ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Optimal Asset Market Operations

Authors	Yu-Ting Chiang, and Piotr Zoch
Working Paper Number	2025-014A
Creation Date	June 2025
Citable Link	https://doi.org/10.20955/wp.2025.014
Suggested Citation	Chiang, Y.-T., Zoch, P., 2025; Optimal Asset Market Operations, Federal Reserve Bank of St. Louis Working Paper 2025-014. URL https://doi.org/10.20955/wp.2025.014

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Optimal Asset Market Operations

Yu-Ting Chiang[†] Piotr Zoch[‡]

June 16, 2025

Abstract

We characterize governments' optimal responses to asset market disturbances across a broad class of models with financial frictions. We show that the Ramsey plan can be achieved by a policy rule targeting a specific relationship between asset returns, regardless of the underlying disturbances. This relationship is determined by asset supply and demand elasticities that can be estimated empirically with standard identification strategies. Absent financial frictions, the optimal policy stabilizes spreads across all assets. However, in the presence of financial frictions, the optimal rule prescribes time-varying spreads to facilitate financial intermediation. We apply our framework to study the optimal design of asset purchase and lending programs, as implied by key empirical estimates of asset supply and demand elasticities.

Keywords: monetary policy, financial frictions, asset demand estimation, sufficient statistics
JEL code: E2, E6, H3, H6

[†]Federal Reserve Bank of St. Louis, contact email: yu-ting.chiang@stls.frb.org.
[‡]University of Warsaw and FAME|GRAPE, contact email: p.zoch@uw.edu.pl
We thank Fernando Alvarez, Adrien Auclert, Markus Brunnermeier, Zhiyu Julie Fu, Chen Lian, Mikhail Golosov, Andreas Schaab, Harald Uhlig, and Mark Wright for valuable discussions, and seminar and conference participants for helpful comments. Mick Dueholm provided excellent research assistance. Zoch acknowledges the support of National Center for Science (2024/55/B/HS4/01765). Views in this paper do not necessarily represent those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

1 Introduction

How should governments respond to disturbances in asset markets? Since the Great Financial Crisis, governments have deployed increasingly sophisticated "asset market operations" beyond traditional monetary policy, including Quantitative Easing, Operation Twist, and the Troubled Asset Relief Program. However, to design these policies, existing macroeconomic models mostly rely on detailed microfoundations of financial frictions and focus on specific shocks to the economy. Policy prescriptions from such analysis inevitably depend on particular structural assumptions, raising questions about their broader applicability. Since deep parameters in these models are difficult to measure and shocks hard to observe, implementing their prescriptions proves challenging in practice. In this paper, we aim to provide a policy framework to overcome these challenges.

We study a broad class of models with financial frictions in which the government can trade assets in response to arbitrary shocks to the economy. We show that the Ramsey plan can be achieved (up to first order) by maintaining a target relationship between asset returns. This relationship depends on agents' asset supply and demand elasticities. These elasticities are sufficient statistics (Chetty, 2009; Kleven, 2021) that capture the invariant features of the economy—preferences, technology, and financial frictions. In equilibrium, asset returns fluctuate to reflect underlying economic disturbances under the target relationship. Crucially, the target relationship links observable asset returns with measurable elasticities and does not depend directly on underlying disturbances. Policymakers can therefore implement optimal policy by trading assets according to the target relationship without observing the specific shocks driving market conditions.

Our framework features a production economy with a general form of financial intermediation. A representative household has preferences over various forms of financial assets, while the financial sector issues (or holds) these financial assets and holds claims on physical capital. Financial frictions are described by a generic mapping between expected returns and the issuance (or holdings) of financial assets, similar to Chiang and Zoch (2023). This "financial block" allows for frictions from agency problems (Bernanke et al., 1999; Gertler and Karadi, 2011), interbank search and bargaining (Afonso and Lagos, 2015; Bianchi and Bigio, 2022), and heterogeneous intermediaries with complex interactions, similar to that in Bellifemine et al. (2022). We use the sequence space approach (Auclert et al., 2023, 2025; McKay and Wolf, 2022; Dávila and Schaab, 2023) to abstract from these complications, and instead focus on analyzing how financial frictions shape the government's optimal policy design in general equilibrium.

We study a one-time unexpected shock that shifts preferences, technology, and financial frictions in arbitrary ways. The government can implement long-run macroprudential policy in the form of taxes on assets, but can only respond to shocks by trading financial assets—performing asset market operations. This captures the realistic constraint that policymakers often cannot rapidly adjust regulatory frameworks in response to unexpected events, unlike the state-contingent taxes in Farhi and Werning (2016) and Dávila and Korinek (2018), and instead rely on direct market interventions. We solve a Ramsey problem to characterize the government's optimal asset market operations in response to shocks around the steady state, providing a theory that complements the literature's focus on macroprudential taxes.

We show that the Ramsey plan can be achieved by a target rule under which the government trades assets so that returns on financial assets equal weighted averages of returns on capital and net worth, with weights determined by agents' asset supply and demand elasticities. The government implements this by committing to supply assets perfectly elastically at target returns, allowing quantities to adjust endogenously while maintaining the optimal relationship between returns. Under the rule, returns on capital and net worth adjust to incorporate information about the underlying disturbances, while returns on financial assets adjust accordingly to address externalities from financial frictions.

Our approach combines the design of policy rules from the New Keynesian tradition (Giannoni and Woodford, 2003) with insights from the sequence space approach, formulating the Ramsey problem as a planner facing an asset supply and demand system, similar to Auclert et al. (2024). In contrast to New Keynesian policy rules, where targets such as output gaps can be difficult to observe, our target rule specifies the relationship between readily observable asset returns with weights determined by measurable elasticities—sufficient statistics that capture underlying preferences, technology, and financial frictions.

We illustrate how financial frictions shape optimal policy through a series of examples. In our framework, the Ramsey planner faces a key trade-off between satiating household asset demand and addressing pecuniary externalities—due to financial frictions, the planner can create welfare gains from influencing asset returns and financial intermediation (Lorenzoni, 2008). Without financial frictions, we show that the planner equalizes returns on all assets, implementing a version of the Friedman rule. However, financial frictions motivate the planner to deviate from this rule: When returns on capital are high, the planner lowers funding costs to maintain an optimal level of intermediation. Hence, the optimal policy rule prescribes time-varying spreads rather than equalizing returns as in a neoclassical benchmark. We show how this mechanism alters the optimal target rule under various frictions, including those arising from costly state verification and asset diversion problems. This emphasis on financial frictions distinguishes our work from studies emphasizing fiscal considerations for the optimal debt structures, including Angeletos (2002), Buera and Nicolini (2004), Aparisi de Lannoy et al. (2022), and Angeletos et al. (2023).

To demonstrate the applicability of our approach, we use empirical methods from the asset demand estimation literature (Gabaix and Koijen, 2021, 2024; Chaudhary et al., 2025) to estimate key elasticities required for the optimal target rule. We focus on the U.S. banking sector and use balance sheets from the Call Reports to construct holdings across three asset categories: liquid assets (e.g., reserves, deposits, Treasury securities), illiquid assets (e.g., corporate bonds, agency securities), and capital (e.g., loans, goodwill). We apply the Granular Instrumental Variables (GIV) approach to isolate idiosyncratic shocks and construct instruments that allow us to identify the banking sector's asset supply and demand elasticities. Our application demonstrates how empirical elasticities from existing methods can directly inform optimal policy through our framework.

Building on our empirical estimates, we compute the optimal target rule that implements the Ramsey plan. The rule prescribes that when expected returns on capital are high, the government should target lower returns on liquid assets while simultaneously targeting higher returns on illiquid assets. We demonstrate how the economy behaves under the rule by analyzing the optimal response to a financial sector net worth shock—a shock that resembles a financial crisis. The optimal policy involves the government purchasing illiquid assets, reducing these assets' returns relative to no intervention, while simultaneously providing liquid lending to reduce the liquid rate. This intervention reduces the funding cost for the financial sector, enabling banks to maintain higher liquid asset issuance and mitigating capital liquidation and the sell-off of illiquid holdings compared to the laissez-faire outcomes.

Finally, our work complements recent literature on monetary transmission through financial markets, including Piazzesi et al. (2019), Vayanos and Vila (2021), Pflueger and Rinaldi (2022), Kekre et al. (2024), and Caballero et al. (2024). While these works provide insights into specific transmission mechanisms, such as nominal rigidities and time-varying risk premia, their structural approach ties policy prescriptions to particular modeling assumptions and shock processes. Though our framework does not incorporate all the mechanisms in these models, our work is a first step towards providing policy guidance that is less dependent on specific structural assumptions.

2 Model

2.1 Setup

Time is discrete, with $t \in \{0, \ldots, \infty\}$. The economy comprises a representative household, a financial sector, and the government. There are three classes of assets: productive assets, financial assets, and the net worth of the financial sector. Productive assets consist of the value of physical capital, which can only be held by the financial sector, yielding returns R_t^A. Financial assets, denoted by set $\mathcal{B}$, are in zero net supply. They can be traded by all agents and generate returns $\boldsymbol{R}_t^B$. The net worth of the financial sector can be held only by the household, and its returns are denoted by R_t^N.

We study how the government can best respond to a one-time unexpected shock at time 0, $\{\boldsymbol{\omega}_t\}$, which is a sequence of aggregate states that moves preferences, technology, and the financial sector's ability to intermediate assets: With probability one, $\boldsymbol{\omega}_t = \bar{\boldsymbol{\omega}}$; with probability zero, $\boldsymbol{\omega}_t$ follows a known path given by

$$\boldsymbol{\omega}_t = \bar{\boldsymbol{\omega}} + \hat{\boldsymbol{\omega}}_t, \quad \forall t \geq 0.$$

Throughout the paper, we use the convention that a generic variable $\boldsymbol{x}_t$ is a column vector and braces $\{x_s\}$ denote sequences from $s = 0, \ldots, \infty$. We use $\mathbf{0}_{\mathcal{B}}$ and $\mathbf{1}_{\mathcal{B}}$ to represent vectors of zeros and ones of dimension $|\mathcal{B}|$.

Household

The representative household derives utility from consumption and holdings of financial assets and equity, $\{c_t, \boldsymbol{b}_t, n_t\}$, where asset holdings are denoted in real values with the consumption goods as numeraire. The household solves the following maximization problem:

$$\max_{\{\boldsymbol{b}_t, n_t, c_t\}_{t=0}^{\infty}} \mathbb{E} \sum_{t=0}^{\infty} \beta^t u\left(\{c_s, \boldsymbol{b}_s, n_s\}_{s\leq t}; \boldsymbol{\omega}_t\right) \tag{1}$$

subject to budget constraints

$$c_t + \mathbf{1}_{\mathcal{B}}^{\intercal}\boldsymbol{b}_t + n_t + \Psi\left(\{\boldsymbol{b}_s, n_s\}_{s\leq t}; \boldsymbol{\omega}_t\right) = w_t + (\boldsymbol{R}_t^B - \boldsymbol{\tau}_t^B)^{\intercal}\boldsymbol{b}_{t-1} + (R_t^N - \tau_t^N)n_{t-1} - T_t.$$

The household supplies one unit of labor inelastically for income w_t and pays lump-sum taxes T_t. The household uses income for consumption and for savings in various assets. Asset holdings are taxed at rate $\boldsymbol{\tau}_t^B$ and τ_t^N, and they can potentially incur a portfolio adjustment cost $\Psi\left(\{\boldsymbol{b}_s, n_s\}_{s\leq t}; \boldsymbol{\omega}_t\right)$. We allow the aggregate state $\boldsymbol{\omega}_t$ to shift the household's preferences and costs of holding assets. These shifts can generate "asset demand shocks" such as a flight to liquidity, which allows us to study policy responses to those scenarios. We use $\mathbb{E}$ to denote expectation before the realization of $\boldsymbol{\omega}_t$.

Although we allow quite arbitrary preferences and technology in holding assets, we maintain an assumption that there exists an asset (or a set of assets) that allows us to "span" the household's intertemporal valuation:

Assumption 1 *There exists a function $\boldsymbol{\ell}(\cdot)$ such that $R_{t+1}^* = \boldsymbol{\ell}(\boldsymbol{R}_{t+1}^B, \boldsymbol{R}_{t+1}^N)$, and*

$$\lambda_t = \beta\lambda_{t+1}R_{t+1}^*,$$

where $\{\lambda_t\}$ are the Lagrange multipliers on the household's budget constraint.

Intuitively, the assumption holds when a "standard" Euler equation exists for the household with respect to some "neutral rate" R_{t+1}^*. This can occur if holding a particular asset (say, B_0) incurs no utility or portfolio adjustment costs. In this case, we have $R_{t+1}^* = \boldsymbol{R}_{t+1}^{B_0}$. Similarly, if holding equity is costless, then $R_{t+1}^* = \boldsymbol{R}_{t+1}^N$. Even when no single asset satisfies a standard Euler equation, the neutral rate may exist for a "synthetic" asset with return $R^*_{t+1} = \boldsymbol{\alpha}^\intercal \boldsymbol{R}^B_{t+1}$ for some portfolio weight $\boldsymbol{\alpha}$, as we demonstrate in Appendix B.3.

Another important class of models that satisfy this assumption trivially is when the household has quasi-linear preferences:

$$u\left(\{c_s, \boldsymbol{b}_s, n_s\}_{s\le t}; \boldsymbol{\omega}_t\right) = c_t + \nu(\boldsymbol{b}_t; \boldsymbol{\omega}_t). \tag{2}$$

This class of models is common in the money search frameworks, and it is a useful simplification in models of financial frictions, as in Bianchi and Bigio (2022). In this case, $\lambda_t \equiv 1$ and $R^*_{t+1} = \beta^{-1}$, which satisfies Assumption 1.

Production

Production of output uses labor h_t and capital k_{t-1} subject to a constant-return-to-scale production function $F(h_t, k_{t-1}; \boldsymbol{\omega}_t)$. The aggregate state $\boldsymbol{\omega}_t$ shifts the production of output, allowing disturbances such as TFP shocks. A representative firm chooses labor input and capital rental to maximize profit, taking wage w_t and the rental rate of capital ϱ_t as given:

$$\max_{h_t, k_{t-1}} F(h_t, k_{t-1}; \boldsymbol{\omega}_t) - w_t h_t - \varrho_t k_{t-1}. \tag{3}$$

The accumulation of capital is governed by a capital production function $\Phi(\cdot)$:

$$k_t = \Phi(\iota_t; \boldsymbol{\omega}_t) k_{t-1}, \tag{4}$$

where ι_{t+1} is the investment rate per unit of capital and the aggregate state $\boldsymbol{\omega}_t$ shifts capital production, allowing for disturbances such as capital destruction shocks. Holding capital from period $t-1$ to period t is taxed at rate τ^A_t and generates returns R^A_t:

$$R^A_{t+1} + \tau^A_{t+1} \coloneqq \frac{1}{q_t}\Big[\varrho_{t+1} + \max_{\iota_{t+1}} \big\{q_{t+1}\Phi(\iota_{t+1}; \boldsymbol{\omega}_{t+1}) - \iota_{t+1}\big\}\Big]. \tag{5}$$

We denote the value of capital in period t by

$$a^K_t \coloneqq q_t k_t \tag{6}$$

which comprises the supply of productive assets in the economy. Tax τ_t^A allows the government to control the supply of capital in the economy.

The Financial Sector

The financial sector holds productive assets, a_t^F. It can issue or hold financial assets, $\boldsymbol{b}_t^F$. We adopt the convention that positive entries of $\boldsymbol{b}_t^F$ indicate issuance and, therefore, constitute the supply of financial assets to the rest of the economy and liabilities of the financial sector, for example, deposits issued by banks. Negative entries correspond to financial assets held by the financial sector, for example, Treasury bills or reserves at the central bank. Given the equity of the financial sector, n_t, the balance sheet is given by:

$$a_t^F = \mathbf{1}_{\mathcal{B}}^{\intercal}\boldsymbol{b}_t^F + n_t.$$

Frictions in the financial sector are described by how its asset holdings and issuance depend on returns $\{R_{s+1}^A, \boldsymbol{R}_{s+1}^B\}$ per unit of equity:

$$a_t^F = \theta_t n_t, \quad \boldsymbol{b}_t^F = \boldsymbol{\phi}_t n_t, \tag{7}$$

where

$$\boldsymbol{\phi}_t \coloneqq \boldsymbol{\phi}(\{R_{s+1}^A, \boldsymbol{R}_{s+1}^B\}, \boldsymbol{\omega}_t)$$

is a function of returns sequences. The financial sector balance sheet implies that $\theta_t = \mathbf{1}_{\mathcal{B}}^{\intercal}\boldsymbol{\phi}_t + 1$.

The function $\boldsymbol{\phi}_t$ captures how the financial sector's ability to intermediate and supply assets depend on expected returns in any equilibrium. As we discuss in Section 2.3, this formulation of the financial sector allows us to nest a large class of financial intermediation models. The aggregate state $\boldsymbol{\omega}_t$ represents disturbances affecting the financial sector's ability to supply different assets.

Financial sector equity n_t represents the residual claim on holdings on its balance sheet, and the return on equity R_t^N satisfies:

$$R_t^N = R_t^A\theta_{t-1} - \boldsymbol{R}_t^{B\intercal}\boldsymbol{\phi}_{t-1}. \tag{8}$$

Government

The government can issue or hold financial assets with positions $\boldsymbol{b}_t^G$, where a positive entry stands for issuance and a negative entry for holdings. The total resource the government receives from the (net) issuance of assets is $\mathbf{1}_{\mathcal{B}}^{\intercal}\boldsymbol{b}_t^G$. The government budget constraints are given by:

$$\mathbf{1}_{\mathcal{B}}^{\intercal}\boldsymbol{b}_t^G = \boldsymbol{R}_t^{B\intercal}\boldsymbol{b}_{t-1}^G - \tau_t^A a_{t-1}^K - \boldsymbol{\tau}_t^{B\intercal}\boldsymbol{b}_{t-1} - \tau_t^N n_{t-1} - T_t. \tag{9}$$

We assume that the government can operate in the asset markets by adjusting $\boldsymbol{b}_t^G$ contingent on the realizations of $\{\hat{\boldsymbol{\omega}}_t\}$. These operations include traditional tools, like open market operations, and unconventional tools such as quantitative easing and tightening, the Troubled Asset Relief Program, and Operation Twist. We allow the government to adjust T_t contingent on $\{\hat{\boldsymbol{\omega}}_t\}$, but restrict the taxes $\tau_t^A, \boldsymbol{\tau}_t^B, \tau_t^N$ so that they cannot respond to the shocks. Taxes $\tau_t^A, \boldsymbol{\tau}_t^B, \tau_t^N$ act as macroprudential policies as they allow the government to influence the incentive to issue and hold assets, confronting frictions in the financial sector. The restriction that these taxes are not state-contingent captures the fact that many macroprudential regulations might not be adjusted timely in response to unexpected shocks. The restriction is analogous to this on product market subsidies present in the New Keynesian framework – these subsidies correct monopolistic distortions in the steady state but cannot respond to shocks, thus making room for monetary policy.

Definition of Equilibrium

An equilibrium consists of allocations $\{c_t, \boldsymbol{b}_t, n_t, h_t, k_t, \iota_t, a_t^K, a_t^F, \boldsymbol{b}_t^F\}$, prices $\{w_t, \varrho_t, q_t\}$, returns $\{R_t^A, \boldsymbol{R}_t^B, R_t^N\}$, and policies $\{\boldsymbol{b}_t^G, T_t, \tau_t^A, \boldsymbol{\tau}_t^B, \tau_t^N\}$ such that: $\{c_t, \boldsymbol{b}_t, n_t\}$ solves the households' problem (1) given returns and policies; $\{h_t, k_t, \iota_t\}$ solves the firm's problem (3) given wage, rental rate, and capital price; supply of productive asset is given by (6); the financial sector's supply and demand of assets, $\{a_t^F, \boldsymbol{b}_t^F\}$, is given by (7); the labor and capital rental market clear; returns on productive assets and equity are (5) and (8); asset markets clear:

$$a_t^K = a_t^F, \quad \boldsymbol{b}_t^G + \boldsymbol{b}_t^F = \boldsymbol{b}_t;$$

and the government budget constraint (9) holds. The goods market clears as a result of Walras' law.

2.2 A Sequence Space Representation

We represent the private sector's consumption, saving, and production behavior as an intertemporal system of supply and demand for assets. This representation is useful because, given the system, any competitive equilibrium is characterized by prices and policies that clear markets and satisfy government budget constraints. As a result, this representation builds a connection between a Ramsey planner's constraints and the properties of the supply and demand system that can be measured empirically.

Lemma 1 *(i) Given $\{R^A_{s+1}, \tau^A_{s+1}; \boldsymbol{\omega}_s\}$, there exist functions $w_t(\cdot), \varrho_t(\cdot), k_t(\cdot)$, and $q_t(\cdot)$ that satisfy production optimality and clear factor markets. The capital supply is given by*

$$a^K_t(\{R^A_{s+1}, \tau^A_{s+1}; \boldsymbol{\omega}_s\}) \coloneqq q_t(\{R^A_{s+1}, \tau^A_{s+1}; \boldsymbol{\omega}_s\})\ k_t(\{R^A_{s+1}, \tau^A_{s+1}; \boldsymbol{\omega}_s\}).$$

(ii) Given returns and taxes $\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}, \tau^A_{s+1}, \boldsymbol{\tau}^B_{s+1}, \tau^N_{s+1}, T_s; \boldsymbol{\omega}_s\}$, there exist asset demand functions

$$\begin{gathered} \boldsymbol{b}_t(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}, \tau^A_{s+1}, \boldsymbol{\tau}^B_{s+1}, \tau^N_{s+1}, T_s; \boldsymbol{\omega}_s\}), \\ n_t(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}, \tau^A_{s+1}, \boldsymbol{\tau}^B_{s+1}, \tau^N_{s+1}, T_s; \boldsymbol{\omega}_s\}) \end{gathered}$$

that solve the household's problem with welfare

$$v(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}, \tau^A_{s+1}, \boldsymbol{\tau}^B_{s+1}, \tau^N_{s+1}, T_s; \boldsymbol{\omega}_s\}).$$

These functions allow us to express production and household decisions as functions of returns and taxes. As these functions incorporate all decisions of agents in the private sector, we can characterize an equilibrium with the following excess asset supply functions:

$$\boldsymbol{\epsilon}_t(\cdot) \coloneqq \boldsymbol{b}^G_t + \boldsymbol{\phi}_t(\cdot) n_t(\cdot) - \boldsymbol{b}_t(\cdot), \quad \kappa_t(\cdot) \coloneqq a^K_t(\cdot) - \theta_t(\cdot) n_t(\cdot). \tag{10}$$

Implementability

Given these excess asset supply functions, an equilibrium requires that

$$\boldsymbol{\epsilon}_t = \mathbf{0}_{\mathcal{B}}, \quad \kappa_t = 0,$$

and that the government budget holds. These conditions form the constraints faced by a Ramsey planner when choosing policies—the market-clearing conditions correspond to the implementability constraint, and the government budget constraint replaces the feasibility constraint.

Our characterization of the planner's constraints as a requirement to clear an intertemporal asset supply and demand is similar to Auclert et al. (2024) in the context of a heterogeneous-agent model, where the complexity of the economy can be succinctly summarized by aggregate functions. In our context, the same idea allows us to represent a wide variety of financial frictions with potentially complex structures, extending the result in Chiang and Zoch (2023). We provide a few examples in the following section.

2.3 Nested Models of Financial Intermediation

We now illustrate how our sequence space representation can accommodate a wide variety of financial frictions. This flexibility illustrates the key advantage of our approach: complex financial intermediation structures can be succinctly summarized by the aggregate function $\boldsymbol{\phi}(\cdot)$, allowing us to remain agnostic about specific microfoundations while capturing their economic interaction with the aggregate economy. The exact microfoundation becomes irrelevant for policy analysis as long as we can measure properties of the asset supply and demand functions.

For canonical macroeconomic models in which financial intermediaries issue only one asset ($|\mathcal{B}| = 1$), Chiang and Zoch (2023) show that the formulation contains intermediation frictions originating from asset diversion (Gertler and Karadi, 2011), costly state verification (Bernanke et al., 1999), costly leverage (Uribe and Yue, 2006), and collateral constraints similar to Kiyotaki and Moore (1997) and Bianchi and Mendoza (2018).

This paper extends the framework to allow for an arbitrary number of financial assets

$\mathcal{B}$. The extension provides a general policy framework to study how the government can best operate in various asset markets. In Appendix C, we show that the formulation allows for key frictions in recent frameworks developed to study these policies. For example, by specializing to $|\mathcal{B}| = 3$, our formulation of the financial sector nests intermediation frictions in Sims and Wu (2021), Bianchi and Bigio (2022), Piazzesi et al. (2019), where the assets in $\mathcal{B}$ contain reserves, deposits, and government bonds. Similarly, by specializing to $|\mathcal{B}| = 2$, our framework nests the type of financial intermediation studied in Gertler et al. (2012) where assets in $\mathcal{B}$ contain deposits and outside equity.

One common feature of the frameworks above is that financial intermediaries do not feature persistent heterogeneity, and, as a result, the distribution of intermediary net worth does not have aggregate effects. However, our representation of the financial sector is not restricted by this property: the aggregate function $\boldsymbol{\phi}(\cdot)$ can also accommodate cases where heterogeneity within the financial sector does matter, as emphasized by Bellifemine et al. (2022). In Appendix C.4, we provide an example where financial intermediaries have heterogeneous abilities to hold and issue assets, and where the net worth distribution affects the aggregate financial sector response. Nevertheless, we can still derive an aggregate function $\boldsymbol{\phi}(\cdot)$ that encapsulates the complex and heterogeneous frictions across different financial intermediaries. As welfare does not depend directly on the distribution of net worth but only through the aggregate function $\boldsymbol{\phi}$, our policy analysis remains unchanged regardless of the underlying heterogeneity.

3 The Ramsey Problem

We consider the policy problem of the government under full commitment. The government is assumed to have the same objective function as the household, and it sets policies to choose an equilibrium that maximizes the value of the objective function. We use the intertemporal asset supply and demand system in Section 2.2 to formulate the Ramsey problem. As we show in Section 4, this formulation allows us to characterize the solution to the planner's problem in terms of observable sufficient statistics.

3.1 The Planner's Problem

At time $t = 0$, before observing the realization of $\{\boldsymbol{\omega}_t\}$, the planner announces policies $\{\boldsymbol{b}_t^{\boldsymbol{G}}, T_t, \tau_t^A, \boldsymbol{\tau}_t^B, \tau_t^N\}$. As discussed in Section 2.1, we allow $\boldsymbol{b}_t^G$ and T_t to be state-contingent—that is, they can depend on realizations of shocks $\{\boldsymbol{\omega}_t\}$—while we restrict τ_t^A, $\boldsymbol{\tau}_t^B$, and τ_t^N to remain constant across all possible shock realizations.

Acting as a Ramsey planner, the government solves the following optimization problem:[1]

$$\min_{\boldsymbol{\gamma}_t, \eta_t, \xi_t} \max_{R_{t+1}^A, \boldsymbol{R}_{t+1}^B, \tau_{t+1}^A, \boldsymbol{\tau}_{t+1}^B, \tau_{t+1}^N, \boldsymbol{b}_t^G, T_t} v(\{R_{s+1}^A, \boldsymbol{R}_{s+1}^B, \tau_{s+1}^A, \boldsymbol{\tau}_{s+1}^B, \tau_{s+1}^N, T_s; \boldsymbol{\omega}_s\}) \tag{11}$$

$$+ \mathbb{E} \sum_{t=0}^{\infty} \beta^t \Big\{ \underbrace{\boldsymbol{\gamma}_t^\intercal \boldsymbol{\epsilon}_t + \eta_t \kappa_t}_{\text{asset market clearing}} + \underbrace{\xi_t \big(\mathbf{1}_{\mathcal{B}}^\intercal \boldsymbol{b}_t^G + T_t + \tau_t^A a_{t-1}^K + \boldsymbol{\tau}_t^{B\intercal} \boldsymbol{b}_{t-1} + \tau_t^N n_{t-1} - \boldsymbol{R}_t^{B\intercal} \boldsymbol{b}_{t-1}^G \big)}_{\text{government budget constraint}} \Big\},$$

subject to initial conditions $\{\boldsymbol{R}_0^B, \boldsymbol{b}_{-1}^G, \boldsymbol{b}_{-1}, n_{-1}, \theta_{-1}, \boldsymbol{\phi}_{-1}, k_{-1}\}$.

Here, $\boldsymbol{\gamma}_t$ and η_t denote the Lagrange multipliers associated with the asset market clearing conditions, representing the social value of excess asset supply. The multiplier ξ_t corresponds to the government budget constraint and represents the shadow value of governmental resources.

The government's concern when designing its asset market operation is represented by the first-order condition with respect to $\boldsymbol{b}^G$:

$$\beta \xi_{t+1} \boldsymbol{R}_{t+1}^B - \xi_t \mathbf{1}_{\mathcal{B}} = \boldsymbol{\gamma}_t. \tag{12}$$

The left-hand side represents the fiscal cost the government incurs from issuing the assets: the cost of repayment in period $t+1$ with returns $\boldsymbol{R}_{t+1}^B$, net of the proceeds from issuance in period t.[2] The government trades off the fiscal cost with the social value of generating excess asset supply, which is represented by $\boldsymbol{\gamma}_t$ on the right-hand side.

[1] We keep market clearing constraints and treat returns as choice variables, unlike the standard dual approach, which substitutes out prices. This links the planner's optimality conditions to measurable properties of asset supply and demand.

[2] Return $\boldsymbol{R}^B$ reflects household's preferences over assets and portfolio adjustment cost. For example, high marginal utility of holding assets would appear as low $\boldsymbol{R}^B$.

Social Value Accounting

This brings us to the question of how we can account for the social value of excess asset supply. To understand the welfare impact of an extra unit of asset supply, we use the first-order condition with respect to returns to link the social values $\{\boldsymbol{\gamma}_t, \eta_t\}$ to how changes in returns affect welfare. For example, consider the first-order condition with respect to $\boldsymbol{R}^B_{t+1}$:[3]

$$\sum_{s=0}^{\infty} \beta^s \begin{bmatrix} \boldsymbol{\gamma}_s \\ \eta_s \end{bmatrix}^\intercal \begin{bmatrix} -\boldsymbol{\epsilon}_{s,\boldsymbol{R}^B_{t+1}} \\ -\kappa_{s,\boldsymbol{R}^B_{t+1}} \end{bmatrix} = v_{\boldsymbol{R}^B_{t+1}} + s_{\boldsymbol{R}^B_{t+1}}, \tag{13}$$

where $v_{\boldsymbol{R}^B_{t+1}}$ and $s_{\boldsymbol{R}^B_{t+1}}$ represent how changes in asset returns affect the household's welfare and the government budget:

$$v_{\boldsymbol{R}^B_{t+1}} := \beta^{t+1}\lambda_{t+1}\boldsymbol{b}_t^{G\intercal} + \sum_{s=0}^{\infty} \beta^{s+1}\lambda_{s+1}(\underbrace{R^A_{s+1}\theta_{s,\boldsymbol{R}^B_{t+1}} - {\boldsymbol{R}^B_{s+1}}^\intercal \boldsymbol{\phi}_{s,\boldsymbol{R}^B_{t+1}}}_{\text{pecuniary externality due to financial frictions}})n_s,$$

$$s_{\boldsymbol{R}^B_{t+1}} := -\beta^{t+1}\xi_{t+1}\boldsymbol{b}_t^{G\intercal} + \sum_{s=0}^{\infty} \beta^{s+1}\xi_{s+1}(\tau^A_{s+1}a^K_{s,\boldsymbol{R}^B_{t+1}} + {\boldsymbol{\tau}^B_{s+1}}^\intercal \boldsymbol{b}_{s,\boldsymbol{R}^B_{t+1}} + \tau^N_{s+1}n_{s,\boldsymbol{R}^B_{t+1}}).$$

To understand the formula, consider what happens when asset returns $\boldsymbol{R}^B_{t+1}$ increase by one unit in period $t+1$. First, when returns increase, the household gains on its holdings of assets issued by the government, while the government incurs losses, these effects are captured by the first terms in $v_{\boldsymbol{R}^B_{t+1}}$ and $s_{\boldsymbol{R}^B_{t+1}}$.[4]

However, as asset returns change, they also trigger endogenous responses in the economy. For example, the second term in $v_{\boldsymbol{R}^B_{t+1}}$ captures one such response: how the financial sector adjusts its asset holdings and issuance when returns change. Since households own the financial sector, these adjustments directly affect household welfare. In principle, one might have expected other endogenous responses to matter as well. For example, the household might change its saving decisions when asset returns change. However, by the envelope theorem, such responses have no first-order effect on welfare because households had already made these decisions optimally. By

[3]The first-order condition for returns is an accounting identity, not a characterization of optimal policy: it holds in any equilibrium, and not a tradeoff faced by the planner.

[4]Redistribution between the household and the financial sector is inconsequential because the household is the residual claimant on the financial sector's net worth.

contrast, due to the existence of financial frictions embedded in functions $\boldsymbol{\phi}$ and θ, changes in asset returns can potentially introduce welfare gain. The existence of this term symbolizes the presence of *pecuniary externality* in the economy. Finally, the second term in $s_{\boldsymbol{R}^B_{t+1}}$ represents how endogenous responses to asset returns affect the tax base and, consequently, tax revenue, which has the shadow value ξ_t.[5]

These effect of asset returns on household welfare and government budget allows us to recover the social value of excess asset supply, $\boldsymbol{\gamma}_t$ and η_t: On the left-hand side of Equation (13), $\boldsymbol{\epsilon}_{s,\boldsymbol{R}^B_{t+1}}$ and $\kappa_{s,\boldsymbol{R}^B_{t+1}}$ are sensitivities that describe the effects of asset returns on excess asset supply. Therefore, by "inverting" these sensitivities, we can derive the impact of excess supply on asset returns, and consequently, recover the social value of excess asset supply through its impact on asset returns.

Given these considerations, the government designs the issuance of assets $\boldsymbol{b}^G_t$ along with macroprudential taxes τ^A_t, $\boldsymbol{\tau}^B_t$, and $\boldsymbol{\tau}^N_t$ to maximize household welfare. In the following sections, we first characterize the Ramsey steady state that the government can achieve with the full set of tools in the absence of aggregate shocks. Afterwards, we characterize the government's optimal asset market operations in response to shocks when macroprudential taxes cannot respond.

3.2 Ramsey Steady State

Absent aggregate shocks, we define a Ramsey steady state as follows:

Definition 1 *Given $\boldsymbol{\omega}_t = \bar{\boldsymbol{\omega}}$ for all $t \geq 0$, a Ramsey steady state consists of returns $\bar{R}^A, \bar{\boldsymbol{R}}^B$, policies $\bar{\boldsymbol{b}}^G, \bar{T}, \bar{\tau}^A, \bar{\boldsymbol{\tau}}^B, \bar{\tau}^N$, and multipliers $\bar{\boldsymbol{\gamma}}, \bar{\eta}$, and $\bar{\xi}$, which solve the saddle-point problem (11) for all periods t, with the initial condition given by the steady-state returns, policies, and allocations.*

The following lemma characterizes the Ramsey steady state for the planner's problem described in (11):

Lemma 2 *In a Ramsey steady state, the multipliers and taxes are such that*

$$\bar{\boldsymbol{\gamma}} - \beta\bar{\xi}\bar{\boldsymbol{\tau}}^B = \mathbf{0}_{\mathcal{B}}, \quad \bar{\eta} + \beta\bar{\xi}\bar{\tau}^A = 0, \quad \textit{(Pigou tax)} \tag{14}$$

[5] This term is often referred to as the *fiscal externality* in the sufficient statistics literature, see Kleven (2021) for details.

and $\bar{\xi} = \bar{\lambda}$. *Asset returns satisfy*

$$\beta(\bar{\boldsymbol{R}}^B - \bar{\boldsymbol{\tau}}^B) = \mathbf{1}_{\mathcal{B}}, \quad \beta(\bar{R}^A + \bar{\tau}^A) = 1, \quad \beta(\bar{R}^N - \bar{\tau}^N) = 1. \quad \textit{(Friedman rule)} \tag{15}$$

Proof. See Appendix A. □

The optimal macroprudential tax policy balances welfare benefits against fiscal costs in Equation (14). When the planner increases tax to reduce excess asset demand by one unit, they incur a direct fiscal cost of $\beta\bar{\xi}\bar{\boldsymbol{\tau}}^B$ from losing the tax base. However, this reduction generates a social welfare gain $\bar{\boldsymbol{\gamma}}$, which captures how changing asset demand affects asset returns and welfare through two channels. First, changes in asset returns affect welfare due to pecuniary externalities from financial frictions. Second, these changes in returns create indirect fiscal effects as households adjust their asset demand in response. The optimality condition $\bar{\boldsymbol{\gamma}} - \beta\bar{\xi}\bar{\boldsymbol{\tau}}^B = \mathbf{0}_{\mathcal{B}}$ reflects the standard Pigou tax logic: optimal taxes equate the welfare gain from correcting externalities with the net fiscal cost, including the direct and indirect fiscal effects.

With the externality addressed by macroprudential taxes, the values of resources are equalized between the household and the government, $\bar{\xi} = \bar{\lambda}$, as a lump-sum tax $\bar{T}$ has no additional welfare effects. Moreover, all returns are effectively equalized with β^{-1} after adjusting for taxes, featuring a generalized version of the Friedman rule, as the government meets the household's demand for assets with the issuance of $\boldsymbol{b}^G$. However, as we show in the following section, this simple prescription about asset returns will no longer be optimal in the presence of aggregate shocks when the government cannot adjust macroprudential taxes. Without macroprudential taxes, the government will need to rely on asset market operations $\boldsymbol{b}^G$ to address any pecuniary externality that emerges as the consequence of the shocks.

4 Optimal Asset Market Operations

We analyze the government's optimal responses to shock $\hat{\boldsymbol{\omega}}_t$ around the Ramsey steady state up to first-order approximation. We denote deviations from the steady state as $\hat{\boldsymbol{x}}_t \coloneqq \boldsymbol{x}_t - \bar{\boldsymbol{x}}_t$ and use $\boldsymbol{f}_{t,\boldsymbol{x}_s}$ for the Jacobian of $\boldsymbol{f}_t(\boldsymbol{x}_s)$ with respect to $\boldsymbol{x}_s$ at steady state. We use column vectors discounted at rate β to represent multipliers and returns such as $\hat{\boldsymbol{\xi}} \coloneqq [\beta^t \hat{\xi}_t]$ and $\hat{\boldsymbol{R}}^B \coloneqq [\beta^{t+1}\hat{\boldsymbol{R}}^B_{t+1}]$, as well as the first difference of sequences such as $\Delta\hat{\boldsymbol{\xi}} \coloneqq [\beta^t(\hat{\xi}_{t+1} - \hat{\xi}_t)]$. We use matrices such as $\boldsymbol{\epsilon}_{\boldsymbol{R}^B} \coloneqq [\boldsymbol{\epsilon}_{t,\boldsymbol{R}^j_{s+1}}]$ to capture the sensitivities of supply and demand with respect to returns, where row t describes how $\boldsymbol{\epsilon}_t$ responds to returns $\boldsymbol{R}^B_{s+1}$ across different periods s.

The core objective of the planner is to address the reemergence of pecuniary externalities that arise due to aggregate shocks in the presence of financial frictions, using asset market operations as their instrument.

In principle, one might think that the planner needs to observe the underlying shocks and know about the detailed structure of the economy to design appropriate policy responses. However, this knowledge is unnecessary— for the class of models we study, asset returns contain all relevant information about the underlying shocks, and asset supply and demand elasticities contain everything the planner needs to know about the underlying structure of the economy. It is therefore sufficient for the planner to target a specific relationship between these assets to optimally mitigate the externality.

To see how it works, consider the optimality condition for asset market operations $\boldsymbol{b}^G$, Equation (12), which we approximate up to first order and stack in vector form. By adding the household's intertemporal valuation $\Delta\hat{\boldsymbol{\lambda}}$ to both sides of the equation and rearranging, we can represent the optimality condition as:

$$\Delta\hat{\boldsymbol{\lambda}} \otimes \mathbf{1}_{\mathcal{B}} + \bar{\lambda}\hat{\boldsymbol{R}}^B = \underbrace{\tilde{\boldsymbol{\gamma}}}_{\text{pecuniary externality gap}} + \underbrace{(\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}})}_{\text{government valuation gap}} \otimes \mathbf{1}_{\mathcal{B}}, \tag{16}$$

where $\otimes$ denotes the Kronecker product, and

$$\tilde{\boldsymbol{\gamma}} \coloneqq \left[\beta^t(\hat{\boldsymbol{\gamma}}_t - \beta\hat{\xi}_{t+1}\bar{\boldsymbol{\tau}}^B)\right].$$

The left-hand side represents the extent to which the household's intertemporal valuation deviates from asset returns. This deviation captures illiquidity effects from portfolio adjustment costs, net of any liquidity benefits embedded in household preferences. On the right-hand side, $\tilde{\boldsymbol{\gamma}}$ measures the welfare effects of increasing asset supply, which creates a pecuniary externality. This terms emerges because the government cannot respond to shocks with macroprudential taxes— the steady-state tax rates are no longer sufficient to internalize the pecuniary externality when the economy faces unexpected disturbances. The final term, $\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}}$, represents a valuation gap between households and the government that also arises from this pecuniary externality.

Our key result is to show that both the pecuniary externality gap $\tilde{\boldsymbol{\gamma}}$ and the government valuation gap $\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}}$ can be expressed in terms of observable asset returns weighted by specific combinations of asset supply and demand elasticities. Combined with the condition $\Delta\hat{\boldsymbol{\lambda}} + \bar{\lambda}\hat{\boldsymbol{R}}^* = 0$ from Assumption 1, the optimality condition in Equation (16) provides us with an optimal target rule that relates asset returns to each other. Given the optimal target rule, we will demonstrate how the government can implement the Ramsey plan using this rule without requiring direct observation of the underlying shocks.

4.1 Social Value Accounting

We follow the logic of social value accounting in Section 3.1 to build a link between asset returns, the externality gap $\tilde{\boldsymbol{\gamma}}$, and the valuation gap $\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}}$.

Intuitively, the pecuniary externality arises from the fact that intermediation is subject to financial frictions and may deviate from its optimal level. Given asset returns, the severity of the externality depends on the gain from changing the level of intermediation. Consequently, the size of the externality and valuation gaps is closely tied to how a specific policy can address the externality by affecting the level of intermediation.

Lemma 3 shows we can account for these considerations by expressing $\tilde{\boldsymbol{\gamma}}$ and $\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}}$ as averages of asset returns, weighted by asset supply and demand elasticities.

Lemma 3 *Under the Ramsey plan, pecuniary externality per unit of excess asset supply and the government valuation gap are given by:*

$$\tilde{\boldsymbol{\gamma}} = \bar{\lambda} \begin{bmatrix} \boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} \\ -\boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} \\ -\boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}} \end{bmatrix}^\intercal \begin{bmatrix} \hat{\boldsymbol{R}}^A \\ \hat{\boldsymbol{R}}^B \\ \hat{\boldsymbol{R}}^N \end{bmatrix}, \quad \Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}} = \bar{\lambda} \begin{bmatrix} \boldsymbol{\Gamma}^A_{\Delta\boldsymbol{T}} \\ -\boldsymbol{\Gamma}^B_{\Delta\boldsymbol{T}} \\ -\boldsymbol{\Gamma}^N_{\Delta\boldsymbol{T}} \end{bmatrix}^\intercal \begin{bmatrix} \hat{\boldsymbol{R}}^A \\ \hat{\boldsymbol{R}}^B \\ \hat{\boldsymbol{R}}^N \end{bmatrix}$$

where $\boldsymbol{\Gamma}^j_{\boldsymbol{\epsilon}}$ and $\boldsymbol{\Gamma}^j_{\Delta\boldsymbol{T}}, \forall j \in \{A, B, N\}$ are loading matrices comprised of price and income elasticities evaluated at the Ramsey steady state, given in Appendix A.5.

Proof. See Appendix A. □

To understand how the loading matrices reflect the underlying structure of the economy through asset supply and demand elasticities, consider a special case where the capital supply is perfectly elastic so that there is no feedback from the changes in returns on capital. In this case, we have

$$\boldsymbol{\Gamma}_{\boldsymbol{\epsilon}}^{A} = \tilde{\boldsymbol{a}}_{\boldsymbol{R}^B}^{F}(-\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B})^{-1},$$

where $\tilde{\boldsymbol{a}}_{\boldsymbol{R}^B}^{F}$ and $\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B}$ are Hicksian elasticities of the financial sector's demand for capital and the excess financial asset supply, respectively. In both of these elasticities, the household's demand for net worth and for financial assets is compensated for the income effect resulting from changes in asset returns. Intuitively, the loading matrix $\boldsymbol{\Gamma}_{\boldsymbol{\epsilon}}^{A}$ represents how much the government can influence the financial sector's holding of capital by supplying an additional unit of financial asset. The term $(-\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B})^{-1}$ captures the effect of asset supply on returns, while $\tilde{\boldsymbol{a}}_{\boldsymbol{R}^B}^{F}$ describes how the financial sector would adjust in response.

With these loading matrices, $\tilde{\boldsymbol{\gamma}}$ accounts for the extent to which the government can generate welfare gain by supplying assets, given asset returns disturbances $\hat{\boldsymbol{R}}^A$, $\hat{\boldsymbol{R}}^B$, $\hat{\boldsymbol{R}}^N$. Similarly, the valuation gap reflects the welfare effect that the government generates from balancing the fiscal cost of the asset market operation with taxes $\Delta \boldsymbol{T}$. These welfare effects exist due to the presence of pecuniary externality in the financial sector.

One key feature of the accounting in Lemma 3 is the *absence* of shocks $\hat{\boldsymbol{\omega}}$, conditional on asset returns. Although disturbances $\hat{\boldsymbol{\omega}}$ alter the transmission of policy, these alterations (i.e. their effects on $\boldsymbol{\Gamma}_{\boldsymbol{\epsilon}}^{j}$ and $\boldsymbol{\Gamma}_{\Delta \boldsymbol{T}}^{j}$) do not have a first-order impact on welfare. This is because the level of financial intermediation around the Ramsey steady state is already optimal in the absence of shocks. As a result, all relevant information about the underlying shocks themselves is incorporated into asset return disturbances $\hat{\boldsymbol{R}}^A$, $\hat{\boldsymbol{R}}^B$, and $\hat{\boldsymbol{R}}^N$. Together with loading matrices $\boldsymbol{\Gamma}_{\boldsymbol{\epsilon}}^{A}$ and $\boldsymbol{\Gamma}_{\Delta \boldsymbol{T}}^{j}$, they contain all relevant information about how the government's asset market operations affect welfare.

4.2 The Optimal Target Rule

Using the accounting result in Lemma 3, the optimality condition for asset market operations in Equation (16), along with Assumption 1, immediately imply an optimal target rule for asset returns:

Theorem 1 *The Ramsey plan can be achieved by a target rule of the following form:*

$$\hat{\boldsymbol{R}}^B - \hat{\boldsymbol{R}}^* \otimes \mathbf{1}_{\mathcal{B}} = \begin{bmatrix} \boldsymbol{\Gamma}^A \\ -\boldsymbol{\Gamma}^B \\ -\boldsymbol{\Gamma}^N \end{bmatrix}^\intercal \begin{bmatrix} \hat{\boldsymbol{R}}^A \\ \hat{\boldsymbol{R}}^B \\ \hat{\boldsymbol{R}}^N \end{bmatrix}, \tag{17}$$

where $\boldsymbol{\Gamma}^j := \boldsymbol{\Gamma}^j_{\boldsymbol{\epsilon}} + \boldsymbol{\Gamma}^j_{\Delta \boldsymbol{T}} \otimes \mathbf{1}_{\mathcal{B}}, \forall j \in \{A, B, N\}$ *are loading matrices.*

Proof. See Appendix A. □

The left-hand side of the target rule contains the spreads between financial asset returns and the household's neutral rate, representing their intertemporal valuation of resources. Any spread above the neutral rate represents a welfare cost for the household to increase their asset holding, as a result of the portfolio adjustment cost, net of liquidity service.

The right-hand side is the sum of the externality gap $\tilde{\boldsymbol{\gamma}}$, and the valuation gap $\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}}$. Together, they capture the total welfare gain from mitigating the presence of pecuniary externality with an increase in asset supply.

The rule described in Theorem 1 has two properties that make it especially useful. First, the rule specifies a fixed relationship between asset returns, *regardless* of the underlying disturbances $\hat{\boldsymbol{\omega}}$. These asset returns, which are readily observable in the markets, summarize all relevant information for the prescription of optimal policy regardless of underlying disturbances. Second, this relationship, captured by loading matrices $\boldsymbol{\Gamma}^j$, is fully described by the price and income elasticities. These elasticities reflect the *invariant* properties of the economy, including the underlying preferences, technology, and frictions. They can be estimated using standard identification strategies. The exact microfoundations of the model are irrelevant to the optimal policy rule as long as they imply the same matrices $\boldsymbol{\Gamma}^j$. This represents the spirit of the sufficient statistics approach (Chetty, 2009; Kleven, 2021).

4.3 Implementation

We now show that, given these supply and demand elasticities, Theorem 1 allows the government to implement the Ramsey plan with a simple asset supply schedule based on the target rule without any knowledge about the underlying disturbances, following the spirit of "robust policy rules" discussed in Giannoni and Woodford (2003) and Woodford (2011).

We start by explaining why one might think that implementation could be a challenge. Given a specific sequence of asset market operations $\hat{\boldsymbol{b}}^G$, an equilibrium $\{\hat{\boldsymbol{R}}^A, \hat{\boldsymbol{R}}^B, \hat{\boldsymbol{T}}\}$ satisfies

$$\boldsymbol{\kappa}_{R^A}\hat{\boldsymbol{R}}^A + \boldsymbol{\kappa}_{R^B}\hat{\boldsymbol{R}}^B + \boldsymbol{\kappa}_T\hat{\boldsymbol{T}} + \boldsymbol{\kappa}_\omega\hat{\boldsymbol{\omega}} = 0, \tag{18}$$

$$\hat{\boldsymbol{b}}^G + \boldsymbol{\epsilon}_{R^A}\hat{\boldsymbol{R}}^A + \boldsymbol{\epsilon}_{R^B}\hat{\boldsymbol{R}}^B + \boldsymbol{\epsilon}_T\hat{\boldsymbol{T}} + \boldsymbol{\epsilon}_\omega\hat{\boldsymbol{\omega}} = 0, \tag{19}$$

and the government budget constraint. One way to think about the rule is to solve $\hat{\boldsymbol{R}}^A$ and $\hat{\boldsymbol{R}}^B$ as functions of shocks $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{b}}^G$, calculate the implied $\hat{\boldsymbol{R}}^N$ and $\hat{\boldsymbol{R}}^*$, and find $\hat{\boldsymbol{b}}^G$ that satisfies the optimal target rule. But this is a bad idea because it requires the government to observe how all shocks from various sources disturb asset markets ($\boldsymbol{\kappa}_\omega\hat{\boldsymbol{\omega}}$ and $\boldsymbol{\epsilon}_\omega\hat{\boldsymbol{\omega}}$). Measuring all possible shocks and identifying how they affect asset markets is likely infeasible.

However, the target rule in Theorem 1 provides us with a simple way to circumvent this challenge. In Theorem 1, the target rule specifies a target relationship between asset returns that has to be satisfied under the Ramsey plan, which *does not* depend on the underlying shocks. Now, suppose the government is willing to commit to an asset supply schedule, conditional on returns $\hat{\boldsymbol{R}}^A$, $\hat{\boldsymbol{R}}^N$ and $\hat{\boldsymbol{R}}^*$, that supplies financial assets perfectly elastically at returns $\hat{\boldsymbol{R}}^B$, as given by the formula in Theorem 1. Under this asset supply schedule, the equilibrium will be given by $\{\hat{\boldsymbol{R}}^A, \hat{\boldsymbol{R}}^B, \hat{\boldsymbol{T}}, \hat{\boldsymbol{b}}^G\}$ that satisfies market clearing conditions (18) and (19), the government budget constraint, and Equation (17), given arbitrary realization of $\hat{\boldsymbol{\omega}}$. In other words, given the supply schedule, the quantities of assets $\hat{\boldsymbol{b}}^G$ will adjust to absorb the shocks to achieve the target rule and implement the Ramsey plan for any possible shocks that can affect the economy.

5 The Role of Financial Frictions

We now apply Theorem 1 to a series of examples to illustrate how financial frictions shape the design of the optimal policy rule.

5.1 Neoclassical Benchmark

First, we provide a benchmark case where key features of the financial sector are absent. Specifically, we consider an economy where (1) financial intermediation does not respond to changes in returns: $\boldsymbol{\phi}(\cdot) \equiv \bar{\boldsymbol{\phi}}$ for a constant vector $\bar{\boldsymbol{\phi}}$, implying no pecuniary externality in the financial sector, and (2) the household derives no utility and faces no cost in holding financial sector net worth, which means that returns on net worth equal their intertemporal valuation: $\Delta\hat{\boldsymbol{\lambda}} = -\bar{\lambda}\hat{\boldsymbol{R}}^N$.

Under these assumptions, Theorem 1 prescribes a simple rule: the planner should maintain a constant spread between all returns for any realization of shocks $\hat{\boldsymbol{\omega}}$, thereby preserving the spirit of the Friedman rule:

Corollary 1 *If $\boldsymbol{\phi}(\cdot) \equiv \bar{\boldsymbol{\phi}}$ and $\Delta\hat{\boldsymbol{\lambda}} = -\bar{\lambda}\hat{\boldsymbol{R}}^N$, then a target rule*

$$\hat{\boldsymbol{R}}^B = \hat{\boldsymbol{R}}^A \otimes \mathbf{1}_{\mathcal{B}}$$

achieves the Ramsey plan, and $\tilde{\boldsymbol{\gamma}} = \mathbf{0}$ and $\hat{\boldsymbol{R}}^N = \hat{\boldsymbol{R}}^A$ in equilibrium.

When all returns on financial assets move in tandem with returns on capital, we have $\hat{\boldsymbol{R}}^N = \hat{\boldsymbol{R}}^A$. In this case, households effectively control the capital stock through their net worth holdings, as in a neoclassical growth model. At the same time, the planner issues the necessary amount of each type of financial asset to satiate household demand for these assets arising from any additional motives.

Corollary 1 allows us to isolate the role of financial frictions in designing the optimal policy rule: In the absence of these frictions, the planner would maintain constant spreads between all assets. The presence of financial frictions upends this prescription, requiring the planner to implement a target rule that facilitates intermediation by maintaining variations between spreads in response to shocks.

5.2 Frictions in Financial Intermediation

To highlight the role of financial frictions in designing optimal policy, consider a special case where (1) the household has quasi-linear preferences with liquidity service from holding assets given by a concave function $\nu(\cdot)$:

$$u\left(\{c_s, \boldsymbol{b}_s, n_s\}_{s\le t}; \boldsymbol{\omega}_t\right) = c_t + \nu(\boldsymbol{b}_t; \boldsymbol{\omega}_t),$$

(2) the capital supply from the production sector is perfectly elastic, and the returns on capital $\{\hat{R}^A_{t+1}\}$ are given exogenously, and (3) the financial sector net worth is fixed exogenously: $n_t = \bar{n}$.

Corollary 2 *With quasi-linear preference, perfectly elastic capital, and fixed net worth, the optimal target rule is given by*

$$\hat{\boldsymbol{R}}^B = \left(\bar{n}\boldsymbol{\theta}_{\boldsymbol{R}^B}\boldsymbol{b}^{-1}_{\boldsymbol{R}^B}\right)^{\intercal}\hat{\boldsymbol{R}}^A.$$

To build intuition, we further simplify to the case where there is a single financial asset, $|\mathcal{B}| = 1$, and the level of intermediation per net worth is given by a function $\varphi(\cdot)$ of the spread between R^A_{t+1} and R^B_{t+1}:

$$\phi_t = \varphi(R^A_{t+1} - R^B_{t+1}), \quad \varphi'(\cdot) > 0.$$

Financial frictions of this form arise from microfoundations such as costly state verification (Bernanke et al., 1999) or costly leverage (Cúrdia and Woodford, 2011). Since the level intermediation depends only on next-period returns, the financial sector's asset supply reduces to a static system: $\phi_{s,R^j_{t+1}} = \mathbf{0}$ if $s \neq t$, and the optimal target rule in Corollary 2 reduces to

$$\hat{R}^B_{t+1} = -\frac{\bar{n}\varphi'}{b_{R^B}}\hat{R}^A_{t+1}, \tag{20}$$

where φ' and b_{R^B} are evaluated at the steady state.

Because the financial sector intermediates more assets when the spread is larger ($\varphi' > 0$) and the household holds more assets when returns are higher ($b_{R^B} > 0$), the policy rule prescribes that the planner targets a negative relationship between $\hat{R}^A_{t+1}$ and $\hat{R}^B_{t+1}$. When $\hat{R}^A_{t+1} > 0$, pecuniary externality emerges because the endogenous response of $\hat{R}^B_{t+1}$ does not take into account the fact that a lower $\hat{R}^B_{t+1}$ would allow the financial sector to increase its holding of capital. As a result of the externality, the planner can create welfare gains by decreasing excess asset supply, lowering the financial sector's funding cost, and thereby increasing the level of intermediation. To decide how much excess asset supply to absorb, the planner trades off the welfare gain from addressing the pecuniary externality against the loss of liquidity service that the household values. A decrease in excess asset supply has a large impact on the level of intermediation when (1) the financial sector's intermediation capacity is sensitive to its funding cost (high φ'), and (2) the household asset demand is inelastic (low b_{R^B}). In these cases, the optimal target rule prescribes a larger cut in $\hat{R}^B_{t+1}$.

Forward-Looking Frictions

A similar logic extends to the case in which the financial sector features an intertemporal asset supply system, where the level of intermediation in one period depends on asset returns in all future periods. As an example, suppose that the level of financial friction is given by $\phi_t(\{R^A_{s+1} - R^B_{s+1}\}_{s\geq t})$, with responses to asset returns around the steady state given by:

$$\phi_{t,R^A_{s+1}} = -\phi_{t,R^B_{s+1}} = \phi_R \times \Xi^{s-t}, \quad \forall s \geq t,$$

and zero for all $s < t$, for some $\phi_R > 0$ and $\Xi \in [0, 1)$. As shown in Chiang and Zoch (2023), financial frictions of this form resemble those in asset diversion models as in Gertler and Kiyotaki (2010).

With the intertemporal asset supply system, Corollary 2 implies an optimal target rule of the form:

$$\hat{R}^B_{t+1} = -\frac{\bar{n}\phi_R}{b_{R^B}} \sum_{s=0}^{t} \Xi^{t-s} \hat{R}^A_{s+1}.$$

Unlike the static case in Equation (20), a decrease in the financial sector's funding cost in period t, R^B_{t+1}, enhances the financial sector's intermediation capacity in all periods before t due to the forward-looking friction. As a result, the optimal target for $\hat{R}^B_{t+1}$ reflects disturbances in returns on capital in all periods before t: $\hat{R}^A_{s+1}$ for all $s \leq t$, with the weight Ξ^{s-t} diminishing with time horizons between t and s as the impact of R^B_{t+1} on the level of intermediation in period s decays with the horizon.

6 Taking the Model to the Data

Given the central role of financial frictions in shaping the optimal target rule in Theorem 1, we demonstrate how we can use established empirical methods to obtain the financial sector's asset supply and demand elasticities.

We measure the sensitivities of the function $\boldsymbol{\phi}$ for the U.S. banking sector, which describes how its issuance and holdings of assets respond to returns. We use the Call Reports for their balance sheet information and obtain expected returns for corresponding assets. To address the simultaneity bias, we identify relevant elasticities with the Granular Instrumental Variables (GIV) approach proposed by Gabaix and Koijen (2024). In addition, we use the granular instrumental variables we construct to estimate the capital supply elasticities.

6.1 Asset Supply and Demand Estimation

Empirical Specification

The financial sector consists of J entities. Let $b^F_{j,\ell,t}$ denote financial asset $\ell \in \mathcal{B}$ issued by the entity j in period t. The net worth of entity j is denoted by $n_{j,t}$. Issuance and holdings of assets per unit of net worth are defined as $\boldsymbol{\phi}_{j,t} \coloneqq \{\phi_{j,\ell,t}\}_{\ell=1}^{\mathcal{B}}$ where $\phi_{j,\ell,t} \coloneqq b^F_{j,\ell,t}/n_{j,t}$.

Entity j's holdings and issuance of assets per unit of net worth in period t are given by a supply and demand system:

$$\boldsymbol{\phi}_{j,t} = \bar{\boldsymbol{\phi}}_j(t) + \boldsymbol{\phi}_{\boldsymbol{R}^B} \boldsymbol{S}_{t+1} + \boldsymbol{\Lambda}_j \boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_{j,t}, \tag{21}$$

where $\bar{\boldsymbol{\phi}}_j(t)$ denotes entity-specific time trends of holdings and issuance, the sensitivity matrix $\boldsymbol{\phi}_{\boldsymbol{R}^B}$ parameterizes how they respond to spreads between asset returns (also calculated as deviations of from their trends):

$$\boldsymbol{S}_{t+1} \coloneqq \mathbb{E}_t[\boldsymbol{R}^B_{t+1}] - \mathbb{E}_t[R^A_{t+1}] \otimes \mathbf{1}_{\mathcal{B}},$$

$\boldsymbol{\eta}_t$ are unobserved aggregate shocks on which entity j has loadings $\boldsymbol{\Lambda}_j$, and $\boldsymbol{\varepsilon}_{j,t}$ are idiosyncratic shocks i.i.d. by entity-asset-time.

The specification above imposes several types of restrictions. First, to apply the standard granular instrumental variable approach, we assume entities in the financial sector respond only to expected returns one period ahead. Second, to reduce the number parameters to be estimated, we assume that the portfolio choice is determined by the spreads between $\mathbb{E}_t[\boldsymbol{R}^B_{t+1}]$ and $\mathbb{E}_t[R^A_{t+1}]$, rather than the returns separately. Finally, while we allow entities to have different loadings on the unobserved aggregate shocks, we assume that the sensitivity matrix is homogeneous across entities.

Identification

To obtain consistent estimates of $\boldsymbol{\phi}_{\boldsymbol{R}^B}$ in the presence of unobservable shocks $\boldsymbol{\eta}_t$ and $\boldsymbol{\varepsilon}_{j,t}$ that determine equilibrium values of spreads, we need instrumental variables. We follow the approach in Gabaix and Koijen (2024) to construct granular instrumental variables from idiosyncratic shocks $\boldsymbol{\varepsilon}_{j,t}$ which are "large enough" to generate non-trivial movements in equilibrium spreads. We provide details of our implementation in Appendix D.2.

Data

We construct our dataset using Call Report data for U.S. banking sector balance sheets, supplemented with CRSP data for market values of bank equity and TIC data for foreign countries' asset holdings in our GIV construction. Our sample spans 2003Q2-2023Q4. We build the GIV measure from both bank and foreign holdings data, focusing our analysis on estimating the banking sector's sensitivity matrix $\boldsymbol{\phi}_{\boldsymbol{R}^B}$. We provide an overview below, with full details in Appendices D.1 and D.2.

Entities: An entity corresponds to a bank holding company in the Call Report-CRSP data and a country in the TIC data. After sample selection, our data contains balanced panels of 130 bank holding companies and 48 countries.

Assets: We set $|\mathcal{B}| = 2$ and refer to the two groups of financial assets (or liabilities) as "liquid" and "illiquid." Liquid assets include reserves, deposits, federal funds, security repurchase agreements, and Treasury securities. Illiquid assets are comprised of agency- and GSE-backed securities, municipal securities, and corporate bonds. We calculate the capital holdings of a bank ($\mathcal{A}$) residually from its balance sheet as its market capitalization plus total liabilities net of financial assets.

Returns: We use the real 1-year Treasury yield as the return on the liquid asset and the real BAA yield as the return on the illiquid asset. To calculate $\mathbb{E}_t[R^A_{t+1}]$ we proceed in two steps. For each bank, we construct the expected return on net worth, $\mathbb{E}_t[R^N_{t+1}]$, by regressing the realized excess return on a set of aggregate and bank-specific predictors. We then recover $\mathbb{E}_t[R^A_{t+1}]$ from each bank's balance sheet using Equation (8).

GIV: To isolate the idiosyncratic shocks for banks and countries, we proceed as follows: First, we remove entity-asset-specific quadratic trends and common time fixed effects (which contain responses to expected returns). Next, we extract principal components and their loadings, and calculate the residuals as $\boldsymbol{\varepsilon}_{j,t}$. We apply the procedure for banks and countries separately, constructing the granular instrumental variables as the weighted average of $\boldsymbol{\varepsilon}_{j,t}$ for each group. For banks, we use their net worth share as weights and construct GIV for their holdings of the two $\mathcal{B}$ assets jointly, using three principal components to capture aggregate shocks in our baseline specification. For foreign countries, we use their holding share as weights, focusing on their liquid asset holdings, and absorb aggregate shocks with one principal component.

Asset Supply and Demand of the U.S. Banking Sector

Table 1 presents our estimates of banking sector sensitivities. The first two columns show how net issuance and holdings of liquid and illiquid assets respond to spread changes per unit of net worth. The third column reports the implied sensitivities of capital holdings per unit of net worth, derived from balance sheet constraints.

Table 1: Effect of Spreads on Bank Portfolio

	Liquid	Illiquid	Capital
$R^B_{\text{liquid}} - R^A$	-1.148	0.255	-0.893
	(0.292)	(0.070)	
$R^B_{\text{illiquid}} - R^A$	0.890	-0.205	0.685
	(0.326)	(0.076)	

Note: Responses of $\boldsymbol{\phi}_t$ and θ_t to a one percentage point increase in annualized spreads. Robust standard errors in parentheses. Capital responses are implied by balance sheet constraints.

The estimated sensitivities have intuitive signs. When expected returns on an asset increase, the financial sector reduces its net issuance of that asset (equivalently, increases its holdings). Conversely, when the return on capital R^A rises, the financial sector increases its net issuance of liquid and illiquid assets (reduces its holdings) while increasing its capital holdings.

Capital Supply of the Production Sector

We estimate the elasticity of capital supply, focusing on the banking sector's capital holdings. Using bank holding company balance sheets, we calculate and aggregate total real capital holdings as:

$$k_t = \int (\mathbf{1}_{\mathcal{B}}^{\intercal}\boldsymbol{\phi}_{j,t} + 1) n_{j,t} dj.$$

We estimate the semi-elasticity of k_t with respect to R^A_{t+1} in a 2SLS regression, employing the GIV constructed in Section 6.1 as a source of exogenous variation in capital demand. The semi-elasticity of capital supply with respect to annualized return on capital is -0.01 with a standard error of 0.0020.

6.2 Calibration

Household

We assume the household's utility function has the following form

$$u\left(c_t, \boldsymbol{b}_t; \boldsymbol{\omega}_t\right) = c_t \quad - \sum_{i \in \{\text{liquid,illiquid}\}} \varsigma_i \big(b_{i,t} - \bar{b}_i(\boldsymbol{\omega_t})\big)^2,$$

and there is no portfolio adjustment cost: $\Gamma(\cdot) \equiv 0$. We set $\beta = 0.99$ and calibrate parameters of the utility function such that in the steady state the semi-elasticities with respect to annualized expected returns are equal to 0.1. We assume that the net worth of the financial sector is exogenous: $n_t = n(\boldsymbol{\omega}_t)$. The disturbance $\boldsymbol{\omega}_t$ shifts net worth only.

Steady State

We calculate the long-run averages of returns $\bar{\boldsymbol{R}}^B$ and $\bar{R}^A$, the portfolio shares of the financial sector $\bar{\boldsymbol{\phi}}$ and $\bar{\theta}$, along with household's holdings, as described in Appendix D.3. Treating these values as the Ramsey steady state, we compute the taxes $\bar{\tau}^A$, $\bar{\boldsymbol{\tau}}^B$, and $\bar{\tau}^N$ that are consistent with these asset return values in the Ramsey steady state, assuming the household's holdings are at their bliss points. We normalize the net worth of the financial sector to one. Table 2 shows the steady-state values of key variables.

Table 2: Steady State Values

$\bar{R}^B_{\text{liquid}}$	$\bar{R}^B_{\text{illiquid}}$	$\bar{R}^A$	$\bar{\phi}_{\text{liquid}}$	$\bar{\phi}_{\text{illiquid}}$	$\bar{\theta}$	$\bar{b}_{\text{liquid}}$	$\bar{b}_{\text{illiquid}}$
1.004	1.014	1.002	4.774	-1.700	4.074	6.470	-1.700

Note: Asset returns are expressed as gross quarterly rates. Calculations of returns, portfolio shares, and holdings are provided in Appendix D.3.

7 Policy Implications

Building on our empirical estimates, we examine their implications for optimal policy design. We begin by deriving the optimal target rule that implements the Ramsey plan, which prescribes the relationship between asset returns that the government should maintain in response to economic disturbances. We then demonstrate how this rule operates in practice by analyzing the economy's response to a financial sector net worth shock—a disturbance that resembles a financial crisis—contrasting the optimal policy response with laissez-faire outcomes.

7.1 Optimal Target Rule

Using the asset supply and demand elasticities from our estimation and household calibration, Theorem 1 shows that the government can implement the Ramsey plan through the following target rule:

$$\begin{bmatrix} \hat{R}^B_{\text{liquid},t+1} \\ \hat{R}^B_{\text{illiquid},t+1} \end{bmatrix} = \begin{bmatrix} -0.34 \\ 0.76 \end{bmatrix} \hat{R}^A_{t+1}.$$

This optimal rule prescribes that when expected returns on capital are high, the government should target lower returns on liquid assets while simultaneously targeting higher returns on illiquid assets.

The intuition behind these contrasting prescriptions stems directly from the asset supply and demand elasticities estimated in Section 6.1. The first row of Table 1 shows that higher liquid rates cause the financial sector's balance sheet to contract, reducing both liquid asset issuance and illiquid asset holdings. Since liquid asset issuance responds more strongly than illiquid asset holdings, capital holdings decline as liquid rates rise. In contrast, the second row of Table 1 shows that higher illiquid rates stimulate both liquid asset issuance and illiquid asset holdings, with liquid asset issuance again dominating the portfolio adjustment. This leads to increased capital holdings—indicating that capital holdings and illiquid assets are complementary.

The optimal target rule also informs us about the externality gap $\tilde{\boldsymbol{\gamma}}$. Under the assumption of quasilinear preferences for the household, $\Delta\hat{\boldsymbol{\lambda}}$ and the government valuation gap $\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}}$ are both zeros. Therefore, Equation (16) implies that returns on financial assets are equal to the externality gap $\tilde{\boldsymbol{\gamma}}$ under optimal policy. The sign of the optimal target rule reflects how $\tilde{\boldsymbol{\gamma}}$ comoves with returns on capital: it is positive for illiquid assets and negative for liquid assets.

7.2 Responses to Financial Net Worth Shocks

To understand how the optimal policy shapes the economy's behavior in equilibrium, we analyze responses to a shock that reduces financial sector net worth. This shock resembles the "equity shock" studied in Christiano et al. (2014) and is often considered a key driver of the Great Financial Crisis (Gertler and Kiyotaki, 2010; Gertler and Karadi, 2011).[6]

We parameterize the net worth shock as follows. Given our utility function specification, net worth evolves according to:

$$n_t(\omega_t) = n(\bar{\omega}) - \hat{\omega}_t.$$

The shock $\hat{\omega}_t$ follows an AR(1) process: $\hat{\omega}_t = \rho_\omega \hat{\omega}_{t-1}$, where the value of $\hat{\omega}_0$ corresponds to a destruction of 30% of steady-state net worth, and the persistence parameter is $\rho_\omega = 0.75$.

From Equation (10), the impact of net worth shocks on the excess supply of capital

[6]Such shocks are typically modeled as events that increase capital depreciation rates or reduce capital quality, as in Merton (1973).

and financial assets is given by:

$$\boldsymbol{\kappa}_{\omega}\hat{\boldsymbol{\omega}} = \left\{\ \bar{\theta}\hat{\omega}_t\ \right\}, \quad \hat{\boldsymbol{\epsilon}}_{\omega}\hat{\boldsymbol{\omega}} = \left\{\ -\bar{\boldsymbol{\phi}}\hat{\omega}_t\ \right\}.$$

To illustrate how the government's optimal asset market operations alter equilibrium outcomes, we compare two policy regimes: (1) the government implements the optimal target rule and supplies assets accordingly to meet underlying disturbances, which we refer to as *optimal policy*, and (2) the government does not intervene in asset markets, setting $\hat{\boldsymbol{b}}^G_t = \mathbf{0}_{\mathcal{B}}$ for all t, which we refer to as *laissez faire*.

Figure 1 illustrates how the optimal policy implements the target rule in response to the financial sector net worth shock. The top row shows asset return responses, while the bottom row displays the externality gap $\tilde{\boldsymbol{\gamma}}_t$. Under the optimal policy, the liquid rate decreases with returns on capital while the illiquid rate increases, consistent with the prescription of the target rule. However, both rates remain lower relative to the laissez-faire outcome.

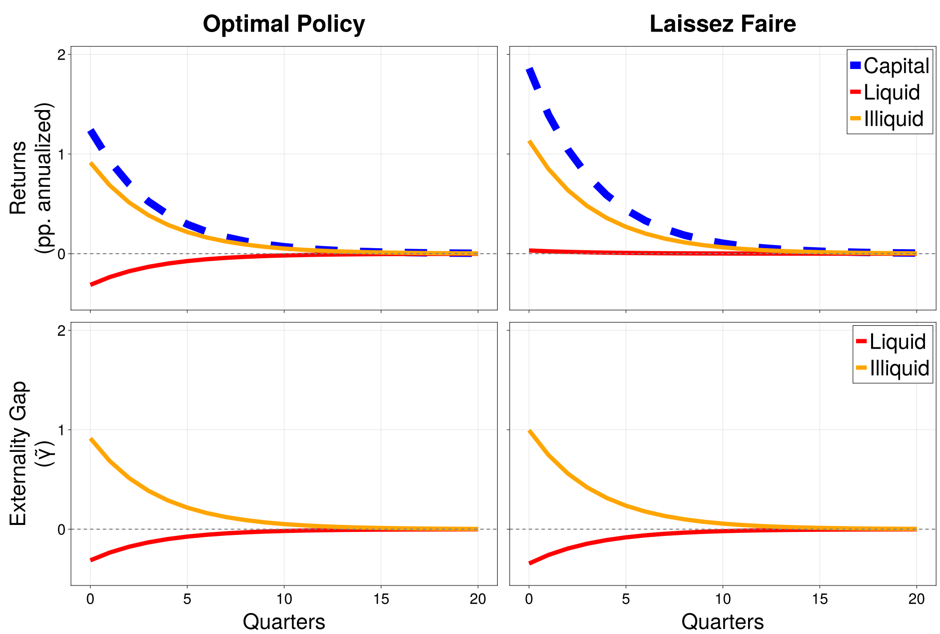

Figure 1: Equilibrium returns and externality gap $\tilde{\boldsymbol{\gamma}}_t$ under the optimal policy and under no asset market operations ($\hat{\boldsymbol{b}}^G = \mathbf{0}$), Responses (annualized) expressed in percentage points.

We can understand the planner's trade-off by comparing asset returns under laissez-faire with the externality gap (right column). Without policy intervention, returns for both assets exceed $\tilde{\boldsymbol{\gamma}}_t$. This deviation from the externality gap indicates suboptimal outcomes, creating an opportunity for the planner to generate welfare gains by reducing returns for both assets. Under the optimal policy (left column), returns align precisely with the externality gap, satisfying the optimality condition in Equation (16), which, in this case, reduces to:

$$\hat{\boldsymbol{R}}^B = \tilde{\boldsymbol{\gamma}}.$$

To move returns downward, the planner increases holdings of both types of assets. Figure 2 shows how the quantities adjust under optimal policy: the government performs a dual intervention—simultaneously purchases illiquid assets to absorb the excess holdings from the households and provides liquid lending to reduce the funding costs for the financial sector.

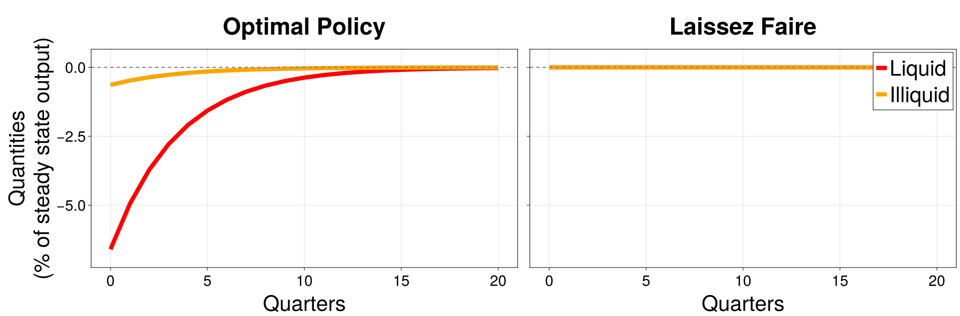

Figure 2: Quantities of assets traded by the government under the optimal policy and under no asset market operations ($\hat{\boldsymbol{b}}^G = \mathbf{0}$), expressed as % of annual steady state output.

Figure 3 shows the financial sector's response to the net worth shock under both policy regimes. Following the shock, the financial sector liquidates capital and sells off illiquid asset holdings (i.e., net issuance increases), though this is partially offset by an increase in liquid asset issuance. However, the optimal policy mitigates these adverse effects compared to laissez-faire. By reducing the financial sector's funding costs, optimal policy moderates the sell-off of illiquid holdings and prevents excessive capital liquidation relative to the laissez-faire outcome.

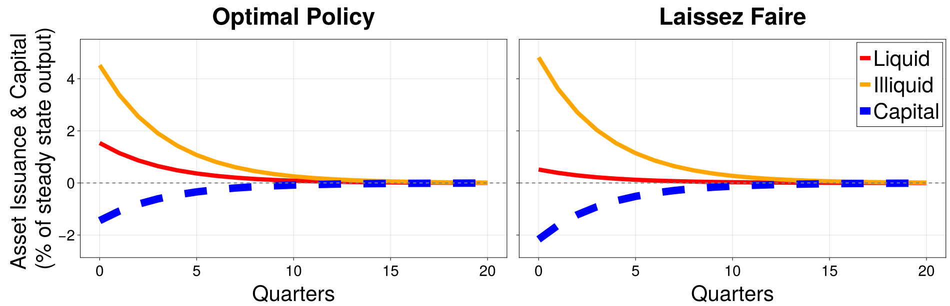

Figure 3: Financial sector's issuance of assets and quantity of capital under the optimal policy and under no asset market operations ($\hat{\boldsymbol{b}}^G = \mathbf{0}$), expressed as % of annual steady state output.

8 Conclusion

In this paper, we characterize governments' optimal responses to asset market disturbances. We show that in a large class of models with financial frictions, the Ramsey plan entails a policy rule that targets a specific relationship between asset returns, regardless of the underlying disturbances. The relationship is determined by asset supply and demand elasticities, which are measurable sufficient statistics capturing features of the economy most relevant for the design of optimal policy.

Our framework omits some key transmission mechanisms, such as time-varying risk premia or nominal rigidities. However, if these features can be characterized by properties of asset supply and demand, we believe it is possible to extend our approach to accommodate them. Another important avenue for future work concerns the empirical structure of the asset supply and demand elasticities. Current empirical works often impose strong restrictions on these elasticities—for example, assuming a static structure, which is not without loss of generality. Developing suitable empirical structures for these elasticities represents a promising avenue for bridging theoretical optimal policy with practical implementation.

References

Afonso, Gara and Ricardo Lagos, "Trade dynamics in the market for federal funds," *Econometrica*, 2015, *83* (1), 263–313.

Angeletos, George-Marios, "Fiscal policy with noncontingent debt and the optimal maturity structure," *The Quarterly Journal of Economics*, 2002, *117* (3), 1105–1131.

— **, Fabrice Collard, and Harris Dellas**, "Public Debt as Private Liquidity: Optimal Policy," *Journal of Political Economy*, 2023, *131* (11), 3233–3264.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub, "The Intertemporal Keynesian Cross," Technical Report 2023.

— **,** — **, and** — , "Fiscal and monetary policy with heterogeneous agents," *Annual Review of Economics*, 2025, *17.*

— **, Michael Cai, Matthew Rognlie, and Ludwig Straub**, "Optimal long-run fiscal policy with heterogeneous agents," 2024.

Bellifemine, Marco, Rustam Jamilov, and Tommaso Monacelli, "Hbank: Monetary policy with heterogeneous banks," 2022.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework," in "Handbook of Macroeconomics," Vol. 1, Elsevier, December 1999, chapter 21.

Bertaut, Carol C. and Ruth A. Judson, "Estimating U.S. Cross-Border Securities Flows: Ten Years of the TIC SLT," FEDS Notes 2022-02-18-2, Board of Governors of the Federal Reserve System (U.S.) February 2022.

Bianchi, Javier and Enrique G Mendoza, "Optimal time-consistent macroprudential policy," *Journal of Political Economy*, 2018, *126* (2).

— **and Saki Bigio**, "Banks, liquidity management, and monetary policy," *Econometrica*, 2022, *90* (1), 391–454.

Buera, Francisco and Juan Pablo Nicolini, "Optimal maturity of government debt without state contingent bonds," *Journal of Monetary Economics*, 2004, *51* (3), 531–554.

Caballero, Ricardo J, Tomás E Caravello, and Alp Simsek, "Financial Conditions Targeting," Technical Report, National Bureau of Economic Research 2024.

Chaudhary, Manav, Zhiyu Fu, and Haonan Zhou, " Anatomy of the Treasury Market: Who Moves Yields?," Technical Report March 2025.

Chetty, Raj, "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 2009, *1* (1), 451–488.

Chiang, Yu-Ting and Piotr Zoch, "Financial Intermediation and Aggregate Demand: A Sufficient Statistics Approach," Technical Report 2023.

Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno, "Risk Shocks," *American Economic Review*, January 2014, *104* (1), 27–65.

Cúrdia, Vasco and Michael Woodford, "The central-bank balance sheet as an instrument of monetarypolicy," *Journal of Monetary Economics*, 2011, *58* (1), 54–79. Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010.

Dávila, Eduardo and Andreas Schaab, "Optimal monetary policy with heterogeneous agents: Discretion, commitment, and timeless policy," Technical Report, National Bureau of Economic Research 2023.

_ and Anton Korinek, "Pecuniary externalities in economies with financial frictions," *The Review of Economic Studies*, 2018, *85* (1), 352–395.

de Lannoy, Leo Aparisi, Anmol Bhandari, David Evans, Mikhail Golosov, and Thomas J Sargent, "Managing Public Portfolios," Technical Report, National Bureau of Economic Research 2022.

Farhi, Emmanuel and Iván Werning, "A theory of macroprudential policies in the presence of nominal rigidities," *Econometrica*, 2016, *84* (5), 1645–1704.

Gabaix, Xavier and Ralph SJ Koijen, "In search of the origins of financial fluctuations: The inelastic markets hypothesis," Technical Report, National Bureau of Economic Research 2021.

_ and _ , "Granular instrumental variables," *Journal of Political Economy*, 2024, *132* (7), 000–000.

Gertler, Mark and Nobuhiro Kiyotaki, "Chapter 11 - Financial Intermediation and Credit Policy in Business Cycle Analysis," in "Handbook of Monetary Economics," Elsevier, 2010.

— **and Peter Karadi**, "A model of unconventional monetary policy," *Journal of monetary Economics*, 2011, *58* (1).

— **, Nobuhiro Kiyotaki, and Albert Queralto**, "Financial crises, bank risk exposure and government financial policy," *Journal of monetary economics*, 2012, *59*, S17–S34.

Giannoni, Marc P and Michael Woodford, "Optimal interest-rate rules: I. General theory," 2003.

Kekre, Rohan, Moritz Lenel, and Federico Mainardi, "Monetary policy, segmentation, and the term structure," Technical Report, National Bureau of Economic Research 2024.

Kiyotaki, Nobuhiro and John Moore, "Credit Cycles," *Journal of Political Economy*, 1997, *105* (2), 211–248.

Kleven, Henrik J, "Sufficient statistics revisited," *Annual Review of Economics*, 2021, *13* (1), 515–538.

Lorenzoni, Guido, "Inefficient credit booms," *The Review of Economic Studies*, 2008, *75* (3), 809–833.

McKay, Alisdair and Christian Wolf, "Optimal policy rules in hank," 2022.

Merton, Robert C., "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 1973, *41* (5), 867–887.

Pflueger, Carolin and Gianluca Rinaldi, "Why does the Fed move markets so much? A model of monetary policy and time-varying risk aversion," *Journal of Financial Economics*, 2022, *146* (1), 71–89.

Piazzesi, Monika, Ciaran Rogers, and Martin Schneider, "Money and banking in a New Keynesian model," *Stanford WP*, 2019.

Sims, Eric and Jing Cynthia Wu, "Evaluating central banks' tool kit: Past, present, and future," *Journal of Monetary Economics*, 2021, *118*, 135–160.

Uribe, Martin and Vivian Z. Yue, "Country spreads and emerging countries: Who drives whom?," *Journal of International Economics*, June 2006, *69* (1).

Vayanos, Dimitri and Jean-Luc Vila, "A preferred-habitat model of the term structure of interest rates," *Econometrica*, 2021, *89* (1), 77–112.

Woodford, Michael, *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2011.

A Proofs and Derivations

A.1 Aggregate Functions

The Production Sector

From the optimality of capital rental of the firms' problem, we have

$$\varrho_t = F_k(h_t, k_{t-1}, \boldsymbol{\omega}_t),$$

which defines a function $\mathring{\varrho}(h_t, k_{t-1}, \boldsymbol{\omega}_t)$.

The optimality of investment gives us

$$q_t \Phi'(\iota_t, \boldsymbol{\omega}_t) = 1,$$

which allows us to define function $\tilde{\iota}(q_t, \boldsymbol{\omega}_t)$.

From the definition of R^A_{t+1}, the law of motion for capital, and the optimality condition with respect to labor, we have

$$\begin{aligned}
&(R^A_{t+1} + \tau^A_{t+1})q_t = \varrho(h_{t+1}, k_t, \boldsymbol{\omega}_{t+1}) + [q_{t+1}\Phi(\tilde{\iota}(q_{t+1}, \boldsymbol{\omega}_{t+1}), \boldsymbol{\omega}_{t+1}) - \tilde{\iota}(q_{t+1}, \boldsymbol{\omega}_{t+1})],\\
&k_t = \Phi(\tilde{\iota}(q_t, \boldsymbol{\omega}_t), \boldsymbol{\omega}_t)k_{t-1},\\
&w_t = F_h(h_t, k_{t-1}, \boldsymbol{\omega}_t), \quad \forall t \geq 0.
\end{aligned}$$

Given initial condition k_{-1} and terminal condition $\lim_{T\to\infty}\prod_{s=0}^{T}(R^A_{s+1})^{-1}q_T = 0$, we can obtain the following functions as the solution of the system:

$$\mathring{k}_t(\{R^A_{s+1} + \tau^A_{s+1}, w_s; \boldsymbol{\omega}_s\}), \quad \mathring{q}_t(\{R^A_{s+1} + \tau^A_{s+1}, w_s; \boldsymbol{\omega}_s\}), \quad \mathring{h}_t(\{R^A_{s+1} + \tau^A_{s+1}, w_s; \boldsymbol{\omega}_s\}).$$

From labor market clearing

$$\mathring{h}_t(\{R^A_{s+1} + \tau^A_{s+1}, w_s; \boldsymbol{\omega}_s\}) = 1, \tag{22}$$

we can solve for a wage function

$$w_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}).$$

Define

$$\begin{aligned}
q_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}) &= \mathring{q}_t(\{R^A_{s+1} + \tau^A_{s+1}, w_s(\{R^A_{\ell+1} + \tau^A_{\ell+1}, \boldsymbol{\omega}_\ell\}); \boldsymbol{\omega}_s\}),\\
k_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}) &= \mathring{k}_t(\{R^A_{s+1} + \tau^A_{s+1}, w_s(\{R^A_{\ell+1} + \tau^A_{\ell+1}, \boldsymbol{\omega}_\ell\}); \boldsymbol{\omega}_s\}),
\end{aligned}$$

then we have the rental rate of capital

$$\varrho_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}) = \mathring{\varrho}(1, k_{t-1}(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\})(\cdot), \boldsymbol{\omega}_t),$$

and the value of capital supply

$$a^K_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}) \coloneqq q_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}) k_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}).$$

The Financial Sector

For $t > 0$, from the definition

$$R^N_t \coloneqq R^A_t \theta_{t-1} - \boldsymbol{R}^{B\intercal}_t \boldsymbol{\phi}_{t-1},$$

together with the definition $\boldsymbol{\phi}_t \coloneqq \boldsymbol{\phi}(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}\}, \boldsymbol{\omega}_t)$, and that $\theta_t = \mathbf{1}^\intercal_{\mathcal{B}} \boldsymbol{\phi}_t + 1$, we can write

$$R^N_{t+1} = R^N_{t+1}(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}\}, \boldsymbol{\omega}_{t+1}).$$

For $t = 0$, we have

$$(R^A_0 + \tau^A_0) q_{-1} = \varrho(1, k_{-1}, \boldsymbol{\omega}_0) + \big[q_0 \Phi(\mathring{\imath}(q_0, \boldsymbol{\omega}_0), \boldsymbol{\omega}_0) - \mathring{\imath}(q_0, \boldsymbol{\omega}_0)\big],$$

Using function $q_t(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\})$, we can obtain

$$R^A_0(\{R^A_{s+1} + \tau^A_{s+1}; \boldsymbol{\omega}_s\}).$$

Suppose that the initial returns for asset $\mathcal{B}$ are given by $R_0^B(\{R_{s+1}^B - \tau_{s+1}^B\})$, then we have

$$R_0^N(\{R_{s+1}^A + \tau_{s+1}^A, \boldsymbol{R}_{s+1}^B - \boldsymbol{\tau}_{s+1}^B; \boldsymbol{\omega}_s\}).$$

The Household Sector

Directly from the household problem, we can solve for the welfare function

$$\mathring{v}(\{\boldsymbol{R}_s^B - \boldsymbol{\tau}_s^B, R_s^N - \tau_s^N, w_s, T_s, \boldsymbol{\omega}_s\}),$$

and their asset demand functions

$$\mathring{\boldsymbol{b}}_t(\{\boldsymbol{R}_s^B - \boldsymbol{\tau}_s^B, R_s^N - \tau_s^N, w_s, T_s, \boldsymbol{\omega}_s\}), \quad \mathring{n}_t(\{\boldsymbol{R}_s^B - \boldsymbol{\tau}_s^B, R_s^N - \tau_s^N, w_s, T_s, \boldsymbol{\omega}_s\}),$$

the Lagrange multipliers on their budget constraints:

$$\mathring{\lambda}_t(\{\boldsymbol{R}_s^B - \boldsymbol{\tau}_s^B, R_s^N - \tau_s^N, w_s, T_s, \boldsymbol{\omega}_s\}).$$

Using the wage function $w_t(\cdot)$ and the return function $R_{t+1}^N(\cdot)$ and $R_0^N(\cdot)$, we can obtain

$$v(\{R_s^A, \boldsymbol{R}_s^B, \tau_s^A, \boldsymbol{\tau}_s^B, \tau_s^N, T_s, \boldsymbol{\omega}_s\}) = \mathring{v}(\{\boldsymbol{R}_s^B - \boldsymbol{\tau}_s^B, R_s^N(\cdot) - \tau_s^N, w_s(\cdot), T_s, \boldsymbol{\omega}_s\}),$$

and similarly, we can obtain functions $\boldsymbol{b}_t(\cdot)$ and $n_t(\cdot)$, and $\lambda_t(\cdot)$.

A.2 Optimality Conditions

The FOC for returns $\boldsymbol{R}_{t+1}^j, \forall j \in \{\mathcal{A}, \mathcal{B}\}$ and $\boldsymbol{b}_t^G$ are given by:

$$\begin{aligned}
[\boldsymbol{R}_{t+1}^j]: \quad & v_{\boldsymbol{R}_{t+1}^j} - (\beta^{t+1}\xi_{t+1}\boldsymbol{b}_t^{G\intercal} + \xi_0 \boldsymbol{b}_{-1}^{G\ \intercal} \boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}})\mathbf{1}_{\{j\in\mathcal{B}\}} \\
& + \sum_{s=0}^{\infty} \beta^s \Big[\boldsymbol{\gamma}_s^\intercal \boldsymbol{\epsilon}_{s,\boldsymbol{R}_{t+1}^j} + \eta_s \kappa_{s,\boldsymbol{R}_{t+1}^j} + \beta\xi_{s+1}(\boldsymbol{\tau}_{s+1}^{B\ \intercal} \boldsymbol{b}_{s,\boldsymbol{R}_{t+1}^j} + \tau_{s+1}^N n_{s,\boldsymbol{R}_{t+1}^j} + \tau_{s+1}^A a^K_{s,\boldsymbol{R}_{t+1}^j})\Big] = 0, \\
[b_t^G]: \quad & \xi_t \mathbf{1}_{\mathcal{B}}^\intercal - \beta\xi_{t+1}(\boldsymbol{R}_{t+1}^B)^\intercal + \boldsymbol{\gamma}_t^\intercal = 0.
\end{aligned}$$

By the envelope theorem, the effects of policies on the household's welfare are:

$$v_{T_t} = -\beta^t \lambda_t, \quad v_{\tau_t^A} = \sum_{s=0}^{\infty} \beta^s \lambda_s w_{s,\tau_t^A} + \lambda_0 a_{-1}^K R_{0,\tau_t^A}^A,$$
$$v_{\boldsymbol{\tau}_t^B} = -\beta^t \lambda_t \boldsymbol{b}_{t-1}{}^\intercal + \lambda_0 \boldsymbol{b}_{-1}^G{}^\intercal \boldsymbol{R}_{0,\boldsymbol{\tau}_t^B}^B, \quad v_{\tau_t^N} = -\beta^t \lambda_t n_{t-1},$$

and the effects of returns on the household's welfare are given by:

$$v_{\boldsymbol{R}_{t+1}^j} = \left(\beta^{t+1} \lambda_{t+1} \boldsymbol{b}_t{}^\intercal + \lambda_0 \boldsymbol{b}_{-1}{}^\intercal \boldsymbol{R}_{0,\boldsymbol{R}_{t+1}^j}^B\right) \mathbf{1}_{\{j\in\mathcal{B}\}} + \sum_{s=0}^{\infty} \beta^s \lambda_s \left(R_{s,\boldsymbol{R}_{t+1}^j}^N n_{s-1} + w_{s,R_{t+1}^A} \mathbf{1}_{\{j\in\mathcal{A}\}}\right),$$

where

$$R_{s,\boldsymbol{R}_{t+1}^j}^N = \tilde{R}_{s,\boldsymbol{R}_{t+1}^j}^N + \left[\theta_t \mathbf{1}_{\{j\in\mathcal{A}\}} - \boldsymbol{\phi}_t^\intercal \mathbf{1}_{\{j\in\mathcal{B}\}}\right] \mathbf{1}_{\{s=t+1\}} + \left[\theta_{-1} R_{0,R_{t+1}^A}^A \mathbf{1}_{\{j\in\mathcal{A}\}} - \boldsymbol{\phi}_{-1}^\intercal \boldsymbol{R}_{0,\boldsymbol{R}_{t+1}^B}^B \mathbf{1}_{\{j\in\mathcal{B}\}}\right] \mathbf{1}_{\{s=0\}},$$

and

$$\tilde{R}_{s,\boldsymbol{R}_{t+1}^j}^N := R_s^A \theta_{s-1,\boldsymbol{R}_{t+1}^j} - \boldsymbol{R}_s^{B\intercal} \boldsymbol{\phi}_{s-1,\boldsymbol{R}_{t+1}^j}.$$

The FOC for taxes $\mathcal{T}_t \in \{T_t, \tau_t^A, \boldsymbol{\tau}_t^B, \tau_t^N\}$ are given by:

$$[\mathcal{T}_t]: \quad v_{\mathcal{T}_t} + \beta^t \xi_t (\mathbf{1}_{\{\mathcal{T}=T\}} + \mathbf{1}_{\{\mathcal{T}=\boldsymbol{\tau}_t^B\}} \boldsymbol{b}_{t-1}{}^\intercal + \mathbf{1}_{\{\mathcal{T}=\tau^A\}} a_{t-1}^K + \mathbf{1}_{\{\mathcal{T}=\tau^N\}} n_{t-1}) - \xi_0 \boldsymbol{b}_{-1}^G{}^\intercal \boldsymbol{R}_{0,\boldsymbol{\tau}_t^B}^B \mathbf{1}_{\{\mathcal{T}=\boldsymbol{\tau}_t^B\}}$$
$$+ \sum_{s=0}^{\infty} \beta^s \left[\boldsymbol{\gamma}_s^\intercal \boldsymbol{\epsilon}_{s,\mathcal{T}_t} + \eta_s \kappa_{s,\mathcal{T}_t} + \beta \xi_{s+1} (\boldsymbol{\tau}_{s+1}^B{}^\intercal \boldsymbol{b}_{s,\mathcal{T}_t} + \tau_{s+1}^N n_{s,\mathcal{T}_t} + \tau_{s+1}^A a_{s,\mathcal{T}_t}^K)\right] = 0.$$

From the household and production problem, we have

$$\boldsymbol{b}_{s,\boldsymbol{R}_{t+1}^j} = -\boldsymbol{b}_{s,\boldsymbol{\tau}_{t+1}^B} \mathbf{1}_{\{j\in\mathcal{B}\}} + \boldsymbol{b}_{s,\tau_{t+1}^A} \mathbf{1}_{\{j\in\mathcal{A}\}} - \sum_{\ell=0}^{\infty} \boldsymbol{b}_{s,\tau_{\ell+1}^N} R_{\ell+1,\boldsymbol{R}_{t+1}^j}^N,$$
$$n_{s,\boldsymbol{R}_{t+1}^j} = -n_{s,\boldsymbol{\tau}_{t+1}^B} \mathbf{1}_{\{j\in\mathcal{B}\}} + n_{s,\tau_{t+1}^A} \mathbf{1}_{\{j\in\mathcal{A}\}} - \sum_{\ell=0}^{\infty} n_{s,\tau_{\ell+1}^N} R_{\ell+1,\boldsymbol{R}_{t+1}^j}^N.$$

Similarly, for the financial sector,

$$\boldsymbol{b}_{s,\boldsymbol{R}_{t+1}^j}^F = -\boldsymbol{b}_{s,\boldsymbol{\tau}_{t+1}^B}^F \mathbf{1}_{\{j\in\mathcal{B}\}} + \boldsymbol{b}_{s,\tau_{t+1}^A}^F \mathbf{1}_{\{j\in\mathcal{A}\}} + \boldsymbol{\phi}_{s,\boldsymbol{R}_{t+1}^j} n_s - \sum_{\ell=0}^{\infty} \boldsymbol{b}_{s,\tau_{\ell+1}^N}^F R_{\ell+1,\boldsymbol{R}_{t+1}^j}^N,$$
$$a_{s,\boldsymbol{R}_{t+1}^j}^F = -a_{s,\boldsymbol{\tau}_{t+1}^B}^F \mathbf{1}_{\{j\in\mathcal{B}\}} + a_{s,\tau_{t+1}^A}^F \mathbf{1}_{\{j\in\mathcal{A}\}} + \theta_{s,\boldsymbol{R}_{t+1}^j} n_s - \sum_{\ell=0}^{\infty} a_{s,\tau_{\ell+1}^N}^F R_{\ell+1,\boldsymbol{R}_{t+1}^j}^N.$$

Together, we have

$$\boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} = -\boldsymbol{\epsilon}_{s,\boldsymbol{\tau}^B_{t+1}}\mathbf{1}_{\{j\in\mathcal{B}\}} + \boldsymbol{\epsilon}_{s,\tau^A_{t+1}}\mathbf{1}_{\{j\in\mathcal{A}\}} + \boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}} n_s - \sum_{\ell=0}^{\infty} \boldsymbol{\epsilon}_{s,\tau^N_{\ell+1}} R^N_{\ell+1,\boldsymbol{R}^j_{t+1}},$$

$$\kappa_{s,\boldsymbol{R}^j_{t+1}} = -\kappa_{s,\boldsymbol{\tau}^B_{t+1}}\mathbf{1}_{\{j\in\mathcal{B}\}} + \kappa_{s,\tau^A_{t+1}}\mathbf{1}_{\{j\in\mathcal{A}\}} - \theta_{s,\boldsymbol{R}^j_{t+1}} n_s - \sum_{\ell=0}^{\infty} \kappa_{s,\tau^N_{\ell+1}} R^N_{\ell+1,\boldsymbol{R}^j_{t+1}}.$$

Substituting the elasticities with respect to returns, the FOCs for taxes, and the expressions for $v_{\tau^A_{t+1}}, v_{\boldsymbol{\tau}^B_{t+1}}, v_{\tau^N_{t+1}}$, we have

$$\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & \lambda_0\big[n_{-1}R^N_{0,\boldsymbol{R}^j_{t+1}} - {\boldsymbol{b}^F_{-1}}^{\intercal}\boldsymbol{R}^B_{0,\boldsymbol{\tau}^B_{t+1}}\mathbf{1}_{\{j\in\mathcal{B}\}} - a^K_{-1}R^A_{0,\tau^A_{t+1}}\mathbf{1}_{\{j\in\mathcal{A}\}}\big] \\
& + \sum_{s=0}^{\infty}\beta^s\Big[\boldsymbol{\gamma}_s^{\intercal}(\boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}} n_s) - \eta_s(\theta_{s,\boldsymbol{R}^j_{t+1}} n_s)\Big] \\
& + \beta^{t+1}\xi_{t+1}({\boldsymbol{b}^F_t}^{\intercal}\mathbf{1}_{\{j\in\mathcal{B}\}} - a^K_t\mathbf{1}_{\{j\in\mathcal{A}\}}) + \sum_{\ell=0}^{\infty}\beta^{\ell+1}\xi_{\ell+1}n_\ell R^N_{\ell+1,\boldsymbol{R}^j_{t+1}} = 0.
\end{aligned}$$

Because $R^N_{0,\boldsymbol{R}^j_{t+1}} = a^K_{-1}R^A_{0,R^A_{t+1}} - {\boldsymbol{b}^F_{-1}}^{\intercal}\boldsymbol{R}^B_{0,R^A_{t+1}}$ and

$$\beta^{t+1}\xi_{t+1}({\boldsymbol{b}^F_t}^{\intercal}\mathbf{1}_{\{j\in\mathcal{B}\}} - a^K_t\mathbf{1}_{\{j\in\mathcal{A}\}}) + \sum_{\ell=0}^{\infty}\beta^{\ell+1}\xi_{\ell+1}n_\ell R^N_{\ell+1,\boldsymbol{R}^j_{t+1}} = \sum_{\ell=0}^{\infty}\beta^{\ell+1}\xi_{\ell+1}n_\ell \tilde{R}^N_{\ell+1,\boldsymbol{R}^j_{t+1}},$$

we have

$$[\boldsymbol{R}^j_{t+1}]: \quad \sum_{s=0}^{\infty}\beta^s\Big[\boldsymbol{\gamma}_s^{\intercal}(\boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}} n_s) - \eta_s(\theta_{s,\boldsymbol{R}^j_{t+1}} n_s)\Big] + \sum_{\ell=0}^{\infty}\beta^{\ell+1}\xi_{\ell+1}n_\ell \tilde{R}^N_{\ell+1,\boldsymbol{R}^j_{t+1}} = 0.$$

Use the expression for $\tilde{R}^N_{\ell,\boldsymbol{R}^j_{t+1}}$,

$$[\boldsymbol{R}^j_{t+1}]: \quad \sum_{s=0}^{\infty}\beta^s\Big[\boldsymbol{\gamma}_s^{\intercal}\boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}} - \eta_s\theta_{s,\boldsymbol{R}^j_{t+1}} + \beta\xi_{s+1}(R^A_{s+1}\theta_{s,\boldsymbol{R}^j_{t+1}} - {\boldsymbol{R}^B_{s+1}}^{\intercal}\boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}})\Big]n_s = 0.$$

From FOC $[b^G_t]$, we have

$$\beta\xi_{t+1}{\boldsymbol{R}^B_{t+1}}^{\intercal} - \boldsymbol{\gamma}_t^{\intercal} = \xi_t\mathbf{1}^{\intercal}_{\mathcal{B}},$$

and

$$[R^A_{t+1}]: \quad \sum_{s=0}^{\infty} \beta^s \Big[-\xi_s \mathbf{1}_{\mathcal{B}}^{\intercal} \boldsymbol{\phi}_{s,R^A_{t+1}} + (\beta\xi_{s+1}R^A_{s+1} - \eta_s)\theta_{s,R^A_{t+1}} \Big] n_s = 0.$$

From the financial sector's balance sheet, we have $\mathbf{1}_{\mathcal{B}}^{\intercal} \boldsymbol{\phi}_{s,R^A_{t+1}} = \theta_{s,R^A_{t+1}}$ and

$$[R^A_{t+1}]: \quad \sum_{s=0}^{\infty} \beta^s \Big[-\xi_s + (\beta\xi_{s+1}R^A_{s+1} - \eta_s) \Big] \theta_{s,R^A_{t+1}} n_s = 0.$$

Suppose that $\boldsymbol{\theta}_{\boldsymbol{R}^A} = [\theta_{s,R^A_{t+1}}]$ is full-rank, then we have

$$\beta\xi_{s+1}R^A_{s+1} - \eta_s = \xi_s, \tag{23}$$

and FOC $[\boldsymbol{R}^B_{t+1}]$ holds trivially.

From the FOCs for taxes and (23), we have

$$\xi_t = \lambda_t,\ \boldsymbol{\tau}^B_{t+1} = (\beta\xi_{t+1})^{-1}\boldsymbol{\gamma}_t,\ \tau^A_{t+1} = -(\beta\xi_{t+1})^{-1}\eta_t,\ \tau^N_{t+1} = (\beta\xi_{t+1})^{-1}(\eta_t^{\intercal}\theta_t - \boldsymbol{\gamma}_t^{\intercal}\boldsymbol{\phi}_t).$$

A.3 Ramsey steady state

The Ramsey steady state satisfies:

$$\begin{gathered}\bar{\xi} = \bar{\lambda},\ \bar{\boldsymbol{\tau}}^B = (\beta\bar{\xi})^{-1}\bar{\boldsymbol{\gamma}},\ \bar{\tau}^A = -(\beta\bar{\xi})^{-1}\bar{\eta},\ \bar{\tau}^N = (\beta\bar{\xi})^{-1}(\bar{\eta}\bar{\theta} - \bar{\boldsymbol{\gamma}}^{\intercal}\bar{\boldsymbol{\phi}}),\\ \beta(\bar{R}^A + \bar{\tau}^A) = 1, \quad \beta(\bar{\boldsymbol{R}}^B - \bar{\boldsymbol{\tau}}^B) = \mathbf{1}_{\mathcal{B}},\end{gathered}$$

and

$$\begin{aligned} [fin\ mkt]: &\quad \bar{\boldsymbol{b}}^G + \bar{\boldsymbol{b}}^F = \bar{\boldsymbol{b}},\\ [cap\ mkt]: &\quad \bar{a}^K = \bar{a}^F,\\ [gov\ bgt]: &\quad \mathbf{1}_{\mathcal{B}}^{\intercal}\bar{\boldsymbol{b}}^G + \bar{T}_t + \bar{\boldsymbol{\tau}}^{B\intercal}\bar{\boldsymbol{b}} + \bar{\tau}^N\bar{n} + \bar{\tau}^A\bar{a}^K = \bar{\boldsymbol{R}}^{B\intercal}\bar{\boldsymbol{b}}^G. \end{aligned}$$

A.4 First Order Responses

$$[b^G_t]: \quad \hat{\xi}_t\mathbf{1}_{\mathcal{B}}^{\intercal} - \beta\hat{\xi}_{t+1}\bar{\boldsymbol{R}}^{B\intercal} - \beta\bar{\xi}\hat{\boldsymbol{R}}^{B}_{t+1}{}^{\intercal} + \hat{\boldsymbol{\gamma}}_t^{\intercal} = 0$$

$$[T_t]: \quad \beta^t(-\hat{\lambda}_t + \hat{\xi}_t) + \sum_{s=0}^{\infty} \beta^s \Big[\hat{\boldsymbol{\gamma}}_s^\intercal \boldsymbol{\epsilon}_{s,T_t} + \hat{\eta}_s \kappa_{s,T_t} + \boldsymbol{\gamma}^\intercal \hat{\boldsymbol{\epsilon}}_{s,T_t} + \eta \hat{\kappa}_{s,T_t}\Big]$$

$$+ \sum_{s=0}^{\infty} \beta^{s+1} \Big[\hat{\xi}_{s+1}(\boldsymbol{\tau}^{B\intercal} \boldsymbol{b}_{s,T_t} + \tau^N n_{s,T_t} + \tau^A a^K_{s,T_t}) + \xi(\boldsymbol{\tau}^{B\intercal} \hat{\boldsymbol{b}}_{s,T_t} + \tau^N \hat{n}_{s,T_t} + \tau^A_{s+1} \hat{a}^K_{s,T_t})\Big] = 0.$$

$$[\boldsymbol{R}^j_{t+1}]: \quad \hat{v}_{\boldsymbol{R}^j_{t+1}} - \beta^{t+1}(\hat{\xi}_{t+1} \bar{\boldsymbol{b}}_t^{G\intercal} + \hat{\xi}_0 {\boldsymbol{b}^G_{-1}}^\intercal \boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}} + \bar{\xi}_{t+1} \hat{\boldsymbol{b}}_t^{G\intercal} + \bar{\xi}_0 {\boldsymbol{b}^G_{-1}}^\intercal \hat{\boldsymbol{R}}^B_{0,\boldsymbol{R}^B_{t+1}}) \mathbf{1}_{\{j\in\mathcal{B}\}}$$

$$+ \sum_{s=0}^{\infty} \beta^s \Big[\hat{\boldsymbol{\gamma}}_s^\intercal \boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + \hat{\eta}_s \kappa_{s,\boldsymbol{R}^j_{t+1}} + \boldsymbol{\gamma}^\intercal \hat{\boldsymbol{\epsilon}}_{s,\boldsymbol{R}^j_{t+1}} + \eta \hat{\kappa}_{s,\boldsymbol{R}^j_{t+1}}\Big]$$

$$+ \sum_{s=0}^{\infty} \beta^{s+1} \hat{\xi}_{s+1}(\boldsymbol{\tau}^{B\intercal} \boldsymbol{b}_{s,\boldsymbol{R}^j_{t+1}} + \tau^N n_{s,\boldsymbol{R}^j_{t+1}} + \tau^A a^K_{s,\boldsymbol{R}^j_{t+1}})$$

$$+ \sum_{s=0}^{\infty} \beta^{s+1} \xi(\boldsymbol{\tau}^{B\intercal} \hat{\boldsymbol{b}}_{s,\boldsymbol{R}^j_{t+1}} + \tau^N \hat{n}_{s,\boldsymbol{R}^j_{t+1}} + \tau^A \hat{a}^K_{s,\boldsymbol{R}^j_{t+1}}) = 0,$$

where

$$\hat{v}_{\boldsymbol{R}^j_{t+1}} = \beta^{t+1} \bigg[\Big(\hat{\lambda}_{t+1} \boldsymbol{b}^{G\intercal} + \lambda \hat{\boldsymbol{b}}_t^{G\intercal}\Big) \mathbf{1}_{\{j\in\mathcal{B}\}} + \Big(\hat{\lambda}_{t+1} a^K + \lambda \hat{a}^K_t + \sum_{s=0}^{\infty} \beta^s (\hat{\lambda}_s w_{s,R^A_{t+1}} + \lambda \hat{w}_{s,R^A_{t+1}})\Big) \mathbf{1}_{\{j\in\mathcal{A}\}}\bigg]$$

$$+ \Big(\lambda \boldsymbol{b}^{G\intercal} \hat{\boldsymbol{R}}^B_{0,\boldsymbol{R}^B_{t+1}} + \hat{\lambda}_0 \boldsymbol{b}^{G\intercal} \boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}}\Big) \mathbf{1}_{\{j\in\mathcal{B}\}} + \Big(\lambda a^K \hat{R}^A_{0,R^A_{t+1}} + a^K R^A_{0,R^A_{t+1}}\Big) \mathbf{1}_{\{j\in\mathcal{A}\}}$$

$$+ \sum_{s=0}^{\infty} \beta^{s+1} \Big(\hat{\lambda}_{s+1} \tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}} n + \lambda \hat{\tilde{R}}^N_{s+1,\boldsymbol{R}^j_{t+1}} n + \lambda \tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}} \hat{n}_s\Big).$$

Using $\beta(R^A + \tau^A) = 1, \quad \beta(\boldsymbol{R}^B - \boldsymbol{\tau}^B) = \mathbf{1}_{\mathcal{B}}$ and

$$\begin{aligned}
&{\boldsymbol{\tau}^B_{s+1}}^\intercal \boldsymbol{b}_{s,T_t} + \tau^N_{s+1} n_{s,T_t} + \tau^A_{s+1} a^K_{s,T_t} = \tau^A_s \kappa_{s,T_t} - \boldsymbol{\tau}_s^{B\intercal} \boldsymbol{\epsilon}_{s,T_t},\\
&\boldsymbol{\tau}^{B\intercal} \boldsymbol{b}_{s,\boldsymbol{R}^j_{t+1}} + \tau^N n_{s,\boldsymbol{R}^j_{t+1}} + \tau^A a^K_{s,\boldsymbol{R}^j_{t+1}} = \tau^A_{s+1} \kappa_{s,\boldsymbol{R}^j_{t+1}} - {\boldsymbol{\tau}^B_{s+1}}^\intercal \boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} - \tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}} n_s,
\end{aligned}$$

we have

$$[T_t]: \quad \beta^t(-\hat{\lambda}_t + \hat{\xi}_t)$$

$$+ \sum_{s=0}^{\infty} \beta^s \Big[(\hat{\boldsymbol{\gamma}}_s - \beta \hat{\xi}_{s+1} \boldsymbol{\tau}^B)^\intercal \boldsymbol{\epsilon}_{s,T_t} + (\hat{\eta}_s + \beta \hat{\xi}_{s+1} \tau^A) \kappa_{s,T_t} + (\boldsymbol{\gamma}^\intercal \hat{\boldsymbol{\phi}}_s - \eta \hat{\theta}_s) n_{s,T_t}\Big] = 0,$$

and

$$
\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & \hat{v}_{\boldsymbol{R}^j_{t+1}} - \beta^{t+1}(\hat{\xi}_{t+1}\boldsymbol{b}^{G\intercal} + \hat{\xi}_0 \boldsymbol{b}^{G\intercal}\boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}} + \xi \hat{\boldsymbol{b}}^{G\intercal}_t + \bar{\xi}_0 \boldsymbol{b}^{G\intercal}\hat{\boldsymbol{R}}^B_{0,\boldsymbol{R}^B_{t+1}})\mathbf{1}_{\{j\in\mathcal{B}\}} \\
& + \sum_{s=0}^{\infty}\beta^s\Big[(\hat{\boldsymbol{\gamma}}_s - \beta\hat{\xi}_{s+1}\boldsymbol{\tau}^B)^\intercal \boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + (\hat{\eta}_s + \beta\hat{\xi}_{s+1}\tau^A)\kappa_{s,\boldsymbol{R}^j_{t+1}} - \beta\hat{\xi}_{s+1}\tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}} n_s\Big] \\
& + \sum_{s=0}^{\infty}\beta^s\Big[\boldsymbol{\gamma}^\intercal(\hat{\boldsymbol{\phi}}_{s,\boldsymbol{R}^j_{t+1}} n + \boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}}\hat{n}_s + \hat{\boldsymbol{\phi}}_s n_{s,\boldsymbol{R}^j_{t+1}}) + \eta(-\hat{\theta}_{s,\boldsymbol{R}^j_{t+1}} n - \theta_{s,\boldsymbol{R}^j_{t+1}}\hat{n}_s - \hat{\theta}_s n_{s,\boldsymbol{R}^j_{t+1}})\Big] \\
& + \sum_{s=0}^{\infty}\beta^{s+1}\hat{\xi}_{s+1}(\boldsymbol{\tau}^{B\intercal}\boldsymbol{b}_{s,\boldsymbol{R}^j_{t+1}} + \tau^N n_{s,\boldsymbol{R}^j_{t+1}} + \tau^A a^K_{s,\boldsymbol{R}^j_{t+1}}) = 0.
\end{aligned}
$$

Note that $\beta(R^A + \tau^A) = 1$ and $\beta(\boldsymbol{R}^B - \boldsymbol{\tau}^B) = \mathbf{1}_{\mathcal{B}}$ imply

$$
\begin{aligned}
& (\boldsymbol{\gamma}^\intercal \boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}} - \eta\theta_{s,\boldsymbol{R}^j_{t+1}} + \beta\lambda\tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}})\hat{n}_s = 0, \\
& \boldsymbol{\gamma}^\intercal\hat{\boldsymbol{\phi}}_{s,\boldsymbol{R}^j_{t+1}} - \eta\hat{\theta}_{s,\boldsymbol{R}^j_{t+1}} + \beta\lambda\hat{\tilde{R}}^N_{s+1,\boldsymbol{R}^j_{t+1}} = \beta\lambda(\hat{R}^A_{s+1}\theta_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal \boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}}).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & \Big[\beta^{t+1}(\hat{\lambda}_{t+1} - \hat{\xi}_{t+1})\boldsymbol{b}^{G\intercal} + (\hat{\lambda}_0 - \hat{\xi}_0)\boldsymbol{b}^{G\intercal}\boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}}\Big]\mathbf{1}_{\{j\in\mathcal{B}\}} \\
& + \Big[\beta^{t+1}(\hat{\lambda}_{t+1}a^K + \lambda\hat{a}^K_t) + (\hat{\lambda}_0 R^A_{0,R^A_{t+1}} + \lambda\hat{R}^A_{0,R^A_{t+1}})a^K + \sum_{s=0}^{\infty}\beta^s(\hat{\lambda}_s w_{s,R^A_{t+1}} + \lambda\hat{w}_{s,R^A_{t+1}})\Big]\mathbf{1}_{\{j\in\mathcal{A}\}} \\
& + \sum_{s=0}^{\infty}\beta^{s+1}\Big[(\hat{\lambda}_{s+1} - \hat{\xi}_{s+1})\tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}} n + \beta^{s+1}\lambda(\hat{R}^A_{s+1}\theta_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal\boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}})n\Big] \\
& + \sum_{s=0}^{\infty}\beta^s\Big[(\hat{\boldsymbol{\gamma}}_s - \beta\hat{\xi}_{s+1}\boldsymbol{\tau}^B_{s+1})^\intercal\boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + (\hat{\eta}_s - \beta\hat{\xi}_{s+1}\tau^A_{s+1})\kappa_{s,\boldsymbol{R}^j_{t+1}} + (\boldsymbol{\gamma}^\intercal_s\hat{\boldsymbol{\phi}}_s - \eta_s\hat{\theta}_s)n_{s,\boldsymbol{R}^j_{t+1}}\Big] = 0.
\end{aligned}
$$

Because $R^A\hat{\theta}_s - \boldsymbol{R}^{B\intercal}\hat{\boldsymbol{\phi}}_s = (\beta\xi)^{-1}(\eta\hat{\theta}_s - \boldsymbol{\gamma}^\intercal\hat{\boldsymbol{\phi}}_s)$, we have

$$
\hat{R}^N_{s+1} = \hat{R}^A_{s+1}\theta - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal\boldsymbol{\phi} + R^A\hat{\theta}_s - \boldsymbol{R}^{B\intercal}\hat{\boldsymbol{\phi}}_s = \hat{R}^A_{s+1}\theta - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal\boldsymbol{\phi} + (\beta\xi)^{-1}(\eta\hat{\theta}_s - \boldsymbol{\gamma}^\intercal\hat{\boldsymbol{\phi}}_s),
$$

and

$$
\boldsymbol{\gamma}^\intercal\hat{\boldsymbol{\phi}}_s - \eta\hat{\theta}_s = \beta\xi\Big[\hat{R}^A_{s+1}\theta - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal\boldsymbol{\phi} - \hat{R}^N_{s+1}\Big].
$$

Substitute $\boldsymbol{\gamma}_s^\intercal \hat{\boldsymbol{\phi}}_s - \eta_s \hat{\theta}_s$, note that

$$a^F_{s,\boldsymbol{R}^j_{t+1}} = \theta_{s,\boldsymbol{R}^j_{t+1}} n_s + \theta_s n_{s,\boldsymbol{R}^j_{t+1}},\ \boldsymbol{b}^F_{s,\boldsymbol{R}^j_{t+1}} = \boldsymbol{\phi}_{s,\boldsymbol{R}^j_{t+1}} n_s + \boldsymbol{\phi}_s n_{s,\boldsymbol{R}^j_{t+1}},\ a^F_{s,T_t} = \theta_s n_{s,T_t},\ \boldsymbol{b}^F_{s,T_t} = \boldsymbol{\phi}_s n_{s,T_t},$$

and define

$$\tilde{\boldsymbol{\gamma}}_s \coloneqq \hat{\boldsymbol{\gamma}}_s - \beta \hat{\xi}_{s+1} \boldsymbol{\tau}^B, \quad \tilde{\eta}_s \coloneqq \hat{\eta}_s + \beta \hat{\xi}_{s+1} \tau^A, \quad \tilde{\xi}_t = \hat{\xi}_t - \hat{\lambda}_t.$$

The FOC for returns can be written as

$$\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & -[\beta^{t+1}\tilde{\xi}_{t+1}\boldsymbol{b}^{G\intercal} + \tilde{\xi}_0 \boldsymbol{b}^{G\intercal} \boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}}]\mathbf{1}_{\{j\in\mathcal{B}\}} - \sum_{s=0}^{\infty} \beta^{s+1} \tilde{\xi}_{s+1} \tilde{R}^N_{s+1,\boldsymbol{R}^j_{t+1}} n \\
& + \left[\beta^{t+1}(\hat{\lambda}_{t+1} a^K + \lambda \hat{a}^K_t) + (\hat{\lambda}_0 R^A_{0,R^A_{t+1}} + \lambda \hat{R}^A_{0,R^A_{t+1}}) a^K + \sum_{s=0}^{\infty} \beta^s (\hat{\lambda}_s w_{s,R^A_{t+1}} + \lambda \hat{w}_{s,R^A_{t+1}})\right] \mathbf{1}_{\{j\in\mathcal{A}\}} \\
& + \sum_{s=0}^{\infty} \beta^{s+1} \lambda \left[\hat{R}^A_{s+1} a^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal \boldsymbol{b}^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{R}^N_{s+1} n_{s,\boldsymbol{R}^j_{t+1}}\right] \\
& + \sum_{s=0}^{\infty} \beta^s \left[\tilde{\boldsymbol{\gamma}}_s^\intercal \boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + \tilde{\eta}_s \kappa_{s,\boldsymbol{R}^j_{t+1}}\right] = 0.
\end{aligned}$$

Similarly, we have

$$[T_t]: \quad \beta^t \tilde{\xi}_t + \sum_{s=0}^{\infty} \beta^s \left[\tilde{\boldsymbol{\gamma}}_s^\intercal \boldsymbol{\epsilon}_{s,T_t} + \tilde{\eta}_s \kappa_{s,T_t}\right] + \beta^{s+1} \xi \left[\hat{R}^A_{s+1} a^F_{s,T_t} - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal \boldsymbol{b}^F_{s,T_t} - \hat{R}^N_{s+1} n_{s,T_t}\right] = 0,$$

and

$$[b^G_t]: \quad \tilde{\boldsymbol{\gamma}}_t^\intercal - \beta \bar{\xi} \hat{\boldsymbol{R}}^B_{t+1}{}^\intercal - (\tilde{\xi}_{t+1} - \tilde{\xi}_t)\mathbf{1}^\intercal_{\mathcal{B}} = (\hat{\lambda}_{t+1} - \hat{\lambda}_t)\mathbf{1}^\intercal_{\mathcal{B}}.$$

We now simplify the expression for

$$\beta^{t+1}(\hat{\lambda}_{t+1} a^K + \lambda \hat{a}^K_t) + (\hat{\lambda}_0 R^A_{0,R^A_{t+1}} + \lambda \hat{R}^A_{0,R^A_{t+1}}) a^K + \sum_{s=0}^{\infty} \beta^s (\hat{\lambda}_s w_{s,R^A_{t+1}} + \lambda \hat{w}_{s,R^A_{t+1}})$$

First, for $a^K_{-1} R^A_{0,R^A_{t+1}}$, note that because capital is predetermined in the initial period and labor is inelastic, we have $\varrho_{0,R^A_{t+1}} = 0$ and

$$q_{-1} R^A_{0,R^A_{t+1}} = \boldsymbol{\phi}(\iota_0) q_{0,R^A_{t+1}},$$

which implies

$$a_{-1}^K R^A_{0,R^A_{t+1}} = k_0 q_{0,R^A_{t+1}}.$$

In order to link $q_{0,R^A_{t+1}}$ to future capital income, we use equation (5) and take derivatives with respect to R^A_{t+1}:

$$q_s \mathbf{1}_{\{s=t\}} + (R^A_{s+1} + \tau^A_{s+1}) q_{s,R^A_{t+1}} = \varrho_{s+1,R^A_{t+1}} + \Phi(\iota_{s+1}) q_{s+1,R^A_{t+1}}, \forall s \geq 0.$$

Multiply both sides by $(R^A_{s+1} + \tau^A_{s+1})^{-1} k_s$ and use $\Phi(\iota_{s+1}) k_s = k_{s+1}$, we have

$$(R^A_{s+1} + \tau^A_{s+1})^{-1} k_s q_s \mathbf{1}_{\{s=t\}} + k_s q_{s,R^A_{t+1}} = (R^A_{s+1} + \tau^A_{s+1})^{-1} (k_s \varrho_{s+1,R^A_{t+1}} + k_{s+1} q_{s+1,R^A_{t+1}}).$$

Let $R^A_{0,l+1} = \prod_{s=0}^{l} \left(R^A_{s+1} + \tau^A_{s+1}\right)$ and use $k_l \varrho_{l+1,R^A_{t+1}} = -w_{l+1,R^A_{t+1}}$ from the optimality of production,[7] we have

$$a_{-1}^K R^A_{0,R^A_{t+1}} = -\sum_{s=0}^{\infty} (R^A_{0,s+1})^{-1} \Big(w_{s+1,R^A_{t+1}} + a_t^K \mathbf{1}_{\{s=t\}} \Big),$$

and the following first-order expansion:

$$a_{-1}^K \hat{R}^A_{0,R^A_{t+1}} = \sum_{s=0}^{\infty} \beta^{s+2} \sum_{l=0}^{s} \hat{R}^A_{l+1} \Big(w_{s+1,R^A_{t+1}} + a_t^K \mathbf{1}_{\{s=t\}} \Big) - \sum_{s=0}^{\infty} \beta^{s+1} \Big(\hat{w}_{s+1,R^A_{t+1}} + \hat{a}_t^K \mathbf{1}_{\{s=t\}} \Big).$$

Using these expressions in the FOC for returns, we have:

$$\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & - [\beta^{t+1} \tilde{\xi}_{t+1} \boldsymbol{b}_t^{G\intercal} + \tilde{\xi}_0 \boldsymbol{b}_{-1}^{G\,\intercal} \boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}}] \mathbf{1}_{\{j \in \mathcal{B}\}} - \sum_{s=0}^{\infty} \beta^s \tilde{\xi}_s \tilde{R}^N_{s,\boldsymbol{R}^j_{t+1}} n_{s-1} \\
& + \sum_{s=0}^{\infty} \beta^{s+1} \Big[\Big(\hat{\lambda}_{s+1} - \hat{\lambda}_0 + \beta\lambda \sum_{l=0}^{s} \hat{R}^A_{l+1} \Big) \Big(w_{s+1,R^A_{t+1}} + a_t^K \mathbf{1}_{\{s=t\}} \Big) \Big] \mathbf{1}_{\{j \in \mathcal{A}\}} \\
& + \sum_{s=0}^{\infty} \beta^s \Big[\tilde{\boldsymbol{\gamma}}_s^\intercal \boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + \tilde{\eta}_s \kappa_{s,\boldsymbol{R}^j_{t+1}} + \beta \xi_{s+1} \Big(\hat{R}^A_{s+1} a^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^{B\,\intercal}_{s+1} \boldsymbol{b}^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{R}^N_{s+1} n_{s,\boldsymbol{R}^j_{t+1}} \Big) \Big] = 0.
\end{aligned}$$

Let $\tilde{\boldsymbol{\iota}}_{\boldsymbol{b}}$ be a vector of size $|\mathcal{B}|$ that sums to one, then the FOC for $\boldsymbol{b}^G$ implies

$$[b_t^G]: \quad (\tilde{\boldsymbol{\gamma}}_t - \beta \bar{\xi} \hat{\boldsymbol{R}}^B_{t+1})^\intercal \tilde{\boldsymbol{\iota}}_{\boldsymbol{b}} - (\tilde{\xi}_{t+1} - \tilde{\xi}_t) = \hat{\lambda}_{t+1} - \hat{\lambda}_t,$$

[7] Because F is constant-return-to-scale, we have $w_t = F_t - \varrho_t k_t$ and $w_{s,R^A_{t+1}} = -\varrho_{s,R^A_{t+1}} k_s$ where we use $F_{t,k} = \varrho_t$.

and therefore

$$
\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & -[\beta^{t+1}\tilde{\xi}_{t+1}\boldsymbol{b}^{G\intercal}_t + \tilde{\xi}_0 \boldsymbol{b}^{G}_{-1}{}^{\intercal}\boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}}]\mathbf{1}_{\{j\in\mathcal{B}\}} - \sum_{s=0}^{\infty}\beta^s\tilde{\xi}_s\tilde{R}^N_{s,\boldsymbol{R}^j_{t+1}}n_{s-1} \\
& - \Big[\sum_{s=0}^{\infty}\beta^{s+1}(\tilde{\xi}_{s+1}-\tilde{\xi}_0)\Big(w_{s+1,R^A_{t+1}} + a^K_t\mathbf{1}_{\{s=t\}}\Big)\Big]\mathbf{1}_{\{j\in\mathcal{A}\}} \\
& + \Big[\sum_{s=0}^{\infty}\beta^{s+1}\sum_{l=0}^{s}\Big(\beta\lambda\hat{R}^A_{l+1} - (\beta\lambda\hat{\boldsymbol{R}}^B_{l+1} - \tilde{\boldsymbol{\gamma}}_l)^{\intercal}\tilde{\boldsymbol{\iota}}_{\boldsymbol{b}}\Big)\Big(w_{s+1,R^A_{t+1}} + a^K_t\mathbf{1}_{\{s=t\}}\Big)\Big]\mathbf{1}_{\{j\in\mathcal{A}\}} \\
& + \sum_{s=0}^{\infty}\beta^s\Big[\tilde{\boldsymbol{\gamma}}_s^{\intercal}\boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + \tilde{\eta}_s\kappa_{s,\boldsymbol{R}^j_{t+1}} + \beta\xi_{s+1}\Big(\hat{R}^A_{s+1}a^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^{\intercal}\boldsymbol{b}^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{R}^N_{s+1}n_{s,\boldsymbol{R}^j_{t+1}}\Big)\Big] = 0.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\sum_{s=0}^{\infty}\beta^{s+1}\sum_{l=0}^{s}(\beta\lambda\hat{R}^A_{l+1} - (\beta\lambda\hat{\boldsymbol{R}}^B_{l+1} - \tilde{\boldsymbol{\gamma}}_l)^{\intercal}\tilde{\boldsymbol{\iota}}_{\boldsymbol{b}})(w_{s+1,R^A_{t+1}} + a^K_t\mathbf{1}_{\{s=t\}}) \\
&= \sum_{s=0}^{\infty}(\beta\lambda\hat{R}^A_{s+1} - (\beta\lambda\hat{\boldsymbol{R}}^B_{s+1} - \tilde{\boldsymbol{\gamma}}_s)^{\intercal}\tilde{\boldsymbol{\iota}}_{\boldsymbol{b}})\sum_{l=s}^{\infty}\beta^{l+1}(w_{l+1,R^A_{t+1}} + a^K_t\mathbf{1}_{\{l=t\}}).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & -[\beta^{t+1}\tilde{\xi}_{t+1}\boldsymbol{b}^{G\intercal}_t + \tilde{\xi}_0 \boldsymbol{b}^{G}_{-1}{}^{\intercal}\boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}}]\mathbf{1}_{\{j\in\mathcal{B}\}} - \sum_{s=0}^{\infty}\beta^s\tilde{\xi}_s\tilde{R}^N_{s,\boldsymbol{R}^j_{t+1}}n_{s-1} \\
& - \sum_{s=0}^{\infty}\beta^{s+1}\Big[(\tilde{\xi}_{s+1}-\tilde{\xi}_0)(w_{s+1,R^A_{t+1}} + a^K_t\mathbf{1}_{\{s=t\}}) \\
& \qquad\qquad + (\beta\lambda\hat{R}^A_{s+1} - (\beta\lambda\hat{\boldsymbol{R}}^B_{s+1} - \tilde{\boldsymbol{\gamma}}_s)^{\intercal}\tilde{\boldsymbol{\iota}}_{\boldsymbol{b}})\boldsymbol{z}^K_{s+1,R^A_{t+1}}\Big]\mathbf{1}_{\{j\in\mathcal{A}\}} \\
& + \sum_{s=0}^{\infty}\beta^s\Big[\tilde{\boldsymbol{\gamma}}_s^{\intercal}\boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + \tilde{\eta}_s\kappa_{s,\boldsymbol{R}^j_{t+1}} + \beta\xi_{s+1}\Big(\hat{R}^A_{s+1}a^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^{\intercal}\boldsymbol{b}^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{R}^N_{s+1}n_{s,\boldsymbol{R}^j_{t+1}}\Big)\Big] = 0,
\end{aligned}
$$

where

$$
\tilde{\boldsymbol{z}}^K_{s+1,R^A_{t+1}} \coloneqq -\sum_{l=s}^{\infty}\beta^{l-s}\Big(w_{l+1,R^A_{t+1}} + a^K_t\mathbf{1}_{\{l=t\}}\Big).
$$

Using FOC T_t to solve for $\tilde{\xi}_t$ and to reduce the expression of FOCs for returns and

$\boldsymbol{b}^G$, we have

$$
\begin{aligned}
[\boldsymbol{R}^j_{t+1}]: \quad & \sum_{s=0}^{\infty} \beta^s \Big[\tilde{\boldsymbol{\gamma}}_s^\intercal \tilde{\boldsymbol{\epsilon}}_{s,\boldsymbol{R}^j_{t+1}} + \tilde{\eta}_s \tilde{\kappa}_{s,\boldsymbol{R}^j_{t+1}} + \beta\xi \Big(\hat{R}^A_{s+1} \tilde{a}^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal \tilde{\boldsymbol{b}}^F_{s,\boldsymbol{R}^j_{t+1}} - \hat{R}^N_{s+1} \tilde{n}_{s,\boldsymbol{R}^j_{t+1}} \Big) \Big] \\
& = \Big[\sum_{s=0}^{\infty} \beta^{s+1} (\beta\xi \hat{R}^A_{s+1} - (\beta\lambda \hat{\boldsymbol{R}}^B_{s+1} - \tilde{\boldsymbol{\gamma}}_s)^\intercal \tilde{\boldsymbol{\iota}}_{\boldsymbol{b}}) \tilde{\boldsymbol{z}}^K_{s+1,R^A_{t+1}} \Big] \mathbf{1}_{\{j\in\mathcal{A}\}}, \qquad (24) \\
[T_t]: \quad & \sum_{s=0}^{\infty} \beta^s \Big[\tilde{\boldsymbol{\gamma}}_s^\intercal \tilde{\boldsymbol{\epsilon}}_{s,\Delta T_{t+1}} + \tilde{\eta}_s \tilde{\kappa}_{s,\Delta T_{t+1}} + \beta\xi \Big(\hat{R}^A_{s+1} \tilde{a}^F_{s,\Delta T_{t+1}} - \hat{\boldsymbol{R}}^B_{s+1}{}^\intercal \tilde{\boldsymbol{b}}^F_{s,\Delta T_{t+1}} - \hat{R}^N_{s+1} \tilde{n}_{s,\Delta T_{t+1}} \Big) \Big] \\
& = -\beta^t (\tilde{\xi}_{t+1} - \tilde{\xi}_t), \qquad (25) \\
[b^G_t]: \quad & (\hat{\lambda}_{t+1} - \hat{\lambda}_t) \mathbf{1}_{\mathcal{B}}^\intercal + \beta\bar{\xi} \hat{\boldsymbol{R}}^B_{t+1}{}^\intercal = \tilde{\boldsymbol{\gamma}}_t^\intercal - (\tilde{\xi}_{t+1} - \tilde{\xi}_t) \mathbf{1}_{\mathcal{B}}^\intercal, \qquad (26)
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{\boldsymbol{\epsilon}}_{s,\boldsymbol{R}^j_{t+1}} \coloneqq \quad & \boldsymbol{\epsilon}_{s,\boldsymbol{R}^j_{t+1}} + \sum_{\ell=0}^{\infty} \boldsymbol{\epsilon}_{s,T_\ell} \tilde{R}^N_{\ell,\boldsymbol{R}^j_{t+1}} n_{\ell-1} + \Big[\boldsymbol{\epsilon}_{s,T_{t+1}} \boldsymbol{b}^{G\intercal} + \boldsymbol{\epsilon}_{s,T_0} \boldsymbol{b}^G_{-1}{}^\intercal \boldsymbol{R}^B_{0,\boldsymbol{R}^B_{t+1}} \Big] \mathbf{1}_{\{j\in\mathcal{B}\}} \\
& + \Big[\sum_{\ell=0}^{\infty} (\boldsymbol{\epsilon}_{s,T_{\ell+1}} - \beta^{\ell+1} \boldsymbol{\epsilon}_{s,T_0}) (w_{\ell+1,R^A_{t+1}} + a^K_t \mathbf{1}_{\{\ell=t\}}) \Big] \mathbf{1}_{\{j\in\mathcal{A}\}}, \\
\tilde{\boldsymbol{\epsilon}}_{s,\Delta T_{t+1}} \coloneqq \quad & \beta^{-1} \boldsymbol{\epsilon}_{s,T_{t+1}} - \boldsymbol{\epsilon}_{s,T_t},
\end{aligned}
$$

and similarly for other objects.

A.5 Matrix Form

Define

$$
\tilde{\boldsymbol{\gamma}} \coloneqq [\beta^t \tilde{\boldsymbol{\gamma}}_t], \quad \hat{\boldsymbol{R}}^j \coloneqq [\beta^{t+1} \hat{\boldsymbol{R}}^j_{t+1}], \quad \tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^j} = [\tilde{\boldsymbol{\epsilon}}_{s,\boldsymbol{R}^j_{t+1}}], \quad \tilde{\boldsymbol{\epsilon}}_{\Delta \boldsymbol{T}} = [\tilde{\boldsymbol{\epsilon}}_{s,\Delta T_{t+1}}]
$$

and similarly for other matrices, and let

$$
\Delta\hat{\boldsymbol{\lambda}} \coloneqq [\beta^t (\hat{\lambda}_{t+1} - \hat{\lambda}_t)], \quad \Delta\hat{\boldsymbol{\xi}} \coloneqq [\beta^t (\hat{\xi}_{t+1} - \hat{\xi}_t)].
$$

We have

$$
[R^A_{t+1}]: \quad \tilde{\boldsymbol{\gamma}}^\intercal (\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^A} - \mathcal{J} \tilde{\boldsymbol{z}}^K_{\boldsymbol{R}^A}) + \tilde{\boldsymbol{\eta}}^\intercal \tilde{\boldsymbol{\kappa}}_{\boldsymbol{R}^A} + \xi \big[\hat{\boldsymbol{R}}^{A\intercal} (\tilde{\boldsymbol{a}}^F_{\boldsymbol{R}^A} - \tilde{\boldsymbol{z}}^K_{\boldsymbol{R}^A}) - \hat{\boldsymbol{R}}^{B\intercal} (\tilde{\boldsymbol{b}}^F_{\boldsymbol{R}^A} - \mathcal{J} \tilde{\boldsymbol{z}}^K_{\boldsymbol{R}^A}) - \hat{\boldsymbol{R}}^{N\intercal} \tilde{\boldsymbol{n}}_{\boldsymbol{R}^A} \big] = \mathbf{0},
$$

$$
[\boldsymbol{R}^B_{t+1}]: \quad \tilde{\boldsymbol{\gamma}}^\intercal \tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \tilde{\boldsymbol{\eta}}^\intercal \tilde{\boldsymbol{\kappa}}_{\boldsymbol{R}^B} + \xi \big[\hat{\boldsymbol{R}}^{A\intercal} \tilde{\boldsymbol{a}}^F_{\boldsymbol{R}^B} - \hat{\boldsymbol{R}}^{B\intercal} \tilde{\boldsymbol{b}}^F_{\boldsymbol{R}^B} - \hat{\boldsymbol{R}}^{N\intercal} \tilde{\boldsymbol{n}}_{\boldsymbol{R}^B} \big] = \mathbf{0},
$$

$$[T_t]:\quad \Delta\hat{\boldsymbol{\lambda}}^\intercal - \Delta\hat{\boldsymbol{\xi}}^\intercal = \left[\tilde{\boldsymbol{\gamma}}^\intercal\tilde{\boldsymbol{\epsilon}}_{\Delta\boldsymbol{T}} + \tilde{\boldsymbol{\eta}}^\intercal\tilde{\boldsymbol{\kappa}}_{\Delta\boldsymbol{T}} + \xi\left(\hat{\boldsymbol{R}}^{A^\intercal}\tilde{\boldsymbol{a}}^F_{\Delta\boldsymbol{T}} - \hat{\boldsymbol{R}}^{B^\intercal}\tilde{\boldsymbol{b}}^F_{\Delta\boldsymbol{T}} - \hat{\boldsymbol{R}}^{N^\intercal}\tilde{\boldsymbol{n}}_{\Delta\boldsymbol{T}}\right)\right],$$

and

$$[\boldsymbol{b}^G_t]:\quad (\Delta\hat{\boldsymbol{\lambda}}\otimes\mathbf{1}_{\mathcal{B}})^\intercal + \xi\hat{\boldsymbol{R}}^{B^\intercal} = \tilde{\boldsymbol{\gamma}}^\intercal + ((\Delta\hat{\boldsymbol{\lambda}} - \Delta\hat{\boldsymbol{\xi}})\otimes\mathbf{1}_{\mathcal{B}})^\intercal,$$

$\mathcal{J} \coloneqq \boldsymbol{I}\otimes\tilde{\boldsymbol{\iota}}_{\boldsymbol{b}}$, and $\boldsymbol{I}$ is an identity matrix of the size of the time horizon.

We define the following notation for arbitrary $\tilde{\boldsymbol{\chi}}$:

$$\begin{aligned}
\check{\boldsymbol{\chi}}_{\boldsymbol{R}^B} &\coloneqq \tilde{\boldsymbol{\chi}}_{\boldsymbol{R}^B} + \tilde{\boldsymbol{\chi}}_{\boldsymbol{R}^A}(-\tilde{\boldsymbol{\kappa}}_{\boldsymbol{R}^A})^{-1}\tilde{\boldsymbol{\kappa}}_{\boldsymbol{R}^B},\\
\check{\boldsymbol{\chi}}_{\Delta\boldsymbol{T}} &\coloneqq \tilde{\boldsymbol{\chi}}_{\Delta\boldsymbol{T}} + \tilde{\boldsymbol{\chi}}_{\boldsymbol{R}^A}(-\tilde{\boldsymbol{\kappa}}_{\boldsymbol{R}^A})^{-1}\tilde{\boldsymbol{\kappa}}_{\Delta\boldsymbol{T}}.
\end{aligned}$$

Solving for $\tilde{\boldsymbol{\eta}}$, we have

$$[\boldsymbol{R}^B_{t+1}]:\quad \tilde{\boldsymbol{\gamma}}^\intercal(\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} - \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B}) + \xi\left[\hat{\boldsymbol{R}}^{A^\intercal}(\check{\boldsymbol{a}}^F_{\boldsymbol{R}^B} - \check{\boldsymbol{z}}^K_{\boldsymbol{R}^B}) - \hat{\boldsymbol{R}}^{B^\intercal}(\check{\boldsymbol{b}}^F_{\boldsymbol{R}^B} - \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B}) - \hat{\boldsymbol{R}}^{N^\intercal}\check{\boldsymbol{n}}_{\boldsymbol{R}^B}\right] = \mathbf{0}.$$

Solving for $\tilde{\boldsymbol{\gamma}}$ from FOC $\boldsymbol{R}^B$, we have

$$[\boldsymbol{R}^B_{t+1}]:\quad \tilde{\boldsymbol{\gamma}}^\intercal = \xi\left[\hat{\boldsymbol{R}}^{A^\intercal}\boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{B^\intercal}\boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{N^\intercal}\boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}}\right],$$

where

$$\begin{aligned}
\boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} &\coloneqq (\check{\boldsymbol{a}}^F_{\boldsymbol{R}^B} - \check{\boldsymbol{z}}^K_{\boldsymbol{R}^B})(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B})^{-1},\\
\boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} &\coloneqq (\check{\boldsymbol{b}}^F_{\boldsymbol{R}^B} - \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B})(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B})^{-1},\\
\boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}} &\coloneqq \check{\boldsymbol{n}}_{\boldsymbol{R}^B}(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B})^{-1}.
\end{aligned}$$

Define

$$\boldsymbol{\Gamma}^A_{\Delta\boldsymbol{T}} \coloneqq (\check{\boldsymbol{a}}^F_{\Delta\boldsymbol{T}} - \check{\boldsymbol{z}}^K_{\Delta\boldsymbol{T}}) + \boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}}\check{\boldsymbol{\epsilon}}_{\Delta\boldsymbol{T}},\quad \boldsymbol{\Gamma}^B_{\Delta\boldsymbol{T}} \coloneqq (\check{\boldsymbol{b}}^F_{\Delta\boldsymbol{T}} - \mathcal{J}\check{\boldsymbol{z}}^K_{\Delta\boldsymbol{T}}) + \boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}}\check{\boldsymbol{\epsilon}}_{\Delta\boldsymbol{T}},\quad \boldsymbol{\Gamma}^N_{\Delta\boldsymbol{T}} \coloneqq \check{\boldsymbol{n}}_{\Delta\boldsymbol{T}} + \boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}}\check{\boldsymbol{\epsilon}}_{\Delta\boldsymbol{T}},$$

then we have

$$\Delta\hat{\boldsymbol{\lambda}}^\intercal - \Delta\hat{\boldsymbol{\xi}}^\intercal = \xi\left[\hat{\boldsymbol{R}}^{A^\intercal}\boldsymbol{\Gamma}^A_{\Delta\boldsymbol{T}} - \hat{\boldsymbol{R}}^{B^\intercal}\boldsymbol{\Gamma}^B_{\Delta\boldsymbol{T}} - \hat{\boldsymbol{R}}^{N^\intercal}\boldsymbol{\Gamma}^N_{\Delta\boldsymbol{T}}\right].$$

Let

$$\boldsymbol{\Gamma}^j \coloneqq \boldsymbol{\Gamma}^j_{\boldsymbol{\epsilon}} + \boldsymbol{\Gamma}^j_{\Delta\boldsymbol{T}}\otimes\mathbf{1}^\intercal_{\mathcal{B}},\quad \forall j\in\{A,B,N\},$$

then we have

$$[\boldsymbol{b}_t^G]: \quad \lambda^{-1}(\Delta\hat{\boldsymbol{\lambda}} \otimes \mathbf{1}_{\mathcal{B}})^\intercal + \hat{\boldsymbol{R}}^{B\intercal} = \hat{\boldsymbol{R}}^{A\intercal}\boldsymbol{\Gamma}^A - \hat{\boldsymbol{R}}^{B\intercal}\boldsymbol{\Gamma}^B - \hat{\boldsymbol{R}}^{N\intercal}\boldsymbol{\Gamma}^N.$$

B Benchmarks and Examples

B.1 Neoclassical Growth Benchmark

Because $\Delta\hat{\boldsymbol{\lambda}} = -\lambda\hat{\boldsymbol{R}}^N$, optimality of the policy rule requires

$$\hat{\boldsymbol{R}}^{B\intercal} - \hat{\boldsymbol{R}}^{N\intercal} \otimes \mathbf{1}_{\mathcal{B}}^\intercal = \hat{\boldsymbol{R}}^{A\intercal}\boldsymbol{\Gamma}^A - \hat{\boldsymbol{R}}^{B\intercal}\boldsymbol{\Gamma}^B - \hat{\boldsymbol{R}}^{N\intercal}\boldsymbol{\Gamma}^N.$$

When θ and $\boldsymbol{\phi}$ are constants, we have

$$\boldsymbol{\Gamma}^A = \theta\boldsymbol{\Gamma}^N - \boldsymbol{\Gamma}^z, \quad \boldsymbol{\Gamma}^B = \boldsymbol{\phi}\boldsymbol{\Gamma}^N - \mathcal{J}\boldsymbol{\Gamma}^z,$$

where $\boldsymbol{\Gamma}^z \coloneqq \boldsymbol{\Gamma}^z_{\boldsymbol{\epsilon}} + \boldsymbol{\Gamma}^z_{\Delta\boldsymbol{T}} \otimes \mathbf{1}_{\mathcal{B}}^\intercal$,

$$\boldsymbol{\Gamma}^z_{\boldsymbol{\epsilon}} \coloneqq \check{\boldsymbol{z}}^K_{\boldsymbol{R}^B}(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^K_{\boldsymbol{R}^B})^{-1}, \quad \boldsymbol{\Gamma}^z_{\Delta\boldsymbol{T}} \coloneqq \check{\boldsymbol{z}}^K_{\Delta\boldsymbol{T}} + \boldsymbol{\Gamma}^z_{\boldsymbol{\epsilon}}\check{\boldsymbol{\epsilon}}_{\Delta\boldsymbol{T}}.$$

The formula for the optimal rule reduces to

$$\hat{\boldsymbol{R}}^{B\intercal} - \hat{\boldsymbol{R}}^{N\intercal} \otimes \mathbf{1}_{\mathcal{B}}^\intercal = \hat{\boldsymbol{R}}^{A\intercal}\boldsymbol{\Gamma}^A - \hat{\boldsymbol{R}}^{B\intercal}\boldsymbol{\Gamma}^B - \hat{\boldsymbol{R}}^{N\intercal}\boldsymbol{\Gamma}^N = -(\hat{\boldsymbol{R}}^{A\intercal} - \hat{\boldsymbol{R}}^{B\intercal}\mathcal{J})\boldsymbol{\Gamma}^z.$$

Because θ and $\boldsymbol{\phi}$ are constant, the definition of R^N implies that

$$\hat{\boldsymbol{R}}^{B\intercal} - \hat{\boldsymbol{R}}^{N\intercal} \otimes \mathbf{1}_{\mathcal{B}}^\intercal = -\hat{\boldsymbol{R}}^{A\intercal} \otimes (\theta\mathbf{1}_{\mathcal{B}}^\intercal) + \hat{\boldsymbol{R}}^{B\intercal}[\boldsymbol{I}_{\mathcal{B}} + \boldsymbol{I} \otimes (\boldsymbol{\phi}\mathbf{1}_{\mathcal{B}}^\intercal)].$$

Therefore, the optimality condition reduces to

$$\hat{\boldsymbol{R}}^{B\intercal}[\boldsymbol{I}_{\mathcal{B}} + \boldsymbol{I} \otimes (\boldsymbol{\phi}\mathbf{1}_{\mathcal{B}}^\intercal)] - \hat{\boldsymbol{R}}^{A\intercal} \otimes (\theta\mathbf{1}_{\mathcal{B}}^\intercal) = -(\hat{\boldsymbol{R}}^{A\intercal} - \hat{\boldsymbol{R}}^{B\intercal}\mathcal{J})\boldsymbol{\Gamma}^z.$$

Note that because $\theta = \boldsymbol{\phi}\mathbf{1}_{\mathcal{B}}^\intercal + 1$, one solution to the system is given by

$$\hat{\boldsymbol{R}}^{B\intercal} = \hat{\boldsymbol{R}}^{A\intercal} \otimes \mathbf{1}_{\mathcal{B}}^\intercal.$$

To check for uniqueness, we need to check that the following matrix is invertible:

$$[\boldsymbol{I}_{\mathcal{B}} + \boldsymbol{I} \otimes (\boldsymbol{\phi}\mathbf{1}_{\mathcal{B}}^{\intercal})] - \mathcal{J}\boldsymbol{\Gamma}^{z}.$$

B.2 Quasi-Linear Preference

Suppose that the household has a quasi-linear preference without portfolio adjustment cost:

$$u\left(\{c_s, \boldsymbol{b}_s, n_s\}_{s\leq t}\right) = c_t + \nu(\boldsymbol{b}_t, n_t),$$

then transfers do not affect asset holdings and we have

$$\boldsymbol{\epsilon}_{\Delta \boldsymbol{T}} = \mathbf{0}, \quad \boldsymbol{a}^{F}_{\Delta \boldsymbol{T}} = \mathbf{0}, \quad \boldsymbol{b}^{F}_{\Delta \boldsymbol{T}} = \mathbf{0}, \quad \boldsymbol{n}_{\Delta \boldsymbol{T}} = \mathbf{0}, \quad \lambda_t = 1.$$

As a result,

$$\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} = \boldsymbol{\epsilon}_{\boldsymbol{R}^B}, \quad \tilde{\boldsymbol{a}}^{F}_{\boldsymbol{R}^B} = \boldsymbol{a}^{F}_{\boldsymbol{R}^B}, \quad \tilde{\boldsymbol{b}}^{F}_{\boldsymbol{R}^B} = \boldsymbol{b}^{F}_{\boldsymbol{R}^B}, \quad \tilde{\boldsymbol{n}}_{\boldsymbol{R}^B} = \boldsymbol{n}_{\boldsymbol{R}^B},$$

and

$$\boldsymbol{\Gamma}^{j} = \boldsymbol{\Gamma}^{j}_{\boldsymbol{\epsilon}}, \quad \forall j \in \{A, B, N\}.$$

The formula reduces to

$$[\boldsymbol{b}^{G}_{t}]: \quad \hat{\boldsymbol{R}}^{B\intercal} = \hat{\boldsymbol{R}}^{A\intercal}\boldsymbol{\Gamma}^{A} - \hat{\boldsymbol{R}}^{B\intercal}\boldsymbol{\Gamma}^{B} - \hat{\boldsymbol{R}}^{N\intercal}\boldsymbol{\Gamma}^{N}.$$

where

$$\begin{aligned}
\boldsymbol{\Gamma}^{A} &= (\check{\boldsymbol{a}}^{F}_{\boldsymbol{R}^B} - \check{\boldsymbol{z}}^{K}_{\boldsymbol{R}^B})(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^{K}_{\boldsymbol{R}^B})^{-1},\\
\boldsymbol{\Gamma}^{B} &= (\check{\boldsymbol{b}}^{F}_{\boldsymbol{R}^B} - \mathcal{J}\check{\boldsymbol{z}}^{K}_{\boldsymbol{R}^B})(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^{K}_{\boldsymbol{R}^B})^{-1},\\
\boldsymbol{\Gamma}^{N} &= \check{\boldsymbol{n}}_{\boldsymbol{R}^B}(-\check{\boldsymbol{\epsilon}}_{\boldsymbol{R}^B} + \mathcal{J}\check{\boldsymbol{z}}^{K}_{\boldsymbol{R}^B})^{-1}.
\end{aligned}$$

Moreover, if there is no capital adjustment cost, then $\boldsymbol{z}^{K}_{\boldsymbol{R}^A} = \mathbf{0}$ and we have

$$\boldsymbol{\Gamma}^{A} := \check{\boldsymbol{a}}^{F}_{\boldsymbol{R}^B}(-\check{\boldsymbol{\epsilon}}^{-1}_{\boldsymbol{R}^B}), \quad \boldsymbol{\Gamma}^{B} := \check{\boldsymbol{b}}^{F}_{\boldsymbol{R}^B}(-\check{\boldsymbol{\epsilon}}^{-1}_{\boldsymbol{R}^B}), \quad \boldsymbol{\Gamma}^{N} := \check{\boldsymbol{n}}_{\boldsymbol{R}^B}(-\check{\boldsymbol{\epsilon}}^{-1}_{\boldsymbol{R}^B}).$$

B.3 Synthetic Intertemporal Price

Consider the case where the household has an additively separable preference over c_t, n_t, and $\boldsymbol{b}_t$:

$$u\left(\{c_s, \boldsymbol{b}_s, n_s\}_{s\le t}; \boldsymbol{\omega}_t\right) = \tilde{u}(c_t; \boldsymbol{\omega}_t) + \tilde{v}(n_t; \boldsymbol{\omega}_t) + \nu(\boldsymbol{b}_t), \tag{27}$$

and there is no portfolio adjustment cost. If $\nu_{\boldsymbol{bb}}$ is not full rank, there exists a vector $\boldsymbol{\alpha}$ such that

$$\boldsymbol{\alpha}^\intercal \nu_{\boldsymbol{bb}} = \mathbf{0}, \quad \boldsymbol{\alpha}^\intercal \mathbf{1}_{\mathcal{B}} = 1.$$

With such a vector $\boldsymbol{\alpha}$, Assumption 1 holds with return $\hat{R}^*_{t+1} = \boldsymbol{\alpha}^\intercal \hat{\boldsymbol{R}}^B_{t+1}$:

$$\Delta \hat{\boldsymbol{\lambda}}^\intercal = -\bar{\lambda} \hat{\boldsymbol{R}}^{*\intercal},$$

where $\hat{\boldsymbol{R}}^* \coloneqq [\beta^{t+1} \hat{R}^*_{t+1}]$.

Let $\boldsymbol{\Omega}$ be a set of orthonormal basis such that $\boldsymbol{\Omega}_1 = \boldsymbol{\alpha}$, and define

$$\hat{\boldsymbol{b}}^Z_t \coloneqq \boldsymbol{\Omega}^\intercal \hat{\boldsymbol{b}}_t, \quad \hat{\boldsymbol{R}}^Z_{t+1} = \boldsymbol{\Omega}^\intercal \hat{\boldsymbol{R}}^B_{t+1},$$

where the first elements are given by $\hat{b}^*_t \coloneqq \boldsymbol{\alpha}^\intercal \hat{\boldsymbol{b}}_t$ and $\hat{R}^*_{t+1}$.

The household budget constraint is:

$$\hat{c}_t + (\boldsymbol{\Omega}^\intercal \mathbf{1}_{\mathcal{B}})^\intercal \hat{\boldsymbol{b}}^Z_t + \hat{n}_t = \hat{w}_t + (\boldsymbol{R}^Z - \boldsymbol{\tau}^Z)^\intercal \hat{\boldsymbol{b}}^Z_{t-1} + \hat{\boldsymbol{R}}^Z_t \boldsymbol{b}^Z + R^N \hat{n}_{t-1} + \hat{R}^N_t n - \hat{T}_t.$$

Solve for $\hat{c}_t$, and substitute it in $\Delta \hat{\lambda}_{t+1} = -\bar{\lambda} \hat{R}^*_{t+1}$ with $\hat{\lambda}_t = u''(\bar{c})\hat{c}_t$ from the household's problem. We have

$$\hat{b}^*_{t+1} - \left[1 + \beta^{-1}\right] \hat{b}^*_t + \beta^{-1} \hat{b}^*_{t-1} = \epsilon^*_t, \tag{28}$$

where

$$\begin{aligned}\epsilon^*_t \coloneqq & -\left[\hat{w}_t + \hat{\boldsymbol{R}}^Z_t \boldsymbol{b}^Z + \hat{R}^N_t n - \hat{T}_t\right] + \left[\hat{w}_{t+1} + \hat{\boldsymbol{R}}^Z_{t+1} \boldsymbol{b}^Z + \hat{R}^N_{t+1} n - \hat{T}_{t+1}\right] - \frac{u'}{u''} \beta \hat{R}^*_{t+1} \\ & + \sum_{j=2}^{|\mathcal{B}|} \left[\left(\boldsymbol{\Omega}_j^\intercal \mathbf{1}_{\mathcal{B}} \hat{b}^{Z_j}_t - (\boldsymbol{R}^Z - \boldsymbol{\tau}^Z) \hat{b}^{Z_j}_{t-1}\right) - \left(\boldsymbol{\Omega}_j^\intercal \mathbf{1}_{\mathcal{B}} \hat{b}^{Z_j}_{t+1} - (\boldsymbol{R}^Z - \boldsymbol{\tau}^Z) \hat{b}^{Z_j}_t\right)\right].\end{aligned}$$

From the household's problem, we have

$$\Delta\hat{\lambda}_{t+1}\mathbf{1}_{\mathcal{B}} + \bar{\lambda}\hat{\boldsymbol{R}}^{B}_{t+1} = v_{\boldsymbol{bb}}\hat{\boldsymbol{b}}_t.$$

Because

$$\boldsymbol{\alpha}^{\intercal}v_{\boldsymbol{bb}}\hat{\boldsymbol{b}}_t = \Delta\hat{\lambda}_{t+1} + \boldsymbol{\alpha}^{\intercal}\bar{\lambda}\hat{\boldsymbol{R}}^{B}_{t+1} = 0$$

for any $\hat{\boldsymbol{b}}_t$, it implies that $\boldsymbol{\alpha}^{\intercal}v_{\boldsymbol{bb}} = \mathbf{0}_{\mathcal{B}}$.

Left multiply $\boldsymbol{\Omega}^{\intercal}$, we have

$$\bar{\lambda}(\hat{\boldsymbol{R}}^{Z}_{t+1} - \hat{R}^{*}_{t+1}\boldsymbol{\Omega}^{\intercal}\mathbf{1}_{\mathcal{B}}) = \boldsymbol{\Omega}^{\intercal}v_{\boldsymbol{bb}}\boldsymbol{\Omega} \times \hat{\boldsymbol{b}}^{Z}_t,$$

where the first row and column of $\boldsymbol{\Omega}^{\intercal}v_{\boldsymbol{bb}}\boldsymbol{\Omega}$ are zeros, and its $(1,1)$ minor is non-zero because $\boldsymbol{\alpha}$ is unique.

As a result, $\hat{\boldsymbol{b}}^{Z_j}_t$ can be solved as linear functions of $\hat{\boldsymbol{R}}^{Z}_{t+1}$ for all $j > 1$, and Equation (28) implies

$$\boldsymbol{\Omega}^{\intercal}(\beta^{-1}\boldsymbol{\epsilon}_{s,T_{t+1}} - \boldsymbol{\epsilon}_{s,T_t}) = \begin{pmatrix} -\mathbf{1}_{\{s=t\}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let $\tilde{\boldsymbol{\gamma}}^{*}_t = \boldsymbol{\alpha}^{\intercal}\tilde{\boldsymbol{\gamma}}_t$, then

$$\tilde{\boldsymbol{\gamma}}^{\intercal}_t(\beta^{-1}\boldsymbol{\epsilon}_{s,T_{t+1}} - \boldsymbol{\epsilon}_{s,T_t}) = (\boldsymbol{\Omega}^{\intercal}\tilde{\boldsymbol{\gamma}}_t)^{\intercal}(\beta^{-1}\boldsymbol{\Omega}^{\intercal}\boldsymbol{\epsilon}_{s,T_{t+1}} - \boldsymbol{\Omega}^{\intercal}\boldsymbol{\epsilon}_{s,T_t}) = -\tilde{\boldsymbol{\gamma}}^{*}_t\mathbf{1}_{\{s=t\}},$$

and, in matrix form,

$$\tilde{\boldsymbol{\gamma}}^{\intercal}\boldsymbol{\epsilon}_{\Delta\boldsymbol{T}} = -\tilde{\boldsymbol{\gamma}}^{*\intercal},$$

where $\tilde{\boldsymbol{\gamma}}^{*} \coloneqq [\beta^{t}\tilde{\boldsymbol{\gamma}}^{*}_t]$.

Moreover, because $\hat{n}_t$ satisfies

$$\Delta\hat{\lambda}_{t+1} + \hat{R}^{N}_{t+1} = \hat{R}^{N}_{t+1} - \hat{R}^{*}_{t+1} = \tilde{v}_{nn}\hat{n}_t + \tilde{v}_{n\boldsymbol{\omega}}\hat{\boldsymbol{\omega}}_t,$$

we have $\boldsymbol{n}_{\Delta \boldsymbol{T}} = \boldsymbol{0}$, which implies that

$$\boldsymbol{a}^F_{\Delta \boldsymbol{T}} = \boldsymbol{0}, \quad \boldsymbol{b}^F_{\Delta \boldsymbol{T}} = \boldsymbol{0}, \quad \boldsymbol{\kappa}_{\Delta \boldsymbol{T}} = \boldsymbol{0}.$$

Therefore, $\check{\boldsymbol{\epsilon}}_{\Delta \boldsymbol{T}} = \boldsymbol{\epsilon}_{\Delta \boldsymbol{T}}$ and

$$\boldsymbol{\Gamma}^A_{\Delta \boldsymbol{T}} \coloneqq \boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} \boldsymbol{\epsilon}_{\Delta \boldsymbol{T}}, \quad \boldsymbol{\Gamma}^B_{\Delta \boldsymbol{T}} \coloneqq \boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} \boldsymbol{\epsilon}_{\Delta \boldsymbol{T}}, \quad \boldsymbol{\Gamma}^N_{\Delta \boldsymbol{T}} \coloneqq \boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}} \boldsymbol{\epsilon}_{\Delta \boldsymbol{T}}.$$

As a result, Lemma 3 implies

$$\Delta \hat{\boldsymbol{\xi}}^\intercal = \Delta \hat{\boldsymbol{\lambda}}^\intercal - \tilde{\boldsymbol{\gamma}}^\intercal \boldsymbol{\epsilon}_{\Delta \boldsymbol{T}} = -\bar{\lambda} \hat{\boldsymbol{R}}^{*\intercal} + \tilde{\boldsymbol{\gamma}}^{*\intercal}.$$

Use it in the first order condition for $\boldsymbol{b}^G$:

$$[\boldsymbol{b}^G_t] : \quad (\Delta \hat{\boldsymbol{\xi}} \otimes \mathbf{1}_{\mathcal{B}})^\intercal + \xi \hat{\boldsymbol{R}}^{B\intercal} = \tilde{\boldsymbol{\gamma}}^\intercal,$$

we have

$$[\boldsymbol{b}^G_t] : \quad \xi (\hat{\boldsymbol{R}}^B - \hat{\boldsymbol{R}}^* \otimes \mathbf{1}_{\mathcal{B}})^\intercal = (\tilde{\boldsymbol{\gamma}} - \tilde{\boldsymbol{\gamma}}^* \otimes \mathbf{1}_{\mathcal{B}})^\intercal.$$

From Lemma 3, we have

$$[\boldsymbol{R}^B_{t+1}] : \quad \tilde{\boldsymbol{\gamma}}^\intercal = \xi \big[\hat{\boldsymbol{R}}^{A\intercal} \boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{B\intercal} \boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{N\intercal} \boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}} \big],$$

and

$$[\boldsymbol{R}^B_{t+1}] : \quad (\tilde{\boldsymbol{\gamma}}^* \otimes \mathbf{1}_{\mathcal{B}})^\intercal = \xi \big[\hat{\boldsymbol{R}}^{A\intercal} \boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{B\intercal} \boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{N\intercal} \boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}} \big] \boldsymbol{I}_{\boldsymbol{\alpha}},$$

where $\boldsymbol{I}_{\boldsymbol{\alpha}} \coloneqq \boldsymbol{I} \otimes (\boldsymbol{\alpha} \mathbf{1}^\intercal_{\mathcal{B}})$.

Using these expressions in the optimality condition for $\boldsymbol{b}^G$, the optimal policy rule is reduced to

$$\hat{\boldsymbol{R}}^{B\intercal} [\boldsymbol{I}_{\mathcal{B}} - \boldsymbol{I}_{\boldsymbol{\alpha}}] = \big[\hat{\boldsymbol{R}}^{A\intercal} \boldsymbol{\Gamma}^A_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{B\intercal} \boldsymbol{\Gamma}^B_{\boldsymbol{\epsilon}} - \hat{\boldsymbol{R}}^{N\intercal} \boldsymbol{\Gamma}^N_{\boldsymbol{\epsilon}} \big] [\boldsymbol{I}_{\mathcal{B}} - \boldsymbol{I}_{\boldsymbol{\alpha}}].$$

Because $rank(\boldsymbol{\alpha} \mathbf{1}^\intercal_{\mathcal{B}}) = 1$ and has 1 as an eigenvalue (with $\mathbf{1}_{\mathcal{B}}$ as an eigenvector), $\boldsymbol{I}_{\mathcal{B}} - \boldsymbol{I}_{\boldsymbol{\alpha}}$ has rank $(|\mathcal{B}| - 1) \times T$.

C Nested Models of Financial Intermediation

We show how our framework nests some commonly used models of financial frictions by appropriately choosing functions $\boldsymbol{\phi}_t(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}\}^\infty_{s=t})$ and $\theta_t(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}\}^\infty_{s=t})$.

C.1 Bianchi and Bigio (2022)

Bianchi and Bigio (2022) develop a state-of-the-art model of the banking sector that issues deposits, holds reserves, government debt, and productive capital. Inside the banking sector, banks face idiosyncratic liquidity shocks and need to settle balances with reserves in an interbank market with search frictions. The presence of interbank market frictions allows the government to affect relative returns on assets by changing the relative supply of bonds and reserves, and influence aggregate outcomes.

Below, we briefly describe their model and show that through the lens of our framework, their banking sector corresponds to a specific function $\boldsymbol{\phi}$.

There is a continuum of financial intermediaries (banks) with relative risk aversion $\gamma \geq 0$. Financial intermediaries choose portfolio shares of loans $\bar{\theta} \geq 0$, government debt $\bar{\phi}^{\text{gov}} \leq 0$, reserves $\bar{\phi}^{\text{res}} \leq 0$ and deposits $\bar{\phi}^{\text{dep}} \geq 0$.[8] The real gross rates of return are respectively R^A_{t+1},$R^B_{\text{gov},t+1}$,$R^B_{\text{res},t+1}$, and $R^B_{\text{dep},t+1}$. Intermediaries face idiosyncratic withdrawal shocks $\omega \in [\omega_{\min}, \infty]$, with $\omega_{\min} \geq -1$, $\mathbb{E}[\omega] = 0$, and a cumulative distribution function F with density f. Intermediaries have to raise enough reserves to settle transactions. In case of a negative realization of ω they first sell their holdings of government debt for reserves and if there is still a shortfall they borrow from the discount window at the rate R^W_{t+1} and from the the interbank market. Intermediaries with positive realizations of ω lend in the interbank market. The interbank market rate is determined by the interbank market tightness. Proposition 7 in Bianchi and

[8]The quantity of reserves held by an intermediary is thus $-\bar{\phi}^{\text{res}} n \geq 0$. We use the same convention for government debt.

Bigio (2022) shows the optimal portfolio solves the following problem[9]

$$\max_{\bar{\theta},\bar{\phi}^{\text{gov}},\bar{\phi}^{\text{res}},\bar{\phi}^{\text{dep}}} \mathbb{E}_{\omega}\left[\left(R^A_{t+1}\bar{\theta} - R^B_{\text{res},t+1}\left(\bar{\phi}_{\text{gov}}+\bar{\phi}_{\text{res}}\right) - R^B_{\text{dep},t+1}\bar{\phi}_{\text{dep}} + \chi_{t+1}(\bar{\phi}_{\text{gov}}+\bar{\phi}_{\text{res}},\bar{\phi}_{\text{dep}},\omega)\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma}}$$

s.t. $\bar{\theta} = 1 + \bar{\phi}_{\text{gov}} + \bar{\phi}_{\text{res}} + \bar{\phi}_{\text{dep}}$, $\bar{\phi}_{\text{dep}} \leq \kappa$, and

$$\chi_{t+1}(\bar{\phi}_{\text{gov}}+\bar{\phi}_{\text{res}},\bar{\phi}_{\text{dep}},\omega) = \begin{cases}\chi^+_{t+1}s, & s>0\\ \chi^-_{t+1}s, & s<0\end{cases}, \quad s = -\left(\bar{\phi}_{\text{gov}}+\bar{\phi}_{\text{res}}\right) + \left(\frac{R^B_{\text{dep},t+1}}{R^B_{\text{res},t+1}}\right)\omega\bar{\phi}_{\text{dep}}.$$

Let $\bar{\phi}_{\text{liq}} \coloneqq \bar{\phi}_{\text{res}} + \bar{\phi}_{\text{gov}}$. Proposition 8 in Bianchi and Bigio (2022) shows that if $\bar{\theta}, \bar{\phi}_{\text{dep}} > 0$ and $\bar{\phi}_{\text{liq}} < 0$ solve the portfolio problem, we have:

$$\begin{aligned}
R^A_{t+1} - R^B_{\text{res},t+1} &= \chi^+_{t+1} + \left(\chi^-_{t+1} - \chi^+_{t+1}\right)\cdot F\left(\omega^*\right)\cdot\frac{\mathbb{E}_{\omega}\left[\left(R^e_{t+1}\right)^{-\gamma} \mid \omega < \omega^*\right]}{\mathbb{E}_{\omega}\left[\left(R^e_{t+1}\right)^{-\gamma}\right]},\\
R^B_{\text{gov},t+1} - R^B_{\text{res},t+1} &= \chi^+_{t+1},\\
R^B_{\text{dep},t+1} - R^B_{\text{res},t+1} &= R^A_{t+1} - R^B_{\text{res},t+1} + \left(\frac{R^B_{\text{dep},t+1}}{R^B_{\text{res},t+1}}\right)\frac{\text{Cov}_{\omega}\left[\left(R^e_{t+1}\right)^{-\gamma}\cdot\chi_{t+1},\omega\right]}{\mathbb{E}_{\omega}\left[\left(R^e_{t+1}\right)^{-\gamma}\right]} - \mu,
\end{aligned}$$

where $\mu \geq 0$ is the multiplier on the leverage constraint, and $\mu(\bar{\phi}_{\text{dep}} - \kappa) = 0$. ω^* is the level of withdrawal shock ω such that $s = 0$:

$$\omega^* = -\frac{-\bar{\phi}_{\text{liq}}/\bar{\phi}_{\text{dep}}}{R^B_{\text{dep},t+1}R^B_{\text{res},t+1}}.$$

R^e_{t+1} is the gross return on unit of net worth:

$$R^e_{t+1} \coloneqq R^A_{t+1}\bar{\theta} - R^B_{\text{res},t+1}\bar{\phi}_{\text{liq}} - R^B_{\text{dep},t+1}\bar{\phi}_{\text{dep}} + \chi_{t+1}(\bar{\phi}_{\text{liq}},\bar{\phi}_{\text{dep}},\omega)$$

Appendix A in Bianchi and Bigio (2022) shows that slopes of the liquidity yield $\chi_{t+1}(\bar{\phi}_{\text{liq}},\bar{\phi}_{\text{dep}},\omega)$ are functions of $R^B_{\text{res},t+1}$, interbank market tightness ϑ_t, and discount

[9] We focus on a case with no regulatory reserve requirement, i.e. after the realization of withdrawal shocks and trade in the interbank market it is enough if reserve holdings are nonnegative. Our results are not affected by this simplification. Moreover, without a loss of generality we focus on an equilibrium with holdings of government bonds by intermediaries in deficit not exceeding the surplus of reserves. This is the case on which Bianchi and Bigio (2022) focus.

window rate R^W_{t+1}:

$$\chi^+_{t+1} = \chi^+(R^B_{\text{res},t+1}, \vartheta_t; R^W_{t+1}) \coloneqq (R^W_{t+1} - R^B_{\text{res},t+1}) \left(\frac{\bar{\vartheta}(\vartheta_t)}{\vartheta_t}\right)^\eta \left(\frac{\vartheta_t^\eta \bar{\vartheta}(\vartheta_t)^{1-\eta} - \vartheta_t}{\bar{\vartheta}(\vartheta_t) - 1}\right),$$
$$\chi^-_{t+1} = \chi^-(R^B_{\text{res},t+1}, \vartheta_t; R^W_{t+1}) \coloneqq (R^W_{t+1} - R^B_{\text{res},t+1}) \left(\frac{\bar{\vartheta}(\vartheta_t)}{\vartheta_t}\right)^\eta \left(\frac{\vartheta_t^\eta \bar{\vartheta}(\vartheta_t)^{1-\eta} - 1}{\bar{\vartheta}(\vartheta_t) - 1}\right),$$

where

$$\bar{\vartheta}(\vartheta_t) = \begin{cases} 1 + (\vartheta_t - 1)\exp(\lambda), & \vartheta_t \geq 1 \\ \frac{1}{1+(1/\vartheta_t - 1)\exp(\lambda)}, & \vartheta_t < 1. \end{cases}$$

Market tightness satisfies the interbank market clearing condition:

$$\vartheta_t = \vartheta(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}) \coloneqq -\frac{-F(\omega^*)\bar{\phi}_{\text{liq}} + \int_{\omega_{\min}}^{\omega^*} \omega f(\omega) d\omega \left(\frac{R^B_{\text{dep},t+1}}{R^B_{\text{res},t+1}}\right) \bar{\phi}_{\text{dep}}}{-(1-F(\omega^*))\bar{\phi}_{\text{liq}} + \int_{\omega^*}^{\infty} \omega f(\omega) d\omega \left(\frac{R^B_{\text{dep},t+1}}{R^B_{\text{res},t+1}}\right) \bar{\phi}_{\text{dep}} + \bar{\phi}_{\text{gov}}},$$

where we already used the fact that in a symmetric equilibrium portfolio choices of all intermediaries are the same.

Define

$$\tilde{\chi}^+(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}; R^W_{t+1}) \coloneqq \chi^+(R^B_{\text{res},t+1}, \vartheta(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}); R^W_{t+1})$$
$$\tilde{\chi}^-(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}; R^W_{t+1}) \coloneqq \chi^-(R^B_{\text{res},t+1}, \vartheta(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}); R^W_{t+1})$$
$$\tilde{\chi}^\Delta(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}; R^W_{t+1}) \coloneqq \tilde{\chi}^-(\cdot) - \tilde{\chi}^+(\cdot).$$
$$\tilde{\chi}(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}, \omega; R^W_{t+1}) = \begin{cases} \tilde{\chi}^+(\cdot)s, & s > 0 \\ \tilde{\chi}^-(\cdot)s, & s < 0 \end{cases}, \quad s = -\bar{\phi}_{\text{liq}} + \left(\frac{R^B_{\text{dep},t+1}}{R^B_{\text{res},t+1}}\right)\omega\bar{\phi}_{\text{dep}}.$$
$$\omega^*\left(R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}\right) \coloneqq \frac{\bar{\phi}_{\text{liq}}/\bar{\phi}_{\text{dep}}}{R^B_{\text{dep},t+1} R^B_{\text{res},t+1}}$$
$$\tilde{R}^e\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, \bar{\theta}, \bar{\phi}_{\text{liq}}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}; R^W_{t+1}\right) \coloneqq R^A_{t+1}\bar{\theta} - R^B_{\text{res},t+1}\bar{\phi}_{\text{liq}} - R^B_{\text{dep},t+1}\bar{\phi}_{\text{dep}}$$
$$+ \tilde{\chi}^+(\cdot)$$

We now use the above formulas in the first order conditions of the intermediary to solve for optimal portfolio shares as functions of returns only.

Binding leverage constraint

If the leverage constraint is binding

$$
\begin{aligned}
R^A_{t+1} - R^B_{\text{res},t+1} &= \tilde{\chi}^+(\cdot; R^W_{t+1}) + \tilde{\chi}^\Delta(\cdot; R^W_{t+1}) \cdot F\left(\omega^*\left(\cdot\right)\right) \cdot \frac{\mathbb{E}_\omega\left[\left(R^e\left(\cdot; R^W_{t+1}\right)\right)^{-\gamma} \mid \omega < \omega^*\left(\cdot\right)\right]}{\mathbb{E}_\omega\left[\left(R^e\left(\cdot; R^W_{t+1}\right)\right)^{-\gamma}\right]}, \\
R^B_{\text{gov},t+1} - R^B_{\text{res},t+1} &= \tilde{\chi}^+(\cdot; R^W_{t+1}), \\
\bar{\theta} &= 1 + \bar{\phi}_{\text{liq}} + \bar{\phi}_{\text{dep}}, \\
\bar{\phi}_{\text{dep}} &= \kappa.
\end{aligned}
$$

Slack leverage constraint

If the leverage constraint is slack

$$
\begin{aligned}
R^A_{t+1} - R^B_{\text{res},t+1} &= \tilde{\chi}^+(\cdot; R^W_{t+1}) + \tilde{\chi}^\Delta(\cdot; R^W_{t+1}) \cdot F\left(\omega^*\left(\cdot\right)\right) \cdot \frac{\mathbb{E}_\omega\left[\left(R^e\left(\cdot; R^W_{t+1}\right)\right)^{-\gamma} \mid \omega < \omega^*\left(\cdot\right)\right]}{\mathbb{E}_\omega\left[\left(R^e\left(\cdot; R^W_{t+1}\right)\right)^{-\gamma}\right]}, \\
R^B_{\text{gov},t+1} - R^B_{\text{res},t+1} &= \tilde{\chi}^+(\cdot; R^W_{t+1}), \\
R^B_{\text{dep},t+1} - R^B_{\text{res},t+1} &= R^A_{t+1} - R^B_{\text{res},t+1} + \left(\frac{R^B_{\text{dep},t+1}}{R^B_{\text{res},t+1}}\right) \frac{\text{Cov}_\omega\left[\left(R^e\left(\cdot; R^W_{t+1}\right)\right)^{-\gamma} \cdot \tilde{\chi}\left(\cdot\right), \omega\right]}{\mathbb{E}_\omega\left[\left(R^e\left(\cdot; R^W_{t+1}\right)\right)^{-\gamma}\right]} - \mu, \\
\bar{\theta} &= 1 + \bar{\phi}_{\text{liq}} + \bar{\phi}_{\text{dep}}.
\end{aligned}
$$

In both cases it is a system of four equations with nine variables: four quantities $\bar{\theta}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}, \bar{\phi}_{\text{liq}}$ and five returns $R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}, R^W_{t+1}$. We use $\bar{\phi}_{\text{res}} = \bar{\phi}_{\text{liq}} - \bar{\phi}_{\text{gov}}$ and solve for $\bar{\theta}, \bar{\phi}_{\text{dep}}, \bar{\phi}_{\text{gov}}, \bar{\phi}_{\text{res}}$ as functions of returns. We denote the solution as

$$
\begin{aligned}
\bar{\theta} &= \theta\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right), \\
\bar{\phi}_{\text{dep}} &= \phi_{\text{dep}}\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right), \\
\bar{\phi}_{\text{res}} &= \phi_{\text{res}}\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right), \\
\bar{\phi}_{\text{gov}} &= \phi_{\text{gov}}\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right).
\end{aligned}
$$

We have

$$\boldsymbol{R}^A_{t+1} = \left[R^A_{t+1}\right]$$

$$\boldsymbol{R}^B_{t+1} = \begin{bmatrix} R^B_{\text{dep},t+1} \\ R^B_{\text{res},t+1} \\ R^B_{\text{gov},t+1} \end{bmatrix}$$

$$\theta_t(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}\}^\infty_{s=t}) = \left[\theta\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right)\right]$$

$$\boldsymbol{\phi}_t(\{R^A_{s+1}, \boldsymbol{R}^B_{s+1}\}^\infty_{s=t}) = \begin{bmatrix} \phi_{\text{dep}}\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right) \\ \phi_{\text{res}}\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right) \\ \phi_{\text{gov}}\left(R^A_{t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}, R^B_{\text{gov},t+1}; R^W_{t+1}\right) \end{bmatrix}$$

Note that we treat R^W_{t+1} as exogenous as in Bianchi and Bigio (2022): the government supplies discount window loans perfectly elastically at R^W_{t+1}. The quantity of discount window loans does not appear in the balance sheet of intermediaries at the stage when they decide on their portfolio allocation. It only enters the government budget constraint: we can either treat it as a part of lump-sum tax T_t or model it explicitly by adding an extra term that depends on other returns on assets (because demand for discount window loans is a function of these returns).

Given the derivation above and the calibration in Bianchi and Bigio (2022), the financial sector's holdings and issuance of assets and liabilities respond to next-period expected returns as in Table 3:

	ϕ_{reserve}	ϕ_{deposit}	ϕ_{govbnd}	θ
R^B_{reserve}	-5.5	0	-460	-470
R^B_{deposit}	0.5	0	-9.8	-9.2
R^B_{govbnd}	310	0	-4200	-3900
R^A	-315	0	5100	4800

Table 3: The financial sector's responses to returns in Bianchi and Bigio (2022). We use the extended version of the model calibrated to match 2006 as the steady state. See Section 5 of Bianchi and Bigio (2022) for details. Note: positive values for sensitivities of ϕ mean that liabilities of this type increase (holdings of this type decrease). Columns do not sum up to up due to zero due to rounding.

C.2 Gertler, Kiyotaki, Queralto (2012)

Gertler et al. (2012) extend the Gertler-Karadi-Kiyotaki (GKK) framework by allowing intermediaries to issue outside equity. Outside equity affects the ability of intermediaries to divert assets: it is more difficult to divert assets funded by deposits than outside equity. Let $\bar{\theta}$ be the portfolio share of loans, $\bar{\phi}_{\text{dep}}$ the porfiolio share of deposits, and $\bar{\phi}_{\text{oeq}}$ the share of outside equity. Intermediary can divert fraction χ of assets and run away, where

$$\chi = \chi_0 \left(1 + \chi_1 \frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}} + \frac{\chi_2}{2} \left(\frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}} \right)^2 \right),$$

The rates of return are $R^A_{t+1}, R^B_{\text{dep},t+1}$, and $R^B_{\text{oeq},t+1}$.

Bellman equation of an intermediary is[10]

$$\begin{aligned}\eta_t = \max_{\bar{\phi}_{\text{oeq}},\bar{\theta}} \Lambda_{t,t+1} \left(f + (1-f)\,\eta_{t+1}\right) \left[\bar{\theta}\left(R^A_{t+1} - R^B_{\text{dep},t+1}\right) - \bar{\phi}_{\text{oeq}}\left(R^B_{\text{oeq},t+1} - R^B_{\text{dep},t+1}\right) + R^B_{\text{dep},t+1}\right] \\ + \lambda_t \left[\eta_t - \chi_0 \left(1 + \chi_1 \frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}} + \frac{\chi_2}{2} \left(\frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}} \right)^2 \right) \bar{\theta} \right],\end{aligned}$$

where $\Lambda_{t,t+1}$ is the discount factor, η_t is the marginal value of net worth, and λ_t is the Lagrange multiplier on the incentive compatibility constraint. We assume the discount factor is a function of returns only, $\Lambda_{t,t+1} = \Lambda(R^A_{t+1}, R^B_{\text{oeq},t+1}, R^B_{\text{dep},t+1}$.

If the incentive compatibility constraint is binding, first order conditions with respect to portfolio shares result in

$$\frac{R^A_{t+1} - R^B_{\text{dep},t+1}}{R^B_{\text{oeq},t+1} - R^B_{\text{dep},t+1}} = - \frac{1 - \frac{\chi_2}{2}\left(\frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}}\right)^2}{\chi_1 \left(\frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}}\right) + \chi_2 \left(\frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}}\right)^2}$$

which determines the ratio $\bar{\phi}_{\text{oeq}}/\bar{\theta}$ as a function of returns, $x(R^A_{t+1}, R^B_{\text{oeq},t+1}, R^B_{\text{dep},t+1})$. This allows us to write

$$\bar{\phi}_{\text{oeq}} = x(R^A_{t+1}, R^B_{\text{oeq},t+1}, R^B_{\text{dep},t+1}) \cdot \bar{\theta}.$$

[10]See Chiang and Zoch (2023) for a detailed derivation of a similar result in a basic GKK framework.

The above result together with $\eta_t = \chi_0 \left(1 + \chi_1 \frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}} + \frac{\chi_2}{2} \left(\frac{\bar{\phi}_{\text{oeq}}}{\bar{\theta}}\right)^2\right) \bar{\theta}$ gives us

$$\eta_t = \Lambda_{t,t+1} \left(f + (1-f)\, \eta_{t+1}\right) \left[\frac{\left(R^A_{t+1} - R^B_{\text{dep},t+1} - x\left(\cdot\right) \left(R^B_{\text{oeq},t+1} - R^B_{\text{dep},t+1}\right)\right)}{\chi_0 \left(1 + \chi_1 x\left(\cdot\right) + \frac{\chi_2}{2} x\left(\cdot\right)^2\right)} \eta_t + R^B_{\text{dep},t+1} \right].$$

This links η_t to η_{t+1}, $\Lambda_{t,t+1}$ and returns. Given our assumption that $\Lambda_{t,t+1}$ is a function of returns only, we have

$$\eta_t = \eta \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1} R^B_{\text{dep},s+1} \right\}_{s \geq t} \right).$$

By using the incentive compatibility constraint, the optimal relationship between $\bar{\theta}$ and $\bar{\phi}_{\text{oeq}}$, and the balance sheet, we have

$$\begin{aligned}
\bar{\theta} &= \frac{1}{\chi_0 \left(1 + \chi_1 x\left(\cdot\right) + \frac{\chi_2}{2} x\left(\cdot\right)^2\right)} \eta \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1} R^B_{\text{dep},s+1} \right\}_{s \geq t} \right), \\
\bar{\phi}_{\text{oeq}} &= \frac{x\left(\cdot\right)}{\chi_0 \left(1 + \chi_1 x\left(\cdot\right) + \frac{\chi_2}{2} x\left(\cdot\right)^2\right)} \eta \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1} R^B_{\text{dep},s+1} \right\}_{s \geq t} \right), \\
\bar{\phi}_{\text{dep}} &= \frac{1 - x\left(\cdot\right)}{\chi_0 \left(1 + \chi_1 x\left(\cdot\right) + \frac{\chi_2}{2} x\left(\cdot\right)^2\right)} \eta \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1} R^B_{\text{dep},s+1} \right\}_{s \geq t} \right) - 1,
\end{aligned}$$

which shows that $\bar{\theta}, \bar{\phi}_{\text{oeq}}, \bar{\phi}_{\text{dep}}$ can be expressed as functions of returns only:

$$\begin{aligned}
\bar{\theta} &= \theta \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1}, R^B_{\text{dep},s+1} \right\}_{s \geq t} \right), \\
\bar{\phi}_{\text{oeq}} &= \phi_{\text{oeq}} \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1}, R^B_{\text{dep},s+1} \right\}_{s \geq t} \right), \\
\bar{\phi}_{\text{dep}} &= \phi_{\text{dep}} \left(\left\{ R^A_{s+1}, R^B_{\text{oeq},s+1}, R^B_{\text{dep},s+1} \right\}_{s \geq t} \right).
\end{aligned}$$

C.3 Sims and Wu (2021)

Sims and Wu (2021) extend the GKK framework by introducing interest-bearing reserves and allowing for a difference between diversion rates of private loans and government bonds. Let $\bar{\theta} \geq 0$ be the portfolio share of private loans and $\bar{\phi}_{\text{gov}}, \bar{\phi}_{\text{res}} \leq 0, \bar{\phi}_{\text{dep}} \geq 0$ the portfolio shares of government bonds, reserves, and deposits. The balance sheet equation is

$$\bar{\theta} = \bar{\phi}_{\text{gov}} + \bar{\phi}_{\text{res}} + \bar{\phi}_{\text{dep}} + 1.$$

The real gross rates of return are respectively R^A_{t+1},$R^B_{\mathrm{gov},t+1}$,$R^B_{\mathrm{res},t+1}$, and $R^B_{\mathrm{dep},t+1}$.

Intermediary can divert a fraction $\chi < 1$ of private loans and $\chi \cdot \Delta$ of government bonds, where $0 \leq \Delta \leq 1$. The reserve requirement is

$$\bar{\phi}_{\mathrm{res}} \leq -\varsigma \bar{\phi}_{\mathrm{dep}}.$$

Bellman equation is

$$\begin{aligned}
\eta_t = \max_{\bar{\theta}\geq 0, \bar{\phi}_{\mathrm{gov}} \leq 0, \bar{\phi}_{\mathrm{res}} \leq 0} & \Lambda_{t,t+1} \left(f + (1-f)\,\eta_{t+1}\right) \left[\bar{\theta}\left(R^A_{t+1} - R^B_{\mathrm{dep},t+1}\right)\right. \\
& - \bar{\phi}_{\mathrm{gov}}\left(R^B_{\mathrm{gov},t+1} - R^B_{\mathrm{dep},t+1}\right) - \bar{\phi}_{\mathrm{res}}\left(R^B_{\mathrm{res},t+1} - R^B_{\mathrm{dep},t+1}\right) + \left.R^B_{\mathrm{dep},t+1}\right] \\
& + \lambda_t\left[\eta_t - \chi\bar{\theta} + \chi\Delta\bar{\phi}_{\mathrm{gov}}\right] - \mu_t\left[\bar{\phi}_{\mathrm{res}} + \varsigma(\bar{\theta} - \bar{\phi}_{\mathrm{gov}} - \bar{\phi}_{\mathrm{res}} - 1)\right]
\end{aligned}$$

where $\Lambda_{t,t+1}$ is the discount factor, η_t is the marginal value of net worth, λ_t is the Lagrange multiplier on the incentive compatibility (IC) constraint, and μ_t is the multiplier on the reserve requirement (RR) constraint. We assume the discount factor is a function of returns only, $\Lambda_{t,t+1} = \Lambda(R^A_{t+1}, R^B_{\mathrm{gov},t+1}, R^B_{\mathrm{res},t+1}, R^B_{\mathrm{dep},t+1})$.

The program is linear in portfolio shares. Therefore, unless returns satisfy certain relations, the solution will involve corner solutions.

The reserve requirement constraint is binding whenever $R^B_{\mathrm{dep},t+1} > R^B_{\mathrm{res},t+1}$. The incentive compatibility constraint is binding whenever

$$\max\left\{R^A_{t+1}, R^B_{\mathrm{gov},t+1}\right\} \geq R_{\mathrm{fund},t+1}, \quad R_{\mathrm{fund},t+1} := R^B_{\mathrm{dep},t+1} + \frac{\varsigma}{1-\varsigma}\left(R^B_{\mathrm{dep},t+1} - R^B_{\mathrm{res},t+1}\right).$$

We will assume this is the case and also that $R^B_{\mathrm{dep},t+1} \geq R^B_{\mathrm{res},t+1}$. If

$$\frac{R^B_{\mathrm{gov},t+1} - \Delta R^A_{t+1}}{1-\Delta} = R_{\mathrm{fund},t+1}$$

the intermediary is indifferent between holding government debt and loans. Let

$$R^*\left(R^A_{t+1}, R^B_{\mathrm{gov},t+1}, R^B_{\mathrm{dep},t+1}, R^B_{\mathrm{res},t+1}\right) := \begin{cases} R^A_{t+1}, & R^B_{\mathrm{gov},t+1} \leq (1-\Delta)\,R_{\mathrm{fund},t+1} + \Delta R^A_{t+1} \\ R^B_{\mathrm{gov},t+1,} & R^B_{\mathrm{gov},t+1} > (1-\Delta)\,R_{\mathrm{fund},t+1} + \Delta R^A_{t+1} \end{cases}$$

We can rewrite the Bellman equation as

$$\eta_t = \Lambda_{t,t+1}\left(f + (1-f)\,\eta_{t+1}\right)\left[\left(R^*\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}\right) - R_{\text{fund},t+1}\right)\frac{\eta_t}{\chi} + R_{\text{fund},t+1}\right].$$

Because we assumed $\Lambda_{t,t+1}$ is a function of returns only, we have

$$\eta_t = \eta\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{res},t+1}, \eta_{t+1}\right).$$

We can therefore use the binding incentive compatibility constraint to get

$$\bar{\theta} - \Delta\bar{\phi}_{\text{gov}} = \frac{1}{\chi}\eta\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{res},t+1}, \eta_{t+1}\right)$$

The exact portfolio composition depends on returns. For example, if the incentive compatibility constraint is binding and $R_{\text{dep},t+1} \geq R_{\text{res},t+1}$, we can express demand for loans as

$$\theta\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}\right) \coloneqq \begin{cases} \frac{1}{\chi}\eta\left(\cdot\right), & (1-\Delta)\,R_{\text{fund},t+1} + \Delta R^A_{t+1} > R^B_{\text{gov},t+1}, \\ \left[0, \frac{1}{\chi}\eta\left(\cdot\right)\right], & (1-\Delta)\,R_{\text{fund},t+1} + \Delta R^A_{t+1} = R^B_{\text{gov},t+1}, \\ 0, & (1-\Delta)\,R_{\text{fund},t+1} + \Delta R^A_{t+1} < R^B_{\text{gov},t+1}. \end{cases}$$

The portfolio share of government debt is

$$\phi_{\text{gov}}\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}\right) = \frac{1}{\Delta}\left(\theta\left(\cdot\right) - \frac{1}{\chi}\eta\left(\cdot\right)\right).$$

If $R_{\text{dep},t+1} > R_{\text{res},t+1}$ we can use the reserve requirement together with the balance sheet to get

$$\begin{aligned} \phi_{\text{res}}\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}\right) &= -\frac{\varsigma}{1-\zeta}\left(\theta\left(\cdot\right) - \phi_{\text{gov}}\left(\cdot\right) - 1\right) \\ \phi_{\text{dep}}\left(R^A_{t+1}, R^B_{\text{gov},t+1}, R^B_{\text{dep},t+1}, R^B_{\text{res},t+1}\right) &= \frac{1}{1-\zeta}\left(\theta\left(\cdot\right) - \phi_{\text{gov}}\left(\cdot\right) - 1\right). \end{aligned}$$

If $R_{\text{dep},t+1} = R_{\text{res},t+1}$, then only the sum $\bar{\phi}_{\text{res}} + \bar{\phi}_{\text{dep}}$ is determined and equals

$$\theta\left(\cdot\right) - \phi_{\text{gov}}\left(\cdot\right) - 1.$$

In this model, if rates of return are such that none of the portfolio shares is zero, demand for assets will be perfectly elastic. If rates of return are such that some asset has a portfolio share equal to zero demand for this asset will be inelastic. For example, if the return on loans sufficiently dominates the return on government debt, demand for debt will be inelastic. In this case, demand for loans will have an intermediate elasticity.

C.4 Heterogeneity in The Financial Sector

There is a continuum of financial intermediaries indexed by $j \in [0,1]$ with net worth $n_{j,t}$. For each unit of net worth, the intermediary can hold $\theta_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right)$ units of capital and issue $\boldsymbol{\phi}_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right)$ units of financial assets. Suppose that the net worth of intermediary j relative to the total financial sector is given by density $\mu_{j,t}$:

$$\frac{n_{j,t}}{n_t} \coloneqq \mu_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right),$$

where

$$\int \mu_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) dj = 1.$$

We define the following aggregate functions:

$$\begin{aligned}
\theta_t\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) &\coloneqq \int \theta_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) \mu_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) dj, \\
\boldsymbol{\phi}_t\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) &\coloneqq \int \boldsymbol{\phi}_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) \mu_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) dj.
\end{aligned}$$

Given these aggregate functions, we have

$$\begin{aligned}
\theta_{t,\boldsymbol{R}_{k+1}^B}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) &= \int \theta_{j,t,\boldsymbol{R}_{k+1}^B}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) \mu_{j,t} + \theta_{j,t}\mu_{j,t,\boldsymbol{R}_{k+1}^B}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) dj, \\
\boldsymbol{\phi}_{t,\boldsymbol{R}_{k+1}^B}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) &= \int \boldsymbol{\phi}_{j,t,\boldsymbol{R}_{k+1}^B}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) \mu_{j,t} + \boldsymbol{\phi}_{j,t}\mu_{j,t,\boldsymbol{R}_{k+1}^B}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) dj,
\end{aligned}$$

and the return on net worth is given by:

$$\begin{aligned}
R_t^N &= \int R_{j,t}^N \mu_{j,t} dj \\
&= R_t^A \int \theta_{j,t}\mu_{j,t} dj - \boldsymbol{R}_t^{B\intercal} \int \boldsymbol{\phi}_{j,t}\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) \mu_{j,t} dj \\
&= R_t^A \theta_t\left(\{R_s^A, \boldsymbol{R}_s^B\}\right) - \boldsymbol{R}_t^{B\intercal} \boldsymbol{\phi}_t\left(\{R_s^A, \boldsymbol{R}_s^B\}\right).
\end{aligned}$$

D Empirical Appendix

D.1 Data Construction

D.1.1 Balance Sheets of Banks

Our primary data source is the Reports of Condition and Income (Call Reports) U.S. Call Reports provided by the Federal Financial Institutions Examination Council. The dataset contains quarterly data on the income statements and balance sheets of all U.S. commercial banks. We use data from 2001Q1 to 2024Q4.

For each reporting entity in the dataset we calculate the following items on the balance sheet:

- "Total assets" (RCFD2170);
- "Equity" (RCFD3210);
- "Cash" (sum of codes RCFD0071 "Interest Bearing Balances" and RCFD0081 "Non-interest Bearing Balances and Coin");
- "Net loans and leases" (RCFDB528);
- "Fed funds sold and securities purchased under agreement to sell" (sum of RCFDB987 "Federal Funds Sold" and RCFDB989 "Securities Purchased Under Agreement to Sell");
- "Trading assets" (RCFD3545);
- "Trading Treasury securities" (RCON3531);
- "Total securities holdings" (sum of RCFD1773 "Securities Available for Sale" and RCFD1754 "Held-to-maturity Securities", substituting RCFDJJ34 for RCFD1754 after December 31, 2018, when regulatory reporting codes changed);
- "Treasury securities holdings" (sum of codes RCON1287 "Fair Value of Available-for-Sale U.S. Treasury Securities" and RCON0211 "Amortized Cost of Held-to-Maturity U.S. Treasury Securities");
- "Deposits" (sum of codes RCON2200 "Domestic Deposits" and RCFN2200 "Foreign Deposits");

- “Fed funds purchased and securities sold under agreement to repurchase” (sum of RCFDB993 “Federal Funds Sold” and RCFDB995 “Securities Sold Under Agreement to Repurchase”);
- “Trading liabilities” (RCFD3548);
- “Other borrowed money” (RCFD3190).

Banks without foreign offices have missing values of the RCFD codes. In this case, we use the RCON codes instead (i.e., RCON3190 for other borrowed money).

We merge Call Reports data with Center of Research in Security Prices (CRSP) on stock prices, dividends and returns (1980-2024) by using a crosswalk provided by the Federal Reserve Bank of New York. We retain only entities with a 1-1 mapping between “permco” in CRSP data and “bhcid” in Call Reports data.

For each “permco” we calculate products of the number of outstanding shares (code “SHROUT”) and prices (code “PRC”), which we then label as “TCAP”. Since one entity can have more than one type of stock outstanding, we calculate a sum of all “TCAP” corresponding to a given entity at each date. We define “Market value” as the first sum of “TCAP” of an entity in a given quarter.

We calculate “Other securities holdings” as “Total securities holdings” - “Treasury securities holdings.” We calculate “Other trading securities” as “Trading assets” - “Trading Treasury securities.”

We define “Financial assets” as: “Cash” + “Net loans and leases” + “Fed funds sold and securities purchased under agreement to sell” + “Other trading securities” + “Trading Treasury securities” + “Other securities holdings” + “Treasury securities holdings.”

We define “Other liabilities” as: “Liabilities” - “Deposits” - “Fed funds purchased and securities sold under agreement to repurchase” - “Trading liabilities” - “Other borrowed money.”

We define “Non-financial assets” as: “Market value“ + “Liabilities” - “Financial assets.”

We now group assets net of liabilities into three categories: “Liquid assets”, “Illiquid assets”, and “Capital”:

- We define “Liquid assets” as: “Cash” + “Treasury securities holdings” + “Fed funds sold and securities purchased under agreement to sell” + “Trading Treasury securities” - “Deposits” - “Fed funds purchased and securities sold under agreement to repurchase;”
- We define “Illiquid assets” as: “Other securities holdings” + “Other trading securities” - “Trading liabilities;”
- We define “Capital” as: ”Non-financial assets” + “Loans” - “Other borrowed money” - “Other liabilities.”

D.1.2 Foreign

To construct foreign holdings of Treasury securities, we follow Chaudhary et al. (2025) and use information from the Treasury International Capital system (TIC) to obtain country-level holdings of Treasury securities. For long-term Treasury securities we use estimates from Bertaut and Judson (2022). For short-term holdings, we use data from the banking liabilities survey.

D.1.3 Returns

We use the 3-month Treasury Bills yield as expected nominal return on “Liquid assets” and Moody’s Seasoned BAA Corporate Bond yield as expected nominal return on “Illiquid assets.” We denote annualized rates of return as $E_t\left[i_{t+1}^{\text{Liquid}}\right]$ and $E_t\left[i_{t+1}^{\text{Illiquid}}\right]$.

In CRSP data we calculate daily dividends on each stock of an entity as returns (code “RET”) minus returns without dividends (code “RETX”) times the market value of a stock (“TCAP”). To calculate the total dividends of an entity we sum over all types of stocks. To calculate total returns of an entity, we sum “RET” multiplied by “TCAP” over all stocks of an entity and divide by the sum of “TCAP”.

Because CRSP data is daily, this gives us a daily series of dividends and returns. To obtain quarterly variables from CRSP data, we use the following procedure. First, we define “Market value” as the first sum of “TCAP” of an entity within a given quarter. To calculate “Quarterly dividends,” we sum all daily dividends of an entity in a quarter. To calculate “Quarterly returns,” we compute the product of daily returns. We denote nominal; "Quarterly returns" of an entity j in a quarter t as $i^N_{j,t+1}$. To obtain the "Dividend to price ratio" in a quarter t we divide the sum of "Quarterly dividends" in the four preceding quarters by the "Market value" in the first day of the quarter t. We denote the resulting series as $dp_{j,t}$.

We follow the same procedure with the aggregate S&P500 index to obtain the series of "Dividend to price ratio" $dp_{S\&P500,t}$

To obtain nominal expected quarterly returns $E_t\left[i^N_{j,t+1}\right]$ for each entity, we estimate:

$$\begin{aligned}\text{Excess return}_{j,t+1} &= \alpha_{j,0} + \beta_{j,1} dp_{S\&P500,t} + \beta_{j,2} dp_{j,t} \\ &+ \beta_{j,3}\text{Term premium}_t + \beta_{j,4}\log \text{Market Value}_{j,t} + \epsilon_{j,t+1},\end{aligned}$$

where $\text{Excess return}_{j,t+1} := i^N_{j,t+1} - E_t\left[i^{\text{Liquid}}_{t+1}\right]$, Term premium_t is the difference between the 10-year and 3-month Treasury yields. We use the predicted values of $\text{Excess return}_{j,t+1}$, $\widehat{\text{Excess return}}_{j,t+1}$, to construct

$$E_t\left[i^N_{j,t+1}\right] := \widehat{\text{Excess return}}_{j,t+1} + E_t\left[i^{\text{Liquid}}_{t+1}\right].$$

To recover the nominal expected return on capital, $E_t\left[i^{\text{Capital}}_{j,t+1}\right]$ we use

$$\begin{aligned}E_t\left[i^N_{j,t+1}\right] &= \frac{\text{Capital}_{j,t}}{\text{Market Value}_{i,t}} E_t\left[i^{\text{Capital}}_{j,t+1}\right] \\ &+ \frac{\text{Liquid Assets}_{j,t}}{\text{Market Value}_{j,t}} E_t\left[i^{\text{Liquid}}_{j,t+1}\right] + \frac{\text{Illiquid Assets}_{j,t}}{\text{Market Value}_{j,t}} E_t\left[i^{\text{Illiquid}}_{j,t+1}\right].\end{aligned}$$

We then calculate the aggregate return $E_t\left[i^{\text{Capital}}_{t+1}\right]$ as

$$E_t\left[i^{\text{Capital}}_{t+1}\right] := \sum_{j=1}^{J} \frac{\text{Capital}_{j,t}}{\sum_{i=1}^{J} \text{Capital}_{j,t}} E_t\left[i^{\text{Capital}}_{j,t+1}\right]$$

We calculate real expected returns by subtracting 1-year expected inflation series from Cleveland Fed.

Finally, we define:

$$E_t\left[R^A_{t+1}\right] \coloneqq E_t\left[1+\frac{i^{\text{Capital}}_{t+1}}{400}\right]$$
$$E_t\left[R^B_{\text{Liquid},t+1}\right] \coloneqq E_t\left[1+\frac{i^{\text{Liquid}}_{t+1}}{400}\right]$$
$$E_t\left[R^B_{\text{Illiquid},t+1}\right] \coloneqq E_t\left[1+\frac{i^{\text{Illiquid}}_{t+1}}{400}\right].$$

D.2 Construction of GIV

We describe how we construct Granular Instrumental Variables (GIVs) in our setting. We follow the procedure outlined in Gabaix and Koijen (2024).

We first obtain an instrument associated with foreign holdings. We work with logarithms of country-level holdings of Treasury securities, $\log \text{Treasury}_{c,t}$, where c represents a country and t is a time period. We postulate

$$\log \text{Treasury}_{c,t} = \overline{\log \text{Treasury}}_c(t) + \boldsymbol{\Lambda}_c \boldsymbol{\eta}^C_t + u_{c,t}.$$

Here $\overline{\log \text{Treasury}}_c(t)$ is a country-specific quadratic trend, $\boldsymbol{\eta}^C_t$ is a vector of aggregate factors, and $\boldsymbol{\Lambda}_c$ are the loadings on them. The last term, $u_{c,t}$, represents idiosyncratic i.i.d. country-level shocks. For each country, we remove a country-specific quadratic trend and a common time fixed effect (we assume the loadings on the first element of $\boldsymbol{\eta}^C_t$ are equal for all countries). We then remove the first principal component from the panel of detrended country-level holdings. We call the residuals $\hat{u}_{c,t}$. They are proxies for $u_{c,t}$. We define

$$z^{\text{TIC}}_t \coloneqq \sum_c s_c \hat{u}_{c,t}, \quad s_c \coloneqq \frac{\sum_t \text{Treasury}_{c,t}}{\sum_c \sum_t \text{Treasury}_{c,t}}.$$

We next construct GIVs in Call Reports data. Let

$$\boldsymbol{\phi}_{j,t} \coloneqq -\frac{1}{\text{Market Value}_{j,t}}\begin{bmatrix} \text{Liquid Assets}_{j,t} \\ \text{Illiquid Assets}_{j,t} \end{bmatrix}.$$

We assume portfolio shares $\boldsymbol{\phi}_{j,t}$ are given by Equation (21), i.e:

$$\boldsymbol{\phi}_{j,t} = \bar{\boldsymbol{\phi}}_j(t) + \boldsymbol{\phi}_{\boldsymbol{R}^B}\boldsymbol{S}_{t+1} + \boldsymbol{\Lambda}_j\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_{j,t}, \tag{29}$$

where $\bar{\boldsymbol{\phi}}_j(t)$ are entity-specific quadratic time trends, $\boldsymbol{\phi}_{\boldsymbol{R}^B}$ parameterizes responses to deviations of spreads from trends,

$$\boldsymbol{S}_{t+1} := \mathbb{E}_t[\boldsymbol{R}^B_{t+1}] - \mathbb{E}_t[R^A_{t+1}] \otimes \mathbf{1}_{\mathcal{B}},$$

$\boldsymbol{\eta}_t$ are unobserved aggregate shocks on which entity j has loadings $\boldsymbol{\Lambda}_j$, and $\boldsymbol{\varepsilon}_{j,t}$ are idiosyncratic shocks i.i.d. by entity-asset-time.

Our goal is to obtain a proxy for $\boldsymbol{\varepsilon}_{j,t}$. We remove entity-asset specific quadratic trends and an asset-specific time fixed effect common for all entities. We normalize the resulting detrended portfolio shares by dividing them by their standard deviation (common for all entities). We then remove three principal components from the panel of detrended and normalized entity-level portfolio shares. We are left with a vector of residuals $\hat{\boldsymbol{\varepsilon}}_{j,t}$ for each entity. We construct GIVs as

$$\boldsymbol{z}_t^{\text{Call Reports}} := \sum_j s_j \hat{\boldsymbol{\varepsilon}}_{j,t}, \quad s_j := \frac{\sum_t \text{Market Value}_{j,t}}{\sum_j \sum_t \text{Market Value}_{j,t}}.$$

We now estimate $\boldsymbol{\phi}_{\boldsymbol{R}^B}$ in Equation (21). The key difficulty is the endogeneity of $\boldsymbol{S}_{t+1}$–these returns are determined by equilibrium conditions. However, $\boldsymbol{\varepsilon}_{j,t}$ are by construction i.i.d.. We use three instruments for $\boldsymbol{S}_{t+1}$: one from TIC and two from Call Reports,

$$\boldsymbol{z}_t := \begin{bmatrix} z_t^{\text{TIC}} \\ \boldsymbol{z}_t^{\text{Call Reports}} \end{bmatrix}.$$

We calculate simple cross-sectional averages of $\boldsymbol{\phi}_{j,t}$. We regress them on $\boldsymbol{S}_{t+1}$, with $\boldsymbol{z}_t$ as instruments, controlling for aggregate factors recovered by using principal component analysis (1 from TIC data and 3 from Call Reports data) and quadratic trend. We do it separately for liquid and illiquid assets separately. Table 1 in Section 6.1 shows the estimates.

Our instruments are strong: the value of the effective F-statistic is 28.1 and 11.5 for the spread between R^B_{liquid} and R^A and between R^B_{illiquid} and R^A, respectively.

D.3 Model Calibration

To calculate the steady-state portfolio shares $\bar{\boldsymbol{\phi}}$ and $\bar{\theta}$, we first compute portfolio shares for each period. We sum the negative of "Liquid Assets" and "Illiquid Assets" across all entities j in the Call Reports and divide by the sum of "Market Value." These variables are defined in Appendix D.1.1. This yields $\boldsymbol{\phi}_t$ for each period. We then calculate the sample average of these ratios from 2001Q1 to 2024Q4 to obtain $\bar{\boldsymbol{\phi}}$. We compute $\bar{\theta}$ using the balance sheet identity: $\bar{\theta} = \bar{\phi}\text{liquid} + \bar{\phi}\text{illiquid} + 1$. We normalize the financial sector's net worth, $\bar{n}$, to one.

The steady-state real returns on assets are obtained as long-run averages of returns calculated in Appendix D.1.3.

To determine steady-state household holdings of liquid assets, $\bar{b}_{\text{liquid}}$, we calculate the long-run average of the financial sector's liquid *liabilities* using Call Report data. This category comprises two balance sheet positions: "Deposits" and "Fed funds purchased and securities sold under agreement to repurchase."

By construction, the sum of these liabilities exceeds the $\bar{\phi}_{\text{liquid}}$ calculated for the financial sector. The portfolio share $\bar{\phi}_{\text{liquid}}$ represents liquid asset issuance net of liquid asset holdings and therefore does not appropriately measure the value of assets held by households. The difference between these measures largely corresponds to "Cash," "Treasury security holdings," and "Fed funds sold and securities purchased under agreement to sell," which we attribute to $\bar{b}^G_{\text{liquid}}$.

We set $\bar{b}_{\text{illiquid}}$ equal to $\bar{\phi}_{\text{liquid}}$, which implies that $\bar{b}^G_{\text{illiquid}}$ equals zero. This assumption is justified because our measure of "Illiquid assets" excludes Treasury securities.

Steady-state levels of household asset holdings matter only insofar as they affect household asset demand sensitivities, given our target semielasticities with respect to returns of 0.1.

Finally, we assume that these long-run averages of returns and asset positions correspond to the Ramsey steady state. Under this assumption, the household asset positions calculated through this method correspond to the bliss point.