set of wedges between the CE and CEA is similar to the flexible-price case. The financial wedges remain unchanged, while the real wedges are now affected by monopoly power and nominal rigidities. The latter reduce the extent of between-agent risk sharing in the CEA compared to the FCEA, although it remains strong quantitatively. The fact that financial wedges are not affected by nominal rigidities has two important implications. First, the fundamental nature of Ramsey-optimal monetary policy is not affected by the availability of the complete set of fiscal instruments needed to decentralize the CEA. Second, the implications of optimal monetary policy are similar to the basic New Keynesian environment: price stability is optimal in the long run, even if there is an effective lower bound (ELB) on the policy rate that does not exceed the steady-state real interest rate. In the short run, the optimal inflation rate is characterized by an Euler equation with different compensating mechanisms: the inflation rate is not necessarily zero but remains close to zero quantitatively. In the presence of an ELB, the Ramsey allocation under optimal monetary policy highlights an additional aggregate demand externality not internalized by the private agents in the CE.

Using a social-welfare consumption-equivalent measure, conditional on choosing a worker-biased vector of Pareto weights, the FCEA constitutes 98.9% of the first best, compared to 86.2% in the FCE; the sticky-price CEA provides 94% of the first best, compared to 78.1% in the CE. The flexible-price Ramsey allocation with optimal leverage limits and labor taxation—but not other fiscal instruments—gives 94.1% of the first best, while an analogous sticky-price Ramsey allocation with optimal monetary policy stands at 90.7%. The FCEA and CEA have perfect consumption risk sharing between bankers and entrepreneurs; the risk sharing with workers is not exactly perfect, but the correlation between the marginal utilities is close to unity. Most of the magnitude and variance of the wedges is explained by the components that arise from consumer type heterogeneity; therefore, the ability to improve between-agent risk sharing is the main source of welfare gains from the FCEA and CEA. Nominal rigidities do have a notable impact on the real wedges. In the FCEA and CEA, bank leverage is suboptimal from the planner's perspective, and the entrepreneur's leverage is lower than in the market allocations. Consequently, the FCE and CE have both overborrowing and overlending by the banking sector.

Finally, I compare the dynamics in the decentralized FCE and CE economies with the centralized FCEA, CEA, and Ramsey allocations around financial crises in the flexible-price settings and the episodes of hitting the zero lower bound (ZLB) on the policy rate in the sticky-price environments. A financial crisis is defined as an event that satisfies two conditions: the collateral constraint is slack for at least four quarters before the crisis and is binding for at least five quarters since the start of the crisis. An event defined this way is observed in the FCE with a relative frequency of 3.2 crises per century, consistent with the data. In the FCE, such crises follow a boom-bust pattern: output, credit, and the collateral asset price increase ahead of the crisis, followed by a sharp and persistent fall when the collateral constraint binds. In the FCEA, the collateral constraint remains slack during the whole crisis window, and the dynamics of real and financial variables resemble usual business cycle fluctuations. When the intertemporal distortions cannot be addressed but leverage limits are set optimally, the dynamics are more similar to the FCE, although the amplitude of the fluctuations is reduced.

The ZLB crises are identified similarly as events where the ZLB is slack during the year before the start of a crisis and is binding for at least three quarters, implying a simulated frequency of 2.5 crises per century in the CE. Compared to financial crises, ZLB crises have a different pattern: before the ZLB binds, the economy is already in a recession or stagnation, and inflation is below the target. When the ZLB binds, the recession deepens, and inflation decreases further, followed by an increase due to the rise in the marginal cost. When the ZLB becomes slack, the recovery in investment and the asset price is faster than after financial crises, but the recovery in output and credit is slow. As with financial crises, the CEA dynamics are much smoother, and the ZLB is not hit. The dynamics in the Ramsey allocation with optimal labor taxation, leverage limits, and monetary policy are somewhere between the CE and CEA, and the planner typically just avoids the ZLB. The optimal bank capital and LTV ratios have countercyclical dynamics around financial crises and ZLB episodes.

This paper is related to different sets of the literature. The theoretical model is in the class of New Keynesian economies with consumer type heterogeneity (Iacoviello, 2005; Andrés et al., 2013). The banking sector is based on Iacoviello (2015), while the entrepreneurial and retail sectors have features of Kiyotaki and Moore (1997), Bernanke et al. (1999), and Iacoviello (2005). The focus on the CEA in the normative analysis follows Lorenzoni (2008). Similar to Lorenzoni (2008), Benigno et al. (2016), Bianchi and Mendoza (2018), Dávila and Korinek (2018), and Jeanne and Korinek (2019), the competitive equilibrium is inefficient due to a pecuniary externality present in the collateral constraint. Unlike in most of these papers, the pecuniary externality is associated with borrowing in the domestic banking sector at an endogenous interest rate in the current paper. Moreover, the pecuniary externality is not the only externality that leads to constrained inefficiency. Due to consumer type heterogeneity, multiple wedges stem from multiple price externalities. Farhi and Werning (2016), Korinek and Simsek (2016), and Schmitt-Grohé and Uribe (2016) emphasize aggregate demand externalities that arise in the presence of constraints on monetary policy, fixed exchange rates, or downward sticky wages. The definition of the CEA used in this paper specifies that the social planner faces the same constraints as private agents. Hence, the CEA social planner does not internalize any monetary policy constraints. On the other hand, the Ramsey planner that determines the optimal monetary policy is generally subject to an ELB constraint. If such a constraint is present, the CE allocation has an aggregate demand externality compared to the Ramsey allocation.

By characterizing optimal monetary policy in the presence of financial frictions, this paper is related to Bean et al. (2010), Andrés et al. (2013), Cúrdia and Woodford (2016), Farhi and Werning (2016), Collard et al. (2017), De Paoli and Paustian (2017), Ferrero et al. (2018), Leduc and Natal (2018), and Van der Ghote (2021). The closest set-ups to the current paper are in Andrés et al. (2013) and Ferrero et al. (2018), who also allow for consumer type heterogeneity, collateral constraints, and financial intermediation. Both these papers have a housing market with an inelastic supply that provides collateral for entrepreneurs, while this paper considers capital stock as collateral, and the supply side is endogenous. Moreover, as in Iacoviello (2015), this paper considers bankers as generally risk-averse consumers, allowing for an additional degree of heterogeneity. In terms of the normative analysis, Andrés et al. (2013) and Ferrero et al. (2018) adopt a linear-quadratic approach accurate in the neighborhood of the steady state. At the same time, this paper characterizes globally optimal constrained efficient and Ramsey allocations, respecting occasionally binding constraints in the theoretical derivations, as in Bianchi and Mendoza (2018). Consistent with Andrés et al. (2013) and Ferrero et al. (2018), this paper finds that optimal monetary policy does not entail perfect consumption insurance between consumers. However, this paper provides conditions under which quantitatively perfect consumption insurance is observed in the CEA. The analysis in Andrés et al. (2013) is limited to separable preferences logarithmic in consumption, while Ferrero et al. (2018) restrict attention to exponential preferences. In contrast, this paper conducts normative analysis with general preferences and technology.

By proving that the optimal long-run inflation rate in the absence of uncertainty is zero even with financial frictions, this paper is consistent with Cúrdia and Woodford (2016), who came to an identical conclusion in the case of a credit spread friction. In this paper, an endogenous credit spread arises from the bank leverage constraint. Andrés et al. (2013) and Collard et al. (2017) have also argued that zero steady-state inflation is optimal, albeit quantitatively.

The rest of the paper is organized as follows. Section 3.2 describes the model and defines and characterizes the competitive equilibrium. Section 3.3 conducts a normative analysis of the flexible-price and sticky-price economies. Section 3.4 presents quantitative results. Section 3.5 concludes. An Appendix provides proofs of theoretical results.

3.2 Model

Consider an infinite-horizon discrete-time economy populated by consumers—workers (w), bankers (b), and entrepreneurs (e)—and producers of capital, retail, and final goods. Conditional on the type $i \in \mathcal{I} \equiv \{b, e, w\}$, there is a unit measure of identical risk-averse consumers. Workers are infinitely lived with certainty, but each period, a constant share of bankers and entrepreneurs exit the economy, being replaced by new consumers of the same measure who inherit the assets and liabilities of the former. As noted by Andrés et al. (2013), a trivial life-cycle structure of this sort facilitates a tractable normative analysis. The differences in survival rates result in the differences in effective patience: workers apply a discount factor $\beta \in (0, 1)$, while bankers and entrepreneurs use $\beta_b \leq \beta$ and $\beta_e \leq \beta$, respectively.

Workers solve a standard consumption-saving problem and are owners of firms that produce capital, retail, and final goods. Bankers manage banks that issue deposits to workers and supply loans to entrepreneurs. Entrepreneurs manage firms that supply wholesale goods to the retail sector that operates subject to nominal rigidities, similar to Bernanke et al. (1999) and Iacoviello (2005).

Capital goods are produced using a nonlinear technology as in Lucas and Prescott (1971).

Following Gertler and Karadi (2011), we will assume that financial assets—deposits and loans—are contracted in real terms. This assumption allows increasing the tractability of the normative analysis, since our baseline economy will have a well-defined special case of a flexible-price economy with perfectly competitive markets. Consequently, it will be easier to decipher the roles of financial frictions, consumer type heterogeneity, and nominal rigidities for the efficiency of a competitive equilibrium allocation.

Corresponding to our economy, for each $t \geq 0$, there is a set Z^t of histories of states of nature $z^t \in Z^t$. To save on notation, the dependence on histories will be hidden, but one should be aware that a variable x_t will typically correspond to a number $x_t(z^t)$, $\{x_t\}$ will denote a sequence $\{x_t\}_{t=0}^{\infty}$ of Borel measurable functions $x_t : Z^t \to \mathbb{R}$ for all $t \geq 0$, and $\{x_{1,t}, \ldots, x_{n,t}\}$ will denote a list of n such sequences.

3.2.1 Workers

A worker's decision problem involves choosing consumption C_t^w, savings in one-period bank deposits D_t at a risk-free gross real interest rate R_t, and labor supply N_t given a wage rate W_t. The worker's income is augmented by the aggregate profits Ξ_t from the ownership of retail and capital good producing firms. The final good is the numeraire, so the budget constraint is

$$C_t^w + D_t \leq W_t N_t + R_{t-1} D_{t-1} + \Xi_t.$$

The worker's preferences are represented by $\mathbb{E}_0[\sum_{t=0}^{\infty} \beta^t U^w(C_t^w, N_t)]$, where $U^w : \mathbb{R}_+^2 \to \mathbb{R}$ is twice continuously differentiable and strictly concave with $U_C^w(C, N) > 0$ and $U_N^w(C, N) < 0$ for all $(C, N) \in \mathbb{R}_{++}^2$, and $\lim_{C \to 0} U_C^w(C, N) = \infty$ for all $N \geq 0$. Define a stochastic discount factor $\Lambda_{t,s} \equiv \beta^{s-t} \frac{U_{C,s}^w}{U_{C,t}^w}$, where $s \geq t \geq 0$. The necessary conditions for optimality include the budget constraint holding as equality, the labor supply condition (3.1) postulating the equality between the wage and the marginal rate of substitution of consumption for leisure, and the Euler equation (3.2) that prices bank deposits:

$$W_t = -\frac{U_{N,t}^w}{U_{C,t}^w}, \tag{3.1}$$

$$1 = \mathbb{E}_t(\Lambda_{t,t+1})R_t. \tag{3.2}$$

3.2.2 Bankers

Following Iacoviello (2015), consider a simple banking sector where banks issue deposits to workers and use their own net worth to extend one-period loans L_t to entrepreneurs at a state-contingent gross real loan rate R_t^L. The bank's net worth is the difference between the ex-post loan repayments from entrepreneurs and deposit repayments to workers, that is, $R_t^L L_{t-1} - R_{t-1}D_{t-1}$. Bankers are specialists in managing the banks and their only owners. The banking business provides a dividend C_t^b, so the banker's budget constraint is

$$C_t^b + L_t \leq R_t^L L_{t-1} - R_{t-1}D_{t-1} + D_t. \tag{3.3}$$

Furthermore, the banker's budget set is limited by a leverage constraint

$$D_t \leq (1-\kappa_t)L_t, \tag{3.4}$$

where $\kappa_t \in [0,1]$ can be interpreted as a bank capital requirement. The leverage constraint (3.4) may reflect agency frictions between workers and bankers or prudential regulation. We will consider κ_t as a policy instrument set by a policymaker.

The banker's preferences are represented by $\mathbb{E}_0[\sum_{t=0}^{\infty} {\beta_b}^t U^b(C_t^b)]$, where $U^b : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable with $U_C^b > 0$ and $U_{CC}^b \leq 0$. Denoting the normalized Lagrange multiplier on (3.4) as λ_t^b, the Karush—Kuhn—Tucker (KKT) conditions associated with the banker's problem include (3.3) as equality, (3.4), the Euler equations for deposits (3.5) and loans (3.6), and the complementary slackness conditions (3.7):

$$U_{C,t}^b = \beta_b R_t \mathbb{E}_t(U_{C,t+1}^b) + \lambda_t^b, \tag{3.5}$$

$$U_{C,t}^b = \beta_b \mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L) + \lambda_t^b(1-\kappa_t), \tag{3.6}$$

$$0 = \lambda_t^b[(1-\kappa_t)L_t - D_t], \qquad \lambda_t^b \geq 0. \tag{3.7}$$

Whenever the leverage constraint is binding, the marginal benefit of issuing deposits and borrowing from workers to consume more at t exceeds the marginal cost of deposit repayments and lower consumption at $t+1$ by the shadow value $\lambda_t^b \geq 0$. If the leverage constraint is slack at t, but there is a positive probability that it will bind at any contingency in the future, the marginal cost of issuing deposits at t is higher than in the absence of the leverage constraint, which can be seen by iterating (3.5) forward. Consequently, bankers would like to decrease borrowing to insure themselves against the future instances of a binding leverage constraint.

Both risk aversion and the leverage constraint lead to a spread between the required expected return on loans and deposits:

$$\mathbb{E}_t(R_{t+1}^L) - R_t = -\operatorname{cov}_t\left[\frac{U_{C,t+1}^b}{\mathbb{E}_t(U_{C,t+1}^b)}, R_{t+1}^L\right] + \frac{\kappa_t \lambda_t^b}{\beta_b \mathbb{E}_t(U_{C,t+1}^b)},$$

which follows from (3.5) and (3.6). The first spread component is a risk premium for holding an asset with procyclical payoffs, present only if bankers are risk averse. The second component arises from the leverage constraint and is positive if and only if $\kappa_t \lambda_t^b > 0$. This component becomes larger when bankers are more constrained: either directly due to a higher capital requirement κ_t or indirectly due to a higher value of the Lagrange multiplier λ_t^b.

3.2.3 Entrepreneurs

Similar to Bernanke et al. (1999) and Iacoviello (2005), entrepreneurs manage firms that produce wholesale goods supplied to retailers. The production process requires capital K_t and labor N_t and is affected by two types of exogenous stochastic disturbances: a total factor productivity (TFP) process A_t and a capital quality process ξ_t. As in Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), the capital stock K_{t-1} purchased yesterday has an effective productive value $\xi_t K_{t-1}$ today. The capital quality process serves as an exogenous source of variation in the asset price and the return on capital. The effective factors of production are combined using a Cobb—Douglas technology $F : \mathbb{R}_+^2 \to \mathbb{R}_+$; therefore, the output of the wholesale good is $Y_t^w \equiv A_t F(\xi_t K_{t-1}, N_t)$.

The entrepreneur consumes C_t^e, buys new capital goods at a relative price Q_t, demands labor from workers, sells the produced wholesale good at a price P_t^w, and obtains external financing from the banking sector. Hence, the entrepreneur's budget constraint is

$$C_t^e + Q_t K_t + W_t N_t + R_t^L L_{t-1} \leq P_t^w A_t F(\xi_t K_{t-1}, N_t) + Q_t(1-\delta)\xi_t K_{t-1} + L_t. \tag{3.8}$$

Following Kiyotaki and Moore (1997), external financing requires collateral. Bankers consider the possibility that entrepreneurs may default, in which case the former could recover a fraction of the value of the entrepreneur's effective capital stock $Q_{t+1}\xi_{t+1}K_t$. Since both the value of collateral and the value of repayment are contingent on the state, bankers will be willing to extend loans to entrepreneurs if

$$\mathbb{E}_t(R_{t+1}^L)L_t \leq m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t, \tag{3.9}$$

where $m_t \in [0, 1]$ reflects recovery costs as perceived by the banker or a policymaker. We will use the latter interpretation and assume that m_t is a policy instrument. Moreover, we will restrict attention to equilibria where in all contingencies, the loan rate R_t^L is such that both bankers and entrepreneurs get strictly positive consumption, and no defaults occur ex-post.

Note how capital quality affects the entrepreneur's budget set. An expected decrease in ξ_{t+1} tomorrow directly tightens the collateral constraint today, leading to a decrease in external financing. An income effect causes a decrease in the entrepreneur's spending, including purchasing new capital goods, which depresses Q_t and K_t. The latter further tightens the collateral constraint, and the logic just described repeats, producing a multiplicative effect of the original shock. Moreover, if the capital quality process is persistent, a decrease in ξ_t today would also trigger the described sequence of events due to a decrease in the anticipated capital quality tomorrow. Another source of financial amplification comes from the forward-looking nature of the asset price Q_t, as demonstrated below.

The entrepreneur's preferences are represented by $\mathbb{E}_0[\sum_{t=0}^{\infty} {\beta_e}^t U^e(C_t^e)]$, where $U^e : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable with $U_C^e > 0$ and $U_{CC}^e \leq 0$. Denoting the normalized Lagrange multiplier on (3.9) as λ_t^e, the KKT conditions include (3.8) as equality, (3.9), the labor demand condition (3.10), the Euler equations for loans (3.11) and capital (3.12), and the complementary slackness conditions (3.13):

$$W_t = P_t^w A_t F_{N,t}, \tag{3.10}$$

$$U^e_{C,t} = \beta_e \mathbb{E}_t(U^e_{C,t+1} R^L_{t+1}) + \lambda^e_t \mathbb{E}_t(R^L_{t+1}), \tag{3.11}$$

$$U^e_{C,t} Q_t = \beta_e \mathbb{E}_t\{U^e_{C,t+1}[P^w_{t+1} A_{t+1} F_{K,t+1} + Q_{t+1}(1-\delta)]\xi_{t+1}\} + \lambda^e_t m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1}), \tag{3.12}$$

$$0 = \lambda^e_t [m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1}) K_t - \mathbb{E}_t(R^L_{t+1}) L_t], \qquad \lambda^e_t \geq 0. \tag{3.13}$$

The collateral constraint affects the entrepreneur's Euler equations (3.11) and (3.12) similar to the way the leverage constraint affects the banker's Euler equations (3.5) and (3.6). When the collateral constraint is binding, the marginal benefit of borrowing is greater than the marginal cost by $\lambda^e_t \mathbb{E}_t(R^L_{t+1})$. Moreover, there is self-insurance against the future states when the collateral constraint binds, as reflected by the greater marginal cost of borrowing compared to the economy without the collateral constraint. The capital Euler equation demonstrates that the asset price Q_t is determined by the expected future payoff from capital and the marginal value of capital used as collateral, both of which depend on Q_{t+1}, making the asset price forward looking. Through the future asset prices, the asset price today also reflects the collateral benefits at all future states when the collateral constraint is binding.

Define the gross return on capital

$$R^K_t \equiv \frac{P^w_t A_t F_{K,t} + Q_t(1-\delta)}{Q_{t-1}} \xi_t.$$

Inspecting (3.11)–(3.13), we can derive a premium between the required expected returns on capital and loans:

$$\mathbb{E}_t(R^K_{t+1} - R^L_{t+1}) = -\operatorname{cov}_t \left[\frac{U^e_{C,t+1}}{\mathbb{E}_t(U^e_{C,t+1})}, R^K_{t+1} - R^L_{t+1} \right] + \frac{\lambda^e_t \mathbb{E}_t(R^L_{t+1})}{\beta_e \mathbb{E}_t(U^e_{C,t+1})} \left(1 - \frac{L_t}{Q_t K_t}\right).$$

When entrepreneurs have enough internal financing to support their business so that the collateral constraint is slack, the premium is determined by the covariance between the future marginal utility and the difference in ex-post returns. The latter is numerically small, and thus in expectation, bankers recover approximately the gross return on capital, similar to Gertler and Kiyotaki (2010). When the collateral constraint is binding, entrepreneurs require a higher expected return on capital so that internal financing could compensate for the lack of available external financing. In this case, bankers can expect to get only a share of the return on capital, and this share is more significant when entrepreneurs fund a greater share of their capital purchases using the banking system. When the amount of external financing is enough to fund the purchase of the new capital goods fully, the expected returns on loans and capital are approximately equal, independently of whether the collateral constraint is slack or binding.

3.2.4 Capital, retail, and final good production

Producers of capital goods combine the input of final goods I_t and the aggregate capital stock available at the beginning of the period K_{t-1} to build new capital goods $\Phi\left(\frac{I_t}{K_{t-1}}\right)K_{t-1}$, where $\Phi:\mathbb{R}_+\to\mathbb{R}$, $\Phi'>0$, $\Phi''\leq 0$, $\lim_{x\to 0}\Phi'(x)=\infty$, and $\lim_{x\to\infty}\Phi'(x)=0$, similar to Lucas and Prescott (1971). A capital good producer maximizes the expected discounted profits $\mathbb{E}_0\left\{\sum_{t=0}^{\infty}\Lambda_{0,t}\left[Q_t\Phi\left(\frac{I_t}{K_{t-1}}\right)K_{t-1}-I_t\right]\right\}$ under perfect competition; therefore, the supply of new capital goods is described by

$$Q_t=\left[\Phi'\left(\frac{I_t}{K_{t-1}}\right)\right]^{-1}. \tag{3.14}$$

There is a unit measure of retail varieties produced by retailers. Each retailer has monopolistic power, internalizing the demand curve of the final good produces. The latter, acting under perfect competition, combine retail varieties into the final good according to a production technology with a constant elasticity of substitution $\epsilon>1$. The retail sector is subject to the pricing mechanism of Calvo (1983) and Yun (1996): at any point in time and any contingency, a retailer cannot reset a price with a probability $\theta\in[0,1]$. Standard derivations imply that retailers that can update their prices choose the same new price, and the following equations hold:

$$\widetilde{P}_t=\frac{\epsilon}{\epsilon-1}\frac{\Omega_{1,t}}{\Omega_{2,t}}, \tag{3.15}$$

$$\Omega_{1,t}=P_t^wY_t+\theta\mathbb{E}_t(\Lambda_{t,t+1}\Pi_{t+1}^{\epsilon}\Omega_{1,t+1}), \tag{3.16}$$

$$\Omega_{2,t}=Y_t+\theta\mathbb{E}_t(\Lambda_{t,t+1}\Pi_{t+1}^{\epsilon-1}\Omega_{2,t+1}), \tag{3.17}$$

$$\Pi_t^{1-\epsilon}=\theta+(1-\theta)(\Pi_t\widetilde{P}_t)^{1-\epsilon}, \tag{3.18}$$

$$\Delta_t=\theta\Pi_t^{\epsilon}\Delta_{t-1}+(1-\theta)\widetilde{P}_t^{-\epsilon}, \tag{3.19}$$

where $\widetilde{P}_t$ is the optimal new relative price, $\Omega_{1,t}$ defined by (3.16) reflects the retailer's expected marginal cost, $\Omega_{2,t}$ defined by (3.17) represents the retailer's expected marginal benefit, Y_t is the aggregate output of the final good, Π_t is the gross inflation rate, and Δ_t is a measure of price dispersion. (3.15) shows that the optimal relative price is set with a time-varying markup over the marginal cost, and (3.18) demonstrates that the optimal relative price is an increasing function of the inflation rate. According to (3.19), price dispersion evolves recursively based on the new optimal price and the aggregate inflation rate, and these two forces affect the price dispersion in opposite directions, implying a stationary relationship.

3.2.5 Market clearing

The capital (3.20), wholesale (3.21), and final (3.22) good market-clearing conditions are

$$K_t = (1-\delta)\xi_t K_{t-1} + \Phi\left(\frac{I_t}{K_{t-1}}\right) K_{t-1}, \tag{3.20}$$

$$A_t F(\xi_t K_{t-1}, N_t) = \Delta_t Y_t, \tag{3.21}$$

$$Y_t = C_t^b + C_t^e + C_t^w + I_t. \tag{3.22}$$

3.2.6 Competitive equilibrium

We are now ready to define a competitive equilibrium.

Definition 3.1. *Given exogenous stochastic processes $\{A_t, \xi_t\}$ and boundary conditions, a sequential competitive equilibrium (CE) is a list of allocations $\{C_t^b, C_t^e, C_t^w, D_t, I_t, K_t, L_t, N_t, Y_t\}$, prices $\{\widetilde{P}_t, P_t^w, Q_t, R_t, R_t^L, W_t\}$, Lagrange multipliers $\{\lambda_t^b, \lambda_t^e\}$, auxiliary objects $\{\Delta_t, \Omega_{1,t}, \Omega_{2,t}\}$, and policies $\{\kappa_t, m_t, \Pi_t\}$, such that:*

1. *Given policies and prices, all agents solve their problems, that is,* (3.1)–(3.19) *hold. (Retailers set the prices of individual retail varieties optimally, generating $\widetilde{P}_t$.)*

2. *Prices are such that market-clearing conditions* (3.20)–(3.22) *are satisfied.*

At this point, we have not specified the nature of the policies $\{\kappa_t, m_t, \Pi_t\}$. The normative analysis will explore how to set the policies optimally. To compute the CE, we will assume that the leverage limits κ_t and m_t are constants, and there is a central bank that targets inflation according to a Taylor rule with an effective lower bound (ELB) $\underline{R} > 0$ on the gross nominal interest rate

$R_t^N \equiv R_t \mathbb{E}_t(\Pi_{t+1})$. Let R_t^* denote the nominal rate when the lower bound is slack. The policy rule can be described as follows:

$$R_t^N = \max(R_t^*, \underline{R}), \qquad R_t^* = (R_{t-1}^*)^{\rho_R} \left[\frac{\bar{\Pi}}{\beta} \left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\eta_\pi} \left(\frac{P_t^w}{P^w} \right)^{\eta_y} \right]^{1-\rho_R}, \tag{3.23}$$

where $\rho_R \in [0,1)$, $\bar{\Pi} \geq 1$ is the central bank's gross inflation target, and $(\eta_\pi, \eta_y) \in \mathbb{R}_+^2$ are the response parameters. The deviation of the retailer's marginal cost from the steady state is a proxy for the output gap. (The exact relationship holds in the basic New Keynesian model.) Note that if we use (3.23) to determine $\{\Pi_t\}$, the latter is endogenous to our economy. In the context of definition 3.1, it means that there is an implicit consistency condition that requires $\{\Pi_t\}$ to satisfy (3.23). Although an ELB in (3.23) necessarily generates a multiplicity of equilibria (Benhabib et al., 2001), we will restrict attention to the conventional targeted-inflation regime since it appears to be consistent with the US data (Aruoba et al., 2018). Using specifications similar to (3.23), Braun and Körber (2011) and Fernández-Villaverde et al. (2015) have also argued for selecting a conventional equilibrium. Moreover, we will show in the normative analysis that optimal monetary policy is uniquely determined even in the presence of an ELB.

Let us complete the CE description with two lemmas that characterize the deterministic steady state and give more insight into the optimal decisions of bankers and entrepreneurs. Define

$$\widetilde{\beta}_e \equiv \frac{\beta}{1 + \kappa \left(\frac{\beta}{\beta_b} - 1 \right)}.$$

Lemma 3.1. *Conditional on $\Pi = \bar{\Pi}$, there exists a unique steady state with positive financial flows if and only if $\beta_b < \beta$ and $\beta_e < \widetilde{\beta}_e$. In this steady state, (3.4) and (3.9) are binding.*

The intuition for lemma 3.1 is clearer after we rewrite the inequalities $\beta_b < \beta$ and $\beta_e < \widetilde{\beta}_e$ as $\beta_b R < 1$ and $\beta_e R^L < 1$, which follows from (3.2), (3.5), (3.6), and (3.11). The latter conditions mean that bankers and entrepreneurs would like to borrow in a steady-state equilibrium because the effective rate of time preference exceeds the interest rate. This condition is consistent with the analysis of the income fluctuations problem of Schechtman and Escudero (1977). If $\beta_b = \beta$, any amount of deposits that satisfies the leverage constraint is associated with an unstable steady state. Note that $\beta_b > \beta$ is ruled out by construction. Similarly, if $\beta_e = \widetilde{\beta}_e$, the quantity of loans is indeterminate. If $\beta_e \in (\widetilde{\beta}_e, \beta]$, then entrepreneurs would choose $L \leq 0$. To make the analysis interesting, we will assume strict inequalities in both cases.

Assumption 3.1. *$\beta_b < \beta$ and $\beta_e < \widetilde{\beta}_e$.*

The following lemma shows that net assets equal the lifetime stream of consumption discounted at the agent-specific stochastic discount factor for both bankers and entrepreneurs.

Lemma 3.2. *At the optimum, bank capital satisfies $L_t - D_t = \frac{1}{U^b_{C,t}} \sum_{s=1}^{\infty} \beta_b^s \mathbb{E}_t(U^b_{C,t+s} C^b_{t+s})$. Similarly, the entrepreneur's net assets satisfy $Q_t K_t - L_t = \frac{1}{U^e_{C,t}} \sum_{s=1}^{\infty} \beta_e^s \mathbb{E}_t(U^e_{C,t+s} C^e_{t+s})$.*

Note that we have simple contemporaneous relationships with logarithmic preferences: $L_t - D_t = \frac{\beta_b}{1-\beta_b} C^b_t$ and $Q_t K_t - L_t = \frac{\beta_e}{1-\beta_e} C^e_t$. Since the banker's net worth is $R^L_t L_{t-1} - R_{t-1} D_{t-1}$, using (3.3), we see that the banker allocates the majority of her net worth—a share β_b—for bank capital, while the remaining share $1 - \beta_b$ is allocated for consumption (dividends). The more patient the banker is, the greater is the share of net worth reinvested back into the banking business. Due to the Inada condition, consumption is guaranteed to be positive, which implies that bankers would optimally like to hold a positive amount of bank capital independently of the capital requirement, that is, even if $\kappa_t = 0$.

Similarly, if we define the entrepreneur's net worth as $R^K_t Q_{t-1} K_{t-1} - R^L_t L_{t-1}$ and consider logarithmic preferences, then (3.8), (3.10), and lemma 3.2 imply that the entrepreneur's net assets take a share β_e of net worth, while consumption takes the remaining share $1 - \beta_e$. Since $C^e_t > 0$ due to the Inada condition, and thus $L_t < Q_t K_t$, entrepreneurs fund the purchases of new capital goods with a nontrivial combination of internal and external financing. Consequently, when the collateral constraint is binding, entrepreneurs will require a strictly higher expected return on capital $\mathbb{E}_t(R^K_{t+1})$ than the loan rate $\mathbb{E}_t(R^L_{t+1})$, as follows from the premium derived at the end of section 3.2.3.

Furthermore, when both (3.4) and (3.9) are binding, lemma 3.2 implies that with logarithmic preferences, the consumption ratio of constrained bankers and entrepreneurs can be expressed as a function of policies and prices only:

$$\frac{C^e_t}{C^b_t} = \frac{1-\beta_e}{\beta_e} \frac{\beta_b}{1-\beta_b} \frac{1}{\kappa_t} \left[\frac{Q_t \mathbb{E}_t(R^L_{t+1})}{m_t \mathbb{E}_t(Q_{t+1} \xi_{t+1})} - 1 \right].$$

Other things equal, the more impatient agent tends to consume more. A higher bank capital requirement causes bankers to accumulate more net worth, positively affecting consumption. A greater value of collateral per unit of capital stock makes entrepreneurs use relatively more external financing, leading to lower net assets and consumption. Conversely, a higher expected loan rate decreases the available quantity of bank loans for a given value of collateral, increasing the entrepreneur's share of internal financing, net assets, and consumption. A greater price of capital at t also has a positive partial effect on net assets and consumption. Note that the consumption ratio's dependence on the leverage limits anticipates the latter's ability to enhance risk sharing between constrained bankers and entrepreneurs.

3.3 Normative analysis

The purpose of this section is: first, to demonstrate how endogenous financial constraints, nominal rigidities, and consumer type heterogeneity make the CE allocation inefficient; second, to show how to decentralize the constrained efficient allocation with the appropriate fiscal instruments; and third, to characterize Ramsey-optimal leverage limits and monetary policy both when the above-mentioned fiscal instruments are available to the policymaker and when they are not. To understand the differential role of financial frictions and nominal rigidities, we will start by characterizing efficiency and Ramsey-optimal leverage limits in a flexible-price economy with a perfectly competitive retail sector. We will then study constrained efficiency in the benchmark sticky-price economy and will characterize jointly Ramsey-optimal monetary policy and leverage limits under alternative sets of available fiscal instruments.

To begin with, we must define a welfare objective. Since we have ex-ante heterogeneous consumers—workers, bankers, and entrepreneurs—a benevolent social planner should care about all of them. Due to lemma 3.1, our economy has well-defined local dynamics only when bankers and entrepreneurs are sufficiently impatient relative to workers. Suppose we take a weighted average of the agents' lifetime utility functions as a welfare objective. Due to the differences in patience, a relatively more impatient consumer could get a socially optimal consumption plan that asymptotically converges to zero. Following Andrés et al. (2013), a way to achieve stationarity is to add the lifetime utilities of all future newborn impatient consumers to the welfare objective.

Definition 3.2. *Let* $V_t^i \equiv \mathbb{E}_t(\sum_{s=0}^{\infty} \beta_i^s U_{t+s}^i)$ *denote the lifetime utility of a representative consumer of type* $i \in \mathcal{I}$ *living at* $t \geq 0$*. The social welfare objective at* $t \geq 0$ *is* $\mathcal{W}_t \equiv \sum_{i \in \mathcal{I}} \omega_i \mathcal{W}_t^i$*, where* $\omega_i \geq 0$ *for all* $i \in \mathcal{I}$*, and* $\mathcal{W}_t^i \equiv V_t^i + \frac{\beta - \beta_i}{\beta} \mathbb{E}_t(\sum_{s=1}^{\infty} \beta^s V_{t+s}^i)$*, with* $\beta_w \equiv \beta$*.*

Consider the aggregate welfare of type i consumers $\mathcal{W}_t^i$: it is a sum of the lifetime utility of the representative consumer living at $t \geq 0$ and the discounted expected lifetime utilities of all future newborns. By definition, β_i equals β adjusted for the survival probability. Therefore, the exit probability is $\frac{\beta - \beta_i}{\beta}$, and it equals the measure of newborns. It turns out that $\mathcal{W}_t^i$ has an equivalent representation independent of the type-specific survival probability.

Lemma 3.3. *The aggregate welfare of type* i *consumers satisfies* $\mathcal{W}_t^i = \mathbb{E}_t(\sum_{s=0}^{\infty} \beta^s U_{t+s}^i)$*.*

The intuition for lemma 3.3 is that by adding the welfare of future newborns to the welfare objective, we can exactly compensate for the uncertain survival of the currently living impatient consumers.

3.3.1 Flexible-price economy

In this section, we will consider the flexible-price economy. We will, first, characterize the unconstrained Pareto-optimal allocation that will serve as a reference for welfare comparisons. Second, we will study the constrained efficient allocation and show how to decentralize it in a regulated competitive equilibrium with taxes. Finally, we will explore Ramsey-optimal leverage limits under alternative sets of fiscal instruments available to the Ramsey planner.

The flexible-price economy is a special case of the economy studied in section 3.2 after setting $\theta = 0$ and $\epsilon \to \infty$. In this case, (3.15)–(3.19) imply $\widetilde{P}_t = P_t^w = \Delta_t = 1$, $\Omega_{1,t} = \Omega_{2,t} = Y_t$, and Π_t becomes immaterial. Accordingly, we can revise definition 3.1 to define a competitive equilibrium in such a setting.

Definition 3.3. *Given exogenous stochastic processes* $\{A_t, \xi_t\}$ *and boundary conditions, a flexible-price competitive equilibrium (FCE) is a list of allocations* $\{C_t^b, C_t^e, C_t^w, D_t, I_t, K_t, L_t, N_t, Y_t\}$*, prices* $\{Q_t, R_t, R_t^L, W_t\}$*, Lagrange multipliers* $\{\lambda_t^b, \lambda_t^e\}$*, and policies* $\{\kappa_t, m_t\}$*, such that:*

1. *Given policies and prices, all agents solve their problems:* (3.1)–(3.14) *hold with* $P_t^w = 1$*.*

2. *Prices are such that market-clearing conditions* (3.20)–(3.22) *are satisfied with* $\Delta_t = 1$*.*

First best

As a benchmark for welfare comparisons, consider an unconstrained Pareto-optimal allocation—"first best"—associated with the flexible-price economy. This allocation is an outcome of a planning problem where a benevolent social planner directly allocates consumption and factors of production subject to resource constraints. Conditional on Pareto weights $(\omega_b, \omega_e, \omega_w) \in \mathbb{R}^3_+$, the first-best allocation is a solution to

$$\max_{\{C_t^b, C_t^e, C_t^w, I_t, K_t, N_t\}} \mathbb{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \sum_{i \in \mathcal{I}} \omega_i U_t^i \right)$$

subject to

$$\begin{aligned}
\lambda_t^K : &\quad 0 \leq (1-\delta)\xi_t K_{t-1} + \Phi\left(\frac{I_t}{K_{t-1}}\right) K_{t-1} - K_t, \\
\lambda_t^Y : &\quad 0 \leq A_t F(\xi_t K_{t-1}, N_t) - \sum_{i \in \mathcal{I}} C_t^i - I_t.
\end{aligned}$$

The first-order conditions (FOCs) for C_t^i, N_t, I_t, and K_t can be written as

$$\begin{aligned}
\lambda_t^Y &= \omega_i U_{C,t}^i, \\
-\frac{U_{N,t}^w}{U_{C,t}^w} &= A_t F_{N,t}, \\
\frac{\lambda_t^K}{\lambda_t^Y} &= \left[\Phi'\left(\frac{I_t}{K_{t-1}}\right)\right]^{-1}, \\
\frac{\lambda_t^K}{\lambda_t^Y} U_{C,t}^e &= \beta \mathbb{E}_t \Bigg[U_{C,t+1}^e \Bigg\{ \left[A_{t+1} F_{K,t+1} + \frac{\lambda_{t+1}^K}{\lambda_{t+1}^Y}(1-\delta) \right] \xi_{t+1} \\
&\qquad + \frac{\lambda_{t+1}^K}{\lambda_{t+1}^Y} \left[\Phi\left(\frac{I_{t+1}}{K_t}\right) - \Phi'\left(\frac{I_{t+1}}{K_t}\right) \frac{I_{t+1}}{K_t} \right] \Bigg\} \Bigg].
\end{aligned}$$

At the unconstrained Pareto optimum, we have perfect consumption risk sharing between workers, bankers, and entrepreneurs. By construction, the first-best problem ignores the occupational differences reflected in the individual budget constraints, and bankers and entrepreneurs face no financial constraints. As can be shown numerically, the marginal utility gaps in the FCE are quite significant. If all consumers have separable preferences logarithmic in consumption, workers tend to consume by an order of magnitude more than bankers and entrepreneurs, despite being more patient. Thus, we can anticipate that one of the objectives of a constrained planner in our economy is to improve between-agent consumption insurance.

The labor market equilibrium in the FCE is consistent with the first best, as follows from combining (3.1) and (3.10) and setting $P_t^w = 1$. By defining $Q_t \equiv \frac{\lambda_t^K}{\lambda_t^Y}$, we see that the competitive supply of new capital goods is efficient. On the contrary, the competitive demand for capital is inefficient, as follows from comparing the FOC for K_t to the capital Euler equation (3.12) with $P_t^w = 1$. On the one hand, due to uncertain survival, individual entrepreneurs underestimate the social marginal benefit of capital due to its usefulness for future newborns. On the other hand, entrepreneurs find a marginal benefit in capital stock due to its value as collateral—a motive absent in the planner's problem. Moreover, entrepreneurs do not internalize the impact of their private decisions on the productive capacity of capital good producers. This latter effect is present if and only if the technology Φ is nonlinear.

To summarize, the FCE is generally first-best inefficient, manifested in the lack of between-agent consumption risk sharing and the inefficient demand for capital.

Constrained efficient allocation

Now let us turn to the second-best efficiency. Following Lorenzoni (2008), consider a constrained efficient allocation chosen by a benevolent planner who faces the same constraints as private agents but internalizes the impact of allocations on market prices. In our flexible-price economy, we have four market prices: Q_t, R_t, R_t^L, and W_t. In the corresponding markets for factors of production and financial assets, both the market demand and supply are endogenously determined, which implies that there are multiple concepts of constrained optimality in our framework, with potentially different implications for the welfare and efficiency of the FCE. Since the worker's problem has no financial frictions, while bankers and entrepreneurs face endogenous financial constraints, we will focus on how the planner can improve over the competitive market allocation by making decisions on behalf of bankers and entrepreneurs. We will allow the planner to intervene in all the aforementioned markets, considering the most general set-up. Since our economy features consumer type heterogeneity, the sources of constrained inefficiency may not be limited to pecuniary externalities due to prices affecting the collateral constraint.

On the banker's side, the planner chooses deposits, internalizing the demand curve $R_t =$

$R(U_C^w(C_t^w, N_t), \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})])$ implied by the worker's Euler equation (3.2). Bankers still choose consumption and loans, taking the deposit allocation as given. Hence, the implementability conditions include the banker's (binding) budget constraint (3.3), the leverage constraint (3.4), the Euler equation for loans (3.6), and the complementary slackness conditions (3.7). These conditions can be simplified as follows. Using the budget constraint, we can solve for the loan repayment $B_t \equiv R_t^L L_{t-1} = C_t^b + L_t - D_t + R_{t-1}D_{t-1}$. The Euler equation for loans then implies $\lambda_t^b(1-\kappa_t)L_t = U_{C,t}^b L_t - \beta_b \mathbb{E}_t(U_{C,t+1}^b B_{t+1})$. If $\kappa_t < 1$ and $D_t > 0$, the leverage constraint implies $L_t > 0$, and thus the complementary slackness conditions are equivalent to $0 = \lambda_t^b(1-\kappa_t)L_t[(1-\kappa_t)L_t - D_t]$ and $\lambda_t^b(1-\kappa_t)L_t \geq 0$. If $\kappa_t < 1$ and $D_t = 0$, the leverage constraint is equivalent to $L_t \geq 0$, which is independently implied by the nonnegativity of consumption; therefore, in this case, $\lambda_t^b = 0$, and the complementary slackness conditions are satisfied. If $\kappa_t = 1$, the leverage constraint leaves $D_t = 0$ as the only choice, again implying $\lambda_t^b = 0$. Hence, if $D_t = 0$, we have $U_{C,t}^b L_t = \beta_b \mathbb{E}_t(U_{C,t+1}^b B_{t+1})$.

The planner chooses capital stock, labor, and loans on the entrepreneur's side, internalizing the corresponding prices. The worker's labor supply curve (3.1) determines the wage rate $W_t = W(C_t^w, N_t)$. The capital good producer's supply curve (3.14) defines the price of capital $Q_t = Q(K_{t-1}, K_t, \xi_t)$ after using the capital good market-clearing condition (3.20) to solve for $I_t = I(K_{t-1}, K_t, \xi_t)$. The return on loans must be consistent with the banker's Euler equation, one of the implementability conditions on the banker's side. Entrepreneurs themselves only make consumption decisions, which implies that the budget constraint (3.8) is binding, and entrepreneurs consume the "endowment" determined by the planner's choices. Apart from the binding budget constraint, the planner faces the same collateral constraint (3.9) as the individual entrepreneur.

Based on definition 3.3, the only remaining implementability constraints are the market-clearing conditions (3.21)—with $\Delta_t = 1$—and (3.22), which can be combined in one resource constraint for the final good. The constrained efficient allocation is thus defined as follows.

Definition 3.4. *A flexible-price constrained efficient allocation (FCEA) is a solution to*

$$\max_{\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t\}} \mathbb{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \sum_{i \in \mathcal{I}} \omega_i U_t^i \right)$$

subject to

$$\begin{aligned}
\lambda_t^b: &\quad 0 \leq (1-\kappa_t)L_t - D_t, \\
\lambda_{1,t}^L: &\quad 0 \leq U_C^b(C_t^b)L_t - \beta_b \mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t)], \qquad \textit{equality if } D_t = 0, \\
\lambda_{2,t}^L: &\quad 0 = \{U_C^b(C_t^b)L_t - \beta_b \mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t)]\}[(1-\kappa_t)L_t - D_t], \\
\lambda_t^C: &\quad 0 = A_t F(\xi_t K_{t-1}, N_t) - Q(K_{t-1}, K_t, \xi_t)[K_t - (1-\delta)\xi_t K_{t-1}] - W(C_t^w, N_t)N_t + D_t \\
&\qquad - R_{t-1}D_{t-1} - C_t^b - C_t^e, \\
\lambda_t^e: &\quad 0 \leq m_t \mathbb{E}_t(Q(K_t, K_{t+1}, \xi_{t+1})\xi_{t+1})K_t - \mathbb{E}_t(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t), \\
\lambda_t^Y: &\quad 0 = A_t F(\xi_t K_{t-1}, N_t) - \sum_{i \in \mathcal{I}} C_t^i - I(K_{t-1}, K_t, \xi_t),
\end{aligned}$$

where $R_t = R(U_C^w(C_t^w, N_t), \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})])$*, and the functions* W*,* R*,* Q*, and* I *are defined by* (3.1)*,* (3.2)*,* (3.14)*, and* (3.20)*, respectively.*

Definition 3.4 implies that the FCE is generally constrained inefficient. The collateral constraint has a conventional pecuniary externality due to the price of capital that affects the value of collateral and an externality working through the expected loan rate affected by the banker's loan supply decisions. Moreover, since we have heterogeneous consumers, only one of the budget constraints is redundant, which we chose to be the worker's. The combined budget constraint of bankers and entrepreneurs depends on market prices, resulting in additional externalities that arise even if the collateral constraint is slack with probability one. The bank leverage constraint is not a source of inefficiency since it is independent of prices; however, the associated market complementary slackness conditions combined with the banker's loan supply Euler equation may affect the efficiency of loan demand. If the worker's preferences are not separable in consumption and leisure, the market deposit rate depends on the labor allocation, potentially creating another externality.

Let $\lambda_t^L \equiv \lambda_{1,t}^L + \lambda_{2,t}^L[(1-\kappa_t)L_t - D_t]$. The following proposition formalizes the intuitive discussion above and presents other findings.

Proposition 3.1. *The FCE allocation is constrained inefficient: the right-hand sides of the planner's analogs of* (3.5) *and* (3.10)*–*(3.12) *have additional terms* Ψ_t^D*,* Ψ_t^L*,* Ψ_t^N*, and* Ψ_t^K*. Moreover, the FCEA has the following properties.*

1. *There is generally imperfect consumption insurance. The risk sharing between bankers and entrepreneurs is perfect across the contingencies where* $\lambda^L_t = \lambda^L_{t-1} = \lambda^e_{t-1} = 0$.

2. *Suppose* $U^w(C^w, N) = u(C^w) - v(N)$*, the steady-state profits of capital good producers are zero, and* $\lambda^e = 0$*. The steady-state Pareto-weighted marginal utilities of all consumers are equal—risk sharing is "approximately perfect"—if and only if* $u(\cdot) = \ln(\cdot)$.

3. *There exists* $\bar{D} > 0$*, such that any* $D \in [0, \bar{D}]$ *defines an unstable steady state. The optimal constant plan in the absence of uncertainty—optimal steady state—features* $D = 0$*, provided that* $\lambda^C > 0$ *if* $D > 0$.

Wedges Proposition 3.1 states that the FCE is constrained inefficient due to the additional terms present in the planner's optimality conditions that reflect the wedges between the FCE and FCEA. The derivation of the wedges is provided in the proof, and here we will explore their structure.

The wedge associated with deposit supply (3.5) is

$$\Psi^D_t \equiv (\beta - \beta_b) R_t \mathbb{E}_t(U^b_{C,t+1}) + \frac{\lambda^Y_t - \beta R_t \mathbb{E}_t(\lambda^Y_{t+1})}{\omega_b} + \Gamma^D_t,$$

where Γ^D_t represents all the terms that arise from the market loan supply and complementary slackness conditions, vanishing in the neighborhood of the steady state under the baseline calibration. The term $(\beta - \beta_b) R_t \mathbb{E}_t(U^b_{C,t+1}) > 0$ arises from the uncertain survival of bankers: the social marginal cost of deposit issuance is greater than the private marginal cost, since future newborn bankers will have to honor the liabilities of the exiting ones. The term $\lambda^Y_t - \beta R_t \mathbb{E}_t(\lambda^Y_{t+1})$ reflects the planner's risk-sharing goals and appears because, with heterogeneous consumers, both the consolidated budget constraint of bankers and entrepreneurs and the resource constraint matter to the planner. When resources are scarce, e.g., A_t or ξ_t is low, then λ^Y_t is higher, and the resource constraint is "more binding," so the planner may need to decrease the consumption of all consumers. In such states, it is more costly for bankers to borrow from the planner's perspective because the leverage constraint would require expanding assets, bank capital, and consumption. In the steady state, $\beta R = 1$ from (3.2), so the risk-sharing component is zero.

If $D_t > 0$, the wedge corresponding to the loan demand condition (3.11) is

$$\Psi_t^L = (\beta - \beta_e)\mathbb{E}_t(U^e_{C,t+1}R^L_{t+1}) + \frac{\lambda_t^Y - \beta R_t \mathbb{E}_t(\lambda^Y_{t+1})}{\omega_e} - \mathbb{E}_t\left[\left(\beta U^e_{C,t+1} + \frac{\lambda^e_t}{\omega_e}\right)(R^L_{t+1} - R_t)\right]$$
$$+ \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\frac{\kappa_t}{1-\kappa_t}\frac{\lambda^e_{t-1}}{\omega_e} + \Gamma^L_t,$$

where $\mathbf{1}_{\mathbb{N}}(t)$ equals 1 if $t > 0$, and Γ^L_t reflects the marginal effect of L_t on the bank loan supply and private complementary slackness conditions, vanishing in the neighborhood of the steady state. The first two components of the loan wedge are symmetric to the deposit wedge. The term $-\mathbb{E}_t\left[\left(\beta U^e_{C,t+1} + \frac{\lambda^e_t}{\omega_e}\right)(R^L_{t+1} - R_t)\right] \leq 0$ demonstrates that the private marginal cost of borrowing is inefficiently high when there is a positive credit spread. This component arises because the planner borrows from workers on behalf of entrepreneurs effectively at the deposit interest rate R_t, which is a consequence of aggregating the budget constraints of bankers and entrepreneurs. The term $\frac{\kappa_t}{1-\kappa_t}\frac{\lambda^e_{t-1}}{\omega_e} \geq 0$ demonstrates that if the collateral constraint is binding at $t-1$, the planner would like to increase the marginal cost of borrowing at the continuation histories at t. The lower expected borrowing at t decreases the expected loan rate and relaxes the collateral constraint at $t-1$. The higher the bank capital requirement, the stronger this effect, since the bank balance sheet implies a positive relationship between the return on loans and bank capital.

If $D_t = 0$, the loan demand wedge is

$$\Psi_t^L = (\beta_b - \beta_e)\mathbb{E}_t(U^e_{C,t+1}R^L_{t+1}) - \frac{\beta - \beta_b}{\beta}\frac{\lambda^e_t}{\omega_e}\mathbb{E}_t(R^L_{t+1}) + \Gamma^L_t,$$

where Γ^L_t is generally not identical to the term present when $D_t > 0$ but has a similar interpretation. The component $(\beta_b - \beta_e)\mathbb{E}_t(U^e_{C,t+1}R^L_{t+1})$ reflects potential differences in the survival rates of bankers and entrepreneurs. If entrepreneurs are relatively more patient, the planner wants to decrease the marginal cost of borrowing and allow more external financing, leading to lower net assets and consumption. Due to $-\frac{\beta-\beta_b}{\beta}\frac{\lambda^e_t}{\omega_e}\mathbb{E}_t(R^L_{t+1}) \leq 0$, the marginal cost of borrowing is lower if the banker's survival is more uncertain, and the collateral constraint is binding at t: higher loan demand at t increases the expected loan rate and net worth of newborn bankers.

The wedge relative to the planner's analog of the labor demand condition (3.10) is

$$\Psi_t^N = \frac{(\omega_e U_{C,t}^e - \omega_w U_{C,t}^w - \lambda_t^C) A_t F_{N,t} - \lambda_t^C W_{N,t} N_t}{\omega_w U_{C,t}^w + \lambda_t^C} + \Gamma_t^N,$$

where Γ_t^N reflects the marginal effect of the choice of labor on the interest rate R_t and vanishes if U^w is separable in consumption and leisure. There are two sources of the labor wedge: imperfect consumption risk sharing ($\omega_e U_{C,t}^e \neq \omega_w U_{C,t}^w$) and the positive shadow value of wealth ($\lambda_t^C > 0$). (These two sources also determine the Γ_t^N term, as is clear from the proof of proposition 3.1.) In the first-best allocation, risk sharing is perfect, and only the resource constraint is relevant, that is, $\lambda_t^C = 0$; therefore, the labor wedge is zero, consistent with section 3.3.1. The term $(\omega_e U_{C,t}^e - \omega_w U_{C,t}^w - \lambda_t^C) A_t F_{N,t}$ reflects the differences in the marginal utility valuation of the marginal product of labor by workers and entrepreneurs. If risk sharing is "approximately perfect," then $(\omega_e U_{C,t}^e - \omega_w U_{C,t}^w - \lambda_t^C) A_t F_{N,t} \approx -\lambda_t^C A_t F_{N,t} < 0$. Since W represents the market supply curve, $W_{N,t} > 0$ and $-\lambda_t^C W_{N,t} N_t < 0$, reflecting that individual entrepreneurs do not internalize how their labor demand affects the equilibrium wage. Since $\omega_w U_{C,t}^w + \lambda_t^C > 0$, we have $\Psi_t^N < 0$. The latter means that the planner would like to decrease labor demand for a given wage. At the same time, the planner would like to decrease the wage to redistribute part of the worker's labor income to entrepreneurs and achieve some convergence in Pareto-weighted marginal utilities. By lowering the wage, the planner could support a greater labor demand. Numerically, the FCE wage is inefficiently high, and the quantity of labor is inefficiently low but to a smaller extent; therefore, the FCEA entails an increase in labor supply and a slight decrease in labor demand.

The wedge relative to the planner's analog of the market demand for capital (3.12) satisfies

$$\begin{aligned}\omega_e \Psi_t^K = {} & (\beta - \beta_e)\mathbb{E}_t(\omega_e U_{C,t+1}^e R_{t+1}^K) Q_t + \beta \mathbb{E}_t \left\{ \lambda_{t+1}^Y \left[Q_{t+1} \Phi \left(\frac{I_{t+1}}{K_t} \right) - \frac{I_{t+1}}{K_t} \right] \right\} \\ & - \lambda_t^C Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] - \beta \mathbb{E}_t \{ \lambda_{t+1}^C Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1} K_t] \} \\ & + \lambda_t^e m_t \mathbb{E}_t(Q_{1,t+1}\xi_{t+1}) K_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda_{t-1}^e m_{t-1} Q_{2,t} \xi_t K_{t-1}.\end{aligned}$$

Similar to other Euler equation wedges, $(\beta - \beta_e)\mathbb{E}_t(\omega_e U_{C,t+1}^e R_{t+1}^K) Q_t > 0$ reflects uncertain survival. The component $+\beta \mathbb{E}_t \left\{ \lambda_{t+1}^Y \left[Q_{t+1} \Phi \left(\frac{I_{t+1}}{K_t} \right) - \frac{I_{t+1}}{K_t} \right] \right\}$ demonstrates that entrepreneurs do not internalize how the choice of capital affects the future profits of capital good producers through a nonlinear technology Φ, which, in turn, affects the amount of resources available for all consumers and is valued at the shadow value of output. If the steady-state profits are zero, so is this wedge component, but it is not generally zero in the neighborhood of the steady state. The sign of λ^Y_{t+1} is generally ambiguous but typically positive. The next two terms $-\lambda^C_t Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] \leq 0$ and $-\beta \mathbb{E}_t\{\lambda^C_{t+1} Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1} K_t]\} \geq 0$ reflect the marginal effect of an increase in capital stock at t on the aggregate wealth of bankers and entrepreneurs at t and $t+1$, respectively, transmitted through the price of capital. Finally, there are pecuniary externalities present in the collateral constraint. First, $\lambda^e_t m_t \mathbb{E}_t(Q_{1,t+1}\xi_{t+1})K_t \leq 0$ reflects a lower social marginal benefit of capital at t due to a lower collateral asset price at $t+1$, stemming from the concave capital good technology Φ. On the contrary, $\lambda^e_{t-1} m_{t-1} Q_{2,t}\xi_t K_{t-1} \geq 0$ represents an additional marginal benefit of capital at t due to a higher asset price and the value of collateral expected at $t-1$.

Risk sharing and the optimal steady state Consider the remaining parts of proposition 3.1. First, we do not generally have perfect consumption insurance between all types of consumers at the second best. Across the contingencies where the market loan supply and the banker's private complementary slackness conditions are slack at t and $t-1$, and the collateral constraint is slack at $t-1$, insurance between bankers and entrepreneurs is perfect. In this case, C^b_t and C^e_t affect the planner's budget set in an identical linear way through the consolidated budget constraint of bankers and entrepreneurs and the final good resource constraint.

Second, there is a special case when we have approximately perfect between-agent insurance in the neighborhood of the steady state. The latter holds when workers have separable preferences over consumption and leisure with a unit constant relative risk aversion, capital good producers earn zero profits in the steady state, and the steady-state collateral constraint is slack. A sufficient condition for zero steady-state profits is $Q = \xi = 1$ and $\frac{I}{K} = \delta$—standard normalizations or calibration targets. Although insurance between workers and constrained consumers is only approximately perfect, the correlation between marginal utilities is quantitatively close to one.

Third, the FCEA is locally indeterminate: any $D \in [0, \bar{D}]$ defines a steady state, where $\bar{D}$ corresponds to the case when either the collateral constraint or the bank leverage constraint is binding. The multiplicity is resolved if we consider the optimal constant plan in the absence of uncertainty. Any FCEA steady state satisfies the planner's constraints in the absence of uncertainty, being a feasible constant plan. It turns out that $D = 0$ is part of the optimal plan if the consolidated budget constraint is relevant to the planner, that is, if $\lambda^C > 0$. Intuitively, by decreasing the quantity of deposits, the planner can allocate more consumption to bankers and entrepreneurs because $-(R-1)D \leq 0$. Since L must decrease to satisfy the private complementary slackness conditions, the collateral constraint is relaxed. To satisfy the resource constraint, the planner can increase both labor and the worker's consumption to achieve a Pareto improvement relative to any constant plan with $D > 0$. Thus, we can restrict attention to the steady state corresponding to $D = 0$—the optimal steady state.

Decentralization Consider now how to decentralize the FCEA in a regulated FCE. A natural way to address the wedges is through proportional taxation rebated lump sum, as described in the following proposition.

Proposition 3.2. *The FCEA can be decentralized in a regulated FCE with linear taxes rebated lump sum. Compared to the FCE, the banker's budget constraint is modified as*

$$C_t^b + L_t \leq R_t^L L_{t-1} - R_{t-1} D_{t-1} + (1 - \tau_t^D) D_t + T_t^b,$$

where τ_t^D *and* $T_t^b \equiv \tau_t^D D_t$ *are taken as given by the individual banker. The entrepreneur's budget constraint is modified as*

$$C_t^e + (1+\tau_t^K) Q_t K_t + (1+\tau_t^N) W_t N_t + R_t^L L_{t-1} \leq A_t F(\xi_t K_{t-1}, N_t) + Q_t(1-\delta)\xi_t K_{t-1} + (1-\tau_t^L) L_t + T_t^e,$$

where $(\tau_t^K, \tau_t^N, \tau_t^L)$ *and* $T_t^e \equiv \tau_t^K Q_t K_t + \tau_t^N W_t N_t + \tau_t^L L_t$ *are taken as given by the individual entrepreneur. The taxes defined in terms of the FCEA are*

$$\tau_t^D = \frac{1}{U_{C,t}^b} \left[\frac{\lambda_t^b}{\omega_b} - \frac{U_{C,t}^b - \beta_b \mathbb{E}_t (U_{C,t+1}^b R_{t+1}^L)}{1 - \kappa_t} + \Psi_t^D \right],$$

$$\tau_t^N = \frac{-\Psi_t^N}{W_t}, \qquad \tau_t^L = \frac{\Psi_t^L}{U_{C,t}^e}, \qquad \tau_t^K = \frac{-\Psi_t^K}{U_{C,t}^e Q_t}.$$

Furthermore, the FCEA and $\{\tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K, T_t^d, T_t^e\}$ *defined above constitute the allocation-policy pair chosen by the Ramsey planner that selects the best regulated FCE.*

The taxes applied to entrepreneurs are simple functions of the wedges. The deposit supply tax τ_t^D reflects potential differences between the normalized social Lagrange multiplier on the bank leverage constraint $\frac{\lambda_t^b}{\omega_b}$ and the private Lagrange multiplier $\frac{U^b_{C,t}-\beta_b \mathbb{E}_t(U^b_{C,t+1}R^L_{t+1})}{1-\kappa_t}$ expressed based on (3.6). Since $\Psi_t^N < 0$, it must be that $\tau_t^N > 0$: it is optimal to tax the entrepreneur's labor demand. The signs of the other wedges and taxes are generally ambiguous, necessitating quantitative analysis. The policy that decentralizes the FCEA is Ramsey optimal. Moreover, additional taxation instruments cannot improve over the second-best optimum unless the planner can directly set prices instead of internalizing the price functions arising in the competitive markets.

Optimal leverage limits

Since $\{\kappa_t, m_t\}$ are exogenous to the FCE, we have considered them as given so far. Let us now study how to set these policies optimally. We will focus on two cases based on whether the Ramsey planner can address all distortions with the complete set of taxes $\{\tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K, T_t^d, T_t^e\}$ or the planner can only account for the intratemporal labor wedge with $\{\tau_t^N, T_t^e\}$. Loosely speaking, the first case corresponds to finding the best FCEA by setting the leverage limits optimally. In the second case, the regulated FCE is constrained inefficient, and we can explore the merits of state-contingent leverage limits in mitigating the Euler equation distortions.

By proposition 3.2, conditional on $\{\kappa_t, m_t\}$, setting $\{\tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K, T_t^d, T_t^e\}$ optimally amounts to solving for the FCEA. Suppose the Ramsey planner can also optimize with respect to $\{\kappa_t, m_t\}$. Since the leverage limits determine the strictness of inequality constraints, the optimal $\{\kappa_t, m_t\}$ are generally not unique: if a leverage constraint is slack at a specific leverage limit, it is also slack at any other feasible looser limit. However, the associated Ramsey allocation is typically unique and can be characterized using the primal approach, as stated in the following lemma.

Lemma 3.4. *An allocation $\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t\}$ and policy $\{\kappa_t, m_t, \tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K, T_t^d, T_t^e\}$ are part of a Ramsey equilibrium associated with the regulated FCE of proposition 3.2 if and only if the allocation $\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t\}$ is a solution to a relaxed problem based on definition 3.4 but with $\kappa_t = 0$, $m_t = 1$, and no constraint corresponding to $\lambda^L_{2,t}$. Conditional on the allocation, the policy is defined as follows. Set $\kappa_t \equiv 1 - \frac{D_t}{L_t}$ if $U^b_{C,t} > \beta_b \mathbb{E}_t(U^b_{C,t+1}R^L_{t+1})$; otherwise, choose any $\kappa_t \in \left[0, 1-\frac{D_t}{L_t}\right]$; choose any $m_t \in \left[\frac{\mathbb{E}_t(R^L_{t+1})L_t}{\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t}, 1\right]$; set the taxes to satisfy the regulated analogs* *of* (3.5) *and* (3.10)–(3.12)*, rebating them lump sum.*

According to lemma 3.4, without loss of generality, we can focus on the leverage limits that make the market leverage constraints binding: we can always set $\kappa_t \equiv 1 - \frac{D_t}{L_t}$ and $m_t \equiv \frac{\mathbb{E}_t(R^L_{t+1})L_t}{\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t}$. The construction of $\{\kappa_t\}$ ensures that the $\lambda^L_{2,t}$ constraint of definition 3.4 is satisfied. The relaxed problem of lemma 3.4 has a larger feasible set than the problem of definition 3.4; therefore, the Ramsey allocation with optimal leverage limits weakly dominates any FCEA associated with a given policy $\{\kappa_t, m_t\}$ unless the leverage constraints under $\{\kappa_t, m_t\}$ are slack with probability one. It is straightforward to show that the Ramsey allocation of lemma 3.4 has the risk-sharing and steady-state properties described in proposition 3.1.

Consider the second case when only $\{\kappa_t, m_t, \tau^N_t, T^e_t\}$ are available. Now we cannot dispense with the Euler equations (3.5), (3.11), and (3.12). To simplify the problem, we can use (3.5) and (3.11) to solve for the private Lagrange multipliers and then use the private complementary slackness conditions to rearrange the Euler equations (3.6) and (3.12), expressing them in terms of allocations and price functions. As in lemma 3.4, we can characterize the Ramsey problem entirely in terms of choosing allocations.

Lemma 3.5. *An allocation* $\{C^b_t, C^e_t, C^w_t, D_t, K_t, L_t, N_t\}$ *and policy* $\{\kappa_t, m_t, \tau^N_t, T^e_t\}$ *are part of a Ramsey equilibrium associated with the regulated FCE of proposition 3.2—after imposing* $\tau^D_t = \tau^L_t = \tau^K_t = 0$*—if and only if the allocation* $\{C^b_t, C^e_t, C^w_t, D_t, K_t, L_t, N_t\}$ *is a solution to a relaxed problem stated in appendix 3.A.7. Conditional on the allocation, the policy is defined as follows. If* $U^b_{C,t} > \beta_b R_t \mathbb{E}_t(U^b_{C,t+1})$*, set* $\kappa_t \equiv 1 - \frac{D_t}{L_t}$*; otherwise, choose any* $\kappa_t \in \left[0, 1 - \frac{D_t}{L_t}\right]$*. If* $U^e_{C,t} > \beta_e \mathbb{E}_t(U^e_{C,t+1} R^L_{t+1})$*, set* $m_t \equiv \frac{\mathbb{E}_t(R^L_{t+1})L_t}{\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t}$*; otherwise, choose any* $m_t \in \left[\frac{\mathbb{E}_t(R^L_{t+1})L_t}{\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t}, 1\right]$*. Set* $\tau^N_t \equiv \frac{A_t F_{N,t}}{W_t} - 1$ *and* $T^e_t \equiv \tau^N_t W_t N_t$*.*

As shown in appendix 3.A.7, the constraints corresponding to the banker's and entrepreneur's problems have a certain symmetry: in both cases, we have a leverage constraint, an asset Euler equation expressed in terms of allocations, and a constraint that requires the private Lagrange multiplier on the leverage constraint to be nonnegative; finally, we have a consolidated budget constraint. The symmetry is imperfect: while the banker's Euler equation implies that bank capital is equal to the expected discounted value of the stream of consumption, as in lemma 3.2, the entrepreneur's Euler equation does not produce a similar relationship, provided there is a nontrivial labor wedge addressed by the tax τ_t^N. Compared to lemma 3.4, the construction of the LTV ratio in lemma 3.5 must be consistent with the entrepreneur's private complementary slackness conditions. The following proposition summarizes some implications of the Ramsey problem in lemma 3.5.

Proposition 3.3. *An optimal allocation-policy pair in the Ramsey problem of lemma 3.5 generally has imperfect consumption insurance. There is approximately perfect risk sharing between bankers and entrepreneurs if the relaxed collateral constraint is slack in the steady state. If, moreover,* $U^w(C^w, N) = \ln(C^w) - v(N)$ *and the steady-state profits of capital good producers are zero, there is approximate insurance across all consumers. A steady state is generally unique.*

Similar to the FCEA and the Ramsey allocation of lemma 3.4, there is generally imperfect consumption insurance, but it is approximately perfect under the same conditions. A difference from the former allocations is that even if the relaxed collateral constraint is slack in the neighborhood of the steady state, risk sharing between bankers and entrepreneurs is only approximate. At the same time, the relaxed leverage constraints generate a larger feasible set of leverage ratios for bankers and entrepreneurs, potentially enhancing risk sharing relative to the FCE. In contrast to the FCEA and the Ramsey allocation of lemma 3.4, the allocation of lemma 3.5 generally has a unique steady state with $D > 0$, which is a consequence of respecting the intertemporal Euler equations and the arguments related to the proof of lemma 3.1.

3.3.2 Sticky-price economy

Consider now the general environment with nominal rigidities. Given the analysis in section 3.3.1, the exposition can be significantly simplified. Apart from exploring the implications of nominal rigidities for constrained efficiency, the main objective of this subsection is to characterize jointly optimal leverage limits and monetary policy.

Constrained efficient allocation

Compared to the flexible-price economy, we have two additional markets: wholesale goods and retail varieties. Retailers act as monopolists, internalizing the demand curve of the final good producers; hence, the only additional way to achieve an improvement over the CE allocation is to intervene in the competitive market for wholesale goods.

A social planner, making decisions on behalf of the entrepreneur, internalizes the determination of the wholesale good price P_t^w from the retailer's optimality conditions. Like the individual agents, the planner takes policies $\{\kappa_t, m_t, \Pi_t\}$ as given. If inflation is pinned down by a Taylor rule (3.23) in the CE, it must be so in the centralized allocation. In this case, (3.23) must not be part of the planner's implementability conditions: instead, it augments the planner's optimality conditions. Note that (3.18) yields a conditional solution for the retailer's optimal relative price $\widetilde{P}_t = \widetilde{P}(\Pi_t)$, which allows constructing the price dispersion sequence $\{\Delta_t\}$ recursively based on $\{\Pi_t, \widetilde{P}_t\}$ and an initial condition Δ_{-1}, using (3.19). Hence, effectively, the planner takes as given $\{\Delta_t, \Pi_t, \widetilde{P}_t\}$. Using (3.15) and (3.21), we can solve for the measure of the retailer's marginal benefit $\Omega_{2,t} = \frac{\epsilon}{\epsilon-1}\frac{\Omega_{1,t}}{\widetilde{P}_t}$ and final good output $Y_t = \frac{A_t}{\Delta_t}F(\xi_t K_{t-1}, N_t)$. Then (3.16) defines the retailer's demand curve for wholesale goods:

$$P_t^w = \frac{\Delta_t}{A_t F(\xi_t K_{t-1}, N_t)} \left\{ \Omega_{1,t} - \frac{\beta\theta \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})\Pi_{t+1}^{\epsilon}\Omega_{1,t+1}]}{U_C^w(C_t^w, N_t)} \right\}.$$

Relative to the FCEA problem, we have one additional control variable—a measure of the retailer's marginal cost $\Omega_{1,t}$. Similarly, there is one additional implementability condition—the recursive definition of the retailer's marginal benefit (3.17). Since $\Omega_{1,t}$ is an auxiliary variable, the set of potential wedges does not change. The constrained efficient allocation can then be defined as follows.

Definition 3.5. *A constrained efficient allocation (CEA) is a solution to*

$$\max_{\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t, \Omega_{1,t}\}} \mathbb{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \sum_{i\in\mathcal{I}} \omega_i U_t^i \right)$$

subject to the same constraints as in definition 3.4—with the consolidated budget constraint of bankers and entrepreneurs and the resource constraint modified as shown below—and (3.17). The modified and additional constraints are

$$\lambda_t^C: \quad 0 = \Delta_t \left\{ \Omega_{1,t} - \frac{\beta\theta \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})\Pi_{t+1}^{\epsilon}\Omega_{1,t+1}]}{U_C^w(C_t^w, N_t)} \right\} - Q(K_{t-1}, K_t, \xi_t)[K_t - (1-\delta)\xi_t K_{t-1}]$$
$$- W(C_t^w, N_t)N_t + D_t - R_{t-1}D_{t-1} - C_t^b - C_t^e,$$

$$\lambda_t^Y: \quad 0 = \frac{A_t}{\Delta_t} F(\xi_t K_{t-1}, N_t) - \sum_{i \in \mathcal{I}} C_t^i - I(K_{t-1}, K_t, \xi_t),$$

$$\lambda_t^\Omega: \quad 0 = \frac{\epsilon - 1}{\epsilon} \widetilde{P}_t \frac{A_t}{\Delta_t} F(\xi_t K_{t-1}, N_t) - \Omega_{1,t} + \frac{\beta\theta \mathbb{E}_t \left[U_C^w(C_{t+1}^w, N_{t+1}) \Pi_{t+1}^{\epsilon-1} \frac{\widetilde{P}_t}{\widetilde{P}_{t+1}} \Omega_{1,t+1} \right]}{U_C^w(C_t^w, N_t)},$$

where $\{\Delta_t, \Pi_t, \widetilde{P}_t\}$ *are taken as given.*

For the convenience of the reader using the electronic version of this paper, the modifications relative to the flexible-price economy are in color. The implementability conditions on the banker's side and the collateral constraint on the entrepreneur's side are identical to the flexible-price case. Consequently, the financial wedges corresponding to deposit supply and loan demand will be identical to those in the FCEA. The real wedges corresponding to the entrepreneur's demand for factors of production do change in the sticky-price economy. First, the planner internalizes how the optimal retail pricing affects the relative price of wholesale goods P_t^w, which directly affects the entrepreneur's revenue and the combined income of constrained consumers. Second, there is an output loss due to price dispersion $\Delta_t \geq 1$, limiting all consumers' consumption. (In the flexible-price economy, we have $P_t^w = \Delta_t = 1$ for all $t \geq 0$.) The following proposition formalizes this discussion and compares the CEA and FCEA.

Proposition 3.4. *The CE allocation is constrained inefficient, reflected in additional terms* Ψ_t^D, Ψ_t^L, Ψ_t^N, *and* Ψ_t^K, *as in proposition 3.1. The financial wedges* Ψ_t^D *and* Ψ_t^L *are identical to those in the flexible-price economy. There is perfect consumption insurance between bankers and entrepreneurs when the collateral constraint is slack. The CEA is locally indeterminate, and the optimal steady state features* $D = 0$*. The CEA can be decentralized in a regulated CE with linear taxes* $\{\tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K, T_t^d, T_t^e\}$ *identically to proposition 3.2, and the policy is Ramsey optimal.*

Unlike in the FCEA, we do not have a special case of approximate full risk sharing in the CEA because we would need to have zero steady-state profits of retailers. Since retailers act as monopolists, setting a time-varying markup over the marginal cost, their steady-state profits are positive. The other CEA properties are identical to those of the FCEA, except for the differences in real wedges.

The labor wedge is now

$$\Psi_t^N = \frac{[(\omega_e U^e_{C,t} + \lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda_t^C)(P_t^w \Delta_t)^{-1} - \omega_w U^w_{C,t} - \lambda_t^C] P_t^w A_t F_{N,t} - \lambda_t^C W_{N,t} N_t}{\omega_w U^w_{C,t} + \lambda_t^C} + \Gamma_t^N,$$

where Γ_t^N is identical to the one in the FCEA. The marginal product of labor is now priced at $P_t^w < 1$, and the marginal utility gap between workers and entrepreneurs is affected by nominal rigidities. The term $\lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda_t^C$ arises because the consolidated budget constraint does not directly contain the entrepreneur's output—it is now present in the constraint that reflects the retailer's marginal benefit. In the steady state, $\lambda^\Omega \widetilde{P} = \lambda^C$, and thus typically $\lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda_t^C < 0$. Moreover, there is a multiplicative factor $(P_t^w \Delta_t)^{-1}$, greater than unity under a reasonable calibration. Quantitatively, the second effect dominates and $(\omega_e U^e_{C,t} + \lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda_t^C)(P_t^w \Delta_t)^{-1} > \omega_e U^e_{C,t}$, implying that the difference between the marginal utility gap and the shadow value of wealth tends to decrease, as does the magnitude of the labor wedge.

The capital wedge now satisfies

$$\begin{aligned}
\omega_e \Psi_t^K = {} & (\beta - \beta_e)\mathbb{E}_t(\omega_e U^e_{C,t+1} R^K_{t+1}) Q_t + \beta \mathbb{E}_t \left\{ \lambda^Y_{t+1} \left[Q_{t+1} \Phi\left(\frac{I_{t+1}}{K_t}\right) - \frac{I_{t+1}}{K_t} \right] \right\} \\
& - \lambda_t^C Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] - \beta \mathbb{E}_t\{\lambda^C_{t+1} Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1} K_t]\} \\
& + \lambda_t^e m_t \mathbb{E}_t(Q_{1,t+1}\xi_{t+1}) K_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda^e_{t-1} m_{t-1} Q_{2,t} \xi_t K_{t-1} \\
& + \beta \mathbb{E}_t \left\{ \left[\omega_e U^e_{C,t+1}(1 - P^w_{t+1}\Delta_{t+1}) + \lambda^\Omega_{t+1} \frac{\epsilon - 1}{\epsilon} \widetilde{P}_{t+1} - \lambda^C_{t+1} \right] \frac{A_{t+1}}{\Delta_{t+1}} F_{K,t+1} \xi_{t+1} \right\}.
\end{aligned}$$

The effects of nominal rigidities parallel those for the labor wedge but are applied to the effective marginal product of capital. First, $P^w_{t+1} < 1$ affects the future return on capital. Second, $\lambda^\Omega_{t+1} \frac{\epsilon-1}{\epsilon}\widetilde{P}_{t+1} - \lambda^C_{t+1} < 0$. Third, $\omega_e U^e_{C,t+1}(1 - P^w_{t+1}\Delta_{t+1}) > 0$. Quantitatively, the magnitude of the wedge tends to increase.

Optimal monetary policy and leverage limits

Consider now how to set $\{\kappa_t, m_t, \Pi_t\}$ optimally. In section 3.3.1, we argued that the optimal leverage limits are not unique, but the corresponding allocation is uniquely determined as a solution to a relaxed planning problem. We can use a similar approach here, except inflation $\{\Pi_t\}$ will be a control variable. Although (3.19) allows us to construct $\{\Delta_t\}$ conditional on $\{\Pi_t\}$ and an initial condition, any Δ_t will be history dependent, complicating the optimization with respect to inflation. It is more tractable to add $\{\Delta_t\}$ to the set of controls and (3.19) to the implementability conditions. Furthermore, we will allow for an ELB. Our relaxed problems will thus feature the additional constraints[2]

$$\lambda_t^{\Delta}: \quad 0 = \theta \Pi_t^{\epsilon} \Delta_{t-1} + (1-\theta)(\widetilde{P}(\Pi_t))^{-\epsilon} - \Delta_t,$$
$$\lambda_t^{R}: \quad 0 \leq R_t \mathbb{E}_t(\Pi_{t+1}) - \underline{R}.$$

As in section 3.3.1, consider two alternative Ramsey problems. The first problem—case 1—allows the Ramsey planner to set $\{\kappa_t, m_t, \Pi_t, \tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K, T_t^d, T_t^e\}$ optimally. By proposition 3.4, conditional on $\{\kappa_t, m_t, \Pi_t\}$, we get a CEA if the ELB constraint is slack with probability one. Therefore, the case 1 Ramsey allocation approximately corresponds to the best CEA. The constraint set of the relaxed problem is formed by taking the constraints from lemma 3.4, modifying the λ_t^C and λ_t^Y constraints and adding the λ_t^{Ω} constraint as in definition 3.5, and adding the λ_t^{Δ} and λ_t^R constraints above. The ELB affects the worker's consumption through the real interest rate R_t, but there is no direct effect on bankers and entrepreneurs; therefore, the case 1 Ramsey allocation has the same risk-sharing and steady-state properties as the CEA. Let us postpone the characterization of optimal monetary policy until after we have described our alternative problem.

The second problem—case 2—has only $\{\kappa_t, m_t, \Pi_t, \tau_t^N, T_t^e\}$ as policy instruments. The case 2 Ramsey allocation is thus constrained inefficient. The constraint set in the relaxed problem is formed by taking the constraints from lemma 3.5, modifying or adding the λ_t^C, λ_t^Y, λ_t^{Ω}, λ_t^{Δ}, and

[2]Note that the ELB constraint highlights an aggregate demand externality that was absent in the comparison of the CE and CEA allocations. This externality is similar to that emphasized in Farhi and Werning (2016) and Korinek and Simsek (2016). Unlike the Ramsey planner, the individual agents do not internalize that their consumption-saving choice affects the strictness of the ELB constraint through the worker's Euler equation (3.2). A thorough analysis of this externality in our environment is left for future research.

λ_t^R constraints identically to case 1, and modifying the λ_t^K constraint as follows:

$$\begin{aligned}\lambda_t^K: \quad 0 = \beta_e \mathbb{E}_t \Bigg\{ & U_C^e(C_{t+1}^e) \Bigg[\alpha \Delta_{t+1} \left\{ \Omega_{1,t+1} - \frac{\beta\theta \mathbb{E}_{t+1}[U_C^w(C_{t+2}^w, N_{t+2}) \Pi_{t+2}^{\epsilon} \Omega_{1,t+2}]}{U_C^w(C_{t+1}^w, N_{t+1})} \right\} \\ & + Q(K_t, K_{t+1}, \xi_{t+1})(1-\delta)\xi_{t+1} K_t - C_{t+1}^b - L_{t+1} + D_{t+1} - R_t D_t \Bigg] \Bigg\} \\ & - U_C^e(C_t^e)(Q(K_{t-1}, K_t, \xi_t) K_t - L_t),\end{aligned}$$

where $\alpha \equiv \frac{F_K(\xi_t K_{t-1}, N_t)\xi_t K_{t-1}}{F(\xi_t K_{t-1}, N_t)}$ is the capital share. (It is constant because F is Cobb—Douglas.) The modified term corresponds to $P_{t+1}^w A_{t+1} F_{K,t+1} \xi_{t+1} K_t$, reflecting the determination of P_t^w from the retailer's problem. As in proposition 3.3, the case 2 Ramsey allocation has partial risk sharing between bankers and entrepreneurs in the neighborhood of the steady state if the relaxed collateral constraint is slack, and there typically exists a unique steady state with $D > 0$.

Note that inflation $\{\Pi_t\}$ affects the planner's constraints in both problems identically with one exception: in case 2, future inflation affects the future return on capital in the λ_t^K constraint through P_{t+1}^w. However, we can define an auxiliary variable that captures the combined shadow value of the effect through P_{t+1}^w with the effect through P_t^w, where the latter is common to both problems. Conditional on this auxiliary variable, the optimal monetary policy has identical long-run and short-run characteristics. The following proposition summarizes and formalizes our discussion.

Proposition 3.5. *The case 1 and 2 Ramsey allocations have the risk-sharing and steady-state properties of propositions 3.4 and 3.3, respectively, except for the special case of approximate full insurance.*

In both cases, the optimal monetary policy is characterized as follows. The long-run gross inflation rate in the absence of uncertainty is uniquely determined as

$$\Pi = \begin{cases} 1 & \text{if } \underline{R} \le \frac{1}{\beta} \\ \beta \underline{R} & \text{if } \underline{R} > \frac{1}{\beta} \end{cases}.$$

The short-run inflation behavior is represented by the Euler equation

$$0 = \lambda_t^{\Omega}\widetilde{P}'(\Pi_t)\left[\frac{\epsilon-1}{\epsilon}Y_t + \frac{\beta\theta\mathbb{E}_t\left(U^w_{C,t+1}\Pi_{t+1}^{\epsilon-1}\frac{\Omega_{1,t+1}}{\widetilde{P}_{t+1}}\right)}{U^w_{C,t}}\right] + \lambda_t^{\Delta}\epsilon\left[\theta\Pi_t^{\epsilon-1}\Delta_{t-1} - (1-\theta)\frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t^{\epsilon+1}}\right]$$
$$- \mathbf{1}_{\mathbb{N}}(t)\theta\Pi_t^{\epsilon-1}\Omega_{1,t}\frac{U^w_{C,t}}{U^w_{C,t-1}}\left[\widetilde{\lambda}^C_{t-1}\Delta_{t-1}\epsilon - \lambda^{\Omega}_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\left(\frac{\epsilon-1}{\Pi_t} - \frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t}\right)\right] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda^R_{t-1}R_{t-1},$$

where $\widetilde{\lambda}^C_t \equiv \lambda^C_t$ *in case 1, and* $\widetilde{\lambda}^C_t \equiv \lambda^C_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda^K_{t-1}\beta_e U^e_{C,t}\alpha$ *in case 2.*

Since an ELB typically satisfies $\underline{R} \leq 1 < \frac{1}{\beta}$, by proposition 3.5, the long-run price stability is optimal independently of whether the planner can address the intertemporal distortions. One can demonstrate that steady-state price stability is optimal even if inflation is the only policy instrument. Stable prices eliminate output losses due to price dispersion $\Delta \geq 1$, and the optimality of price stability is consistent with the normative analyses of the basic New Keynesian economies (Schmitt-Grohé and Uribe, 2010; Woodford, 2010). Moreover, Cúrdia and Woodford (2016) obtained a similar result in a model with a credit spread friction that may be viewed as an approximation of the endogenous credit spread that arises in our model due to the leverage constraint. Coibion et al. (2012) found that a slightly positive steady-state inflation could be optimal in the context of Taylor rules with an ELB in the basic New Keynesian model, but they showed that inflation is close to zero quantitatively under the optimal policy with commitment. In our economy, depending on $\underline{R}$, positive inflation might arise in a stochastic steady state due to the planner's precautionary motive to insure against the binding ELB.

It is worth emphasizing that the steady-state inflation rate under the optimal monetary policy is uniquely determined, unlike in the case of an ad hoc Taylor rule with an ELB, where multiple equilibria are an inherent property of the functional form (Benhabib et al., 2001). The Ramsey planner is not subject to functional form restrictions and chooses a state-contingent plan subject to an ELB inequality constraint. At the same time, Armenter (2018) has shown that multiple Markov equilibria might arise under an optimal discretionary policy with an ELB.

In the short run, stabilizing prices state-by-state is not generally optimal, and the inflation dynamics are characterized by an Euler equation that balances different forces. Note that if $\Pi_t \approx 1$, then $\widetilde{P}_t = \frac{1}{\Pi_t}\left(\frac{\Pi_t^{1-\epsilon}-\theta}{1-\theta}\right)^{\frac{1}{1-\epsilon}} \approx 1$ and $\widetilde{P}'(\Pi_t) = \frac{\widetilde{P}_t}{\Pi_t}\left(\frac{\widetilde{P}_t^{\epsilon-1}}{1-\theta} - 1\right) > 0$. First, a greater inflation rate positively affects welfare by raising the retailer's marginal benefit through the higher optimal relative price, reflected in the term $\lambda_t^{\Omega}\widetilde{P}'(\Pi_t) > 0$. Second, inflation affects price dispersion: positively by expanding the price dispersion inherited from the previous period, $\theta\Pi_t^{\epsilon-1}\Delta_{t-1} > 0$, and negatively by raising the retailer's optimal price, reflected in $-(1-\theta)\frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t^{\epsilon+1}} < 0$. In the steady state, the net effect is zero. Third, by raising inflation at t, the planner affects the expectation of retailers at $t-1$ regarding the marginal cost at t, having a negative effect $-\theta\Pi_t^{\epsilon-1}\Omega_{1,t}\frac{U^w_{C,t}}{U^w_{C,t-1}}\widetilde{\lambda}^C_{t-1}\Delta_{t-1}\epsilon < 0$. Fourth, higher inflation at t also affects the retailer's marginal benefit expected at $t-1$, which has a positive effect $\theta\Pi_t^{\epsilon-1}\Omega_{1,t}\frac{U^w_{C,t}}{U^w_{C,t-1}}\lambda^{\Omega}_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\left(\frac{\epsilon-1}{\Pi_t} - \frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t}\right) > 0$ since $\frac{\epsilon-1}{\Pi_t} - \frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t} = \frac{1}{\Pi_t}\left(\epsilon - \frac{\widetilde{P}_t^{\epsilon-1}}{1-\theta}\right) > 0$, provided that the elasticity of substitution ϵ is sufficiently greater than the price duration $\frac{1}{1-\theta}$. Fifth, higher inflation at t raises the expected inflation at $t-1$ and relaxes the ELB at $t-1$.

None of the inflation Euler equation components are directly related to financial constraints. In our economy, the Tinbergen separation principle applies: the implications of the Ramsey allocations with optimal leverage limits are similar to those in the flexible-price environment, and the approximate price stability is optimal as in the basic New Keynesian model. Collard et al. (2017) also found support for such independence of policy goals exploring jointly optimal bank capital requirements and monetary policy in the absence of collateral constraints and consumer type heterogeneity but with nominal contracts. The latter indicates that our choice to proceed with real contracts is mostly without loss of generality. Indeed, one can show that the optimal monetary policy would still have steady-state price stability and similar short-run dynamics.

3.4 Quantitative results

This section describes the model calibration and computation, quantifies the welfare losses due to the constrained inefficiency and welfare benefits of optimal policies, compares the extent of consumption insurance observed in the different types of decentralized and centralized allocations, and studies the economic dynamics around financial crises and binding ELB events.

3.4.1 Calibration

To simplify the interpretation of quantitative results, we will assume that all consumers have logarithmic preferences over period consumption. With $U^b(C^b) = \ln\big(C^b\big)$ and $U^e(C^e) = \ln(C^e)$, lemma 3.2 implies that the banker's and entrepreneur's net assets are proportional to consumption. The worker's preferences are separable in consumption and leisure, taking the form $U^w(C^w, N) = \ln(C^w) - \chi\frac{N^{1+\phi}}{1+\phi}$, where $\chi > 0$ is the labor disutility scale, and $\phi \geq 0$ is the inverse of the Frisch elasticity of labor supply. Thus, the worker's preferences are consistent with the special case of full insurance of propositions 3.1 and 3.3. As for technology, the entrepreneur's output is produced according to $F(\xi K, N) = (\xi K)^\alpha N^{1-\alpha}$ with $\alpha \in (0,1)$, and capital goods are built using $\Phi(x) = \zeta + \kappa_1 x^\psi$ with $\zeta \in \mathbb{R}$, $\kappa_1 > 0$, and $\psi \in (0,1]$. The logarithms of the exogenous stochastic processes $\{A_t\}$ and $\{\xi_t\}$ are independent Gaussian AR(1) with autocorrelations (ρ_a, ρ_ξ) and shock standard deviations (σ_a, σ_ξ), respectively, implying a steady-state normalization $A = \xi = 1$.

Table 3.1 reports the model parameter values. To calibrate the structural parameters, we need

Table 3.1. Parameter values

Parameter	Value	Target
Baseline policy		
$\bar{\kappa}$	0.105	Basel III total capital requirement + conservation buffer
$\bar{m}$	0.7	FDIC LTV limits for raw land (65%) and land development (75%)
$\bar{\Pi}$	1.005	annual inflation = 2%
$\underline{R}$	1	zero lower bound
Preferences and technology		
α	0.404	average nonfarm labor share ≈ 59.6%
β	0.995	annualized real interest rate = 2%
β_b	0.972	annual NAICS 52 establishment exit rate ≈ 9.1%
β_e	0.974	annual NAICS 31–33 establishment exit rate ≈ 8.2%
δ	0.02	annual depreciation rate ≈ 7.6%
ϵ	9.093	average retail markup = 1.125
ζ	-0.002	$\frac{I}{K} = \delta$ and $Q = 1$
θ	0.75	average price duration = 4 quarters
κ_1	0.781	$\frac{I}{K} = \delta$ and $Q = 1$
ϕ	0.625	microfounded aggregate Frisch elasticity = 1.6
χ	0.94	$N = 1$ in the FCE
ψ	0.75	panel data evidence
Exogenous stochastic processes		
ρ_a	0.918	
ρ_ξ	0.935	First-step MSM estimation based on the FCE, targeting
σ_a	0.005	$\text{corr}(\widehat{Y}_t, \widehat{Y}_{t-1})$, $\text{sd}(\widehat{I}_t)$, $\text{sd}(\widehat{Y}_t)$, and $\text{corr}(\widehat{I}_t, \widehat{Y}_t)$.
σ_ξ	0.003	
Taylor rule		
ρ_R	0.897	Second-step MSM estimation based on the CE, targeting
η_π	3.366	$\text{corr}(\widehat{Y}_t, \widehat{Y}_{t-1})$, $\text{sd}(\widehat{Y}_t)$, $\text{corr}(\widehat{\Pi}_t, \widehat{\Pi}_{t-1})$, $\text{sd}(\widehat{\Pi}_t)$, $\text{corr}(\widehat{\Pi}_t, \widehat{Y}_t)$, and
η_y	3.104	$\Pr\left(R_t^N = \underline{R}\right)$.

Note. $\widehat{X}_t$ denotes the cyclical component of $\ln(X_t)$ extracted using the HP filter with $\lambda = 1600$.

to determine the baseline policies taken as given by the private agents. The leverage limits are set to constant values $\kappa_t = \bar{\kappa}$ and $m_t = \bar{m}$ for all $t \geq 0$. The capital requirement $\bar{\kappa}$ corresponds to the Basel III minimum total capital requirement that includes the conservation buffer. The LTV ratio $\bar{m}$ is set to the average of the Federal Deposit Insurance Corporation's (FDIC) recommended maximum LTV limits for raw land and land development—a proxy for commercial loans. The inflation target $\bar{\Pi}$ corresponds to the annual target of 2%, and the effective lower bound $\underline{R}$ is the zero lower bound (ZLB).

The structural parameters that affect the steady state are either based on micro evidence or target various long-run moments in the US quarterly—or annual if not available—data for 1990–2019 or the largest available subset. The remaining parameters are estimated using the method of simulated moments (MSM) of McFadden (1989).

From the preference parameters, the discount factor β corresponds to the annualized real interest rate of 2%. The effective discount factors of bankers and entrepreneurs are based on the average annual establishment exit rates in finance and insurance (NAICS 52) and manufacturing (NAICS 31–33), respectively, using Business Dynamics Statistics data. The inverse of the Frisch elasticity of the worker's labor supply ϕ targets the average of the microfounded estimates of the aggregate Frisch elasticity for males (Erosa et al., 2016) and females (Attanasio et al., 2018). The labor disutility scale χ is set to normalize $N = 1$ in the FCE.

Turning to the technology parameters, the capital share α targets the average labor share in the nonfarm business sector based on US Bureau of Labor Statistics data. The depreciation rate δ is based on the average depreciation rate of the current-cost net stock of private fixed assets and consumer durables in Bureau of Economic Analysis data. The capital good technology elasticity ψ is based on the panel data evidence (Gertler et al., 2020a). Conditional on ψ, there is a one-to-one correspondence between the location and scale parameters (ζ, κ_1) and a steady-state pair $(\frac{I}{K}, Q)$. Using (3.14) and (3.20), we get $\kappa_1 = \frac{1}{\psi Q}\left(\frac{I}{K}\right)^{1-\psi}$ and $\zeta = 1 - (1-\delta)\xi - \kappa_1 \left(\frac{I}{K}\right)^{\psi}$. We have already normalized $\xi = 1$. By targeting $\frac{I}{K} = \delta$ and normalizing $Q = 1$, the steady-state profits of capital good producers are zero, and thus the calibration is consistent with the special case of perfect insurance in propositions 3.1 and 3.3. The Calvo price stickiness parameter targets the average price duration $\frac{1}{1-\theta}$, and the elasticity of substitution between retail varieties ϵ is mapped to the markup $\frac{1}{P^w}$ in retail, solving a steady-state equation $P^w = \frac{\epsilon-1}{\epsilon}\frac{1-\beta\theta\Pi^{\epsilon}}{1-\beta\theta\Pi^{\epsilon-1}}\left(\frac{1-\theta}{1-\theta\Pi^{\epsilon-1}}\right)^{\frac{1}{\epsilon-1}}$ that follows from combining (3.15)–(3.18). The targets are consistent with the micro evidence as in Galí (2015).

To account for multiple occasionally binding constraints in simulations of both competitive equilibria and centralized allocations, I use the piecewise linear perturbation approach of Guerrieri and Iacoviello (2015), extending it to handle an arbitrary number of regime-switching constraints.[3] In some exercises, I use a standard second-order perturbation, taking advantage of the possibility to approximate theoretical moments when the system stays close to the steady state. To get a locally unique approximation for the FCEA, CEA, the Ramsey allocation of lemma 3.4, and the case 1 Ramsey allocation of proposition 3.5, I fix the quantity of deposits at the optimal steady-state value of zero. The welfare benefits of the corresponding allocations are thus generally underestimated.

3.4.2 Welfare comparison

Starting from this subsection, we will use additional notation, referring to the Ramsey allocations of lemmas 3.4 and 3.5 as FCEA OLL and OLL, respectively, where "OLL" means "optimal leverage limits." We will call the case 1 and case 2 Ramsey allocations of proposition 3.5 CEA OLLMP and OLLMP, respectively, where "OLLMP" corresponds to "optimal leverage limits and monetary policy."

Table 3.2 reports the welfare ranking of alternative environments conditional on a Pareto vector $\omega = (\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$. There is a unit measure of all types of consumers in the model, so one might want to choose comparable Pareto weights for all agents. On the other hand, the real-world population of workers is significantly greater than that of bankers or entrepreneurs, which suggests a worker-biased weighting. The chosen Pareto vector reflects these two margins: it is worker biased, but the banker's and entrepreneur's weights are still sizable. As a welfare benchmark, we will consider the first-best allocation. The differences relative to the first best are represented in consumption equivalents. Let $\mathcal{W}^i_{FB}$ and $\mathcal{W}^i$ denote the expected welfare of type i consumers at the first best and an alternative set-up, respectively. We can solve for λ^i that satisfies $\mathcal{W}^i = \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t U^i(\lambda^i C^i_{FB,t})\right]$, where $\{C^i_{FB,t}\}$ is the first-best consumption plan. By construction, $\lambda^i \in (0,1)$ is the proportion of the first-best consumption plan—applied in all contingencies—that

[3]The extension is available at `https://github.com/azaretski/occbin-n`.

Table 3.2. Welfare in consumption equivalents, % of first best

	bankers	entrepreneurs	workers	social welfare
First best	100	100	100	100
FCE	28.8	109.7	95.9	86.2
FCEA	94.2	94.2	100.1	98.9
FCEA OLL	94.2	94.2	100.1	98.9
OLL	71.8	88.6	98.1	94.1
CE	21.0	79.8	91.7	78.1
CEA	78.6	78.6	98.4	94.0
CEA OLLMP	79.0	79.0	98.9	94.5
OLLMP	77.4	60.3	97.4	90.7

Note. Second-order accurate theoretical moments in the neighborhood of the steady state, conditional on a Pareto vector $(\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$. The OLLMP row is based on $\beta_b \approx 0.989$—the nearest neighbor where the Blanchard—Kahn conditions for local uniqueness hold.

yields the same welfare for agent i as the alternative consumption allocation. With logarithmic preferences, we have a closed-form solution $\lambda^i = \exp\left[(1-\beta)(\mathcal{W}^i - \mathcal{W}^i_{FB})\right]$. Similarly, we can get a social welfare ranking by computing $\lambda = \exp[(1-\beta)(\mathcal{W} - \mathcal{W}_{FB})]$, where $\mathcal{W}$ and $\mathcal{W}_{FB}$ denote the expected social welfare of the alternative and first-best allocations, respectively, and $\lambda \in (0, 1)$ is the proportion of the first-best consumption plan—applied in all contingencies and for all consumers—that yields the same value of social welfare as the alternative consumption allocation.

Compared to the first best, constrained bankers and entrepreneurs are more worse off than workers in most environments, reflecting the worker-biased Pareto vector. Due to nominal rigidities, the sticky-price environments tend to be welfare-dominated by their flexible-price counterparts. The welfare gains from constrained efficiency—FCEA over FCE and CEA over CE—are rather significant. The FCEA and FCEA OLL allocations have identical welfare implications because both financial constraints are locally slack in the FCEA: bank leverage is suboptimal, and the optimal entrepreneur's LTV ratio is lower than the calibrated limit. Relaxing the leverage constraints might impact precautionary savings, but we cannot account for this effect using our computation method. The OLL allocation is Pareto dominated by the FCEA, since bankers have positive leverage, and the relaxed collateral constraint is binding. At the same time, the OLL allocation constitutes a significant social welfare gain over the FCE.

Although leverage constraints are locally slack in the CEA, with nominal rigidities, there is a distinction between the CEA and CEA OLLMP allocations since the latter has optimal monetary policy, compared to an ad hoc Taylor rule in the CEA. Similar to the flexible-price case, the OLLMP allocation is between the CE and CEA in social welfare terms, although optimal monetary policy reduces the relative distance to the CEA.

3.4.3 Risk sharing

Table 3.3 reports the correlations between the HP-filtered logged marginal utilities of consumption across consumers in the alternative allocations. The first-best allocation is the only one that

Table 3.3. Consumption risk sharing

	$\text{corr}(\widehat{U}^b_{C,t}, \widehat{U}^e_{C,t})$	$\text{corr}(\widehat{U}^b_{C,t}, \widehat{U}^w_{C,t})$	$\text{corr}(\widehat{U}^e_{C,t}, \widehat{U}^w_{C,t})$
First best	1	1	1
FCE	0.07	0.57	-0.51
FCEA	1	1.0	1.0
FCEA OLL	1	1.0	1.0
OLL	0.92	-0.55	-0.71
CE	-0.1	0.59	-0.57
CEA	1	0.99	0.99
CEA OLLMP	1	0.99	0.99
OLLMP	0.8	-0.85	-0.97

Note. Second-order accurate theoretical correlations in the neighborhood of the steady state, conditional on a Pareto vector $(\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$. The decimal point in 1.0 indicates that the correlation is not exactly 1. $\widehat{X}_t$ denotes the cyclical component of $\ln(X_t)$ extracted using the HP filter with $\lambda = 1600$. The OLLMP row is based on $\beta_b \approx 0.989$—the nearest neighbor where the Blanchard—Kahn conditions for local uniqueness hold.

has perfect consumption insurance. Consistent with propositions 3.1, 3.4, and 3.5, the FCEA, FCEA OLL, CEA, and CEA OLLMP allocations have perfect risk sharing between bankers and entrepreneurs since the collateral constraint is locally slack. The latter is not the case in the OLL and OLLMP allocations. Financial constraints are locally binding in the FCE and CE, and consumption insurance is largely imperfect. Since our calibration is consistent with the special case of proposition 3.1, the FCEA and FCEA OLL allocations have nearly perfect insurance across all consumers: the correlation between the worker's marginal utility and the marginal utility of constrained consumers is near unity. With nominal rigidities, the correlation is only slightly lower.

Although perfect consumption risk sharing is a feature of the first best, stronger risk sharing between consumers is not a prerequisite for higher welfare, as tables 3.2 and 3.3 demonstrate. For example, insurance is much stronger in the CEA than in the OLL allocation, but the latter has no nominal rigidities and has greater social welfare. Conditional on a flexible-price or a sticky-price environment, stronger risk sharing is associated with higher welfare.

3.4.4 Wedges and overborrowing

Table 3.4 quantifies the financial and real wedges. Sections 3.3.1 and 3.3.2 show that each wedge can be decomposed into several components. By definition, the means of components add up to

Table 3.4. Wedges

	FCEA		CEA	
	mean, %	variance, %	mean, %	variance, %
Ψ_t^D, % of $U_{C,t}^b$	2.4	0.1	2.4	0.1
uncertain survival: bankers	99.9	73.6	99.9	65.6
consumer type heterogeneity	0.1	19.6	0.1	29.1
Ψ_t^L, % of $U_{C,t}^e$	-0.3	0.0	-0.3	0.0
survival rate differences: $\beta_b \neq \beta_e$	100	100	100	100
uncertain survival: bankers	0	0	0	0
Ψ_t^N, % of W_t	-8.7	8.6	-17.8	4.4
consumer type heterogeneity	64.0	42.3	24.0	61.3
W-externality	36.0	12.2	61.6	3.1
nominal rigidities	0	0	14.5	0.6
Ψ_t^K, % of $U_{C,t}^e Q_t$	2.1	0.1	2.3	0.1
uncertain survival: entrepreneurs	99.7	67.1	91.1	60.1
Φ-externality	0.6	3.4	0.4	1.7
Q-externality	-0.3	0.9	-1.1	1.4
nominal rigidities	0	0	9.5	1.0

Note. Second-order accurate theoretical moments in the neighborhood of the steady state, conditional on a Pareto vector $(\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$. Components of wedges are in % of the mean or variance of the corresponding wedge. “Consumer type heterogeneity” reflects marginal utility gaps and terms that arise because $\lambda_t^Y \neq \omega_w U_{C,t}^w$. The W-, Φ-, and Q-externalities are the externalities through the wage, the capital good production technology directly, and the capital good price, respectively.

100%. Since the components are generally correlated, the sum of the variances need not be equal to the variance of the corresponding wedge.

The expected value of the deposit wedge Ψ_t^D is almost entirely based on the survival externality and entirely in the steady state. The survival externality is dominant in terms of the variance, but the consumer type heterogeneity component also has a nonnegligible variation. The loan wedge Ψ_t^L is entirely determined by the difference in the survival rates of bankers and entrepreneurs because the collateral constraint is locally slack in both the FCEA and the CEA. By proposition

3.4, nominal rigidities do not affect the expressions of financial wedges, which results in an identical decomposition of means.

In the FCEA, about two-thirds of the expected value of the labor wedge stems from the direct implications of consumer type heterogeneity, and the rest is explained by the wage externality—an indirect consequence of consumer type heterogeneity. In the CEA, the order is reversed, and nominal rigidities play an additional role. Consumer type heterogeneity explains a significant part of the variance in both environments, especially in the CEA. Although the additive term arising from nominal rigidities in the CEA contributes to only 14.5% of the expected value, nominal rigidities also affect the consumer type heterogeneity component multiplicatively, so their impact cannot be easily decoupled. The absolute value of the consumer type heterogeneity component is significantly less in the CEA than in the FCEA, as predicted in section 3.3.2, although the magnitude of the wedge is greater in the CEA due to the other two components.

The uncertain survival of entrepreneurs explains a major part of the expected value and variance of the capital wedge. Nominal rigidities constitute the second strongest direct source of the wedge in the CEA, and they also have an indirect multiplicative effect through the price of wholesale goods that affects the return on capital and the uncertain survival component. The role of the asset-price externality is modest. Since the collateral constraint is locally slack in the FCEA and CEA, the pecuniary externality only has precautionary savings effects. Hence, the asset-price externality works exclusively through the consolidated budget constraint of bankers and entrepreneurs. Although our calibration ensures that the steady-state profits of capital good producers are zero, the expected value is slightly positive, and so is the first-order externality that works through the capital good production technology Φ directly.

As a result of constrained inefficiency, our economy has inefficient borrowing in the financial markets. There are two types of borrowing: banks' borrowing from workers and entrepreneurs' borrowing from banks. Propositions 3.1 and 3.4 show that the constrained efficient bank leverage is zero in the optimal steady state, implying extreme overborrowing by banks in the competitive equilibria. The intertemporal inefficiency of the entrepreneur's borrowing is reflected in the wedge Ψ_t^L. Figure 3.4 shows that the wedge is negative since $\beta_b < \beta_e$. Although the competitive demand for bank loans is inefficiently low, overborrowing by the banking sector results in an inefficiently large supply, making the competitive quantity of bank loans inefficiently large if the Pareto vector is sufficiently worker biased.

Figure 3.1 displays the histograms of bank loans in the FCE and CE compared to the FCEA and CEA, respectively. By construction, in the FCEA and CEA, there is no variation in the quantity of

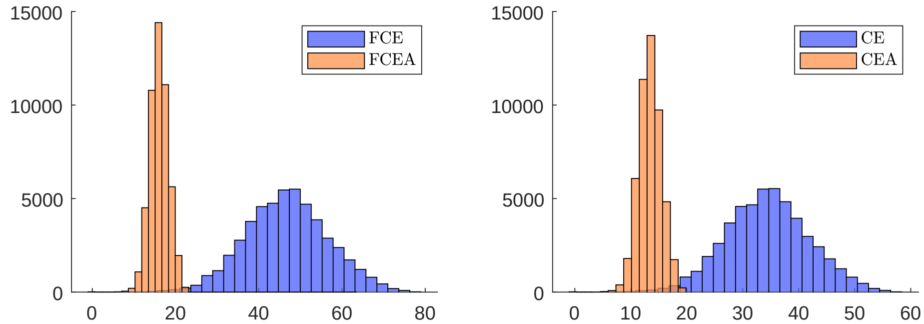

Figure 3.1. Histograms of bank loans. 50,000-period simulation conditional on a Pareto vector $(\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$.

deposits fixed at the optimal steady-state value of zero; consequently, the variance of bank loans is smaller in the FCEA and CEA. The expected values are considerably smaller, reflecting overlending in the FCE and CE. Nominal rigidities tend to decrease economic activity, shifting the distributions of bank loans to the left.

3.4.5 Financial crises

This subsection explores the economic dynamics around financial crises. The focus is on the flexible-price economy to isolate the effect of the occasionally binding collateral constraint. Financial crises are defined similarly as in Mendoza (2010). To be qualified as a financial crisis that starts at t, two conditions must be true: first, the collateral constraint is slack at $[t-4, t-1]$; second, the collateral constraint is binding at $[t, t+4]$. Such an event is observed in the FCE with a frequency of 3.2 crises per century, consistent with the data.

Figure 3.2 illustrates the dynamics around financial crises in alternative environments based on a 50,000-period simulation conditional on an identical sequence of exogenous shocks drawn randomly from the corresponding distributions. The financial crisis events are identified in the FCE simulation, and the identified dates are used to extract the corresponding paths in the FCEA and OLL simulations. The dynamics around identified crises are averaged, and each crisis is normalized to start at $t = 1$, lasting at least until $t = 5$.

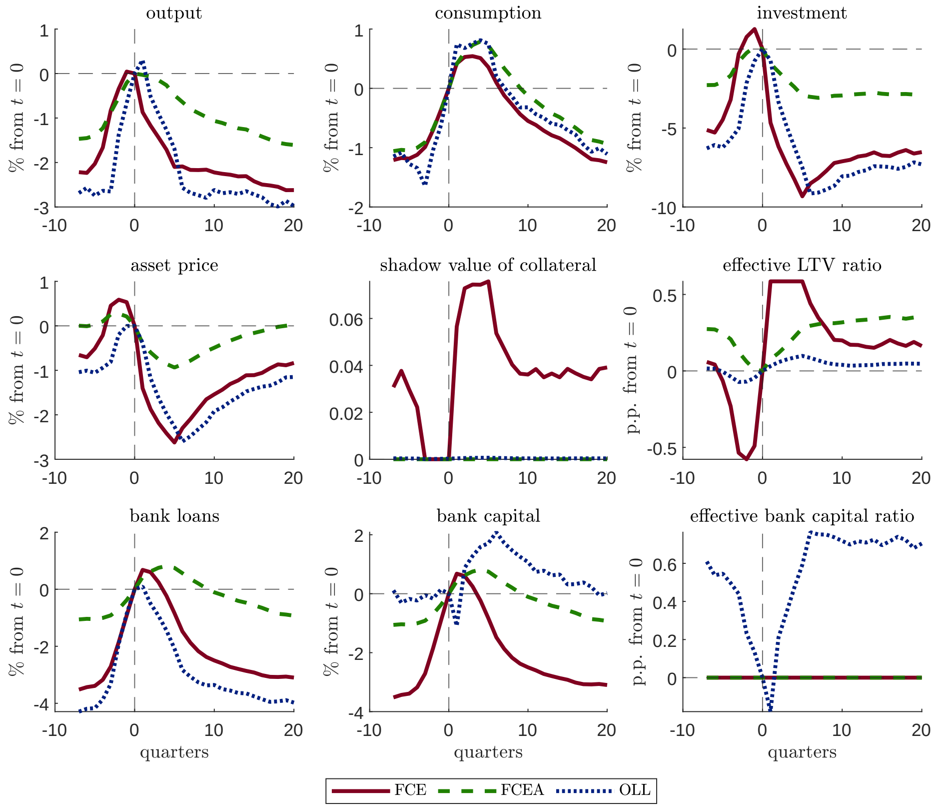

Figure 3.2. Financial crises. Each line is based on an average of 399 crisis episodes over a 50,000-period simulation conditional on a Pareto vector $(\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$. A crisis starts at $t = 1$ and lasts at least five quarters. The shadow value of collateral is in levels. The effective LTV ratio is the ratio of the expected loan repayment to the value of collateral. The effective bank capital ratio is the ratio of bank capital to bank loans. "p.p." is "percentage points."

Ahead of a typical crisis in the FCE, the economy is booming: output, consumption, and investment are increasing, so is bank lending and—for most of the period—the collateral asset price. By construction, the collateral constraint is slack during the year before the crisis, so the shadow value of collateral is zero during that time. The asset price starts to fall a few quarters ahead of the crisis, leading to a decrease in the value of collateral and triggering a switch of the collateral constraint from the slack to the binding regime. Output and bank lending start to drop, while investment starts to fall earlier, responding to a fall in the asset price. As the collateral constraint returns to a slack regime, which occurs at different times in each crisis, the asset price and investment start to recover, and the fall in output and bank lending slows down, plateauing gradually. There is a one percentage point increase in the entrepreneur's LTV ratio just before the crisis until it hits the LTV limit $\bar{m}$ during the crisis. The bank leverage constraint remains binding in the FCE, so the bank capital ratio is constant at $\bar{\kappa}$.

The FCEA and FCEA OLL allocations have identical dynamics since both leverage constraints remain slack in the simulation; therefore, the figure shows only the dynamics in the FCEA. What happens to be a financial crisis in the FCE is reminiscent of a cyclical slowdown in the FCEA, which is a consequence of the fact that the optimal entrepreneur's leverage is smaller and the collateral constraint is slack. The fluctuations in the LTV ratio are small. Since bank leverage is constant at zero in the FCEA, the capital ratio is constant at one.

In the OLL allocation, the dynamics are more similar to the FCE than the FCEA. A fall in real quantities and the asset price tends to be initially smaller, but the eventual decrease is similar to that in the FCE. The amplitude of the relative changes in investment and the asset price is slightly smaller than that in the FCE, while the opposite is true for output and bank lending. The variation in the entrepreneur's optimal LTV ratio is negligible. The Ramsey planner keeps bank capital at a stable level ahead of a crisis and provides additional capital during the crisis. Combined with the credit dynamics, the optimal bank capital ratio decreases ahead of the crisis and increases during the crisis, although the changes are in the range of one percentage point.

3.4.6 Zero lower bound

This subsection considers a different type of crisis that occurs when the ZLB binds. A ZLB crisis that starts at t is an event that satisfies two conditions: the ZLB constraint is slack at $[t-4, t-1]$ and is binding at $[t, t+2]$. An event defined this way is observed in the CE with a frequency of 2.5 crises per century. Figure 3.3 illustrates the dynamics around a typical ZLB crisis and is constructed similar to figure 3.2.

Unlike financial crisis episodes that follow a boom-bust pattern, ZLB events occur when the economy is either already in a recession or a state of stagnation. The latter is reflected in both the real sector variables—output, consumption, and investment—and the financial sector variables, such as the collateral asset price and bank loans. Ahead of a ZLB crisis, the central bank consistently fails to match a 2% annualized inflation target. When the ZLB binds at $t = 1$, there is a further decrease in inflation, followed by a spike reflecting an increase in the retailer's marginal cost due to the drop in the entrepreneur's supply of wholesale goods. In our economy, a ZLB crisis results from a persistent decrease in the TFP and capital quality processes, leading to a sharp drop in

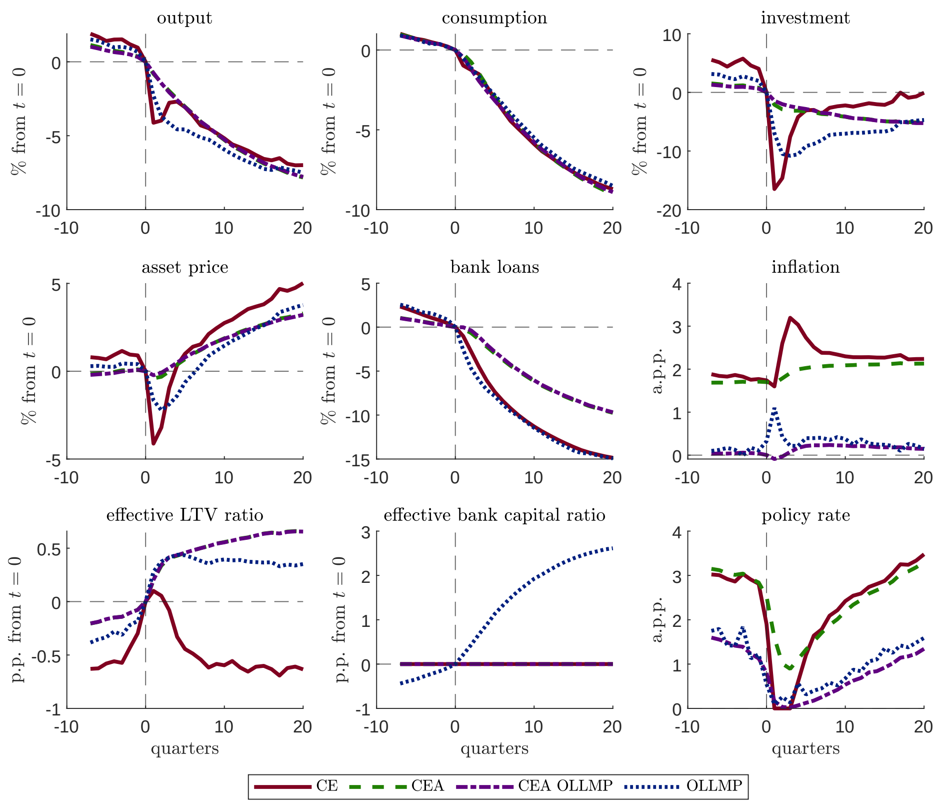

Figure 3.3. Zero-lower-bound crises. Each line is based on an average of 309 crisis episodes over a 50,000-period simulation—with an exception below—conditional on a Pareto vector $(\omega_b, \omega_e, \omega_w)' = (0.1, 0.1, 0.8)'$. A crisis starts at $t = 1$ and lasts at least three quarters. The effective LTV ratio is the ratio of the expected loan repayment to the value of collateral. The effective bank capital ratio is the ratio of bank capital to bank loans. "p.p." is "percentage points," and "a.p.p." is "annualized percentage points." The OLLMP paths are the averages over the crises observed in $0 \leq t \leq 24,108$, after which the simulation algorithm encounters numerical problems. The OLLMP simulation is based on setting $\beta_b = 0.995 < \beta$, the nearest neighbor that satisfies the Blanchard—Kahn conditions for local uniqueness and permits a relatively long simulation.

output, investment, and the asset price. The decrease in consumption and bank loans accelerates. When the ZLB becomes slack, the asset price and investment start to recover, but a decrease in output continues, and the recovery is slow. The collateral constraint is typically binding during a ZLB crisis, and there is a spike in the shadow value of collateral when the ZLB binds, reflected in the rise in the entrepreneur's LTV ratio. The bank leverage constraint remains binding during the whole crisis window, so the bank capital ratio is constant at $\bar{\kappa}$.

Except for the paths of inflation and the policy rate, the dynamics in the CEA and CEA OLLMP allocations are similar, consistent with the flexible-price analysis. Although the economy is stagnating ahead of ZLB crises, followed by a deep recession, there are no sharp changes in output growth, no drop in the asset price, and the investment dynamics are reminiscent of a cyclical decline. Bank loans eventually decrease by about five percentage points less than in the CE. A key reason for these differences is that in the CEA and CEA OLLMP allocations, the collateral constraint remains slack around a ZLB crisis, which allows for an increase in the entrepreneur's LTV ratio, supporting investment and the asset price and smoothing out a decrease in output and credit. Since bank leverage is suboptimal in the CEA, the capital ratio is constant at one. In the CEA, monetary policy is determined by the same Taylor rule as in the CE. However, the ZLB is not hit, and inflation stays close to the target. There is optimal monetary policy in the CEA OLLMP allocation, inflation stays close to the long-run level of zero throughout the crisis window, consistent with proposition 3.5, and the Ramsey planner typically avoids the ZLB.

A long simulation of the OLLMP allocation is prone to numerical problems because it entails accounting for five regime-switching constraints: the private complementary slackness conditions of bankers and entrepreneurs, the corresponding planner's complementary slackness conditions, and the planner's effective lower bound constraint. After increasing the banker's survival rate, a relatively long simulation is possible, but the results are not directly comparable to those in the other environments. Considering this limitation, we see that the dynamics of the real sector variables and the asset price are roughly a convex combination of the CE and CEA dynamics. A drop in investment and the asset price is less than in the CE, since the relaxed collateral constraint allows the Ramsey planner to increase the entrepreneur's LTV ratio. Although inflation is close to zero during most of the crisis window, there is a spike to about one percentage point after the ZLB binds in the CE. By increasing the inflation rate, the Ramsey planner evades the ZLB, which allows the planner to smooth out fluctuations, facilitated by the planner's ability to increase the bank capital ratio.

3.5 Conclusion

Financial constraints combined with consumer type heterogeneity lead to multiple sources of the inefficiency of the CE allocation. The inefficiency is reflected in both the real sector wedges in the demand for factors of production—labor and capital—and the financial sector wedges in the supply of bank deposits and the demand for bank loans. Nominal rigidities affect the real wedges but not the financial wedges. Consequently, optimal monetary policy in the presence of financial constraints and consumer type heterogeneity is reminiscent of the basic New Keynesian economy: stabilizing prices is optimal, exactly in the long run and approximately in the short run.

If a policymaker has the appropriate fiscal instruments to correct the intertemporal and intratemporal distortions in the CE allocation, the resulting CEA entails significant welfare gains. Under certain assumptions, such an allocation is close to an unconstrained Pareto optimum, having quantitatively perfect consumption insurance within consumer types and between types. Furthermore, the CEA has lower leverage in both the banking and the entrepreneurial sectors. These features help eliminate or mitigate the boom-bust financial crises and zero-lower-bound crises observed occasionally in the decentralized economy.

Correcting the Euler equation distortions might constitute an ambitious task. If that is not possible, but the leverage limits can be set optimally, the policymaker can still smooth out fluctuations by making the leverage ratios state contingent. The optimal bank capital and LTV ratios appear to be countercyclical around financial and ZLB crises.

Appendix

3.A Proofs

3.A.1 Lemma 3.1

Note that (3.2) and (3.5) imply $\lambda^b = U_C^b\left(1 - \frac{\beta_b}{\beta}\right)$. Using the latter, (3.6), (3.11), and the definition of $\widetilde{\beta}_e$, we get $\lambda^e = \frac{U_C^e}{R^L}\left(1 - \frac{\beta_e}{\widetilde{\beta}_e}\right)$.

If Suppose $\beta_b < \beta$ and $\beta_e < \widetilde{\beta}_e$. Then $\lambda^b > 0$ and $\lambda^e > 0$, which implies that (3.4) and (3.9) are binding. Straightforward algebraic manipulation of the system of static model equations and inequalities shows that there is a closed-form sequential solution for a unique steady state, where we set $\tau^D = \tau^K = \tau^L = \tau^N = 0$. In this steady state, the binding collateral constraint is used to solve for $L > 0$ conditional on $K > 0$. The binding leverage constraint is then used to solve for $D > 0$ conditional on $L > 0$.

Only if Suppose there exists a unique steady state with $D > 0$ and $L > 0$. Since $\lambda^b \geq 0$ and $\lambda^e \geq 0$, we must have $\beta_b \leq \beta$ and $\beta_e \leq \widetilde{\beta}_e$. If $\beta_b = \beta$, then (3.5) is equivalent to $\lambda^b = 0$. Moreover, the complementary slackness conditions (3.7) are automatically satisfied. Since $L > 0$ by the premise, any $D \in (0, (1-\kappa)L]$ can be part of an unstable steady state, which contradicts uniqueness. It follows that $\beta_b < \beta$. An identical argument applied to (3.9) and (3.11) demonstrates that we must have $\beta_e < \widetilde{\beta}_e$. ■

3.A.2 Lemma 3.2

Bankers Multiply both sides of (3.5) by D_t, multiply both sides of (3.6) by L_t, and subtract the former from the latter:

$$U^b_{C,t}(L_t - D_t) = \beta_b \mathbb{E}_t[U^b_{C,t+1}(R^L_{t+1}L_t - R_t D_t)] + \lambda^b_t[(1-\kappa_t)L_t - D_t].$$

Using (3.3) and (3.7),

$$U^b_{C,t}(L_t - D_t) = \beta_b \mathbb{E}_t(U^b_{C,t+1}C^b_{t+1}) + \beta_b \mathbb{E}_t[U^b_{C,t+1}(L_{t+1} - D_{t+1})].$$

Iterating this equation forward, we obtain

$$L_t - D_t = \frac{1}{U^b_{C,t}} \sum_{s=1}^{\infty} \beta_b^s \mathbb{E}_t(U^b_{C,t+s}C^b_{t+s}).$$

Entrepreneurs The argument is symmetric to the case of bankers. Multiply (3.11) by L_t and (3.12) by K_t, subtract the former from the latter and use (3.8), (3.10), and (3.13), noting that F is Cobb—Douglas, to obtain

$$U^e_{C,t}(Q_t K_t - L_t) = \beta_e \mathbb{E}_t(U^e_{C,t+1}C^e_{t+1}) + \beta_e \mathbb{E}_t[U^e_{C,t+1}(Q_{t+1}K_{t+1} - L_{t+1})].$$

Iterating forward, we get

$$Q_t K_t - L_t = \frac{1}{U^e_{C,t}} \sum_{s=1}^{\infty} \beta_e^s \mathbb{E}_t(U^e_{C,t+s}C^e_{t+s}). \quad ■$$

3.A.3 Lemma 3.3

The definition of $\mathcal{W}_t^i$ implies

$$\begin{aligned}
\frac{\beta}{\beta-\beta_i}(\mathcal{W}_t^i - V_t^i) &= \mathbb{E}_t\left(\sum_{s=1}^{\infty}\beta^s V_{t+s}^i\right) \\
&= \beta\mathbb{E}_t(V_{t+1}^i) + \beta\mathbb{E}_t\left[\mathbb{E}_{t+1}\left(\sum_{s=1}^{\infty}\beta^s V_{t+1+s}^i\right)\right] \\
&= \beta\mathbb{E}_t(V_{t+1}^i) + \beta\mathbb{E}_t\left[\frac{\beta}{\beta-\beta_i}(\mathcal{W}_{t+1}^i - V_{t+1}^i)\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathcal{W}_t^i &= V_t^i - \beta_i\mathbb{E}_t(V_{t+1}^i) + \beta\mathbb{E}_t(\mathcal{W}_{t+1}^i) \\
&= U_t^i + \beta\mathbb{E}_t(\mathcal{W}_{t+1}^i) \\
&= \mathbb{E}_t\left(\sum_{s=0}^{\infty}\beta^s U_{t+s}^i\right). \quad \blacksquare
\end{aligned}$$

3.A.4 Proposition 3.1

Define $\lambda_t^L \equiv \lambda_{1,t}^L + \lambda_{2,t}^L[(1-\kappa_t)L_t - D_t]$. The FOCs are

$$\begin{aligned}
C_t^b: &\quad 0 = \omega_b U_{C,t}^b - \lambda_t^Y - \lambda_t^C + \lambda_t^L U_{CC,t}^b L_t - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L\beta_b(U_{CC,t}^b R_t^L L_{t-1} + U_{C,t}^b) + \lambda_{t-1}^e], \\
C_t^e: &\quad 0 = \omega_e U_{C,t}^e - \lambda_t^Y - \lambda_t^C, \\
C_t^w: &\quad 0 = \omega_w U_{C,t}^w - \lambda_t^Y - \lambda_t^C W_{C,t}N_t - [\lambda_t^L\beta_b\mathbb{E}_t(U_{C,t+1}^b) + \beta\mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e]R_{1,t}U_{CC,t}^w D_t \\
&\qquad - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L\beta_b\mathbb{E}_{t-1}(U_{C,t}^b) + \beta\mathbb{E}_{t-1}(\lambda_t^C) + \lambda_{t-1}^e]R_{2,t-1}U_{CC,t}^w D_{t-1}, \\
D_t: &\quad 0 \geq -\lambda_t^b - \lambda_{2,t}^L[U_{C,t}^b - \beta_b\mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)]L_t + \lambda_t^C - [\lambda_t^L\beta_b\mathbb{E}_t(U_{C,t+1}^b) + \beta\mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e]R_t \\
&\qquad + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda_{t-1}^L\beta_b U_{C,t}^b + \lambda_{t-1}^e), \qquad \text{equality if } D_t > 0, \\
K_t: &\quad 0 = -\lambda_t^C\{Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] + Q_t\} + \lambda_t^e m_t\mathbb{E}_t[(Q_{1,t+1}K_t + Q_{t+1})\xi_{t+1}] - \lambda_t^Y I_{2,t} \\
&\qquad + \beta\mathbb{E}_t[(\lambda_{t+1}^C + \lambda_{t+1}^Y)A_{t+1}F_{K,t+1}\xi_{t+1} + \lambda_{t+1}^C\{Q_{t+1}(1-\delta)\xi_{t+1} - Q_{1,t+1}[K_{t+1} \\
&\qquad - (1-\delta)\xi_{t+1}K_t]\} - \lambda_{t+1}^Y I_{1,t+1}] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda_{t-1}^e m_{t-1}Q_{2,t}\xi_t K_{t-1}, \\
L_t: &\quad 0 = \{\lambda_t^b + \lambda_{2,t}^L[U_{C,t}^b - \beta_b\mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)]L_t\}(1-\kappa_t) + \lambda_t^L U_{C,t}^b - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda_{t-1}^L\beta_b U_{C,t}^b + \lambda_{t-1}^e),
\end{aligned}$$

$$N_t: \quad 0 = \omega_w U^w_{N,t} + (\lambda^C_t + \lambda^Y_t) A_t F_{N,t} - [\lambda^L_t \beta_b \mathbb{E}_t(U^b_{C,t+1}) + \beta \mathbb{E}_t(\lambda^C_{t+1}) + \lambda^e_t] R_{1,t} U^w_{CN,t} D_t$$
$$- \lambda^C_t (W_{N,t} N_t + W_t) - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} [\lambda^L_{t-1} \beta_b \mathbb{E}_{t-1}(U^b_{C,t}) + \beta \mathbb{E}_{t-1}(\lambda^C_t) + \lambda^e_{t-1}] R_{2,t-1} U^w_{CN,t} D_{t-1}.$$

The complementary slackness conditions are

$$0 = \lambda^b_t [(1-\kappa_t) L_t - D_t], \qquad \lambda^b_t \geq 0,$$
$$0 = \lambda^L_{1,t} [U^b_{C,t} L_t - \beta_b \mathbb{E}_t(U^b_{C,t+1} B_{t+1})], \qquad D_t \lambda^L_{1,t} \geq 0,$$
$$0 = \lambda^e_t [m_t \mathbb{E}_t(Q_{t+1} \xi_{t+1}) K_t - \mathbb{E}_t(B_{t+1})], \qquad \lambda^e_t \geq 0.$$

Constrained inefficiency

Follows from inspecting the planner's analogs of (3.5) and (3.10)–(3.12). Consider them one-by-one.

Deposit supply The FOCs for C^b_t and D_t imply

$$U^b_{C,t} \leq \beta_b R_t \mathbb{E}_t(U^b_{C,t+1}) + \frac{\lambda^b_t}{\omega_b} + \Psi^D_t, \qquad \text{equality if } D_t > 0,$$

where

$$\omega_b \Psi^D_t \equiv (\beta - \beta_b) R_t \mathbb{E}_t(\omega_b U^b_{C,t+1}) + \lambda^Y_t - \beta R_t \mathbb{E}_t(\lambda^Y_{t+1}) + \lambda^L_{2,t} [U^b_{C,t} - \beta_b \mathbb{E}_t(U^b_{C,t+1} R^L_{t+1})] L_t$$
$$- \lambda^L_t [U^b_{CC,t} + \beta_b R_t \mathbb{E}_t(U^b_{CC,t+1} R^L_{t+1})] L_t + \beta R_t \mathbb{E}_t(\lambda^L_{t+1} U^b_{CC,t+1} L_{t+1})$$
$$+ \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda^L_{t-1} \beta_b U^b_{CC,t} R^L_t L_{t-1}.$$

Loan demand If $D_t > 0$, the FOCs for C^e_t, D_t, and L_t imply

$$U^e_{C,t} = \beta_e \mathbb{E}_t(U^e_{C,t+1} R^L_{t+1}) + \frac{\lambda^e_t}{\omega_e} \mathbb{E}_t(R^L_{t+1}) + \Psi^L_t,$$

where

$$\omega_e \Psi^L_t = (\beta - \beta_e) \mathbb{E}_t(\omega_e U^e_{C,t+1} R^L_{t+1}) - \mathbb{E}_t[(\beta \omega_e U^e_{C,t+1} + \lambda^e_t)(R^L_{t+1} - R_t)] + \lambda^Y_t - \beta R_t \mathbb{E}_t(\lambda^Y_{t+1})$$
$$- \lambda^L_t \left[\frac{U^b_{C,t}}{1-\kappa_t} - \beta_b R_t \mathbb{E}_t(U^b_{C,t+1}) \right] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \frac{\kappa_t}{1-\kappa_t} (\lambda^L_{t-1} \beta_b U^b_{C,t} + \lambda^e_{t-1}).$$

If $D_t = 0$, we still have $L_t > 0$, so the FOC for L_t holds. To see this, note that the leverage constraint implies $C^b_{t+1} + L_{t+1} - D_{t+1} \geq 0$, and the inequality is strict if $C^b_{t+1} > 0$. Provided that $C^b_{t+1} > 0$ with positive measure, which is guaranteed if bankers are risk averse and the Inada condition holds, $D_t = L_t = 0$ would contradict the constraint associated with $\lambda^L_{1,t}$. Note that the FOCs for C^b_t and C^e_t yield the following general relationship between the marginal utilities

$$\omega_b U^b_{C,t} = \omega_e U^e_{C,t} - \lambda^L_t U^b_{CC,t} L_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda^L_{t-1}\beta_b(U^b_{CC,t}R^L_t L_{t-1} + U^b_{C,t}) + \lambda^e_{t-1}].$$

With $D_t = 0$, we have $U^b_{C,t} = \beta_b \mathbb{E}_t(U^b_{C,t+1}R^L_{t+1})$. The FOC for L_t then implies $\lambda^L_t U^b_{C,t} = \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda^L_{t-1}\beta_b U^b_{C,t} + \lambda^e_{t-1})$ at t. Combining these results, if $D_t = 0$, the wedge satisfies

$$\begin{aligned}
\omega_e \Psi^L_t = {} & (\beta_b - \beta_e)\mathbb{E}_t(\omega_e U^e_{C,t+1}R^L_{t+1}) + \lambda^L_t \Big\{ U^b_{CC,t}L_t - U^b_{C,t} \\
& + \frac{\beta_b^2}{\beta}\mathbb{E}_t[(U^b_{CC,t+1}R^L_{t+1}L_t + U^b_{C,t+1})R^L_{t+1}] \Big\} - \beta_b \mathbb{E}_t(\lambda^L_{t+1}U^b_{CC,t+1}L_{t+1}R^L_{t+1}) \\
& - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda^L_{t-1}\beta_b U^b_{CC,t}R^L_t L_{t-1} - \frac{\beta - \beta_b}{\beta}\lambda^e_t \mathbb{E}_t(R^L_{t+1}).
\end{aligned}$$

Labor demand The FOCs for C^e_t, C^w_t, and N_t combined with the definition of W_t imply

$$W_t = A_t F_{N,t} + \Psi^N_t,$$

where

$$\Psi^N_t = \frac{(\omega_e U^e_{C,t} - \omega_w U^w_{C,t} - \lambda^C_t)A_t F_{N,t} - \lambda^C_t W_{N,t}N_t}{\omega_w U^w_{C,t} + \lambda^C_t} - \frac{U^w_{CN,t}}{U^w_{CC,t}} \frac{\omega_w U^w_{C,t} - \omega_e U^e_{C,t} + \lambda^C_t(1 - W_{C,t}N_t)}{\omega_w U^w_{C,t} + \lambda^C_t}.$$

Capital demand The FOCs for C^e_t and K_t imply

$$U^e_{C,t}Q_t = \beta_e \mathbb{E}_t\{U^e_{C,t+1}[A_{t+1}F_{K,t+1} + Q_{t+1}(1-\delta)]\xi_{t+1}\} + \frac{\lambda^e_t}{\omega_e}m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1}) + \Psi^K_t,$$

where, using the form of I,

$$\omega_e \Psi_t^K = (\beta - \beta_e)\mathbb{E}_t(\omega_e U^e_{C,t+1} R^K_{t+1}) Q_t + \beta \mathbb{E}_t \left\{ \lambda^Y_{t+1} \left[Q_{t+1} \Phi\left(\frac{I_{t+1}}{K_t}\right) - \frac{I_{t+1}}{K_t} \right] \right\}$$
$$- \lambda^C_t Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] - \beta \mathbb{E}_t\{\lambda^C_{t+1} Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1} K_t]\}$$
$$+ \lambda^e_t m_t \mathbb{E}_t(Q_{1,t+1}\xi_{t+1}) K_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda^e_{t-1} m_{t-1} Q_{2,t} \xi_t K_{t-1}.$$

Risk sharing

That consumption insurance is generally imperfect follows immediately from inspecting the FOCs with respect to C_t^b, C_t^e, and C_t^w. The same applies to partial risk sharing between bankers and entrepreneurs. Note that the FOCs for D_t and L_t imply a steady-state relationship $\lambda^e = \lambda^L(\beta - \beta_b) U^b_C(C^b)$. Hence, λ^e and λ^L are either both zero or both positive.

Suppose workers have separable preferences $U^w(C^w, N) = u(C^w) - v(N)$ and $\lambda^e = \lambda^L = 0$. In this case, $\omega_b U^b_C(C^b) = \omega_e U^e_C(C^e) = \lambda^Y + \lambda^C$. Using the definition of functions R and W, we have $R_1 = \frac{R}{u'(C^w)}$, $R_2 = -\frac{\beta R^2}{u'(C^w)}$, $W_C = -W \frac{u''(C^w)}{u'(C^w)}$, and $\beta R = 1$. The FOC for C^w then implies

$$\begin{aligned} 0 &= \omega_w u'(C^w) - \lambda^Y + \lambda^C \frac{u''(C^w)}{u'(C^w)}[WN + (R-1)D] \\ &= \omega_w u'(C^w) - \lambda^Y + \lambda^C \frac{u''(C^w)}{u'(C^w)} C^w, \end{aligned}$$

where the second equality is true if the steady-state profits of capital good producers are zero so that the worker's budget constraint implies $C^w = WN + (R-1)D$. It follows that $\omega_w u'(C^w) = \lambda^Y + \lambda^C$ if and only if $(-C^w)\frac{u''(C^w)}{u'(C^w)} = 1$ if and only if $u(\cdot) = \ln(\cdot)$.

Indeterminacy and optimal steady state

The steady state construction reduces to considering two cases, $\lambda^L = 0$ and $\lambda^L > 0$. If $\lambda^L = 0$, D must satisfy the rearranged collateral constraint:

$$C^b + (R-1)D + \max\left\{ \frac{1}{1-\kappa} D, \frac{\beta_b}{1-\beta_b}[C^b + (R-1)D] \right\} \leq mQ\xi K.$$

If $\lambda^L > 0$, we instead have a rearranged bank leverage constraint: $D \leq \frac{\beta_b(1-\kappa)}{1-\beta_b[1+(1-\kappa)(R-1)]}C^b$. In both cases, there is a generally infinite set of solutions $D \in [0, \bar{D}]$ for some $\bar{D} > 0$. Since there is an uncountable infinity of steady states, each such steady state is unstable, and the FCEA is locally indeterminate. Numerical analysis under the baseline calibration demonstrates that each choice of D yields either a unique solution to a nonlinear system or no solutions, and welfare $\mathcal{W}$ is strictly decreasing in D. The latter is related to the problem of finding an optimal steady state.

Consider the planner's problem with no uncertainty, restricting attention to constant plans. An optimal plan of this sort will define the optimal steady state. In the steady state, $R = \frac{1}{\beta}$, $\frac{I}{K} = \Phi^{-1}[1-(1-\delta)\xi]$, and $Q = \left[\Phi'\left(\frac{I}{K}\right)\right]^{-1}$. Moreover, the constraints associated with λ^b, λ_1^L, and λ_2^L are equivalent to

$$L = \max\left\{\frac{1}{1-\kappa}D, \frac{\beta_b}{1-\beta_b}[C^b + (R-1)D]\right\},$$

conditional on (C^b, D). The optimal steady state is then a solution to

$$\max_{(C^b, C^e, C^w, D, K, N)} \sum_{i \in \mathcal{I}} \omega_i U^i$$

subject to

$$\begin{aligned}
\lambda^C: \quad & 0 = AF(\xi K, N) - Q[1-(1-\delta)\xi]K - W(C^w, N)N - (R-1)D - C^b - C^e, \\
\lambda^e: \quad & 0 \leq mQ\xi K - [C^b + (R-1)D + L], \\
\lambda^Y: \quad & 0 = AF(\xi K, N) - \sum_{i \in \mathcal{I}} C^i - \frac{I}{K}K.
\end{aligned}$$

Conditional on C^b, L is a strictly increasing function of D, differentiable everywhere except at the kink. We can assume without loss of generality that the derivative at the kink is an average of the left and right derivatives. Suppose (C^b, C^e, C^w, D, K, N) is optimal, where $D > 0$. It must satisfy the FOC for D:

$$0 = -\lambda^C(R-1) - \lambda^e\left(R - 1 + \frac{\partial L}{\partial D}\right).$$

Note that $R > 1$, $\lambda^e \geq 0$, and $\frac{\partial L}{\partial D} > 0$. If also $\lambda^C > 0$, we have $-\lambda^C(R-1) - \lambda^e\left(R - 1 + \frac{\partial L}{\partial D}\right) < 0$, which is a contradiction. Therefore, $D = 0$ is optimal.

Intuitively, λ^C must be positive since it is the shadow value of wealth associated with the consolidated budget constraint of bankers and entrepreneurs. Assume separable preferences and combine the FOCs for C^w and N together with the definition of W to obtain

$$\lambda^C = \frac{\omega_w u'(C^w)(AF_N - W)}{(W_C N - 1)AF_N + W_N N + W}.$$

By definition, $W_C > 0$ and $W_N > 0$. Hence, if N and C^w are less than in the first-best allocation, F_N must be greater and W less; therefore, $AF_N - W > 0$. A sufficient—but not necessary—condition for the denominator to be positive is $W_C N \geq 1$. If u has constant relative risk aversion $\gamma_w > 0$, as is the case in the quantitative analysis, $W_C = -W\frac{u''(C^w)}{u'(C^w)} = \frac{W}{C^w}\gamma_w$. If γ_w is large enough, we are done. Alternatively, if $\gamma_w \approx 1$ and $D \approx 0$, then $C_w \approx WN$ and $W_C N \approx 1$; therefore, $(W_C N - 1)AF_N \approx 0$. Since $W_N N + W > 0$, we then have $\lambda^C > 0$. ■

3.A.5 Proposition 3.2

Bankers Note that the form of T_t^b ensures that (3.3) is true in equilibrium. The Euler equation for deposits is now

$$U_{C,t}^b(1 - \tau_t^D) \leq \beta_b R_t \mathbb{E}_t(U_{C,t+1}^b) + \lambda_t^b, \qquad \text{equality if } D_t > 0.$$

Using (3.6)—which remains unchanged relative to the FCE—to solve for λ_t^b, the Euler equation for deposits can be rearranged as

$$U_{C,t}^b \leq \beta_b R_t \mathbb{E}_t(U_{C,t+1}^b) + \frac{U_{C,t}^b - \beta_b \mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)}{1 - \kappa_t} + \tau_t^D U_{C,t}^b, \qquad \text{equality if } D_t > 0.$$

As follows from section 3.A.4, the right-hand side is equivalent to the one in the FCEA if and only if

$$\tau_t^D = \frac{1}{U_{C,t}^b}\left[\frac{\lambda_t^b}{\omega_b} - \frac{U_{C,t}^b - \beta_b \mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)}{1 - \kappa_t} + \Psi_t^D\right].$$

Entrepreneurs The form of T_t^e guarantees that (3.8) holds in equilibrium. Without loss of generality, let $\frac{\lambda_t^e}{\omega_e}$ denote the scaled Lagrange multiplier on the collateral constraint. The modified FOCs are

$$(1 + \tau_t^N)W_t = A_t F_{N,t},$$

$$U^e_{C,t}(1-\tau^L_t) = \beta_e \mathbb{E}_t(U^e_{C,t+1}R^L_{t+1}) + \frac{\lambda^e_t}{\omega_e}\mathbb{E}_t(R^L_{t+1}),$$
$$U^e_{C,t}(1+\tau^K_t)Q_t = \beta_e \mathbb{E}_t\{U^e_{C,t+1}[A_{t+1}F_{K,t+1} + Q_{t+1}(1-\delta)]\xi_{t+1}\} + \frac{\lambda^e_t}{\omega_e}m_t\mathbb{E}_t(Q_{t+1}\xi_{t+1}).$$

Section 3.A.4 then immediately implies that we must set

$$\tau^N_t = \frac{-\Psi^N_t}{W_t}, \qquad \tau^L_t = \frac{\Psi^L_t}{U^e_{C,t}}, \qquad \tau^K_t = \frac{-\Psi^K_t}{U^e_{C,t}Q_t}.$$

Ramsey equilibrium On the banker's side, we can use the regulated deposit supply Euler equation to solve for τ^b_t in terms of allocations and prices. The remaining constraints are identical to those faced by the social planner in the definition of an FCEA. Similarly, on the entrepreneur's side, we can use the regulated demand conditions for labor, loans, and capital to back out the corresponding tax rates τ^N_t, τ^L_t, and τ^K_t. Guessing that the private complementary slackness conditions associated with the collateral constraint are not binding, we are left with the entrepreneur's budget constraint and the collateral constraint—the same set of constraints as in the FCEA definition. After solving for prices and the investment function as in the FCEA, the complete set of constraints faced by the Ramsey planner is identical to the one in the FCEA definition. Therefore, the FCEA is exactly the allocation that is part of the Ramsey equilibrium. Finally, we can verify that the individual entrepreneur's complementary slackness conditions are indeed not binding because they are implied by the planner's analogous complementary slackness conditions. ∎

3.A.6 Lemma 3.4

The relaxed problem is

$$\max_{\{C^b_t, C^e_t, C^w_t, D_t, K_t, L_t, N_t\}} \mathbb{E}_0\left(\sum_{t=0}^{\infty}\beta^t\sum_{i\in\mathcal{I}}\omega_i U^i_t\right)$$

subject to

$$\begin{aligned}
\lambda^b_t:&\quad 0 \le L_t - D_t,\\
\lambda^L_t:&\quad 0 \le U^b_C(C^b_t)L_t - \beta_b\mathbb{E}_t[U^b_C(C^b_{t+1})(C^b_{t+1} + L_{t+1} - D_{t+1} + R_tD_t)], \qquad \text{equality if } D_t = 0,\\
\lambda^C_t:&\quad 0 = A_tF(\xi_tK_{t-1}, N_t) - Q(K_{t-1}, K_t, \xi_t)[K_t - (1-\delta)\xi_tK_{t-1}] - W(C^w_t, N_t)N_t + D_t\\
&\qquad - R_{t-1}D_{t-1} - C^b_t - C^e_t,
\end{aligned}$$

$$\lambda_t^e: \quad 0 \leq \mathbb{E}_t(Q(K_t, K_{t+1}, \xi_{t+1})\xi_{t+1})K_t - \mathbb{E}_t(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t),$$
$$\lambda_t^Y: \quad 0 = A_t F(\xi_t K_{t-1}, N_t) - \sum_{i \in \mathcal{I}} C_t^i - I(K_{t-1}, K_t, \xi_t),$$

where $R_t = R(U_C^w(C_t^w, N_t), \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})])$, the functions W, R, Q, and I are the same as in definition 3.4. An allocation-policy pair is part of a Ramsey equilibrium if—combined with the associated prices and Lagrange multipliers—it constitutes a regulated competitive equilibrium with the maximum level of welfare over all feasible allocation-policy pairs.

Consider a feasible policy $\{\kappa_t, m_t, \tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K\} \subset [0,1]^2 \times \mathbb{R}^4$ and the corresponding regulated FCE allocation $\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t\}$. The policy is consistent with the construction in the lemma. If $U_{C,t}^b > \beta_b \mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)$, then (3.6) implies that $\lambda_t^b > 0$, and thus the leverage constraint is binding, which implies $\kappa_t = 1 - \frac{D_t}{L_t}$; otherwise, $\kappa_t \geq 0$ combined with the leverage constraint is equivalent to $\kappa_t \in \left[0, 1 - \frac{D_t}{L_t}\right]$. The collateral constraint combined with $m_t \leq 1$ is equivalent to $m_t \in \left[\frac{\mathbb{E}_t(R_{t+1}^L)L_t}{\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t}, 1\right]$. The tax rates are consistent with the regulated analogs of (3.5) and (3.10)–(3.12). Moreover, as argued in proposition 3.2, the allocation is feasible for the FCEA problem. Since $D_t \leq (1-\kappa_t)L_t \leq L_t$ and $\mathbb{E}_t(R_{t+1}^L)L_t \leq m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t \leq \mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t$, the allocation is feasible for the relaxed problem.

Conversely, suppose an allocation $\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t\}$ is feasible for the relaxed problem and construct the corresponding policy as described in the lemma. The construction of κ_t ensures that the FCE version of the bank leverage constraint and the private complementary slackness conditions are satisfied. The construction of m_t guarantees that the FCE version of the collateral constraint is respected. The construction of the tax rates makes sure that the regulated analogs of (3.5) and (3.10)–(3.12) hold. The policy is feasible, that is, $\{\kappa_t, m_t, \tau_t^D, \tau_t^N, \tau_t^L, \tau_t^K\} \subset [0,1]^2 \times \mathbb{R}^4$. It follows that the allocation and the constructed policy—combined with the associated prices and Lagrange multipliers—constitute an FCE.

We have established that the two problems have identical feasible sets of allocation-policy pairs. Since the objective functions are equivalent, the two problems yield identical optimal allocation-policy pairs. ■

3.A.7 Lemma 3.5

The relaxed problem is

$$\max_{\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t\}} \mathbb{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \sum_{i \in \mathcal{I}} \omega_i U_t^i \right)$$

subject to

$$\begin{aligned}
\lambda_t^b : \quad & 0 \leq L_t - D_t, \\
\lambda_t^L : \quad & 0 = \beta_b \mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b + L_{t+1} - D_{t+1})] - U_C^b(C_t^b)(L_t - D_t), \\
\lambda_t^D : \quad & 0 \leq U_C^b(C_t^b) - \beta_b R_t \mathbb{E}_t(U_C^b(C_{t+1}^b)), \\
\lambda_t^C : \quad & 0 = A_t F(\xi_t K_{t-1}, N_t) - Q(K_{t-1}, K_t, \xi_t)[K_t - (1-\delta)\xi_t K_{t-1}] - W(C_t^w, N_t)N_t + D_t \\
& \quad - R_{t-1}D_{t-1} - C_t^b - C_t^e, \\
\lambda_t^e : \quad & 0 \leq \mathbb{E}_t(Q(K_t, K_{t+1}, \xi_{t+1})\xi_{t+1})K_t - \mathbb{E}_t(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t), \\
\lambda_t^K : \quad & 0 = \beta_e \mathbb{E}_t[U_C^e(C_{t+1}^e)\{[A_{t+1}F_K(\xi_{t+1}K_t, N_{t+1}) + Q(K_t, K_{t+1}, \xi_{t+1})(1-\delta)]\xi_{t+1}K_t - C_{t+1}^b \\
& \quad - L_{t+1} + D_{t+1} - R_t D_t\}] - U_C^e(C_t^e)(Q(K_{t-1}, K_t, \xi_t)K_t - L_t), \\
\lambda_t^B : \quad & 0 \leq U_C^e(C_t^e)L_t - \beta_e \mathbb{E}_t[U_C^e(C_{t+1}^e)(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t)], \\
\lambda_t^Y : \quad & 0 = A_t F(\xi_t K_{t-1}, N_t) - \sum_{i \in \mathcal{I}} C_t^i - I(K_{t-1}, K_t, \xi_t),
\end{aligned}$$

where $R_t = R(U_C^w(C_t^w, N_t), \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})])$, the functions W, R, Q, and I are as in definition 3.4.

In the absence of taxation on the banker's side, (3.3)–(3.7) must be respected by the planner. As before, we can use (3.3) to solve for $B_t \equiv R_t^L L_{t-1} = C_t^b + L_t - D_t + R_{t-1}D_{t-1}$. Now we can use (3.5) to express $\lambda_t^b = U_{C,t}^b - \beta_b R_t \mathbb{E}_t(U_{C,t+1}^b)$. Multiplying (3.5) by D_t and (3.6) by L_t, subtracting the former from the latter, and using the complementary slackness conditions (3.7), (3.6) can be expressed in terms of allocations only. Hence, the implementability conditions that go to the Ramsey problem from the banker's side are

$$\begin{aligned}
& 0 \leq (1-\kappa_t)L_t - D_t, \\
& 0 = \beta_b \mathbb{E}_t[U_{C,t+1}^b(C_{t+1}^b + L_{t+1} - D_{t+1})] - U_{C,t}^b(L_t - D_t),
\end{aligned}$$

$$0 \leq U^b_{C,t} - \beta_b R_t \mathbb{E}_t(U^b_{C,t+1}),$$

$$0 = [U^b_{C,t} - \beta_b R_t \mathbb{E}_t(U^b_{C,t+1})][(1-\kappa_t)L_t - D_t].$$

Consider the entrepreneur's problem. As before, we can use the regulated analog of (3.10) to solve for $\tau^N_t \equiv \frac{A_t F_{N,t}}{W_t} - 1$. Using (3.11), we can express $\lambda^e_t \mathbb{E}_t(B_{t+1}) = U^e_{C,t} L_t - \beta_e \mathbb{E}_t(U^e_{C,t+1} B_{t+1})$. By multiplying (3.11) by L_t and (3.12) by K_t, subtracting the former from the latter, and using the complementary slackness conditions (3.13), (3.12) can be identically expressed in terms of allocations. Using the definition of B_t based on (3.3), the implementability conditions from the entrepreneur's side are

$$0 = A_t F(\xi_t K_{t-1}, N_t) - Q_t[K_t - (1-\delta)\xi_t K_{t-1}] - W_t N_t + D_t - R_{t-1} D_{t-1} - C^b_t - C^e_t,$$

$$0 \leq m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1}) K_t - \mathbb{E}_t(B_{t+1}),$$

$$0 = \beta_e \mathbb{E}_t[U^e_{C,t+1}\{[A_{t+1} F_K(\xi_{t+1} K_t, N_{t+1}) + Q_{t+1}(1-\delta)]\xi_{t+1} K_t - B_{t+1}\}] - U^e_{C,t}(Q_t K_t - L_t),$$

$$0 \leq U^e_{C,t} L_t - \beta_e \mathbb{E}_t(U^e_{C,t+1} B_{t+1}),$$

$$0 = [U^e_{C,t} L_t - \beta_e \mathbb{E}_t(U^e_{C,t+1} B_{t+1})][m_t \mathbb{E}_t(Q_{t+1}\xi_{t+1}) K_t - \mathbb{E}_t(B_{t+1})].$$

The remaining implementability conditions are constituted in the functions W, R, Q, and I, defined by (3.1), (3.2), (3.14), and (3.20), as well as the resource constraint obtained by combining (3.21) and (3.22).

The equivalence between the feasible sets of allocation-policy pairs that satisfy the implementability conditions above and the constraints of the relaxed problem follows from the arguments that are identical to the proof of lemma 3.4. Now we have only one tax rate τ^N_t that can be constructed from the regulated version of (3.10), and both κ_t and m_t are set such that the private complementary slackness conditions are satisfied. ■

3.A.8 Proposition 3.3

The FOCs for the problem of lemma 3.5 are

$$
\begin{aligned}
C_t^b: \quad & 0 = \omega_b U^b_{C,t} - \lambda^Y_t - \lambda^C_t - [\lambda^L_t(L_t - D_t) - \lambda^D_t]U^b_{CC,t} + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\{[\lambda^L_{t-1}(C^b_t + L_t - D_t) \\
& - \lambda^D_{t-1}R_{t-1}]\beta_b U^b_{CC,t} + \lambda^L_{t-1}\beta_b U^b_{C,t} - \lambda^e_{t-1} - (\lambda^K_{t-1} + \lambda^B_{t-1})\beta_e U^e_{C,t}\}, \\
C_t^e: \quad & 0 = \omega_e U^e_{C,t} - \lambda^Y_t - \lambda^C_t - [\lambda^K_t(Q_tK_t - L_t) - \lambda^B_t L_t]U^e_{CC,t} \\
& + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda^K_{t-1}R^K_t Q_{t-1}K_{t-1} - (\lambda^K_{t-1} + \lambda^B_{t-1})R^L_t L_{t-1}]\beta_e U^e_{CC,t}, \\
C_t^w: \quad & 0 = \omega_w U^w_{C,t} - \lambda^Y_t - \lambda^C_t W_{C,t}N_t - \{\lambda^D_t \beta_b \mathbb{E}_t(U^b_{C,t+1}) + [\beta\mathbb{E}_t(\lambda^C_{t+1}) + \lambda^e_t \\
& + (\lambda^K_t + \lambda^B_t)\beta_e\mathbb{E}_t(U^e_{C,t+1})]D_t\}R_{1,t}U^w_{CC,t} - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\{\lambda^D_{t-1}\beta_b\mathbb{E}_{t-1}(U^b_{C,t}) \\
& + [\beta\mathbb{E}_{t-1}(\lambda^C_t) + \lambda^e_{t-1} + (\lambda^K_{t-1} + \lambda^B_{t-1})\beta_e\mathbb{E}_{t-1}(U^e_{C,t})]D_{t-1}\}R_{2,t-1}U^w_{CC,t}, \\
D_t: \quad & 0 = -\lambda^b_t + \lambda^L_t U^b_{C,t} + \lambda^C_t - [\beta\mathbb{E}_t(\lambda^C_{t+1}) + \lambda^e_t + (\lambda^K_t + \lambda^B_t)\beta_e\mathbb{E}_t(U^e_{C,t+1})]R_t \\
& + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[-\lambda^L_{t-1}\beta_b U^b_{C,t} + \lambda^e_{t-1} + (\lambda^K_{t-1} + \lambda^B_{t-1})\beta_e U^e_{C,t}], \\
K_t: \quad & 0 = -\lambda^C_t\{Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] + Q_t\} + \lambda^e_t\mathbb{E}_t[(Q_{1,t+1}K_t + Q_{t+1})\xi_{t+1}] - \lambda^Y_t I_{2,t} \\
& + \lambda^K_t\{\beta_e\mathbb{E}_t[U^e_{C,t+1}\{[A_{t+1}F_{KK,t+1}\xi_{t+1} + Q_{1,t+1}(1-\delta)]\xi_{t+1}K_t + R^K_{t+1}Q_t\}] \\
& - U^e_{C,t}(Q_{2,t}K_t + Q_t)\} + \beta\mathbb{E}_t[(\lambda^C_{t+1} + \lambda^Y_{t+1})A_{t+1}F_{K,t+1}\xi_{t+1} \\
& + \lambda^C_{t+1}\{Q_{t+1}(1-\delta)\xi_{t+1} - Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1}K_t]\} \\
& - \lambda^K_{t+1}U^e_{C,t+1}Q_{1,t+1}K_{t+1} - \lambda^Y_{t+1}I_{1,t+1}] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda^e_{t-1} + \lambda^K_{t-1}\beta_e U^e_{C,t}(1-\delta)]Q_{2,t}\xi_t K_{t-1}, \\
L_t: \quad & 0 = \lambda^b_t - \lambda^L_t U^b_{C,t} + (\lambda^K_t + \lambda^B_t)U^e_{C,t} + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda^L_{t-1}\beta_b U^b_{C,t} - \lambda^e_{t-1} - (\lambda^K_{t-1} + \lambda^B_{t-1})\beta_e U^e_{C,t}], \\
N_t: \quad & 0 = \omega_w U^w_{N,t} + (\lambda^C_t + \lambda^Y_t)A_tF_{N,t} - \lambda^C_t(W_{N,t}N_t + W_t) - \{\lambda^D_t\beta_b\mathbb{E}_t(U^b_{C,t+1}) \\
& + [\beta\mathbb{E}_t(\lambda^C_{t+1}) + \lambda^e_t + (\lambda^K_t + \lambda^B_t)\beta_e\mathbb{E}_t(U^e_{C,t+1})]D_t\}R_{1,t}U^w_{CN,t} - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\{\lambda^D_{t-1}\beta_b\mathbb{E}_{t-1}(U^b_{C,t}) \\
& + [\beta\mathbb{E}_{t-1}(\lambda^C_t) + \lambda^e_{t-1} + (\lambda^K_{t-1} + \lambda^B_{t-1})\beta_e\mathbb{E}_{t-1}(U^e_{C,t})]D_{t-1}\}R_{2,t-1}U^w_{CN,t} \\
& + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda^K_{t-1}\beta_e U^e_{C,t}A_tF_{KN,t}\xi_t K_{t-1}.
\end{aligned}
$$

The complementary slackness conditions are

$$
0 = \lambda^b_t(L_t - D_t), \qquad \lambda^b_t \geq 0,
$$

$$0 = \lambda_t^D[U_{C,t}^b - \beta_b R_t \mathbb{E}_t(U_{C,t+1}^b)], \qquad \lambda_t^D \geq 0,$$

$$0 = \lambda_t^e[\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t - \mathbb{E}_t(B_{t+1})], \qquad \lambda_t^e \geq 0,$$

$$0 = \lambda_t^B[U_{C,t}^e L_t - \beta_e \mathbb{E}_t(U_{C,t+1}^e B_{t+1})], \qquad \lambda_t^B \geq 0.$$

Inspecting the FOCs for C_t^b, C_t^e, and C_t^w, we see that consumption insurance is generally imperfect.

Consider the steady state. The λ_t^L constraint implies $L - D = \frac{\beta_b}{1-\beta_b}C^b \geq 0$, which makes the relaxed leverage constraint redundant, implying $\lambda^b = 0$. Since $\beta_b < \beta$, we have $\lambda^D = 0$. The FOC for D_t then implies $\lambda^L = 0$. Guessing that C^b is sufficiently small relative to D, since $\beta_e < \beta$, we will have $L > \beta_e B$; therefore, $\lambda^B = 0$. (A sufficient condition is $\beta_e \leq \beta_b$.) The FOC for L_t then implies $\lambda^K = \frac{\lambda^e}{(\beta-\beta_e)U_C^e} \geq 0$. If $\lambda^e = 0$, we have $\omega_b U_C^b = \omega_e U_C^e = \lambda^C + \lambda^Y$; therefore, there is approximately perfect risk sharing between bankers and entrepreneurs. Risk sharing is only approximate because generally $\lambda_t^L \neq 0$ outside of the steady state. If $U^w(C^w, N) = \ln(C^w) - v(N)$ and the steady-state profits of capital good producers are zero, the FOC for C_t^w implies $\omega_w U_C^w = \lambda^C + \lambda^Y$, as in the proof of proposition 3.1.

The construction of the steady state boils down to considering two cases: collateral constraint is slack or binding. Each case can be reduced to solving a system of three nonlinear equations. Conditional on solving a nonlinear system, the sequential solution uniquely determines the steady state. Since the problem reduces to a numerical one, we cannot claim that the steady state is unique. However, it is the case under the baseline calibration and other parameterizations considered in the analysis. ■

3.A.9 Proposition 3.4

The planning problem is

$$\max_{\{C_t^b, C_t^e, C_t^w, D_t, K_t, L_t, N_t, \Omega_{1,t}\}} \mathbb{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \sum_{i \in \mathcal{I}} \omega_i U_t^i \right)$$

subject to

$$\lambda_t^b: \quad 0 \leq (1-\kappa_t)L_t - D_t,$$

$$\lambda_{1,t}^L: \quad 0 \le U_C^b(C_t^b)L_t - \beta_b \mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t)], \qquad \text{equality if } D_t = 0,$$

$$\lambda_{2,t}^L: \quad 0 = \{U_C^b(C_t^b)L_t - \beta_b \mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t)]\}[(1-\kappa_t)L_t - D_t],$$

$$\lambda_t^C: \quad 0 = \Delta_t \left\{ \Omega_{1,t} - \frac{\beta\theta \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})\Pi_{t+1}^{\epsilon}\Omega_{1,t+1}]}{U_C^w(C_t^w, N_t)} \right\} - Q(K_{t-1}, K_t, \xi_t)[K_t - (1-\delta)\xi_t K_{t-1}]$$
$$- W(C_t^w, N_t)N_t + D_t - R_{t-1}D_{t-1} - C_t^b - C_t^e,$$

$$\lambda_t^e: \quad 0 \le m_t \mathbb{E}_t(Q(K_t, K_{t+1}, \xi_{t+1})\xi_{t+1})K_t - \mathbb{E}_t(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t),$$

$$\lambda_t^Y: \quad 0 = \frac{A_t}{\Delta_t} F(\xi_t K_{t-1}, N_t) - \sum_{i\in\mathcal{I}} C_t^i - I(K_{t-1}, K_t, \xi_t),$$

$$\lambda_t^\Omega: \quad 0 = \frac{\epsilon - 1}{\epsilon}\widetilde{P}_t \frac{A_t}{\Delta_t} F(\xi_t K_{t-1}, N_t) - \Omega_{1,t} + \frac{\beta\theta\mathbb{E}_t\left[U_C^w(C_{t+1}^w, N_{t+1})\Pi_{t+1}^{\epsilon-1}\frac{\widetilde{P}_t}{\widetilde{P}_{t+1}}\Omega_{1,t+1}\right]}{U_C^w(C_t^w, N_t)}.$$

As before, define $\lambda_t^L \equiv \lambda_{1,t}^L + \lambda_{2,t}^L[(1-\kappa_t)L_t - D_t]$. The FOCs are

$$C_t^b: \quad 0 = \omega_b U_{C,t}^b - \lambda_t^Y - \lambda_t^C + \lambda_t^L U_{CC,t}^b L_t - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L \beta_b (U_{CC,t}^b R_t^L L_{t-1} + U_{C,t}^b) + \lambda_{t-1}^e],$$

$$C_t^e: \quad 0 = \omega_e U_{C,t}^e - \lambda_t^Y - \lambda_t^C,$$

$$C_t^w: \quad 0 = \omega_w U_{C,t}^w - \lambda_t^Y - \lambda_t^C W_{C,t} N_t - [\lambda_t^L \beta_b \mathbb{E}_t(U_{C,t+1}^b) + \beta\mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e]R_{1,t} U_{CC,t}^w D_t$$
$$- \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L \beta_b \mathbb{E}_{t-1}(U_{C,t}^b) + \beta\mathbb{E}_{t-1}(\lambda_t^C) + \lambda_{t-1}^e]R_{2,t-1}U_{CC,t}^w D_{t-1}$$
$$+ \left[(\lambda_t^C \Delta_t - \lambda_t^\Omega)\Omega_{1,t} - \left(\lambda_t^C P_t^w \Delta_t - \lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t\right) Y_t\right] \frac{U_{CC,t}^w}{U_{C,t}^w}$$
$$- \mathbf{1}_{\mathbb{N}}(t)\left(\lambda_{t-1}^C \Delta_{t-1}\Pi_t - \lambda_{t-1}^\Omega \frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\right)\theta \Pi_t^{\epsilon-1}\Omega_{1,t}\frac{U_{CC,t}^w}{U_{C,t-1}^w},$$

$$D_t: \quad 0 \ge -\lambda_t^b - \lambda_{2,t}^L[U_{C,t}^b - \beta_b\mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)]L_t + \lambda_t^C - [\lambda_t^L\beta_b\mathbb{E}_t(U_{C,t+1}^b) + \beta\mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e]R_t$$
$$+ \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda_{t-1}^L \beta_b U_{C,t}^b + \lambda_{t-1}^e),$$

$$K_t: \quad 0 = -\lambda_t^C\{Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] + Q_t\} + \lambda_t^e m_t \mathbb{E}_t[(Q_{1,t+1}K_t + Q_{t+1})\xi_{t+1}] - \lambda_t^Y I_{2,t}$$
$$+ \beta\mathbb{E}_t\left[\left(\lambda_{t+1}^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_{t+1} + \lambda_{t+1}^Y\right)\frac{A_{t+1}}{\Delta_{t+1}} F_{K,t+1}\xi_{t+1} - \lambda_{t+1}^Y I_{1,t+1} + \lambda_{t+1}^C\right.$$
$$\left. \times \{Q_{t+1}(1-\delta)\xi_{t+1} - Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1}K_t]\}\right] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda_{t-1}^e m_{t-1} Q_{2,t}\xi_t K_{t-1},$$

$$L_t: \quad 0 = \{\lambda_t^b + \lambda_{2,t}^L[U_{C,t}^b - \beta_b\mathbb{E}_t(U_{C,t+1}^b R_{t+1}^L)]L_t\}(1-\kappa_t) + \lambda_t^L U_{C,t}^b - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda_{t-1}^L\beta_b U_{C,t}^b + \lambda_{t-1}^e),$$

$$
\begin{aligned}
N_t: \quad & 0 = \omega_w U^w_{N,t} + \left(\lambda^\Omega_t \frac{\epsilon-1}{\epsilon}\widetilde{P}_t + \lambda^Y_t\right)\frac{A_t}{\Delta_t}F_{N,t} - [\lambda^L_t \beta_b \mathbb{E}_t(U^b_{C,t+1}) + \beta\mathbb{E}_t(\lambda^C_{t+1}) + \lambda^e_t] \\
& \times R_{1,t}U^w_{CN,t}D_t - \lambda^C_t(W_{N,t}N_t + W_t) - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda^L_{t-1}\beta_b\mathbb{E}_{t-1}(U^b_{C,t}) + \beta\mathbb{E}_{t-1}(\lambda^C_t) + \lambda^e_{t-1}] \\
& \times R_{2,t-1}U^w_{CN,t}D_{t-1} + \left[(\lambda^C_t\Delta_t - \lambda^\Omega_t)\Omega_{1,t} - \left(\lambda^C_t P^w_t \Delta_t - \lambda^\Omega_t \frac{\epsilon-1}{\epsilon}\widetilde{P}_t\right)Y_t\right]\frac{U^w_{CN,t}}{U^w_{C,t}} \\
& -\mathbf{1}_{\mathbb{N}}(t)\left(\lambda^C_{t-1}\Delta_{t-1}\Pi_t - \lambda^\Omega_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\right)\theta\Pi^{\epsilon-1}_t\Omega_{1,t}\frac{U^w_{CN,t}}{U^w_{C,t-1}}, \\
\Omega_{1,t}: \quad & 0 = \lambda^C_t\Delta_t - \lambda^\Omega_t - \mathbf{1}_{\mathbb{N}}(t)\left(\lambda^C_{t-1}\Delta_{t-1}\Pi_t - \lambda^\Omega_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\right)\theta\Pi^{\epsilon-1}_t\frac{U^w_{C,t}}{U^w_{C,t-1}}.
\end{aligned}
$$

The complementary slackness conditions are

$$
\begin{aligned}
0 &= \lambda^b_t[(1-\kappa_t)L_t - D_t], \qquad \lambda^b_t \geq 0, \\
0 &= \lambda^L_{1,t}[U^b_{C,t}L_t - \beta_b\mathbb{E}_t(U^b_{C,t+1}B_{t+1})], \qquad D_t\lambda^L_{1,t} \geq 0, \\
0 &= \lambda^e_t[m_t\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t - \mathbb{E}_t(B_{t+1})], \qquad \lambda^e_t \geq 0.
\end{aligned}
$$

Wedges Since the FOCs for C^b_t and D_t are the same as in the FCEA, the deposit wedge Ψ^D_t is too. Since the FOCs for C^e_t and L_t and the $\lambda^L_{1,t}$ constraint are the same as in the FCEA, the loan wedge Ψ^L_t is too.

The FOCs for C^e_t, C^w_t, and N_t combined with the definition of W_t imply

$$
W_t = P^w_t A_t F_{N,t} + \Psi^N_t,
$$

where

$$
\begin{aligned}
\Psi^N_t = & \frac{[(\omega_e U^e_{C,t} + \lambda^\Omega_t \frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda^C_t)(P^w_t\Delta_t)^{-1} - \omega_w U^w_{C,t} - \lambda^C_t]P^w_t A_t F_{N,t} - \lambda^C_t W_{N,t}N_t}{\omega_w U^w_{C,t} + \lambda^C_t} \\
& - \frac{U^w_{CN,t}}{U^w_{CC,t}}\frac{\omega_w U^w_{C,t} - \omega_e U^e_{C,t} + \lambda^C_t(1 - W_{C,t}N_t)}{\omega_w U^w_{C,t} + \lambda^C_t}.
\end{aligned}
$$

The FOCs for C^e_t and K_t imply

$$
U^e_{C,t}Q_t = \beta_e\mathbb{E}_t\{U^e_{C,t+1}[P^w_{t+1}A_{t+1}F_{K,t+1} + Q_{t+1}(1-\delta)]\xi_{t+1}\} + \frac{\lambda^e_t}{\omega_e}m_t\mathbb{E}_t(Q_{t+1}\xi_{t+1}) + \Psi^K_t,
$$

where

$$
\begin{aligned}
\omega_e \Psi_t^K = {} & (\beta - \beta_e)\mathbb{E}_t(\omega_e U^e_{C,t+1} R^K_{t+1}) Q_t + \beta \mathbb{E}_t \left\{ \lambda^Y_{t+1} \left[Q_{t+1} \Phi\left(\frac{I_{t+1}}{K_t}\right) - \frac{I_{t+1}}{K_t} \right] \right\} \\
& - \lambda^C_t Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] - \beta \mathbb{E}_t \{ \lambda^C_{t+1} Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1} K_t] \} \\
& + \lambda^e_t m_t \mathbb{E}_t(Q_{1,t+1}\xi_{t+1}) K_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda^e_{t-1} m_{t-1} Q_{2,t} \xi_t K_{t-1} \\
& + \beta \mathbb{E}_t \left\{ \left[\omega_e U^e_{C,t+1}(1 - P^w_{t+1}\Delta_{t+1}) + \lambda^\Omega_{t+1} \frac{\epsilon - 1}{\epsilon} \widetilde{P}_{t+1} - \lambda^C_{t+1} \right] \frac{A_{t+1}}{\Delta_{t+1}} F_{K,t+1} \xi_{t+1} \right\}.
\end{aligned}
$$

Risk sharing and steady state Risk sharing properties follow from inspecting the FOCs for C^b_t, C^e_t, and C^w_t. In particular, the latter now has the term that reflects the market power of retailers. If $\lambda^e = \lambda^L = 0$ and workers have separable preferences, the FOC for C^w in the steady state is

$$
0 = \omega_w u'(C^w) - \lambda^Y + \lambda^C \frac{u''(C^w)}{u'(C^w)} \left(C^w - \frac{Y}{\epsilon} \right).
$$

Since $\epsilon < \infty$, even with logarithmic preferences, the worker's steady-state Pareto-weighted marginal utility of consumption is less than that of bankers and entrepreneurs.

The steady state construction parallels the FCEA, reducing to two cases—whether $\lambda^L = 0$ or $\lambda^L > 0$. In both cases, the quantity of deposits is indeterminate, but conditional on choosing an admissible value of D, there typically exists a unique steady state. The proof that the optimal steady state has $D = 0$, provided that $\lambda^C > 0$, is identical to the FCEA in proposition 3.1.

Decentralization After replacing A_t with $P^w_t A_t$ in the entrepreneur's problem, the proof is identical to the proof of proposition 3.2. ■

3.A.10 Proposition 3.5

Case 1

The relaxed planning problem is

$$
\max_{\{C^b_t, C^e_t, C^w_t, D_t, K_t, L_t, N_t, \Omega_{1,t}, \Delta_t, \Pi_t\}} \mathbb{E}_0 \left(\sum_{t=0}^{\infty} \beta^t \sum_{i \in \mathcal{I}} \omega_i U^i_t \right)
$$

subject to

$$
\begin{aligned}
\lambda_t^b: &\quad 0 \le L_t - D_t,\\
\lambda_t^L: &\quad 0 \le U_C^b(C_t^b)L_t - \beta_b \mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t)],\\
\lambda_t^C: &\quad 0 = \Delta_t \left\{ \Omega_{1,t} - \frac{\beta\theta \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})\Pi_{t+1}^{\epsilon}\Omega_{1,t+1}]}{U_C^w(C_t^w, N_t)} \right\} - Q(K_{t-1}, K_t, \xi_t)[K_t - (1-\delta)\xi_t K_{t-1}]\\
&\qquad - W(C_t^w, N_t)N_t + D_t - R_{t-1}D_{t-1} - C_t^b - C_t^e,\\
\lambda_t^e: &\quad 0 \le \mathbb{E}_t(Q(K_t, K_{t+1}, \xi_{t+1})\xi_{t+1})K_t - \mathbb{E}_t(C_{t+1}^b + L_{t+1} - D_{t+1} + R_t D_t),\\
\lambda_t^Y: &\quad 0 = \frac{A_t}{\Delta_t} F(\xi_t K_{t-1}, N_t) - \sum_{i \in \mathcal{I}} C_t^i - I(K_{t-1}, K_t, \xi_t),\\
\lambda_t^\Omega: &\quad 0 = \frac{\epsilon - 1}{\epsilon} \widetilde{P}(\Pi_t) \frac{A_t}{\Delta_t} F(\xi_t K_{t-1}, N_t) - \Omega_{1,t} + \frac{\beta\theta \mathbb{E}_t\left[U_C^w(C_{t+1}^w, N_{t+1})\Pi_{t+1}^{\epsilon-1}\frac{\widetilde{P}(\Pi_t)}{\widetilde{P}(\Pi_{t+1})}\Omega_{1,t+1}\right]}{U_C^w(C_t^w, N_t)},\\
\lambda_t^\Delta: &\quad 0 = \theta \Pi_t^\epsilon \Delta_{t-1} + (1-\theta)(\widetilde{P}(\Pi_t))^{-\epsilon} - \Delta_t,\\
\lambda_t^R: &\quad 0 \le R_t \mathbb{E}_t(\Pi_{t+1}) - \underline{R},
\end{aligned}
$$

with $R_t = R(U_C^w(C_t^w, N_t), \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})])$. The FOCs are

$$
\begin{aligned}
C_t^b: &\quad 0 = \omega_b U_{C,t}^b - \lambda_t^Y - \lambda_t^C + \lambda_t^L U_{CC,t}^b L_t - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L \beta_b (U_{CC,t}^b R_t^L L_{t-1} + U_{C,t}^b) + \lambda_{t-1}^e],\\
C_t^e: &\quad 0 = \omega_e U_{C,t}^e - \lambda_t^Y - \lambda_t^C,\\
C_t^w: &\quad 0 = \omega_w U_{C,t}^w - \lambda_t^Y - \lambda_t^C W_{C,t} N_t - [\lambda_t^L \beta_b \mathbb{E}_t(U_{C,t+1}^b) + \beta \mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e] R_{1,t} U_{CC,t}^w D_t\\
&\qquad - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L \beta_b \mathbb{E}_{t-1}(U_{C,t}^b) + \beta \mathbb{E}_{t-1}(\lambda_t^C) + \lambda_{t-1}^e] R_{2,t-1} U_{CC,t}^w D_{t-1} + \Bigg[(\lambda_t^C \Delta_t - \lambda_t^\Omega)\Omega_{1,t}\\
&\qquad - \left(\lambda_t^C P_t^w \Delta_t - \lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t\right) Y_t \Bigg] \frac{U_{CC,t}^w}{U_{C,t}^w} - \mathbf{1}_{\mathbb{N}}(t) \left(\lambda_{t-1}^C \Delta_{t-1}\Pi_t - \lambda_{t-1}^\Omega \frac{\widetilde{P}_{t-1}}{\widetilde{P}_t} \right)\\
&\qquad \times \theta \Pi_t^{\epsilon-1}\Omega_{1,t} \frac{U_{CC,t}^w}{U_{C,t-1}^w} + \left[\lambda_t^R R_{1,t} \mathbb{E}_t(\Pi_{t+1}) + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda_{t-1}^R R_{2,t-1} \mathbb{E}_{t-1}(\Pi_t) \right] U_{CC,t}^w,\\
D_t: &\quad 0 \ge -\lambda_t^b + \lambda_t^C - [\lambda_t^L \beta_b \mathbb{E}_t(U_{C,t+1}^b) + \beta \mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e] R_t + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda_{t-1}^L \beta_b U_{C,t}^b + \lambda_{t-1}^e),\\
K_t: &\quad 0 = -\lambda_t^C \{Q_{2,t}[K_t - (1-\delta)\xi_t K_{t-1}] + Q_t\} + \lambda_t^e \mathbb{E}_t[(Q_{1,t+1}K_t + Q_{t+1})\xi_{t+1}] - \lambda_t^Y I_{2,t}\\
&\qquad + \beta \mathbb{E}_t \Bigg[\left(\lambda_{t+1}^\Omega \frac{\epsilon-1}{\epsilon} \widetilde{P}_{t+1} + \lambda_{t+1}^Y \right) \frac{A_{t+1}}{\Delta_{t+1}} F_{K,t+1}\xi_{t+1} - \lambda_{t+1}^Y I_{1,t+1}\\
&\qquad + \lambda_{t+1}^C \{Q_{t+1}(1-\delta)\xi_{t+1} - Q_{1,t+1}[K_{t+1} - (1-\delta)\xi_{t+1}K_t]\} \Bigg] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda_{t-1}^e Q_{2,t}\xi_t K_{t-1},
\end{aligned}
$$

$$
\begin{aligned}
L_t: \quad & 0 = \lambda_t^b + \lambda_t^L U_{C,t}^b - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}(\lambda_{t-1}^L \beta_b U_{C,t}^b + \lambda_{t-1}^e), \\
N_t: \quad & 0 = \omega_w U_{N,t}^w + \left(\lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t + \lambda_t^Y\right)\frac{A_t}{\Delta_t}F_{N,t} - [\lambda_t^L \beta_b \mathbb{E}_t(U_{C,t+1}^b) + \beta\mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e] \\
& \quad \times R_{1,t}U_{CN,t}^w D_t - \lambda_t^C(W_{N,t}N_t + W_t) - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}[\lambda_{t-1}^L\beta_b\mathbb{E}_{t-1}(U_{C,t}^b) + \beta\mathbb{E}_{t-1}(\lambda_t^C) + \lambda_{t-1}^e] \\
& \quad \times R_{2,t-1}U_{CN,t}^w D_{t-1} + \left[(\lambda_t^C\Delta_t - \lambda_t^\Omega)\Omega_{1,t} - \left(\lambda_t^C P_t^w \Delta_t - \lambda_t^\Omega \frac{\epsilon-1}{\epsilon}\widetilde{P}_t\right)Y_t\right]\frac{U_{CN,t}^w}{U_{C,t}^w} \\
& \quad - \mathbf{1}_{\mathbb{N}}(t)\left(\lambda_{t-1}^C\Delta_{t-1}\Pi_t - \lambda_{t-1}^\Omega\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\right)\theta\Pi_t^{\epsilon-1}\Omega_{1,t}\frac{U_{CN,t}^w}{U_{C,t-1}^w} \\
& \quad + \left[\lambda_t^R R_{1,t}\mathbb{E}_t(\Pi_{t+1}) + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda_{t-1}^R R_{2,t-1}\mathbb{E}_{t-1}(\Pi_t)\right]U_{CN,t}^w, \\
\Omega_{1,t}: \quad & 0 = \lambda_t^C\Delta_t - \lambda_t^\Omega - \mathbf{1}_{\mathbb{N}}(t)\left(\lambda_{t-1}^C\Delta_{t-1}\Pi_t - \lambda_{t-1}^\Omega\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\right)\theta\Pi_t^{\epsilon-1}\frac{U_{C,t}^w}{U_{C,t-1}^w}, \\
\Delta_t: \quad & 0 = \left(\lambda_t^C P_t^w\Delta_t - \lambda_t^\Omega\frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda_t^Y\right)\frac{Y_t}{\Delta_t} - \lambda_t^\Delta + \beta\theta\mathbb{E}_t(\lambda_{t+1}^\Delta\Pi_{t+1}^\epsilon), \\
\Pi_t: \quad & 0 = \lambda_t^\Omega\widetilde{P}'(\Pi_t)\left[\frac{\epsilon-1}{\epsilon}Y_t + \frac{\beta\theta\mathbb{E}_t\left(U_{C,t+1}^w\Pi_{t+1}^{\epsilon-1}\frac{\Omega_{1,t+1}}{\widetilde{P}_{t+1}}\right)}{U_{C,t}^w}\right] + \lambda_t^\Delta\epsilon\left[\theta\Pi_t^{\epsilon-1}\Delta_{t-1} - (1-\theta)\frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t^{\epsilon+1}}\right] \\
& \quad - \mathbf{1}_{\mathbb{N}}(t)\theta\Pi_t^{\epsilon-1}\Omega_{1,t}\frac{U_{C,t}^w}{U_{C,t-1}^w}\left[\lambda_{t-1}^C\Delta_{t-1}\epsilon - \lambda_{t-1}^\Omega\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\left(\frac{\epsilon-1}{\Pi_t} - \frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t}\right)\right] \\
& \quad + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda_{t-1}^R R_{t-1}.
\end{aligned}
$$

The complementary slackness conditions are

$$
\begin{aligned}
0 &= \lambda_t^b(L_t - D_t), \qquad \lambda_t^b \geq 0, \\
0 &= \lambda_t^L[U_{C,t}^b L_t - \beta_b\mathbb{E}_t(U_{C,t+1}^b B_{t+1})], \qquad D_t\lambda_t^L \geq 0, \\
0 &= \lambda_t^e[\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t - \mathbb{E}_t(B_{t+1})], \qquad \lambda_t^e \geq 0, \\
0 &= \lambda_t^R[R_t\mathbb{E}_t(\Pi_{t+1}) - \underline{R}], \qquad \lambda_t^R \geq 0.
\end{aligned}
$$

The risk-sharing and steady-state properties follow from the proof of proposition 3.4 after setting $\lambda_{2,t}^L = 0$, $\kappa_t = 0$, and $m_t = 1$. The short-run inflation behavior is represented by the FOC for Π_t. In the steady state, the FOC for Π is equivalent to

$$
\lambda^R = \frac{\Pi - 1}{\Pi}\frac{\beta\theta\Pi^{\epsilon-1}}{1-\beta\theta\Pi^\epsilon}\frac{(\epsilon-1)\lambda^C + \epsilon\lambda^Y}{1-\theta\Pi^{\epsilon-1}}\beta Y.
$$

Moreover,

$$(\epsilon-1)\lambda^C+\epsilon\lambda^Y=\frac{\epsilon\omega_e U_C^e[\frac{v''(N)}{u'(C^w)}N+W]+\omega_w v'(N)}{\frac{v''(N)}{u'(C^w)}N+W+\frac{1}{\epsilon}\frac{A}{\Delta}F_N}>0.$$

Therefore, $\mathrm{sgn}(\lambda^R)=\mathrm{sgn}(\Pi-1)$. The complementary slackness conditions postulate that $\Pi=\beta\underline{R}$ if $\lambda^R>0$. Hence, if $\underline{R}\leq\frac{1}{\beta}$, then $\Pi=1$; if $\underline{R}>\frac{1}{\beta}$, then $\Pi=\beta\underline{R}$.

Case 2

The planning problem is

$$\max_{\{(C_t^b,C_t^e,C_t^w,D_t,K_t,L_t,N_t,\Omega_{1,t},\Delta_t,\Pi_t)\}}\mathbb{E}_0\left(\sum_{t=0}^{\infty}\beta^t\sum_{i\in\mathcal{I}}\omega_i U_t^i\right)$$

subject to

$$\begin{aligned}
\lambda_t^b:&\quad 0\leq L_t-D_t,\\
\lambda_t^L:&\quad 0=\beta_b\mathbb{E}_t[U_C^b(C_{t+1}^b)(C_{t+1}^b+L_{t+1}-D_{t+1})]-U_C^b(C_t^b)(L_t-D_t),\\
\lambda_t^D:&\quad 0\leq U_C^b(C_t^b)-\beta_b R_t\mathbb{E}_t(U_C^b(C_{t+1}^b)),\\
\lambda_t^C:&\quad 0=\Delta_t\left\{\Omega_{1,t}-\frac{\beta\theta\mathbb{E}_t[U_C^w(C_{t+1}^w,N_{t+1})\Pi_{t+1}^{\epsilon}\Omega_{1,t+1}]}{U_C^w(C_t^w,N_t)}\right\}-Q(K_{t-1},K_t,\xi_t)[K_t-(1-\delta)\xi_t K_{t-1}]\\
&\qquad -W(C_t^w,N_t)N_t+D_t-R_{t-1}D_{t-1}-C_t^b-C_t^e,\\
\lambda_t^e:&\quad 0\leq\mathbb{E}_t(Q(K_t,K_{t+1},\xi_{t+1})\xi_{t+1})K_t-\mathbb{E}_t(C_{t+1}^b+L_{t+1}-D_{t+1}+R_tD_t),\\
\lambda_t^K:&\quad 0=\beta_e\mathbb{E}_t\Bigg\{U_C^e(C_{t+1}^e)\Bigg[\alpha\Delta_{t+1}\left\{\Omega_{1,t+1}-\frac{\beta\theta\mathbb{E}_{t+1}[U_C^w(C_{t+2}^w,N_{t+2})\Pi_{t+2}^{\epsilon}\Omega_{1,t+2}]}{U_C^w(C_{t+1}^w,N_{t+1})}\right\}\\
&\qquad +Q(K_t,K_{t+1},\xi_{t+1})(1-\delta)\xi_{t+1}K_t-C_{t+1}^b-L_{t+1}+D_{t+1}-R_tD_t\Bigg]\Bigg\}\\
&\qquad -U_C^e(C_t^e)(Q(K_{t-1},K_t,\xi_t)K_t-L_t),\\
\lambda_t^B:&\quad 0\leq U_C^e(C_t^e)L_t-\beta_e\mathbb{E}_t[U_C^e(C_{t+1}^e)(C_{t+1}^b+L_{t+1}-D_{t+1}+R_tD_t)],\\
\lambda_t^Y:&\quad 0=\frac{A_t}{\Delta_t}F(\xi_tK_{t-1},N_t)-\sum_{i\in\mathcal{I}}C_t^i-I(K_{t-1},K_t,\xi_t),\\
\lambda_t^{\Omega}:&\quad 0=\frac{\epsilon-1}{\epsilon}\widetilde{P}(\Pi_t)\frac{A_t}{\Delta_t}F(\xi_tK_{t-1},N_t)-\Omega_{1,t}+\frac{\beta\theta\mathbb{E}_t\left[U_C^w(C_{t+1}^w,N_{t+1})\Pi_{t+1}^{\epsilon-1}\frac{\widetilde{P}(\Pi_t)}{\widetilde{P}(\Pi_{t+1})}\Omega_{1,t+1}\right]}{U_C^w(C_t^w,N_t)},\\
\lambda_t^{\Delta}:&\quad 0=\theta\Pi_t^{\epsilon}\Delta_{t-1}+(1-\theta)(\widetilde{P}(\Pi_t))^{-\epsilon}-\Delta_t,
\end{aligned}$$

$$\lambda_t^R: \quad 0 \leq R_t \mathbb{E}_t(\Pi_{t+1}) - \underline{R},$$

with $R_t = R(U_C^w(C_t^w, N_t), \mathbb{E}_t[U_C^w(C_{t+1}^w, N_{t+1})])$. Define $\widetilde{\lambda}_t^C \equiv \lambda_t^C + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda_{t-1}^K \beta_e U_{C,t}^e \alpha$. The FOCs are

$$\begin{aligned}
C_t^b: \quad & 0 = \omega_b U_{C,t}^b - \lambda_t^Y - \lambda_t^C - [\lambda_t^L(L_t - D_t) - \lambda_t^D] U_{CC,t}^b + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \{[\lambda_{t-1}^L(C_t^b + L_t - D_t) \\
& - \lambda_{t-1}^D R_{t-1}] \beta_b U_{CC,t}^b + \lambda_{t-1}^L \beta_b U_{C,t}^b - \lambda_{t-1}^e - (\lambda_{t-1}^K + \lambda_{t-1}^B) \beta_e U_{C,t}^e\}, \\
C_t^e: \quad & 0 = \omega_e U_{C,t}^e - \lambda_t^Y - \lambda_t^C - [\lambda_t^K(Q_t K_t - L_t) - \lambda_t^B L_t] U_{CC,t}^e \\
& + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} [\lambda_{t-1}^K R_t^K Q_{t-1} K_{t-1} - (\lambda_{t-1}^K + \lambda_{t-1}^B) R_t^L L_{t-1}] \beta_e U_{CC,t}^e, \\
C_t^w: \quad & 0 = \omega_w U_{C,t}^w - \lambda_t^Y - \lambda_t^C W_{C,t} N_t - \{\lambda_t^D \beta_b \mathbb{E}_t(U_{C,t+1}^b) \\
& + [\beta \mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e + (\lambda_t^K + \lambda_t^B) \beta_e \mathbb{E}_t(U_{C,t+1}^e)] D_t\} R_{1,t} U_{CC,t}^w - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \{\lambda_{t-1}^D \beta_b \mathbb{E}_{t-1}(U_{C,t}^b) \\
& + [\beta \mathbb{E}_{t-1}(\lambda_t^C) + \lambda_{t-1}^e + (\lambda_{t-1}^K + \lambda_{t-1}^B) \beta_e \mathbb{E}_{t-1}(U_{C,t}^e)] D_{t-1}\} R_{2,t-1} U_{CC,t}^w \\
& + \left[(\widetilde{\lambda}_t^C \Delta_t - \lambda_t^\Omega) \Omega_{1,t} - \left(\widetilde{\lambda}_t^C P_t^w \Delta_t - \lambda_t^\Omega \frac{\epsilon - 1}{\epsilon} \widetilde{P}_t\right) Y_t\right] \frac{U_{CC,t}^w}{U_{C,t}^w} - \mathbf{1}_{\mathbb{N}}(t) \Big(\widetilde{\lambda}_{t-1}^C \Delta_{t-1} \Pi_t \\
& - \lambda_{t-1}^\Omega \frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\Big) \theta \Pi_t^{\epsilon - 1} \Omega_{1,t} \frac{U_{CC,t}^w}{U_{C,t-1}^w} + \left[\lambda_t^R R_{1,t} \mathbb{E}_t(\Pi_{t+1}) + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} \lambda_{t-1}^R R_{2,t-1} \mathbb{E}_{t-1}(\Pi_t)\right] U_{CC,t}^w, \\
D_t: \quad & 0 = -\lambda_t^b + \lambda_t^L U_{C,t}^b + \lambda_t^C - [\beta \mathbb{E}_t(\lambda_{t+1}^C) + \lambda_t^e + (\lambda_t^K + \lambda_t^B) \beta_e \mathbb{E}_t(U_{C,t+1}^e)] R_t \\
& + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} [-\lambda_{t-1}^L \beta_b U_{C,t}^b + \lambda_{t-1}^e + (\lambda_{t-1}^K + \lambda_{t-1}^B) \beta_e U_{C,t}^e], \\
K_t: \quad & 0 = -\lambda_t^C \{Q_{2,t}[K_t - (1 - \delta) \xi_t K_{t-1}] + Q_t\} + \lambda_t^e \mathbb{E}_t[(Q_{1,t+1} K_t + Q_{t+1}) \xi_{t+1}] - \lambda_t^Y I_{2,t} + \lambda_t^K \{\beta_e \\
& \times \mathbb{E}_t[U_{C,t+1}^e (Q_{1,t+1} K_t + Q_{t+1})(1 - \delta) \xi_{t+1}] - U_{C,t}^e (Q_{2,t} K_t + Q_t)\} + \beta \mathbb{E}_t \Bigg[\Bigg(\lambda_{t+1}^\Omega \frac{\epsilon - 1}{\epsilon} \widetilde{P}_{t+1} \\
& + \lambda_{t+1}^Y\Bigg) \frac{A_{t+1}}{\Delta_{t+1}} F_{K,t+1} \xi_{t+1} - \lambda_{t+1}^K U_{C,t+1}^e Q_{1,t+1} K_{t+1} - \lambda_{t+1}^Y I_{1,t+1} + \lambda_{t+1}^C \{Q_{t+1}(1 - \delta) \xi_{t+1} \\
& - Q_{1,t+1}[K_{t+1} - (1 - \delta) \xi_{t+1} K_t]\}\Bigg] + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} [\lambda_{t-1}^e + \lambda_{t-1}^K \beta_e U_{C,t}^e (1 - \delta)] Q_{2,t} \xi_t K_{t-1}, \\
L_t: \quad & 0 = \lambda_t^b - \lambda_t^L U_{C,t}^b + (\lambda_t^K + \lambda_t^B) U_{C,t}^e + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta} [\lambda_{t-1}^L \beta_b U_{C,t}^b - \lambda_{t-1}^e - (\lambda_{t-1}^K + \lambda_{t-1}^B) \beta_e U_{C,t}^e],
\end{aligned}$$

$$
\begin{aligned}
N_t: \quad & 0 = \omega_w U^w_{N,t} + \left(\lambda^\Omega_t \frac{\epsilon-1}{\epsilon}\widetilde{P}_t + \lambda^Y_t\right)\frac{A_t}{\Delta_t}F_{N,t} - \lambda^C_t(W_{N,t}N_t + W_t) - \{\lambda^D_t\beta_b\mathbb{E}_t(U^b_{C,t+1}) \\
& + [\beta\mathbb{E}_t(\lambda^C_{t+1}) + \lambda^e_t + (\lambda^K_t + \lambda^B_t)\beta_e\mathbb{E}_t(U^e_{C,t+1})]D_t\}R_{1,t}U^w_{CN,t} - \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\{\lambda^D_{t-1}\beta_b\mathbb{E}_{t-1}(U^b_{C,t}) \\
& + [\beta\mathbb{E}_{t-1}(\lambda^C_t) + \lambda^e_{t-1} + (\lambda^K_{t-1} + \lambda^B_{t-1})\beta_e\mathbb{E}_{t-1}(U^e_{C,t})]D_{t-1}\}R_{2,t-1}U^w_{CN,t} \\
& + \left[(\widetilde{\lambda}^C_t\Delta_t - \lambda^\Omega_t)\Omega_{1,t} - \left(\widetilde{\lambda}^C_t P^w_t\Delta_t - \lambda^\Omega_t\frac{\epsilon-1}{\epsilon}\widetilde{P}_t\right)Y_t\right]\frac{U^w_{CN,t}}{U^w_{C,t}} - \mathbf{1}_{\mathbb{N}}(t)\Big(\widetilde{\lambda}^C_{t-1}\Delta_{t-1}\Pi_t \\
& - \lambda^\Omega_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\Big)\theta\Pi^{\epsilon-1}_t\Omega_{1,t}\frac{U^w_{CN,t}}{U^w_{C,t-1}} + \left[\lambda^R_t R_{1,t}\mathbb{E}_t(\Pi_{t+1}) + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda^R_{t-1}R_{2,t-1}\mathbb{E}_{t-1}(\Pi_t)\right]U^w_{CN,t}, \\
\Omega_{1,t}: \quad & 0 = \widetilde{\lambda}^C_t\Delta_t - \lambda^\Omega_t - \mathbf{1}_{\mathbb{N}}(t)\left(\widetilde{\lambda}^C_{t-1}\Delta_{t-1}\Pi_t - \lambda^\Omega_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\right)\theta\Pi^{\epsilon-1}_t\frac{U^w_{C,t}}{U^w_{C,t-1}}, \\
\Delta_t: \quad & 0 = \left(\widetilde{\lambda}^C_t P^w_t\Delta_t - \lambda^\Omega_t\frac{\epsilon-1}{\epsilon}\widetilde{P}_t - \lambda^Y_t\right)\frac{Y_t}{\Delta_t} - \lambda^\Delta_t + \beta\theta\mathbb{E}_t(\lambda^\Delta_{t+1}\Pi^\epsilon_{t+1}), \\
\Pi_t: \quad & 0 = \lambda^\Omega_t\widetilde{P}'(\Pi_t)\left[\frac{\epsilon-1}{\epsilon}Y_t + \frac{\beta\theta\mathbb{E}_t\left(U^w_{C,t+1}\Pi^{\epsilon-1}_{t+1}\frac{\Omega_{1,t+1}}{\widetilde{P}_{t+1}}\right)}{U^w_{C,t}}\right] + \lambda^\Delta_t\epsilon\left[\theta\Pi^{\epsilon-1}_t\Delta_{t-1} - (1-\theta)\frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}^{\epsilon+1}_t}\right] \\
& - \mathbf{1}_{\mathbb{N}}(t)\theta\Pi^{\epsilon-1}_t\Omega_{1,t}\frac{U^w_{C,t}}{U^w_{C,t-1}}\left[\widetilde{\lambda}^C_{t-1}\Delta_{t-1}\epsilon - \lambda^\Omega_{t-1}\frac{\widetilde{P}_{t-1}}{\widetilde{P}_t}\left(\frac{\epsilon-1}{\Pi_t} - \frac{\widetilde{P}'(\Pi_t)}{\widetilde{P}_t}\right)\right] \\
& + \frac{\mathbf{1}_{\mathbb{N}}(t)}{\beta}\lambda^R_{t-1}R_{t-1}.
\end{aligned}
$$

The complementary slackness conditions are

$$
\begin{aligned}
& 0 = \lambda^b_t(L_t - D_t), \qquad \lambda^b_t \geq 0, \\
& 0 = \lambda^D_t[U^b_{C,t} - \beta_b R_t\mathbb{E}_t(U^b_{C,t+1})], \qquad \lambda^D_t \geq 0, \\
& 0 = \lambda^e_t[\mathbb{E}_t(Q_{t+1}\xi_{t+1})K_t - \mathbb{E}_t(B_{t+1})], \qquad \lambda^e_t \geq 0, \\
& 0 = \lambda^B_t[U^e_{C,t}L_t - \beta_e\mathbb{E}_t(U^e_{C,t+1}B_{t+1})], \qquad \lambda^B_t \geq 0, \\
& 0 = \lambda^R_t[R_t\mathbb{E}_t(\Pi_{t+1}) - \underline{R}], \qquad \lambda^R_t \geq 0.
\end{aligned}
$$

The risk-sharing and steady-state properties follow from comparing the optimality conditions to those in the proof of proposition 3.3. The special case of approximate full insurance fails for the same reasons as in the proof of proposition 3.4. The short-run inflation behavior is represented by the FOC for Π_t. In the steady state, the FOC for Π is equivalent to

$$
\lambda^R = \frac{\Pi - 1}{\Pi}\frac{\beta\theta\Pi^{\epsilon-1}}{1-\beta\theta\Pi^\epsilon}\frac{(\epsilon-1)\widetilde{\lambda}^C + \epsilon\lambda^Y}{1-\theta\Pi^{\epsilon-1}}\beta Y.
$$

Moreover,

$$
\begin{aligned}
&(\epsilon-1)\widetilde{\lambda}^C+\epsilon\lambda^Y \\
&\quad=\frac{\left\{\omega_e U^e_C+\lambda^K\left[(R-1)(QK-L)U^e_{CC}+\beta_e R U^e_C\alpha\frac{\epsilon-1}{\epsilon}\right]\right\}\epsilon\left(\frac{v''(N)}{u'(C^w)}N+W\right)+\omega_w v'(N)}{\frac{v''(N)}{u'(C^w)}N+W+\frac{1}{\epsilon}\frac{A}{\Delta}F_N}>0.
\end{aligned}
$$

If the relaxed collateral constraint is slack so that $\lambda^K=\lambda^e=0$, the inequality follows immediately; otherwise, it can be verified numerically. Therefore, $\text{sgn}(\lambda^R)=\text{sgn}(\Pi-1)$. The complementary slackness conditions postulate that $\Pi=\beta\underline{R}$ if $\lambda^R>0$. Hence, if $\underline{R}\leq\frac{1}{\beta}$, then $\Pi=1$; if $\underline{R}>\frac{1}{\beta}$, then $\Pi=\beta\underline{R}$. ■

Bibliography

Aguiar, M., S. Chatterjee, H. Cole, and Z. Stangebye, "Quantitative Models of Sovereign Debt Crises," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier, 2016, pp. 1697–1755.

Aguiar, Mark and Gita Gopinath, "Defaultable debt, interest rates and the current account," *Journal of International Economics*, 2006, *69* (1), 64–83.

Akinci, Ozge and Albert Queralto, "Credit Spreads, Financial Crises, and Macroprudential Policy," *American Economic Journal: Macroeconomics*, forthcoming.

Allen, Franklin and Douglas Gale, "Financial Intermediaries and Markets," *Econometrica*, 2004, *72* (4), 1023–1061.

Andrés, Javier, Óscar Arce, and Carlos Thomas, "Banking Competition, Collateral Constraints, and Optimal Monetary Policy," *Journal of Money, Credit and Banking*, 2013, *45* (s2), 87–125.

Arellano, Cristina, "Default Risk and Income Fluctuations in Emerging Economies," *American Economic Review*, June 2008, *98* (3), 690–712.

Ari, Anil, "Gambling traps," Working Paper Series 2217, European Central Bank December 2018.

Armenter, Roc, "The Perils of Nominal Targets," *The Review of Economic Studies*, 01 2018, *85* (1), 50–86.

Aruoba, S. Borağan, Pablo Cuba-Borda, and Frank Schorfheide, "Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries," *The Review of Economic Studies*, 01 2018, *85* (1), 87–118.

Attanasio, Orazio, Peter Levell, Hamish Low, and Virginia Sánchez-Marcos, "Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply," *Econometrica*, 2018, *86* (6), 2049–2082.

Bean, Charles R., Matthias Paustian, Adrian Penalver, and Tim Taylor, "Monetary Policy After the Fall," Macroeconomic Challenges: The Decade Ahead symposium, Federal Reserve Bank of Kansas City 2010.

Begenau, Juliane, Saki Bigio, Jeremy Majerovitz, and Matias Vieyra, "A Q-Theory of Banks," Working Paper 27935, National Bureau of Economic Research October 2020.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe, "The Perils of Taylor Rules," *Journal of Economic Theory*, 2001, *96* (1), 40–69.

Benigno, Gianluca, Huigang Chen, Christopher Otrok, Alessandro Rebucci, and Eric R. Young, "Financial crises and macro-prudential policies," *Journal of International Economics*, 2013, *89* (2), 453–470.

_ , _ , _ , _ , and _ , "Optimal capital controls and real exchange rate policies: A pecuniary externality perspective," *Journal of Monetary Economics*, 2016, *84*, 147 – 165.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist, "The Financial Accelerator in a Quantitative Business Cycle Framework," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1, Elsevier Science B.V., 1999, pp. 1341–1393.

Bianchi, Javier, "Overborrowing and Systemic Externalities in the Business Cycle," *American Economic Review*, December 2011, *101* (7), 3400–3426.

_ , "Efficient Bailouts?," *American Economic Review*, December 2016, *106* (12), 3607–59.

_ and Enrique G. Mendoza, "Optimal Time-Consistent Macroprudential Policy," *Journal of Political Economy*, 2018, *126* (2), 588–634.

Bocola, Luigi, "The Pass-Through of Sovereign Risk," *Journal of Political Economy*, 2016, *124* (4), 879–926.

Boissay, Frédéric, Fabrice Collard, and Frank Smets, "Booms and Banking Crises," *Journal of Political Economy*, 2016, *124* (2), 489–538.

Braun, R. Anton and Lena Mareen Körber, "New Keynesian dynamics in a low interest rate environment," *Journal of Economic Dynamics and Control*, 2011, *35* (12), 2213–2227. Frontiers in Structural Macroeconomic Modeling.

Calvo, Guillermo A., "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 1983, *12* (3), 383–398.

Chari, V. V. and Patrick J. Kehoe, "Sustainable Plans," *Journal of Political Economy*, 1990, *98* (4), 783–802.

_ and _ , "Bailouts, Time Inconsistency, and Optimal Regulation: A Macroeconomic View," *American Economic Review*, September 2016, *106* (9), 2458–93.

Chatterjee, Satyajit and Burcu Eyigungor, "Maturity, Indebtedness, and Default Risk," *American Economic Review*, May 2012, *102* (6), 2674–99.

Coibion, Olivier, Yuriy Gorodnichenko, and Johannes Wieland, "The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?," *The Review of Economic Studies*, 03 2012, *79* (4), 1371–1406.

Collard, Fabrice, Harris Dellas, Behzad Diba, and Olivier Loisel, "Optimal Monetary and Prudential Policies," *American Economic Journal: Macroeconomics*, 2017, *9* (1), 40–87.

Cúrdia, Vasco and Michael Woodford, "Credit Frictions and Optimal Monetary Policy," *Journal of Monetary Economics*, 2016, *84*, 30–65.

Dávila, Eduardo and Anton Korinek, "Pecuniary Externalities in Economies with Financial Frictions," *The Review of Economic Studies*, 2018, *85* (1), 352–395.

De Paoli, Bianca and Matthias Paustian, "Coordinating Monetary and Macroprudential Policies," *Journal of Money, Credit and Banking*, 2017, *49* (2-3), 319–349.

Di Tella, Sebastian, "Optimal Regulation of Financial Intermediaries," *American Economic Review*, January 2019, *109* (1), 271–313.

Diamond, Douglas W. and Philip H. Dybvig, "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 1983, *91* (3), 401–419.

Eaton, Jonathan and Mark Gersovitz, "Debt with Potential Repudiation: Theoretical and Empirical Analysis," *The Review of Economic Studies*, 1981, *48* (2), 289–309.

Erosa, Andrés, Luisa Fuster, and Gueorgui Kambourov, "Towards a Micro-Founded Theory of Aggregate Labour Supply," *The Review of Economic Studies*, 03 2016, *83* (3), 1001–1039.

Farhi, Emmanuel and Iván Werning, "A Theory of Macroprudential Policies in the Presence of Nominal Rigidities," *Econometrica*, 2016, *84* (5), 1645–1704.

_ **and Jean Tirole**, "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts," *American Economic Review*, February 2012, *102* (1), 60–93.

_ **, Mikhail Golosov, and Aleh Tsyvinski**, "A Theory of Liquidity and Regulation of Financial Intermediation," *The Review of Economic Studies*, 07 2009, *76* (3), 973–992.

Fernández-Villaverde, Jesús, Grey Gordon, Pablo Guerrón-Quintana, and Juan F. Rubio-Ramírez, "Nonlinear adventures at the zero lower bound," *Journal of Economic Dynamics and Control*, 2015, *57*, 182–204.

Ferrero, Andrea, Richard Harrison, and Benjamin Nelson, "House Price Dynamics, Optimal LTV Limits and the Liquidity Trap," Discussion Paper DP13400, Centre for Economic

Policy Research December 2018.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications–Second Edition*, Princeton University Press, 2015.

Gertler, Mark and Nobuhiro Kiyotaki, "Financial Intermediation and Credit Policy in Business Cycle Analysis," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Supplement C, Elsevier, 2010, pp. 547–599.

_ and Peter Karadi, "A model of unconventional monetary policy," *Journal of Monetary Economics*, 2011, *58* (1), 17–34.

_ , Nobuhiro Kiyotaki, and Albert Queralto, "Financial Crises, Bank Risk Exposure and Government Financial Policy," *Journal of Monetary Economics*, 2012, *59* (S), 17–34.

_ , _ , and Andrea Prestipino, "A Macroeconomic Model with Financial Panics," *The Review of Economic Studies*, 05 2020, *87* (1), 240–288.

_ , _ , and _ , "Credit booms, financial crises, and macroprudential policy," *Review of Economic Dynamics*, 2020, *37*, S8–S33. The twenty-fifth anniversary of "Frontiers of Business Cycle Research".

Guerrieri, Luca and Matteo Iacoviello, "OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily," *Journal of Monetary Economics*, 2015, *70*, 22–38.

Iacoviello, Matteo, "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle," *The American Economic Review*, 2005, *95* (3), 739–764.

_ , "Financial Business Cycles," *Review of Economic Dynamics*, 2015, *18* (1), 140–164.

Jeanne, Olivier and Anton Korinek, "Managing credit booms and busts: A Pigouvian taxation approach," *Journal of Monetary Economics*, 2019, *107*, 2–17.

Judd, Kenneth L., *Numerical Methods in Economics*, The MIT Press, 1998.

Kehoe, Patrick J. and Fabrizio Perri, "International Business Cycles with Endogenous Incomplete Markets," *Econometrica*, 2002, *70* (3), 907–928.

Kiyotaki, Nobuhiro and John Moore, "Credit Cycles," *The Journal of Political Economy*, 1997, *105* (2), 211–248.

Klein, Paul, Per Krusell, and José-Víctor Ríos-Rull, "Time-Consistent Public Policy," *The Review of Economic Studies*, 07 2008, *75* (3), 789–808.

Korinek, Anton and Alp Simsek, "Liquidity Trap and Excessive Leverage," *American Eco-*

nomic Review, March 2016, *106* (3), 699–738.

Krueger, Dirk and Felix Kubler, "Computing equilibrium in OLG models with stochastic production," *Journal of Economic Dynamics and Control*, 2004, *28* (7), 1411–1436.

Leduc, Sylvain and Jean-Marc Natal, "Monetary and Macroprudential Policies in a Leveraged Economy," *The Economic Journal*, 2018, *128* (609), 797–826.

Leland, Hayne, "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk," Research Program in Finance Working Papers RPF-240, University of California at Berkeley November 1994.

Lorenzoni, Guido, "Inefficient Credit Booms," *The Review of Economic Studies*, 2008, *75* (3), 809–833.

Lucas, Robert E. and Edward C. Prescott, "Investment Under Uncertainty," *Econometrica*, 1971, *39* (5), 659–681.

Malin, Benjamin A., Dirk Krueger, and Felix Kubler, "Solving the multi-country real business cycle model using a Smolyak-collocation method," *Journal of Economic Dynamics and Control*, 2011, *35* (2), 229–239.

Marcet, Albert and Ramon Marimon, "Recursive Contracts," *Econometrica*, 2019, *87* (5), 1589–1631.

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 1989, *57* (5), 995–1026.

Mendoza, Enrique G., "Sudden Stops, Financial Crises, and Leverage," *American Economic Review*, December 2010, *100* (5), 1941–66.

Murphy, Robert G., "Macroeconomic adjustment under alternative lending arrangements," *Journal of Economic Dynamics and Control*, 1991, *15* (1), 103–127.

Na, Seunghoon, Stephanie Schmitt-Grohé, Martín Uribe, and Vivian Yue, "The Twin Ds: Optimal Default and Devaluation," *American Economic Review*, July 2018, *108* (7), 1773–1819.

Perez, Diego J., "Sovereign Debt, Domestic Banks and the Provision of Public Liquidity," 2018. mimeo.

Reinhart, Carmen M. and Kenneth S. Rogoff, *This Time is Different: Eight Centuries of Financial Folly*, Princeton University Press, 2009.

_ and _ , "The Forgotten History of Domestic Debt," *The Economic Journal*, 2011, *121* (552), 319–350.

— **and** — , "From Financial Crash to Debt Crisis," *American Economic Review*, August 2011, *101* (5), 1676–1706.

Schechtman, Jack and Vera L.S. Escudero, "Some results on "an income fluctuation problem"," *Journal of Economic Theory*, 1977, *16* (2), 151–166.

Schmitt-Grohé, Stephanie and Martín Uribe, "Closing small open economy models," *Journal of International Economics*, 2003, *61* (1), 163–185.

— **and** — , "Chapter 13–The Optimal Rate of Inflation," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Elsevier, 2010, pp. 653–722.

— **and** — , "Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment," *Journal of Political Economy*, 2016, *124* (5), 1466–1514.

— **and** — , "Multiple Equilibria in Open Economies with Collateral Constraints," *The Review of Economic Studies*, 2021, *88* (2), 969–1001.

Sosa-Padilla, César, "Sovereign defaults and banking crises," *Journal of Monetary Economics*, 2018, *99*, 88–105.

Van der Ghote, Alejandro, "Interactions and Coordination between Monetary and Macroprudential Policies," *American Economic Journal: Macroeconomics*, January 2021, *13* (1), 1–34.

Woodford, Michael, "Chapter 14–Optimal Monetary Stabilization Policy," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Elsevier, 2010, pp. 723–828.

Yun, Tack, "Nominal price rigidity, money supply endogeneity, and business cycles," *Journal of Monetary Economics*, 1996, *37* (2), 345–370.