Financial Intermediation and Aggregate Demand: A Sufficient Statistics Approach

Yu-Ting Chiang[†] Piotr Zoch[‡]

April 14, 2026

Abstract

We show that the financial sector's asset supply elasticities are sufficient statistics summarizing its macroeconomic effects for a large class of financial frictions. These elasticities are crucial for a wide range of policy questions, ranging from the size of fiscal multipliers to the relative effectiveness of asset purchases targeting the financial sector versus tax cuts targeting households. Workhorse macroeconomic models imply different values of these elasticities, generating output responses to policies that differ by orders of magnitude. We construct empirical measures of these elasticities and evaluate their policy implications in a quantitative model with household heterogeneity and illiquidity.

Keywords: financial frictions, sufficient statistics, HANK, monetary and fiscal policy

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, Anmol Bhandari, Harris Dellas, Bill Dupor, Mikhail Golosov, Marcus Hagedorn, Lars Hansen, Greg Kaplan, Ralph Luetticke, Kurt Mitman, B. Ravikumar, Matthew Rognlie, Wenting Song, Ludwig Straub, Vincent Sterk, Mikayel Sukiasyan, Harald Uhlig, Christian Wolf, and Mark Wright for valuable suggestions. Zoch acknowledges the support of National Science Centre, Poland (2019/35/N/HS4/02189). Views in this paper do not necessarily represent those of the Federal Reserve System.

1 Introduction

Financial intermediation is central to macroeconomics, influencing the transmission of policies and shocks through asset markets. However, models of financial intermediation often rely on various microfoundations with frictions governed by parameters that are intrinsically hard to measure, such as asset diversion rates or monitoring costs. The intricacy of these models makes it challenging to identify and quantify features of the intermediation process that are most relevant for aggregate outcomes. This poses an obstacle to integrating financial frictions into modern quantitative macroeconomic models, which are themselves becoming increasingly complex. Recent approaches address this complexity by identifying key features shared across a wide range of models for each "block" of the economy. For instance, a growing consensus suggests that households' intertemporal marginal propensities to consume effectively summarize their aggregate responses across various models (Auclert et al., 2024). Yet, no similar attempt has been made to derive a counterpart for the financial sector. Our paper addresses this issue.

Our main idea is a simple observation that financial intermediaries are, effectively, suppliers of assets: They take one type of asset, such as loans, and transform it into a different type, such as deposits. As they are suppliers of assets, their supply curves fully describe how they respond to changes in prices and quantities of assets. To the extent that we know the shape of their asset supply, details of the intermediation process are irrelevant.

We build on this observation to derive a set of asset supply elasticities that serve as sufficient statistics to summarize frictions in financial intermediation models. These elasticities allow us to incorporate financial frictions into state-of-the-art quantitative macroeconomic models while remaining agnostic about their microfoundations. We derive formulas to isolate the channels through which these elasticities affect aggregate outcomes, measure them empirically, and compare them to implicit assumptions in standard macroeconomic models. These assumptions significantly affect conclusions about policy issues ranging from the impact of government spending and tax cuts to the effect of asset purchase programs, with output responses varying by up to orders of magnitude.

We derive these elasticities in a general framework that nests models of financial intermediation with various microfoundations. Households consume and save in different assets, with some assets being "liquid" and thus preferable. The financial sector issues liquid assets and holds illiquid capital, supplying liquidity to the economy under frictions.

We show that financial frictions across various microfoundations—asset diversion (Gertler and Karadi, 2011), costly state verification (Bernanke et al., 1999), costly leverage (Cúrdia and Woodford, 2016), and collateral constraints (Kiyotaki and Moore, 1997)—all share a common structure: The financial sector's leverage responds to expected returns over different horizons, with the magnitude governed by two *leverage sensitivity* parameters and the dependency on horizon governed by a *forward-looking component*. These parameters can be mapped directly to data and, in turn, determine the *liquidity supply elasticities* that characterize the financial block of the model.

We embed this financial block into a modern macroeconomic model with household heterogeneity, illiquidity, and nominal rigidities. The government influences aggregate outcomes through spending, taxes, transfers, debt issuance, interest rates, and asset purchases.

We analyze how the financial block affects aggregate outcomes through an intertemporal supply-demand system of goods and assets. The representation is similar to Auclert et al. (2021), Wolf (2025), Dávila and Schaab (2023), McKay and Wolf (2023), and Angeletos et al. (2024), though these works abstract from financial intermediation. In this system, the financial sector interacts with other model components only through asset returns and quantities. As a result, its elasticities are sufficient statistics that contain all relevant information up to first-order approximation. Moreover, our sufficient statistics are "portable" and can be easily used to introduce financial frictions to quantitative frameworks using the sequence space approach.

Financial frictions influence policy transmission through an *asset market channel* that depends on *cross-price elasticities* of liquidity supply with respect to capital returns. Low elasticities indicate intermediaries view liquid assets and capital as less substitutable. Other things equal, an increase in liquidity (e.g., from government debt issuance) then leads to large increases in capital prices, raising aggregate demand through investment and consumption. Despite their crucial role, macroeconomic models assume a broad range of values for these elasticities, from zero to infinity. These assumptions are often embedded in model setups or implicit in functional forms. Our approach makes these assumptions explicit, emphasizing the need for empirical measurement.

We demonstrate how to empirically discipline these elasticities by estimating the leverage sensitivities and the forward-looking component for the U.S. banking sector. To address potential identification threats, we construct instruments from structural shocks using proxies for monetary policy shocks (Bauer and Swanson, 2023) and oil supply shocks (Baumeister and Hamilton, 2019). Our estimates indicate that the U.S. banking sector's liquidity supply elasticities are twice as large as those implied by functional forms in standard financial intermediation models.

Beyond the specific empirical implementation, our theoretical framework has broader applicability for a large literature in empirical finance that aims to estimate similar elasticities (Gabaix and Koijen, 2021, 2024; Chaudhary et al., 2024; Jansen et al., 2024), using different identification assumptions. Our sufficient statistic result provides useful guidance for such work by clarifying how these elasticities matter for aggregate outcomes and how they should be measured—for instance, by highlighting the importance of cross-price elasticities, which existing studies have mostly abstracted from. These elasticities—whether from our estimates or others—provide useful target moments for calibrating financial frictions in quantitative models, including those studying their interactions with complex household consumption-saving behaviors, such as Lee et al. (2020), Bayer et al. (2023), Fernández-Villaverde et al. (2023), Lee (2021), Mendicino et al. (2021), Cui and Sterk (2021), and Ferrante and Gornemann (2022).

We illustrate the quantitative importance of our sufficient statistics with two policy questions: (1) the size of government spending multipliers, and (2) a "Wall Street vs. Main Street" debate: Can asset purchase programs stimulate the economy more effectively than tax cuts to households? We use a standard two-asset heterogeneous agent model calibrated to match asset holdings and consumption responses in the data. Holding all else constant, we vary the financial sector's liquidity supply elasticities from our empirical estimates to values implied by common model assumptions.

Aggregate outcomes under each policy vary significantly across specifications. Government spending multipliers can differ by up to a factor of two, with greater variation under a higher degree of debt financing—as there is a larger impact on asset markets. The "Wall Street vs. Main Street" debate depends even more critically on these elasticities: Output responses to asset purchases differ by orders of magnitude, and the effects of tax cuts vary by a factor of three. Our empirical elasticities are relatively high, implying modest asset market responses compared to standard financial intermediation models. Consequently, they suggest tax cuts targeting households can stimulate output more effectively than asset purchases, which rely heavily on the asset market channel. These results highlight the importance of our sufficient statistics in providing empirical discipline on financial frictions for policy analysis.

Finally, our approach has limitations. While the elasticities we characterize are invariant to policies operating through asset prices and quantities, they may not be invariant to macroprudential regulations that directly affect intermediation frictions. Studying such regulations requires microfounded models, and our elasticities should only serve as calibration targets.

Furthermore, our characterization relies on first-order approximation and cannot capture nonlinear dynamics like those in Brunnermeier and Sannikov (2014). Nevertheless, our approach incorporates financial frictions into quantitative models, addressing policy questions often precluded by simplifying assumptions in nonlinear macro-finance models. By providing a theoretical foundation for measuring financial frictions, it represents a first step toward quantifying these frictions in nonlinear settings.

2 Financial Intermediation

We consider an economy where the financial sector plays a central role in asset intermediation. Time is discrete, $t \in \{0, \dots, \infty\}$. The economy consists of three sectors: households, production, and financial intermediaries.

Households consume a final good, work, and save, choosing between two types of assets, which we refer to as *liquid* and *illiquid*. Liquid assets possess certain features that households prefer. The production sector combines capital and labor to produce final goods. The financial sector issues liquid assets and holds illiquid capital, transforming one type of asset into another under various frictions that depend on prevailing economic conditions.

2.1 The General Form

In each period t, the intermediary holds capital of value ω_t, issues liquid assets d_t, and has net worth n_t. All holdings and issuance are measured in units of final goods in period t. We use Θ_t to denote the intermediary's effective leverage, defined as the value of capital holdings relative to net worth:

$$\omega_t = \Theta_t n_t. \tag{1}$$

Intuitively, Θ_t represents the level of intermediation that can be achieved with each unit of net worth.

The intermediary's holding and issuance of assets are subject to a balance sheet constraint: $\omega_t = d_t + n_t$. From the balance sheet, the net issuance of liquid assets is:

$$d_t = (\Theta_t - 1)n_t.$$

Let r^K_{t+1} and r^B_{t+1} be the returns from investing one unit of final goods in capital and liquid assets from period t to $t+1$. The balance sheet implies that holding each unit of net worth generates return:

$$r^N_{t+1} = \Theta_t(r^K_{t+1} - r^B_{t+1}) + r^B_{t+1}.$$

Intermediation is subject to frictions. We formulate the frictions as a generic function $\Theta(\cdot)$ that maps the path of future returns to leverage:

$$\Theta_t = \Theta(\{r^K_{s+1}, r^B_{s+1}\}_{s\geq t}). \tag{2}$$

Returns $\{r^K_{s+1}, r^B_{s+1}\}_{s\geq t}$ represent the intermediary's future investment opportunities and funding costs, while the function $\Theta(\cdot)$ captures how financial frictions constrain intermediation capacity. As we detail in Section 2.2, this general form nests various microfoundations for financial frictions—such as agency problems originating from asset diversion or monitoring costs—and derives the common implications of these canonical frameworks.

While $\Theta(\cdot)$ describes intermediation per unit of net worth, the evolution of net worth over time is given by:

$$n_t = G(\Theta_{t-1}, r^K_t, r^B_t, n_{t-1}). \tag{3}$$

Function $G(\cdot)$ captures how net worth depends on past leverage Θ_{t-1}, realized returns r^K_t and r^B_t, and past net worth n_{t-1}. For example, higher past leverage can amplify both gains and losses, while the spread between returns on capital and funding costs determines profitability.

In canonical models, $\Theta(\cdot)$ limits the value of assets banks can intermediate per unit of net worth. Net worth evolves according to $G(\cdot)$, a slow-moving process that responds disproportionately to realized returns due to leverage, as in Gertler and Karadi (2011) and Christiano et al. (2014). While these models specify $G(\cdot)$ exogenously in various forms, most are special cases of our formulation in Equation (3), as we discuss in Appendix C.1.

Some papers explicitly model dividend and equity issuance decisions, such as Karadi and Nakov (2021) and Akinci and Queralto (2022). While the net worth process in these models does not take the exact form in Equation (3), we show in Appendix C.3 that they are nested in the following sense: given $\Theta(\cdot)$ and the endogenous dividend and equity issuance problem, we can construct a function $G(\cdot)$ dependent on $\Theta(\cdot)$ that generates identical aggregate responses up to first-order approximations.

2.2 Nested Models of Financial Intermediation

We demonstrate how our framework nests commonly used models of financial frictions with appropriate choices of $\Theta(\{r^K_{s+1}, r^B_{s+1}\}_{s\geq t})$. We focus on an asset diversion model based on Gertler and Kiyotaki (2010) and Gertler and Karadi (2011) (GKK) to illustrate the mapping, then briefly mention other examples.

Example: Asset Diversion Model

There is a continuum of intermediaries indexed by $j \in [0,1]$. Every period, after receiving returns from its asset holdings and repaying for its asset issuance, intermediary j exits with probability f and transfers its retained earnings to its owners. At the same time, new intermediaries enter with a total of m units of initial net worth to operate with. Conditional on survival, intermediary j chooses asset holdings and issuance, $\omega_{j,t}$ and $d_{j,t}$.

An agency problem constrains the level of intermediation. After obtaining funding $d_{j,t}$ and investing in assets $\omega_{j,t}$, the intermediary can divert a fraction $1/\theta$ of assets and run away. If this happens, the intermediary will be forced into bankruptcy and lose its continuation value, $v_{j,t}(n_{j,t})$. The amount of funding an intermediary can receive depends on the franchise value $v_{j,t}(n_{j,t})$, where $n_{j,t}$ is net worth — bank j must be better off continuing instead of running away. The optimization problem is:

$$v_{j,t}(n_{j,t}) = \max_{\omega_{j,t}, d_{j,t}, n_{j,t+1}} \Lambda_{t+1}\left(f n_{j,t+1} + (1-f)\, v_{j,t+1}(n_{j,t+1})\right),$$

such that

$$\omega_{j,t} \leq \theta v_{j,t}(n_{j,t}), \quad \omega_{j,t} = d_{j,t} + n_{j,t}, \quad n_{j,t+1} = \left(1 + r^K_{t+1}\right)\omega_{j,t} - \left(1 + r^B_{t+1}\right) d_{j,t},$$

where the first constraint is the incentive compatibility constraint resulting from the agency problem and $\Lambda_{t+1} = \Lambda(r^K_{t+1}, r^B_{t+1})$ is a function of returns that captures the intermediary's intertemporal valuation.

Linearity of the intermediary's problem in this model implies that $v_{j,t}(n_{j,t}) = \eta_t n_{j,t}$, where η_t is the continuation value per unit of net worth for all intermediaries. The incentive compatibility constraint becomes:

$$\omega_{j,t}/\theta \leq \eta_t n_{j,t}.$$

When the spreads between returns are positive, the constraint will be binding, and the level of intermediation is given by

$$\Theta_t = \theta \eta_t.$$

The optimization problem implies that the continuation value η_t depends on future profitability and survival prospects and takes the following recursive form:

$$\eta_t = \Lambda_{t+1}\left(f + (1-f)\eta_{t+1}\right) \times (1 + r^N_{t+1}),$$

where $r^N_{t+1} = \Theta_t(r^K_{t+1} - r^B_{t+1}) + r^B_{t+1}$ is the return on net worth.

Substituting $\Theta_t = \theta\eta_t$ into the recursive relationship, we have:

$$\Theta_t = \frac{\Lambda_{t+1}\left(f\theta + (1-f)\Theta_{t+1}\right)\left(1 + r^B_{t+1}\right)}{1 - \Lambda_{t+1}\left(f\theta + (1-f)\Theta_{t+1}\right)\left(r^K_{t+1} - r^B_{t+1}\right)}, \tag{4}$$

Iterating forward shows how Θ_t depends on the entire path of future returns $\{r^K_{s+1}, r^B_{s+1}\}_{s\geq t}$, as in our general formulation in Equation (2).

For the evolution of net worth, each intermediary's net worth, conditional on survival, evolves as

$$n_{j,t} = (1 + r^N_t)n_{j,t-1}.$$

In the aggregate, a fraction f of the intermediaries dissolve each period, and a constant initial net worth m flows in from new intermediaries. The aggregate net worth process has the following form:

$$G(\Theta_{t-1}, r^K_t, r^B_t, n_{t-1}) = (1-f)\left[1 + r^B_t + (r^K_t - r^B_t)\Theta_{t-1}\right] n_t + m. \tag{5}$$

Nesting Result

Besides the asset diversion model discussed above *(Model 1)*, our formulation encompasses other commonly used models of financial frictions. While these models differ in their microfoundations, the frictions on intermediation can all be captured by a function $\Theta(\cdot)$ with different properties that reflect the specific microfoundation. We describe three other representative cases below with details provided in Appendix C.1:

Model 2: costly state verification as in Bernanke et al. (1999) and Christiano et al. (2014), where leverage constraints arise from monitoring costs that lenders must incur to verify intermediary outcomes;

Model 3: costly leverage as in Uribe and Yue (2006); Cúrdia and Woodford (2016); Chi et al. (2025), where intermediaries face convex costs that increase with their leverage ratio;

Model 4: collateral constraints similar to Kiyotaki and Moore (1997); Bianchi and Mendoza (2018); Ottonello et al. (2022), where intermediation is limited by a fraction of collateral value backing the capital holdings.

The following lemma formalizes that the microfoundations in these models are special cases of the general formulation in Section 2.1:

Lemma 1 *Given $\{\omega_t, d_t\}$ implied by the intermediation frictions in model $j \in \{1, \ldots, 4\}$ and a net worth process given by (3), there exist $\Theta(\{r^K_{s+1}, r^B_{s+1}\}_{s\geq t})$ such that $\omega_t = \Theta_t n_t$ and*

$d_t = (\Theta_t - 1)n_t$.

Proof. See Appendix C.1. □

2.3 A Common Structure

We show that the models nested in Lemma 1 share a tractable structure for $\Theta(\cdot)$. Having this structure is essential: a fully general function $\Theta(\cdot)$ would not impose sufficient structure to yield sharp analytical results. By focusing on this class of models, we maintain substantial generality—nesting the most commonly used specifications in the literature—while achieving full characterization of the asset supply system.

Lemma 2 characterizes the structure of intermediation frictions for this class:

Lemma 2 *Intermediation frictions $\Theta(\{r^K_{s+1}, r^B_{s+1}\}_{s\geq t})$ implied by model $j \in \{1, \ldots, 4\}$ have the following structure around the steady state:*

$$\frac{\partial \Theta_t}{\partial r^K_{s+1}} = \gamma^{s-t}\,\bar{\Theta}_{r^K}, \quad \frac{\partial \Theta_t}{\partial r^B_{s+1}} = -\gamma^{s-t}\,\bar{\Theta}_{r^B}, \quad \forall s \geq t,$$

where $\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B}, \gamma \geq 0$ are semi-structural parameters given by steady-state variables and parameters in model j; the derivatives $\frac{\partial \Theta_t}{\partial r^K_{s+1}}$ and $\frac{\partial \Theta_t}{\partial r^B_{s+1}}$ are zeros $\forall s < t$.

Proof. See Appendix C.1. □

Parameters $\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B}, \gamma$ describe how the financial sector's intermediation capacity depends on expected returns. Parameter $\bar{\Theta}_{r^K}$ and $\bar{\Theta}_{r^B}$ are *leverage sensitivities* that govern how strongly Θ_t responds to expected returns next period, while γ is a *forward-looking component* that determines how this response decays for returns further in the future. Different microfoundations correspond to different $\Theta(\cdot)$ functions and determine how their derivatives depend on steady-state variables through the three parameters.

For the asset diversion (GKK) models, we can obtain $\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B}, \gamma$ by differentiating Equation (4) and evaluating at the steady state:[1]

$$\bar{\Theta}_{r^K} = \frac{\bar{\Theta}(\bar{\Theta}-1)}{1+r^K}, \quad \bar{\Theta}_{r^B} = \frac{\bar{\Theta}(\bar{\Theta}-1)}{1+r^B}, \quad \gamma = \frac{(1-f)(1+r^B+\left(r^K-r^B\right)\bar{\Theta})^2}{(1+r^K)(1+r^B)}. \tag{6}$$

Intuitively, Θ_t is more sensitive to returns when steady-state leverage $\bar{\Theta}$ is high: if an intermediary is highly leveraged, small return changes have large effects on profits, continuation

[1] The formula depends on the intermediary's discount factor, Λ_{t+1}. We use $1/(1+r^K_{t+1})$ as a baseline, and provide formulas for other alternatives in Appendix C.1.

value, and thus, through the incentive compatibility constraint, on Θ_t. If intermediaries expect to survive longer (lower f), they are more sensitive to returns far in the future, increasing γ.

The GKK framework features a flexible forward-looking component γ governed by the free parameter f, but imposes a tight restriction on the two parameters $\bar{\Theta}_{r^K}$ and $\bar{\Theta}_{r^B}$, linking them to steady-state returns and leverage. This contrasts with costly state verification and costly leverage models.

In costly state verification and costly leverage models, intermediation does not respond to returns beyond the next period ($\gamma = 0$). However, unlike the rigid $\bar{\Theta}_{r^K}$ and $\bar{\Theta}_{r^B}$ in the GKK framework, these models feature an extra degree of freedom to control the leverage sensitivities. These sensitivities are determined by monitoring costs and the distribution of idiosyncratic returns in costly state verification models, and by the curvature of the leverage cost function in costly leverage models.

Collateral constraint models also feature $\gamma = 0$, as collateral value changes are fully captured by changes in r^K_{t+1} and r^B_{t+1}. However, leverage sensitivities $\bar{\Theta}_{r^K}$ and $\bar{\Theta}_{r^B}$ are determined by steady-state leverage and returns, as in asset diversion models.

Beyond nesting existing models, the subclass of models parameterized by $\bar{\Theta}_{r^K}$, $\bar{\Theta}_{r^B}$, and γ also contains some useful generalizations of existing microfoundations. For example, a costly leverage model can be augmented with a dynamic structure similar to asset diversion models to deliver both flexible leverage sensitivities and a forward-looking component $\gamma > 0$. We discuss these generalizations in Appendix C.2.

2.4 Liquidity Supply Elasticities

With the structure in Lemma 2, we characterize the financial sector's asset supply system for this class of models.

We use a liquidity supply function, $\mathcal{D}_t(\{r^K_s, r^B_s\}_{s=0}^{\infty})$, to describe how much liquidity the intermediary supplies given returns. Up to first-order approximation, the liquidity supply function is characterized by its elasticities, which are jointly determined by intermediation frictions $\Theta(\cdot)$ and the net worth process $G(\cdot)$:

Proposition 1 *The cross-price semi-elasticities of liquidity supply around steady state are:*

$$\frac{\partial \mathcal{D}_t / \partial r^K_s}{\mathcal{D}_t} = \begin{cases} \gamma^{s-t-1} \bar{\Theta}_{r^K} \left(\frac{1}{\bar{\Theta}-1} + \gamma \Sigma(t) \right), & s > t, \\ \left(\frac{1}{\bar{n}} \bar{G}_{r^K} + \bar{\Theta}_{r^K} \Sigma(s) \right) \bar{G}_n^{t-s}, & s \leq t, \end{cases}$$

where $\Sigma(s) \coloneqq \frac{1}{\bar{n}}\bar{G}_\Theta \frac{1-\left(\gamma \bar{G}_n\right)^s}{1-\gamma \bar{G}_n}$, *and* $\bar{G}_\Theta, \bar{G}_{r^K}, \bar{G}_n, \bar{n}$ *are derivatives of* $G(\cdot)$ *and net worth at the steady state. The own-price semi-elasticities* $\frac{\partial \mathcal{D}_t/\partial r_s^B}{\mathcal{D}_t}$ *follow the same formula with* $\bar{\Theta}_{r^K}$ *and* $\bar{G}_{r^K}$ *replaced by* $-\bar{\Theta}_{r^B}$ *and* $\bar{G}_{r^B}$.

Proof. See Appendix A.1. □

Intermediation frictions determine how the financial sector's liquidity supply responds to changes in returns. For example, the cross-price elasticities $\frac{\partial \mathcal{D}_t/\partial r_s^K}{\mathcal{D}_t}$ are positive and increasing in $\bar{\Theta}_{r^K}$. High cross-price elasticities indicate that the financial sector is willing to provide more liquidity in response to an increase in r_s^K. In other words, only a small decrease in r_s^K is necessary for the financial sector to absorb a given amount of excess liquid asset supply. In this sense, capital and liquid assets are more substitutable with high cross-price elasticities.

The formula in Proposition 1 describes how liquidity supply at time t responds to changes in returns at time s. For $s > t$, an increase in r_s^K directly increases liquidity supply by relaxing intermediation friction Θ_t with sensitivity $\bar{\Theta}_{r^K}$. Moreover, it relaxes frictions in all periods before t with decay rate γ, increasing liquidity supply through the accumulation of net worth. Function $\Sigma(t)$ summarizes this cumulative effect. For $s \leq t$, an increase in past return r_s^K has no direct effect on Θ_t but affects liquidity supply through net worth propagation: net worth at time s increases directly by $\bar{G}_{r^K}$ plus accumulation from all prior periods given by $\Sigma(s)$, both propagating to period t at rate $\bar{G}_n$.

In the next section, we show that these liquidity supply elasticities are sufficient statistics for how the financial sector affects aggregate outcomes. Our characterization provides a structure for these infinite-dimensional elasticities with several advantages. First, while the elasticities are infinite-dimensional objects that cannot be estimated nonparametrically, we show that for the class of commonly used models with distinct microfoundations, they are governed by only a few parameters. Second, these parameters describe structural relationships between empirically observable objects such as leverage and returns. Therefore, we can obtain an empirical summary of the underlying frictions by estimating these parameters with appropriate identification strategies. Third, the elasticities governed by these parameters are policy-invariant to the extent that policies affect the financial sector through changes in returns. As a result, we can study a broad set of macroeconomic policies without taking a stance on the exact microfoundation of the underlying frictions.[2]

[2] These advantages do not rely on our sequence space approach. Researchers using state-space methods can use the following two equations (up to first order) for the financial block: $\hat{\Theta}_t = \bar{\Theta}_{r^K}\hat{r}_{t+1}^K - \bar{\Theta}_{r^B}\hat{r}_{t+1}^B + \gamma\hat{\Theta}_{t+1}$ and $\hat{n}_t = \bar{G}_\Theta\hat{\Theta}_{t-1} + \bar{G}_{r^K}\hat{r}_t^K + \bar{G}_{r^B}\hat{r}_t^B + \bar{G}_n\hat{n}_{t-1}$, where hats denote deviations from steady state.

3 Financial Intermediation in General Equilibrium

We now embed financial intermediation into a general equilibrium framework. We study perfect foresight equilibria in which there are no aggregate shocks. Equivalently, all results hold as first-order certainty equivalence approximations in the presence of aggregate shocks. We reformulate the economy as an intertemporal supply and demand system to trace how the financial sector affects the propagation of policies and shocks through changes in asset prices and quantities.

3.1 Households

Households are indexed by $i \in [0,1]$ and have time-separable and potentially type-dependent preferences. They derive utility from final goods consumption $c_{i,t}$ and disutility from labor $h_{i,t}$. Each household holds illiquid and liquid assets $a_{i,t}$ and $b_{i,t}$, which pay real returns r_t^A and r_t^B, respectively.[3] We allow liquid holdings to enter the utility, while illiquid holdings incur a cost in budget constraints. These features capture notions such as liquidity services or portfolio adjustment costs, which shape households' asset demand and affect the transmission of policy.

Each household solves:

$$\max_{c_{i,t},a_{i,t},b_{i,t}} \mathbb{E} \sum_{t=0}^{\infty} \beta_i^t \left[u_i\left(c_{i,t}, b_{i,t}\right) - \nu_i(h_{i,t}) \right],$$

subject to budget constraints

$$a_{i,t} + b_{i,t} + c_{i,t} + \Phi(a_{i,t}, (1+r_t^A)a_{i,t-1}) = (1+r_t^A)a_{i,t-1} + (1+r_t^B)b_{i,t-1} + y_{i,t} - \mathcal{T}_t(y_{i,t}),$$

borrowing constraints $a_{i,t} \geq \underline{a}$, $b_{i,t} \geq \underline{b}$. Real labor income $y_{i,t} = z_{i,t}\frac{W_t}{P_t}h_{i,t}$ depends on idiosyncratic earnings shocks $z_{i,t}$, nominal wage per efficiency unit of labor W_t, and the price level P_t. Households form expectations over idiosyncratic shocks $z_{i,t}$. Labor $h_{i,t}$ is taken as exogenous by each household and determined by monopolistically competitive labor unions described below. Income tax is given by a tax function $\mathcal{T}_t(y_{i,t})$.

This formulation of the household sector encompasses standard representative agent models ($z_{i,t} \equiv 1$ with no preference heterogeneity), assets-in-utility models, spender-saver type two-agent models, and Bewley-Huggett-Aiyagari-Imrohoroglu type heterogeneous agent models.

[3] In Appendix G.3, we show how we can modify the model to allow for long-term nominal liquid assets.

3.2 Production

Goods production: A representative firm produces final goods y_t with capital k_{t-1} and differentiated types of labor, $h_{\ell,t}$, supplied by unions indexed by $\ell \in [0,1]$:

$$y_t = k_{t-1}^{\alpha} h_t^{1-\alpha}, \quad h_t = \left(\int h_{\ell,t}^{\frac{\varepsilon_W - 1}{\varepsilon_W}} d\ell \right)^{\frac{\varepsilon_W}{\varepsilon_W - 1}},$$

where $\varepsilon_W > 1$ is the elasticity of substitution between labor types. Given nominal wages $\{W_{\ell,t}\}$ and capital rental rate R_t, the firm chooses capital and labor to maximize profit:

$$\max_{k_{t-1}, \{h_{\ell,t}\}} P_t y_t - R_t k_{t-1} - \int W_{\ell,t} h_{\ell,t} d\ell.$$

Labor supply: We introduce nominal rigidity through monopolistically competitive unions.[4] Each union supplies $h_{\ell,t} = \int z_{i,t} h_{i,\ell,t} di$, following an allocation rule $h_{i,\ell,t} = l(z_{i,t}) h_{\ell,t}$ to combine labor from households, with $\int z_{i,t} l(z_{i,t}) di = 1$. Unions set nominal wage growth $\pi_{W,\ell,t} := \frac{W_{\ell,t}}{W_{\ell,t-1}} - 1$ to maximize utilitarian welfare subject to adjustment costs:

$$\sum_{t=0}^{\infty} \int \beta_i^t \left[u_i \left(c_{i,t}, a_{i,t}, b_{i,t} \right) - \nu_i(h_{i,t}) - \frac{\kappa_W}{2} \pi_{W,\ell,t}^2 d\ell \right] di,$$

where $h_{i,t} = \int h_{i,\ell,t} d\ell$. Wage adjustment cost is borne as disutility by unions and does not affect the resource constraint; $\kappa_W > 0$ controls the degree of nominal wage rigidity. Because unions are symmetric, each household's nominal wages sum to $z_{i,t} W_t h_{i,t}$, where W_t is the ideal wage index.

Capital: The aggregate capital stock has the following law of motion:

$$k_t = \left(1 - \delta + \Gamma\left(\iota_t\right)\right) k_{t-1},$$

where $\iota_t := \frac{x_t}{k_{t-1}}$ denotes the investment rate, x_t is the level of investment, δ is the depreciation rate, and $\Gamma(\cdot)$ captures capital adjustment cost. Let q_t denote the price of capital. Holding capital from periods t to $t+1$ earns a return:

$$r_{t+1}^K = \max_{\iota_{t+1}} \frac{R_{t+1}/P_{t+1} + q_{t+1}\left(1 + \Gamma\left(\iota_{t+1}\right) - \delta\right) - \iota_{t+1}}{q_t} - 1, \tag{7}$$

where the numerator captures rental income, capital gains, and the value created by transforming existing capital into new capital through investment.

Mutual fund:

[4] In Appendix G.1, we provide results for an alternative formulation with price rigidity.

A passive mutual fund supplies illiquid assets in the economy. The mutual fund is composed of the net worth of the financial sector and capital of value $\omega_{M,t}$. Its total value is given by:

$$a_t = \omega_{M,t} + n_t.$$

Given the composition of the fund, returns on illiquid assets are given by the value-weighted average of returns on capital r_t^K and returns on net worth r_t^N:

$$r_{t+1}^A = \frac{1}{a_t}(r_{t+1}^K \omega_{M,t} + r_{t+1}^N n_t). \tag{8}$$

3.3 Government

The government sets a sequence of government purchases g_t, government debt b_t^G, liquid rate target r_t^B, total tax revenue T_t, and illiquid assets holdings a_t^G. The government debt is real and liquid. Monetary policy adjusts the nominal interest rate to achieve the target real liquid rate for all $t > 0$, with r_0^B predetermined. The government collects tax revenue through the tax system $\mathcal{T}_t(y_{i,t}) = y_{i,t} - (1-\tau_t)y_{i,t}^{1-\lambda}$, where λ governs the progressivity. Given $\{y_{i,t}\}$, tax rate τ_t is set to satisfy $T_t = \int \mathcal{T}_t(y_{i,t})di$.

The government faces budget constraints:

$$b_t^G - (1 + r_t^B)b_{t-1}^G = a_t^G - (1 + r_t^A)a_{t-1}^G + g_t - T_t. \tag{9}$$

3.4 Definition of Equilibrium

An equilibrium consists of policy $\{g_t, b_t^G, r_t^B, T_t, a_t^G\}$, prices $\{P_t, R_t, W_{\ell,t}, q_t, r_t^A, r_t^K\}$, and allocations $\{y_t, k_t, h_t, x_t, c_{i,t}, h_{i,\ell,t}, a_{i,t}, b_{i,t}, n_t, d_t, \omega_t, \omega_{M,t}, a_t\}$ such that: (1) households maximize utility; (2) firms maximize profits and investment maximizes capital returns; (3) unions maximize welfare with adjustment costs; (4) intermediaries' holdings and issuance of assets is governed by intermediation frictions and the net worth process; (5) illiquid returns r^A are consistent with mutual fund composition; (6) the government budget constraint holds, and (7) markets clear:

$$\int (c_{i,t} + \Phi_{i,t})\, di + x_t + g_t = y_t, \ \int b_{i,t} di = d_t + b_t^G, \ \int a_{i,t} di + a_t^G = a_t, \ \omega_t + \omega_{M,t} = q_t k_t.$$

These conditions clear the goods market (aggregate demand for goods equals output), the liquid asset market (household holdings equal supply from intermediaries and the government), the illiquid asset market (households' and the government's holdings equal the supply of illiquid assets), and the capital market clearing (intermediaries and the mutual fund's holdings equal the value of aggregate capital stock). Figure 1 in Appendix B.1 summarizes balance sheets in the economy and Appendix B.2 lists all model equations.

4 Aggregate Responses to Policies

We now use the financial sector's asset supply system to study how it interacts with the rest of the economy to determine aggregate responses to policies.

4.1 A Demand-and-Supply Representation

We recast the economy as a demand-and-supply system of goods and assets. Given government policies and key aggregate variables, we solve the optimization problem for each agent to obtain their aggregate behavior along the transition path.

Our result in Section 2.4 shows how the *financial block* of the economy implies a liquidity supply function, $\mathcal{D}_t$, that summarizes how financial intermediaries respond to aggregate conditions through $\{r_s^K, r_s^B\}$. The same logic applies to the *household block* of the model: Given a sequence of output, taxes, returns, and the initial asset position, we can solve the households' consumption-saving problem to obtain a consumption function, $\mathcal{C}_t$, and liquidity demand function, $\mathcal{B}_t$.[5] Similarly, we obtain an investment function, $\mathcal{X}_t$, from the *production block*. Lemma 3 represents the equilibrium of the model as that of a demand-and-supply system.

Lemma 3 *There exist functions $\mathcal{C}_t, \mathcal{B}_t$, and $\mathcal{X}_t$, such that, given policies $\left\{g_s, T_s, r_s^B, b_s^G\right\}_{s=0}^{\infty}$, the equilibrium output and returns on capital $\left\{y_s, r_s^K\right\}_{s=0}^{\infty}$ solve:*

$$\mathcal{C}_t(\{y_s, r_s^A; r_s^B, T_s\}_{s=0}^{\infty}) + \mathcal{X}_t(\{y_s, r_s^K\}_{s=0}^{\infty}) + g_t = y_t,$$

$$\mathcal{B}_t(\{y_s, r_s^A; r_s^B, T_s\}_{s=0}^{\infty}) = \mathcal{D}_t(\{r_s^K, r_s^B\}_{s=0}^{\infty}) + b_t^G,$$

where

$$r_t^A = \mathcal{R}_t^A\left(\{r_s^K; r_s^B, y_s\}_{s=0}^{\infty}; \mathcal{D}_{t-1}(\{r_s^K; r_s^B\}_{s=0}^{\infty})\right).$$

Function $\mathcal{R}_t^A$ corresponds to the accounting identity in Equation (8), and government asset holdings $\left\{a_t^G\right\}$ satisfy the government budget constraint in Equation (9). Moreover, functions $\mathcal{C}_t, \mathcal{B}_t$, and $\mathcal{X}_t$ do not depend on Θ and G.

[5] We define the consumption function, $\mathcal{C}_t$, to include both final goods consumed by the households, $c_{i,t}$, and the portfolio adjustment cost, $\Phi_t(\cdot)$.

Proof. See Appendix A.2. □

Compared to the market-clearing conditions in Section 3.4, the two main equations in Lemma 3 correspond to the goods market and the liquid asset market clearing conditions. We impose capital market clearing to derive the function $\mathcal{R}_t^A$ and drop the illiquid asset market clearing condition, as it is redundant by Walras' law. Given government policies, an equilibrium is described by sequences $\{y_t, r_t^K\}_{t=0}^{\infty}$ such that (1) demand for final goods equals output produced, and (2) liquidity demand equals liquidity supply.

The financial block enters the system *only* through the liquidity supply function $\mathcal{D}_t$. Up to first-order approximation, the liquidity supply elasticities characterized in Proposition 1 are sufficient statistics that summarize how financial frictions affect aggregate outcomes. The exact microfoundations behind the frictions do not matter as long as they generate the same liquidity supply. Since $\mathcal{D}_t$ can be described independently of the household and production sectors, our framework is compatible with a large class of quantitative macroeconomic models, including representative agent frameworks and models with realistic household heterogeneity, such as TANK and HANK models.

The household sector is summarized by the consumption function $\mathcal{C}_t$ and the liquidity demand function $\mathcal{B}_t$. While $\mathcal{C}_t$ plays a key role in the goods market—sharing the same emphasis as Auclert et al. (2024), Auclert et al. (2021), and Wolf (2021)—our result also highlights households' liquidity demand $\mathcal{B}_t$ as an equally important feature. In Appendix E, we characterize several canonical household specifications nested in our framework to illustrate their distinct implications on liquidity demand, including limiting cases ranging from perfectly elastic to perfectly inelastic.

Through the asset market, households' liquidity demand interacts with the financial sector's liquidity supply to jointly determine how policies affect aggregate outcomes $\{y_t, r_t^K\}$. We now discuss the scope and limitations of this approach.

Applicability and Limitations

While our characterization in Section 2.4 focuses on asset supply implied by a particular class of models, the sufficient statistics result in Lemma 3 applies as long as the financial sector delivers a function $\mathcal{D}_t(\{r_s^K, r_s^B\}_{s=0}^{\infty})$, regardless of its specific form. Similarly, our baseline framework focuses on a single liquid asset but can be extended to accommodate multiple assets with heterogeneous characteristics—such as reserves and treasuries with varying maturities. See Chiang and Żoch (2025) for a multi-asset extension.

Our sufficient statistics approach faces more substantial limitations in two contexts. First, the elasticities we characterize are invariant to policies that take effect through changes in prices and quantities of assets, but they may not be invariant to policies such as macroprudential regulations that can alter the shape of the asset supply system. Such regulations directly affect intermediation frictions and change the financial sector's sensitivities to returns, thereby altering the asset supply system itself. Studying these regulations requires microfounded models, and our elasticities should serve as target moments for calibration rather than policy-invariant parameters.

Second, when financial intermediaries are not price-takers, we must incorporate properties of asset demand to describe the financial sector's liquidity supply. Asset demand is likely to depend on modeling choices of the household sector and additional variables such as household income or transfers. This complicates the sufficient statistics representation and undermines the portability of the financial block.

4.2 Aggregate Responses

We study first-order aggregate responses to government policies around the steady state. We focus on policies such that $\{dg_t, dT_t, dr_t^B, db_t^G, da_t^G\}_{t=0}^{\infty}$ converge to zero as $t \to \infty$ and characterize the equilibrium in which aggregate responses also converge to zero. We use bold notation for sequences: $\boldsymbol{y}$ represents $\{y_t\}_{t=0}^{\infty}$ with first-order deviation $d\boldsymbol{y}$; notation for $\boldsymbol{T}, \boldsymbol{b}^G$, and $\boldsymbol{g}$ follows similarly. For returns, $\boldsymbol{r}^K$ represents $\{r_{t+1}^K\}_{t=0}^{\infty}$ with deviation $d\boldsymbol{r}^K$; notation for $\boldsymbol{r}^B$ is analogous. The return sequences start from period 1 because r_0^B is predetermined and r_0^K can be expressed as a function of output and expected returns; $r_0^K(\boldsymbol{y}, \boldsymbol{r}^K)$, as defined in Appendix A.3.

We characterize the equilibrium in two steps. First, we solve the liquid asset market clearing condition for returns on capital $d\boldsymbol{r}^K$, taking policies and output $d\boldsymbol{y}$ as given. Second, we substitute this solution into the goods market clearing condition to solve for equilibrium output responses $d\boldsymbol{y}$.

Excess Liquidity and Asset Markets Responses

To study asset market responses, we define *excess liquidity* as the difference between liquidity supply and demand:

$$\mathcal{E}_t(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{T}, \boldsymbol{b}^G) := \mathcal{D}_t(r_0^K(\boldsymbol{y}, \boldsymbol{r}^K), \boldsymbol{r}^K, \boldsymbol{r}^B) + b_t^G - \mathcal{B}_t(\boldsymbol{y}, \boldsymbol{r}^A(\boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{y}), \boldsymbol{r}^B, \boldsymbol{T}),$$

where $\boldsymbol{r}^A(\boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{y})$ denotes the vector form of functions $\mathcal{R}_t^A(\cdot)$ from Equation (8). We use

$\boldsymbol{\epsilon}_{r^K}$ to denote derivatives of $\mathcal{E}$ with respect to $\boldsymbol{r}^K$, where $\boldsymbol{\epsilon}_{r^K}(t,s)$ represents how excess liquidity in period t responds to r^K_{s+1}. Other derivatives follow the same convention.

Equilibrium in the liquid asset market requires zero excess liquidity. Proposition 2 characterizes how returns on capital adjust to clear the market in response to shifts in excess liquidity due to policies and output.

Proposition 2 *In equilibrium, returns on capital satisfy*

$$d\boldsymbol{r}^K = -\boldsymbol{\epsilon}_{r^K}^{-1}\underbrace{\left(d\boldsymbol{b}^G + \boldsymbol{\epsilon}_T d\boldsymbol{T} + \boldsymbol{\epsilon}_{r^B} d\boldsymbol{r}^B + \boldsymbol{\epsilon}_y d\boldsymbol{y}\right)}_{\textit{shifts in excess liquidity}}. \tag{10}$$

Moreover, if $\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B} \to \infty$ *with* $\bar{\Theta}_{r^B}/\bar{\Theta}_{r^K} \to \varsigma$, *then* $d\boldsymbol{r}^K \to \varsigma d\boldsymbol{r}^B$.

Proof. See Appendix A.4. □

To understand the result, consider a special case where households' liquidity demand $\mathcal{B}_t$ is perfectly inelastic with respect to returns. In this case, $\boldsymbol{\epsilon}_{r^K}$ is determined by the financial sector's cross-price elasticities of liquidity supply (Proposition 1). Suppose further that the financial sector features constant net worth $n_t = \bar{n}$ and no forward-looking component, $\gamma = 0$. Then Proposition 1 implies $\boldsymbol{\epsilon}_{r^K}^{-1} = (\bar{n}\bar{\Theta}_{r^K})^{-1}\mathbf{I}$, and Proposition 2 implies a positive shift in excess liquidity decreases dr^K_{t+1} proportionally to $\bar{\Theta}_{r^K}^{-1}$. For example, an increase in b^G_t raises excess liquidity. Since household liquidity demand is fixed, the financial sector must absorb this increase by holding more government debt and reducing its net liquidity supply. This requires lower expected returns dr^K_{t+1}, accompanied by higher capital prices. This mechanism is stronger when the leverage sensitivity $\bar{\Theta}_{r^K}$ is low (liquidity supply inelastic) and capital and liquid assets are less substitutable for the financial sector.

Two polar cases of the financial sector's liquidity supply provide important benchmarks. When liquidity supply is perfectly elastic ($\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B} \to \infty$), assets are perfect substitutes for the financial sector. The perfect link between asset markets allows the government to fully control returns on capital r^K_{t+1} through monetary policy r^B_{t+1}. Changes in government debt, taxes, and output have no effect on r^K_{t+1}, and households' liquidity demand $\mathcal{B}_t$ plays no role in equilibrium outcomes. Conversely, when liquidity supply is perfectly inelastic ($\bar{\Theta}_{r^K} = \bar{\Theta}_{r^B} = G = 0$), asset market responses are determined entirely by households' liquidity demand $\mathcal{B}_t$.

These limiting cases correspond to common assumptions in models studying fiscal and monetary policies. As we discuss in Appendix G.4, the perfectly elastic benchmark aligns with

Auclert et al. (2024), while the perfectly inelastic benchmark aligns with Kaplan et al. (2018). These assumptions lead to drastically different implications for policy.

The same logic applies to households' liquidity demand. If $\mathcal{B}_t$ is perfectly elastic with respect to returns, financial sector features have no effect on aggregate outcomes. If $\mathcal{B}_t$ is perfectly inelastic, asset market responses are determined entirely by the financial sector. These special cases demonstrate how asset market responses depend on the joint properties of households' liquidity demand and the financial sector's liquidity supply. However, as we show in Section 5, when households' consumption-saving behaviors are calibrated to match standard microdata moments, their liquidity demand is orders of magnitude less elastic than our measures of the financial sector's liquidity supply. Quantitatively, asset market responses are therefore mostly determined by the financial sector's cross-price elasticities derived in Proposition 1.

Aggregate Output Responses

We combine asset market responses with the goods market clearing condition to solve for output responses. To understand the goods market, we define *aggregate demand*, Ψ_t, as the sum of consumption, investment, and government spending:

$$\Psi_t(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{T}, \boldsymbol{g}) := \mathcal{C}_t(\boldsymbol{y}, r^A(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B), \boldsymbol{r}^B, \boldsymbol{T}) + \mathcal{X}_t(\boldsymbol{y}, \boldsymbol{r}^K) + g_t,$$

where $\boldsymbol{\Psi}_{r^K}$ is a matrix of derivatives and $\boldsymbol{\Psi}_{r^K}(t,s)$ represents how aggregate demand in period t responds to r^K_{s+1}. Other derivatives are defined similarly.

Equilibrium in the goods market requires aggregate output to equal aggregate demand. Totally differentiating the goods market clearing condition and substituting the expression for $d\boldsymbol{r}^K$ from Proposition 2 yields the following result:

Theorem 1 *Given $\{d\boldsymbol{g}, d\boldsymbol{T}, d\boldsymbol{r}^B, d\boldsymbol{b}^G\}$, the output response is*

$$d\boldsymbol{y} = \underbrace{\left(\mathbf{I} - \boldsymbol{\Psi}_y - \boldsymbol{\Omega}\boldsymbol{\epsilon}_y\right)^{-1}}_{\text{(3) modified Keynesian cross}} \; \Big(\underbrace{d\boldsymbol{g} + \boldsymbol{\Psi}_T d\boldsymbol{T} + \boldsymbol{\Psi}_{r^B} d\boldsymbol{r}^B}_{\text{(1) goods market channel}} + \underbrace{\boldsymbol{\Omega}\left(d\boldsymbol{b}^G + \boldsymbol{\epsilon}_T d\boldsymbol{T} + \boldsymbol{\epsilon}_{r^B} d\boldsymbol{r}^{\boldsymbol{B}}\right)}_{\text{(2) asset market channel}}\Big),$$

where $\boldsymbol{\Omega} := \boldsymbol{\Psi}_{r^K}(-\boldsymbol{\epsilon}_{r^K}^{-1})$.[6]

Proof. See Appendix A.5. □

Government policies affect output through three channels: (1) The *goods market channel*

[6]We use $(\mathbf{I} - \boldsymbol{\Psi}_y - \boldsymbol{\Omega}\boldsymbol{\epsilon}_y)^{-1}$ to denote a generalized inverse matrix $\mathcal{M}$ such that $\mathcal{M}(\mathbf{I} - \boldsymbol{\Psi}_y - \boldsymbol{\Omega}\boldsymbol{\epsilon}_y) = \mathbf{I}$. See Auclert et al. (2024) for details.

shows how government purchases, taxes, and the liquid rate directly affect aggregate demand, capturing standard Keynesian logic. (2) The *asset market channel* describes how government debt, taxes, and the liquid rate affect aggregate demand through asset markets. (3) A *modified Keynesian cross* captures the feedback between aggregate demand and income.

Transmission through asset markets depends on an *asset market propagation matrix* $\boldsymbol{\Omega}$, which represents how shifts in excess liquidity affect aggregate demand. It consists of two components. Matrix $-\boldsymbol{\epsilon}_{r^K}^{-1}$ describes how shifts in excess liquidity due to policies affect returns on capital $\boldsymbol{r}^K$, representing the asset market responses in Proposition 2. Matrix $\boldsymbol{\Psi}_{r^K}$ describes how changes in $\boldsymbol{r}^K$ affect aggregate demand. For example, a negative entry $\boldsymbol{\Psi}_{r^K}(t,s)$ with $s>t$ reflects how a decrease in expected returns r^K_{s+1} increases consumption and investment at time t through a higher current capital price.

Transmission through asset markets depends on the *asset market propagation matrix* $\boldsymbol{\Omega}$, which represents how shifts in excess liquidity affect aggregate demand. It consists of two components. Matrix $-\boldsymbol{\epsilon}_{r^K}^{-1}$ describes how shifts in excess liquidity due to policies affect returns on capital $\boldsymbol{r}^K$, representing the asset market responses in Proposition 2. For instance, when the government increases liquid asset supply, the price of capital rises relative to liquid assets and therefore has a lower expected return. However, this effect is dampened when the financial sector is more elastic (i.e., has a larger cross-price elasticity of excess liquidity supply). Matrix $\boldsymbol{\Psi}_{r^K}$ then describes how these changes in $\boldsymbol{r}^K$ affect aggregate demand. For example, a negative entry $\boldsymbol{\Psi}_{r^K}(t,s)$ with $s>t$ can reflect how aggregate demand (consumption and investment) increase with a lower expected return r^K_{s+1} through a higher capital price. Overall, when the financial sector features larger cross-price elasticities, the dampened impact on asset returns carries through to a smaller effect on aggregate demand.

The asset market propagation matrix $\boldsymbol{\Omega}$ is shaped by features of the financial sector: Intermediation frictions represented by $\bar{\Theta}_{r^K}$ determine the cross-price elasticities of the financial sector's liquidity supply $\mathcal{D}_t$. Liquidity supply affects the response of excess liquidity through $\boldsymbol{\epsilon}_{r^K}$ and ultimately shapes asset market propagation through $\boldsymbol{\Omega}$.

Besides governing the asset market channel, the asset market propagation matrix also modifies the Keynesian cross feedback. An increase in aggregate income not only affects aggregate demand directly but also shifts excess liquidity through $\boldsymbol{\epsilon}_y$. For example, when households receive higher income, they increase their liquidity demand, reducing excess liquidity and lowering aggregate demand through asset market responses. Therefore, the same force that amplifies demand through the asset market channel can dampen the Keynesian cross feedback.

5 Bringing the Model to the Data

We now bring the model to the data. We first estimate the parameters governing financial frictions and compute the financial sector's liquidity supply elasticities. We then calibrate the remaining model components, discuss the quantitative implications for household liquidity demand, and prepare for a quantitative assessment of how liquidity supply and demand shape aggregate responses to policies.

5.1 Empirical Summary of Financial Frictions

To provide an empirical summary of financial frictions, we estimate functions $\Theta(\cdot)$ and $G(\cdot)$ for the class of models characterized in Section 2 up to first-order approximation, using identified structural shocks.

We assume the following empirical counterparts for $\Theta(\cdot)$ and $G(\cdot)$:

$$d\Theta_t = \sum_{h=1}^{\infty} \tilde{\gamma}^{h-1} \left(\tilde{\Theta}_{r^K} \mathbb{E}_t[dr^K_{t+h}] - \tilde{\Theta}_{r^B} \mathbb{E}_t[dr^B_{t+h}] \right) + \upsilon^{\Theta}_t, \tag{11}$$

$$d \log n_t = \tilde{G}_{r^N} dr^N_t + \tilde{G}_n d \log n_{t-1} + \upsilon^G_t. \tag{12}$$

The residuals υ^{Θ}_t and υ^G_t represent either measurement errors or structural shocks to the level of intermediation or net worth. We restrict the functional form of the net worth process to reduce the number of parameters to be estimated.[7]

We describe how we measure each empirical component below and provide details in Appendix D.1.

Measurement

We focus on the U.S. banking sector. We obtain balance sheet positions from the Call Reports and the market value of equity from CRSP.

Net worth: We use the market value of equity for bank-holding companies to represent financial sector net worth. We aggregate the market value of equity across bank-holding companies at a monthly frequency.

[7] The mapping to derivatives of $G(\cdot)$ is: $\frac{1}{\bar{n}} G_{r^K} = \tilde{G}_{r^N} \bar{\Theta}$, $\frac{1}{\bar{n}} G_{r^B} = -\tilde{G}_{r^N} (\bar{\Theta} - 1)$, and $\frac{1}{\bar{n}} G_{\Theta} = \tilde{G}_{r^N} (r^K - r^B)$.

Leverage: We measure the net liquid asset supply (liquid liabilities minus liquid assets) of bank-holding companies, d_t, and calculate the effective leverage as

$$\text{effective leverage } (\Theta_t) \coloneqq 1 + \frac{\text{net supply of liquid assets } (d_t)}{\text{market value of bank equity } (n_t)}.$$

Liquid liabilities and assets on the banking sector balance sheet include deposits (checkable, time, and savings accounts), money market fund shares, and government liabilities such as cash, reserves, and Treasury debt.

Returns: We compute returns on net worth r_t^N as value-weighted averages of realized returns of bank-holding companies. We use the yield curves on U.S. Treasury bonds to construct liquid rates over different horizons, $\mathbb{E}_t[dr_{t+h}^B]$. For expected returns on capital, $\mathbb{E}_t[dr_{t+h}^K]$, we rely on the yield curves of high-quality corporate bonds (grade A and above). To better represent returns on the banking sector's asset holdings, we adjust the yield curve proportionally so that its fluctuations are similar to Moody's BAA bond yield index at the corresponding horizon.

We adjust net worth for inflation, and nominal yields are converted to real yields using inflation expectations data from the Cleveland Fed. We remove a quadratic trend from all variables (or their logs) and set the steady-state values of each variable to its long-run average:

Steady-state returns and leverage: Liquid rate r^B equals 0%, consistent with the average one-year real Treasury yield in our sample. Return on capital r^K is 3.5% per annum, corresponding to the average real yield on BAA bonds. The average effective leverage $\bar{\Theta}$ is 4.

Threats to Identification

One threat to identification is that the error terms v_t^Θ and v_t^G in Equations (11) and (12) may contain exogenous "leverage shocks" and "net worth shocks" that affect the banking sector's ability to sustain a certain level of leverage and the accumulation of net worth. For example, changes in leverage and net worth can represent changes in macroprudential policies or financial intermediaries' business strategies. Exogenous changes to the idiosyncratic risk profile in Bernanke et al. (1999) studied in Christiano et al. (2014) would also appear in the residual v_t^Θ. Because these shocks affect holdings of capital and liquid assets, expected returns adjust in general equilibrium and lead to omitted variable bias. Therefore, in the presence of these shocks, identification of the semi-structural parameters in Equations (11) and (12) requires instrumental variables.

Identification Strategy

Valid instruments I_t^{Θ} and I_t^{G} for Equations (11) and (12) must satisfy exclusion restrictions:

$$\mathbb{E}[\upsilon_t^{\Theta} \times I_t^{\Theta}] = 0, \quad \mathbb{E}[\upsilon_t^{G} \times I_t^{G}] = 0,$$

and relevance conditions: that they generate non-collinear variations in $\mathbb{E}_t[dr_{t+h}^K], \mathbb{E}_t[dr_{t+h}^B]$ and r_t^N, n_{t-1}, respectively.

We construct instruments from two shock proxies: (1) a proxy for monetary policy shocks from Bauer and Swanson (2023), and (2) a proxy for oil supply shocks from Baumeister and Hamilton (2019).

For the net worth equation (12), we use contemporaneous values and three months of lags of the monetary policy and oil supply shock proxies as instruments I_t^G for r_t^N and n_{t-1}. These instruments are valid under the assumption that changes in equity and dividend issuance (υ_t^G) do not occur within the narrow windows around FOMC announcements or coincide with oil supply changes.

For the leverage equation (11), the identifying assumption is that changes in leverage due to macroprudential policy, banks' idiosyncratic risk profile (Christiano et al., 2014), or their business strategy (υ_t^{Θ}) do not occur during the short windows around FOMC announcements and oil supply changes. To better use the shock proxies to generate variation in expected returns across different horizons, we employ an SVAR model to derive return variations $\mathbb{E}_t[\check{dr}_{t+h}^K]$ and $\mathbb{E}_t[\check{dr}_{t+h}^B]$: We identify structural shocks by assuming they are the only shocks affecting the proxies contemporaneously, extract return variations attributable to these shocks, and then use these variations corresponding to 1, 5, 10, and 30-year horizons as our instrument I_t^{Θ}. The structural shocks account for 10-20% of total variation in expected returns, confirming instrument relevance (see Appendix D.1).

5.2 Estimation Results

We estimate parameters $\tilde{\Theta}_{r^K}, \tilde{\Theta}_{r^B}, \tilde{\gamma}, \tilde{G}_{r^N}$, and $\tilde{G}_n$ with moment conditions of the form:

$$\mathbb{E}[\upsilon_t^{\Theta} \times I_t^{\Theta}] = 0, \quad \mathbb{E}[\upsilon_t^{G} \times I_t^{G}] = 0.$$

For both estimation equations, we consider two alternative specifications. First, as a baseline case, we assume that $\upsilon_t^{\Theta}, \upsilon_t^{G}$ consist purely of measurement errors, and use $I_t^{\Theta} \in \left\{\mathbb{E}_t[dr_{t+h}^K],\ \mathbb{E}_t[dr_{t+h}^B]\right\}$ and $I_t^G \in \left\{r_t^N, n_{t-1}\right\}$ for the estimation. We refer to these specifications as the ordinary least squares (OLS) benchmarks.[8] Second, to address the identification threats that υ_t^Θ and υ_t^G are structural shocks that affect equilibrium outcomes, we use the instruments discussed in Section 5.1. The estimation results are reported in Table 1, and details are provided in Appendix D.1.

Table 1: Estimates of $\tilde{\Theta}_{r^K}, \tilde{\Theta}_{r^B}, \tilde{\gamma}, \tilde{G}_{r^N}$, and $\tilde{G}_n$

	(1) OLS	(2) IV
size of cross-price, $\tilde{\Theta}_{r^K}$	28.91	24.26
	(15.68)	(13.64)
size of own-price, $\tilde{\Theta}_{r^B}$	25.10	25.31
	(19.68)	(19.03)
forward-looking, $\tilde{\gamma}$	0.98	0.98
	(0.01)	(0.01)
response to realized return, $\tilde{G}_{r^N}$	0.478	0.736
	(0.087)	(0.270)
net worth persistence, $\tilde{G}_n$	0.969	0.970
	(0.011)	(0.039)

Note: Estimation of $\tilde{\Theta}_{r^K}, \tilde{\Theta}_{r^B}, \tilde{\gamma}$ uses iterative GMM for optimal weighting matrix; standard errors use heteroskedastic and autocorrelation consistent estimators. Estimation $\tilde{G}_{r^N}, \tilde{G}_n$ uses OLS and TSLS; standard errors use Newey-West estimators. Sample period: January 1999 to December 2019, monthly observation.

The first three rows of the table show the parameters for $\Theta(\cdot)$. Estimates of $\tilde{\Theta}_{r^K}$ and $\tilde{\Theta}_{r^B}$ imply the effective leverage of the banking sector increases by around 25 percentage points in response to one percentage point increase in the spread between the two returns. The forward-looking component $\tilde{\gamma}$ is 0.98, which implies a "half-life" of around two and a half years: the response to a spread increase two and a half years ahead is half as strong as the response to a change in the current period. The total response of banks' effective leverage is a discounted sum of responses to all future spreads. The last two rows of the table show the parameters for $G(\cdot)$. Parameter $\tilde{G}_{r^N}$ is around 0.7, which implies that when returns on net worth are 1% above the steady state, net worth next period will be 0.7% above its long-run trend. The persistence of net worth gives a half-life of around two years.

In Appendix D.1, we generalize our empirical specification to measure the extent to which our estimates vary with the aggregate state of the economy. This provides us with useful information to gauge the situation under which our result is useful. We do not find statistically significant results in support of state dependency, as the standard errors for the state

[8]Strictly speaking, the estimation for $\Theta(\cdot)$ is non-linear and not an OLS regression.

dependency parameter are large.

Finally, while we restrict the functional form for $G(\cdot)$ in our estimation, in Appendix F.3, we allow for a fully flexible G function and demonstrate that our policy conclusion is insensitive to a wide range of values for $\bar{G}_\Theta$, $\bar{G}_{r^K}$, $\bar{G}_{r^B}$, and $\bar{G}_n$.

5.3 Implied Liquidity Supply Elasticities

We convert the empirical estimates into model counterparts in Table 2:

Table 2: Implied Model Parameters

Parameter	$\bar{G}_n$	$\bar{G}_{r^K}$	$\bar{G}_{r^B}$	$\bar{G}_\Theta$	γ	$\bar{\Theta}_{r^K}$	$\bar{\Theta}_{r^B}$
Value	0.912	1.49	−1.11	0.003	0.95	24.26	−25.31

Note: The conversion maps the empirical net worth process to the general function $G(\cdot)$ and adjusts for the frequency from monthly to quarterly to be consistent with the rest of the model.

Figure 1 shows the semi-elasticities of liquidity supply based on these estimates and the formula in Proposition 1. Each line represents how liquidity supply responds to an increase in returns in a different period s. The cross-price elasticities show that liquidity supply increases before period s, reflecting the forward-looking component. After period s, liquidity supply drops sharply but remains elevated due to propagation through net worth. The size of the initial response $\frac{\partial \mathcal{D}_0/\partial r_1^K}{\mathcal{D}_0}$ implies an 8.1% increase in liquidity supply in response to a one percentage point increase in r_1^K. The own-price elasticities have a similar pattern with the opposite sign.

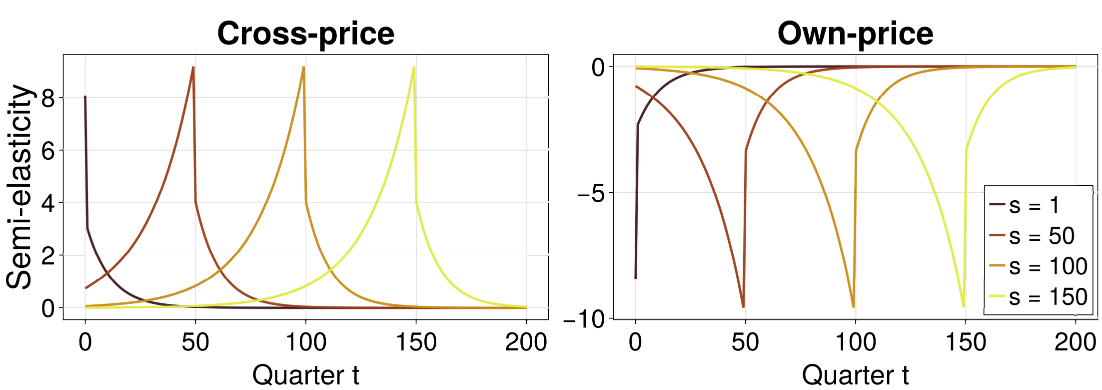

Figure 1: Semi-elasticities of liquidity supply. Left panel: cross-price semi-elasticities with respect to r_s^K. Right panel: own-price semi-elasticities with respect to r_s^B. Each line corresponds to a different period s and shows the semi-elasticity of liquidity supply in quarter t.

Comparison: Common Assumptions in Macro Models

To put our measures of the liquidity supply elasticities into context, we compare them to three common assumptions in workhorse macroeconomic models that feature, respectively,

1. a perfectly inelastic liquidity supply,
2. a perfectly elastic liquidity supply, and
3. liquidity supply implied by the asset diversion frictions as in Gertler-Karadi-Kiyotaki.

The first two benchmarks correspond respectively to the assumptions in Kaplan et al. (2018) and Auclert et al. (2024), as we discussed in Section 4.2. These polar cases are useful benchmarks because they are currently some of the most popular models for quantitative analysis of monetary and fiscal policy.

The third benchmark (GKK and its variants) constitutes the majority of macroeconomic models with financial frictions. As discussed in Section 2.2, GKK imposes a tight restriction on the liquidity supply elasticities, linking them to the intermediary's steady-state returns and leverage (Equation (6)).

Given the steady state values reported in Section 5.1, the implied values of $\bar{\Theta}_{r^K}$ and $\bar{\Theta}_{r^B}$ are, respectively, 11.9 and 12.0. We use the functional form for $G(\cdot)$ implied by the GKK model (Equation (5) in Section 2.2), with the exit rate f calibrated to imply the same value of the forward-looking component γ as our empirical estimate of 0.95.[9] The semi-elasticities implied by these parameters are half of those in our baseline, for example, $\frac{\partial \mathcal{D}_0 / \partial r_1^K}{\mathcal{D}_0}$ equals 3.97. As we show in Section 6, the differences between our empirical measures and these benchmarks are important as they lead to diverging policy conclusions.

Discussion

While our empirical strategy provides discipline for measuring liquidity supply elasticities, several assumptions and potential limitations merit discussion for interpreting our results and guiding future empirical work.

Asset aggregation: Our measure of liquid assets aggregates various types of securities (deposits, money market shares, government liabilities) that may differ in their liquidity properties. The composition of these assets evolves over time, and therefore, the liquidity properties of our liquid asset aggregates may also change.

[9] The calibrated value of f is 0.0955.

Intermediary Heterogeneity: We focus on the banking sector as representative of financial intermediation, but there are other important players in asset markets, including shadow banks, money market funds, and other non-bank financial intermediaries. Their liquidity supply elasticities may differ from those of traditional banks.

Structural shocks and specification: Our instruments address the endogeneity problem originating from unobserved structural shocks in Equations (11) and (12). However, this assumes the functional forms are not misspecified. For example, if the true data-generating process includes other macroeconomic variables that directly enter these functions, our estimates could be biased, as past monetary policy shocks or oil shocks may affect these endogenous variables.

5.4 Production, Government, and Households

Production: The elasticity of output with respect to capital α is set to 0.35. Depreciation rate δ is 5.83% yearly. Capital production function is $\Gamma(\iota_t) = \bar{\iota}_1 \iota_t^{1-\kappa_I} + \bar{\iota}_2$, where $\bar{\iota}_1, \bar{\iota}_2$ generate steady-state investment-to-capital ratio equal to δ and the price of capital equal to 1. We set $\kappa_I = 0.5$ so that the elasticity of investment to capital price is 2. Unions allocate labor uniformly among households: $l(z_{i,t}) \equiv 1/\int z_{i,t} di$. Since monetary policy targets real liquid rates, the slope of the wage Phillips curve does not matter for output responses. Therefore, the exact values of the elasticity of substitution between labor varieties, ε_W, and nominal wage rigidity, κ_W, are inconsequential.

Government: We set steady-state net tax revenue to 15% of output and the tax system's progressivity parameter, λ, to 0.18. Net liquid assets supplied by the government (and held by the private sector) is 21% of steady-state output, consistent with liquid asset positions in the data, as shown in Appendix D. We assume the government holds no illiquid assets in the steady state. The level of government purchases implied by the government budget constraint is 15% of output.

Households

Preferences: There are two types of households, indexed by s with population share μ and $1-\mu$. Period utility functions have the following form:

$$u_s(c) - \nu_s(h) = \frac{c^{1-\sigma_s} - 1}{1-\sigma_s} - \varsigma \frac{h^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}, \quad \sigma_s \geq 0,\ \varphi \geq 0.$$

We introduce heterogeneity in the discount rate β_s and the intertemporal elasticity of substitution $1/\sigma_s$ so that we can match household balance sheet positions, returns, and consumption patterns. Following a simplified version of Aguiar et al. (2024), we set $1/\sigma_s$ to $\{2, 1/2\}$ for $s \in \{1, 2\}$, and calibrate internally the share of agents with high intertemporal elasticity of substitution μ and the discount rates β_s of both types. The Frisch labor supply elasticity, φ, is set to 1. Parameter ς is set so that steady-state average hours worked equal one-third.

Income process: We use a discrete-time version of the income process described in Kaplan et al. (2018), which targets eight moments of the male-earnings distribution from Guvenen et al. (2021). Income process is the same for both household types.

Assets: Adjustment cost of illiquid assets is similar to Auclert et al. (2021):

$$\Phi_t(a_{i,t}, a_{i,t-1}, r_t^A) = \frac{\chi_1}{\chi_2} \left| \frac{a_{i,t} - (1 + r_t^A) a_{i,t-1}}{a_{i,t-1} + \chi_0} \right|^{\chi_2} [a_{i,t-1} + \chi_0] \, .$$

We set χ_0 to 0.1 and calibrate parameters χ_1 and χ_2 internally to determine and slope and curvature of the adjustment cost function. We assume asset positions cannot be negative: $\underline{a} = \underline{b} = 0$.

With the five parameters $\beta_1, \beta_2, \mu, \chi_1, \chi_2$, we target five moments: the steady-state ratios of liquid and illiquid assets to GDP, the shares of wealthy (WHtM) and poor hand-to-mouth (PHtM) agents (25% and 15%), and the first quarter marginal propensity to consume out of \$500 transfer (MPC) (20%). Table 3 shows that the model replicates target moments and reports calibrated parameter values.

Table 3: Households Calibration

Target Moments	Model	Data	Parameter	Value
Net Liquid Holdings to GDP	0.61	0.53	β_1	0.983
Net Illiquid Holdings to GDP	3.37	3.44	β_2	0.943
Poor Hand-to-Mouth	15%	9 - 17%	μ	0.172
Wealthy Hand-to-Mouth	24%	12 - 33%	χ_1	23.34
First quarter MPC	20%	15 - 25%	χ_2	2.0154

Data Source: See Appendix D.2 for liquid assets and illiquid assets positions; shares of HtM households: Table 3 in Kaplan et al. (2014); MPC: Kaplan and Violante (2022).

Our household sector features a canonical two-asset heterogeneous agent model calibrated to match standard targets. Households' consumption responses to an increase in disposable income (MPC) are large, as commonly emphasized in the literature.[10] However, the same

[10]While we target a quarterly MPC of 20% for the first quarter, consumption in subsequent quarters declines

portfolio adjustment frictions that generate illiquidity and large consumption responses for WHtM households also imply that these households face difficulties in adjusting asset positions in response to returns. As a result, with the portfolio adjustment cost calibrated to match empirically plausible MPC, households' consumption responses also inform us about their liquidity demand with respect to returns.

Implied Liquidity Demand

Figure 2 compares household liquidity demand from our calibration to our estimates of the financial sector's liquidity supply, along with the excess liquidity supply. Each line represents responses to an increase in r_s^K, taking into account its effect on the sequence of illiquid returns $\{r_s^A\}$.

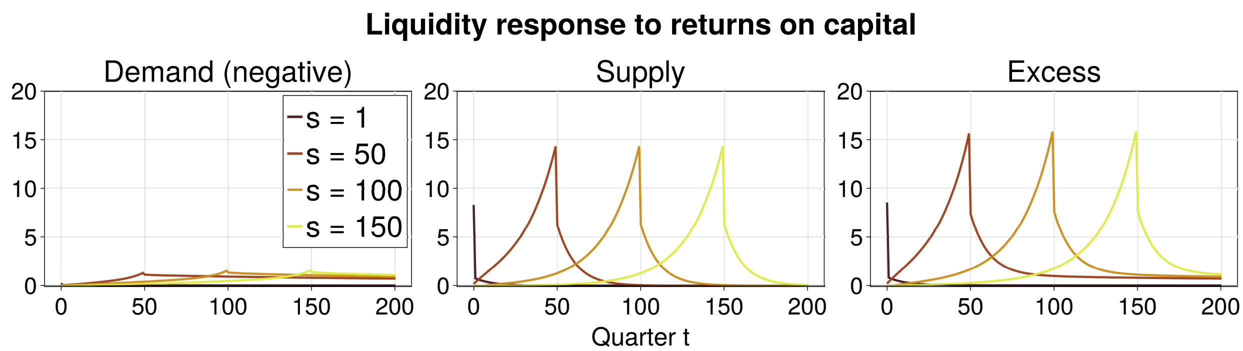

Figure 2: Entries of $-\tilde{\mathbf{B}}_{r^K}, \mathbf{D}_{r^K}$, and $\boldsymbol{\epsilon}_{r^K}$ matrices (see Appendix B.3 for their definitions). Left panel: liquidity demand ($-\tilde{\mathbf{B}}_{r^K}$). Center panel: liquidity supply ($\mathbf{D}_{r^K}$). Right panel: excess liquidity ($\boldsymbol{\epsilon}_{r^K}$). Each line corresponds to a different period s and shows responses in quarter t with respect to r_s^K.

Liquidity demand responses are an order of magnitude smaller than liquidity supply responses. This implication is consistent with existing empirical evidence, such as Gabaix et al. (2024), which suggests that insensitivity and inertia in asset allocations are prominent features of the household sector, including even ultra-rich households. As discussed in Section 4.2, the contrast between the elasticities of liquidity demand and supply has an important implication: When liquidity demand is inelastic with respect to returns, asset market responses are mostly determined by the financial sector through its cross-price elasticities of liquidity supply. Since workhorse macro models feature a wide range of assumptions on these elasticities, they generate substantially different conclusions for a variety of policy questions.

quickly, resulting in annual MPCs of around 30% for the first year and 10% for the second year, lower than the 45% and 14% reported in Auclert et al. (2024).

6 Policy Implications

Our sufficient statistics impose empirical discipline on assumptions about the financial sector and have crucial policy implications. We use two policy questions to demonstrate its importance. The first question concerns the government spending multipliers: How do changes in government spending affect aggregate output? The second question is central to the Wall Street vs. Main Street debate: Can asset purchases stimulate aggregate output more effectively than tax cuts?

Through our sufficient statistics, we systematically compare policy implications of common assumptions about the financial sector, as discussed in Section 5.3, and contrast them with our empirical measures. We calculate aggregate responses to policies under these assumptions, keeping all else equal in our calibrated model.

6.1 Government Spending Multiplier

We show that the size of the government spending multiplier depends crucially on the implicit assumptions about the financial sector. Consider the government implementing a policy path $\{d\hat{b}_t^G, d\hat{g}_t, d\hat{r}_t^B, d\hat{a}_t^G, d\hat{T}_t\}$ with

$$d\hat{g}_t = \eta^t s_0, \quad d\hat{b}_t^G = \rho_{b^G}(d\hat{b}_{t-1}^G + d\hat{g}_t), \quad d\hat{r}_t^B = d\hat{a}_t^G = 0,$$

and tax revenue $d\hat{T}_t$ is set to satisfy the budget constraint. Parameter s_0 controls the size of spending, decaying at rate η; the extent of deficit financing is governed by ρ_{b^G}. We set $\eta = 0.5$ so that most spending is completed in one year; size s_0 is such that debt peaks at 1% of annual GDP.[11] We use $\rho_{b^G} = 0.95$ as a baseline, which implies government debt peaks in one year and is reduced to half in five years. As we discuss below, the main message remains the same for different levels of deficit financing. Figure 3 shows the resulting policy paths.

[11]For ease of comparison, we scale all policy experiments in this section so that they represent the same budgetary commitment: the government issues debt amounting to 1% of annual GDP and allocates the proceeds to either government spending, asset purchases, or tax cuts.

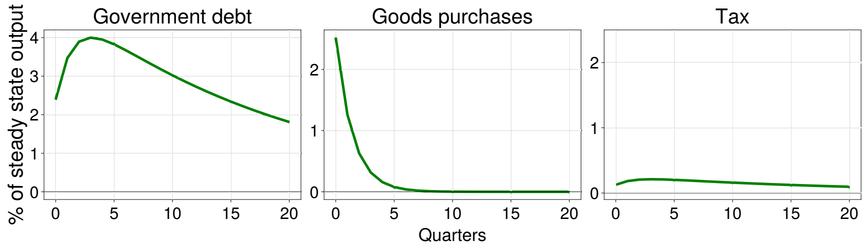

Figure 3: Policy paths for the government spending experiment. Left panel: government debt. Center panel: goods purchases. Right panel: taxes.

Output Responses

Figure 4 shows how output responses to government spending depend on assumptions about the financial sector. Each line represents a version of our model with a different specification of the financial sector's liquidity supply elasticities. The blue and black lines indicate responses with perfectly inelastic and elastic supply. The red line represents output responses with elasticities implied by our empirical estimates for $\Theta(\cdot)$ and $G(\cdot)$. Yellow shades from dark to light represent models with decreasing values for leverage sensitivity $\bar{\Theta}_{r^K}$ from our empirical estimate ($\bar{\Theta}_{r^K} = 24$) to that implied by a Gertler-Karadi-Kiyotaki type model ($\bar{\Theta}_{r^K} = 12$).

Output responses on impact differ substantially across different assumptions about the financial sector, ranging from 2.9% to 4.6% of steady-state output, with inelastic liquidity supply associated with stronger responses. The impact government spending multiplier ranges from 0.9 to 1.8, and the cumulative multiplier from 0.6 to 2.5.[12]

[12]The cumulative multiplier is calculated as $\sum_{t=0}^{\infty}(1+r^B)^{-t}dy_t/\sum_{t=0}^{\infty}(1+r^B)^{-t}dg_t$. If we use r^K instead, the corresponding numbers are 0.7 and 2.5

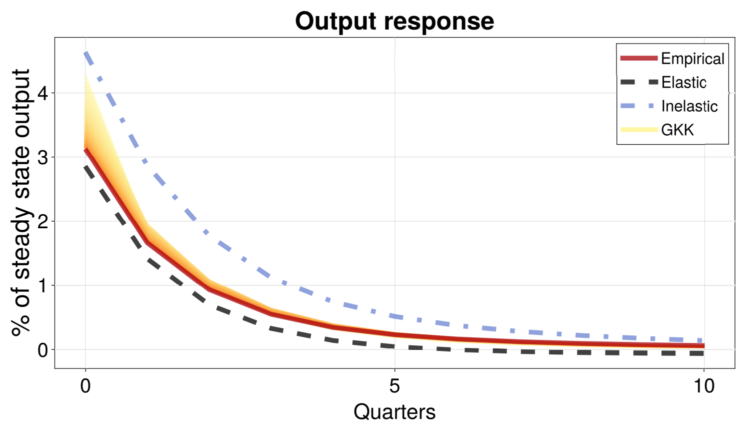

Figure 4: Output response to government goods purchases with $\eta = 0.5, \rho_{b^G} = 0.95$. Red: empirical estimate ($\bar{\Theta}_{r^K} = 24$). Black: perfectly elastic liquidity supply ($\bar{\Theta}_{r^K} \to \infty$). Blue: perfectly inelastic liquidity supply ($\mathcal{D}_{r^K} = 0$). Dark to light yellow: high to low $\bar{\Theta}_{r^K}$ (from $\bar{\Theta}_{r^K} = 24$ to $\bar{\Theta}_{r^K} = 12$ in GKK). Figure 6a in Appendix F.2 shows the responses of consumption and investment.

Decomposition

To understand why assumptions about the financial sector lead to widely different outcomes, we use Theorem 1 to decompose output responses into the goods market channel, the asset market channel, and the modified Keynesian cross, as shown in Figure 5.

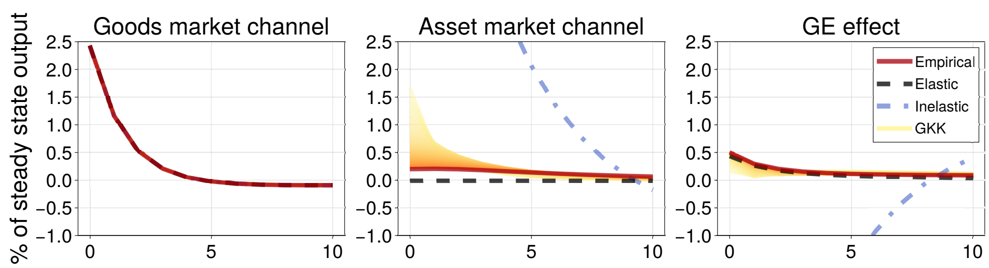

Figure 5: Decomposition of output responses to government goods purchases with $\eta = 0.5, \rho_{b^G} = 0.95$ using Theorem 1. Left panel: goods market channel. Center panel: asset market channel. Right panel: GE effect (the difference between total output responses and the sum of the goods market and asset market channels). Red: empirical estimate ($\bar{\Theta}_{r^K} = 24$). Black: perfectly elastic liquidity supply ($\bar{\Theta}_{r^K} \to \infty$). Blue: perfectly inelastic liquidity supply ($\mathcal{D}_{r^K} = 0$). Dark to light yellow: high to low $\bar{\Theta}_{r^K}$ (from $\bar{\Theta}_{r^K} = 24$ to $\bar{\Theta}_{r^K} = 12$ in GKK).

The goods market channel reflects direct responses of aggregate demand to $d\boldsymbol{g}$ and $d\boldsymbol{T}$. Since these responses do not depend on the financial sector, all specifications generate the same outcome.

By contrast, the same policy generates significant differences in the asset markets due to different assumptions about the financial sector, as shown in the middle panel. As government debt and taxes shift excess liquidity, variation in the financial sector's liquidity supply elasticities generates a wide range of responses through the asset market propagation matrix $\mathbf{\Omega}$. By changing $\bar{\Theta}_{r^K}$ from our empirical estimate to that imposed by a GKK model, aggregate demand rises by an order of magnitude, and the contrast is even more drastic between the perfectly elastic and inelastic benchmarks.

Finally, we present the modified Keynesian cross as the general equilibrium (GE) effect, which is the difference between total output responses and the sum of the first two channels. A prominent aspect of this channel is a strong dampening response when liquidity supply is perfectly inelastic (blue line): While shifts in excess liquidity generate strong aggregate demand through the asset market channel, increases in aggregate income shift up liquidity demand, reduce excess liquidity, and dampen the total output responses. The dampening force reduces the output response on impact by up to 50%. The same dampening force is also present in other specifications, but it is dominated by the standard Keynesian feedback when liquidity supply elasticities are high.

The Size of Multipliers

While Figure 4 shows output responses to government spending with a specific path of government debt and taxes, Figure 6 shows that our main message holds for a wide range of debt financing schemes, as parameterized by ρ_{b^G}.

Across all degrees of debt financing, the government spending multiplier is always significantly larger when the financial sector features an inelastic liquidity supply (blue line) than an elastic supply (black line). In both cases, our model converges to canonical HANK frameworks that are commonly used for analyzing fiscal and monetary policies, as we discussed in Section 4.2. The only difference between the two is their assumptions about the financial sector. Yet, they generate government spending impact multipliers that vary by almost a factor of two as shifts in excess liquidity generate different asset market responses. The differences are even larger for cumulative multipliers. Even in the case without debt financing ($\rho_{b^G} = 0$), assumptions about the financial sector matter as taxes imposed by the government shift households' liquidity demand. With a greater degree of debt financing, assumptions about the financial sector lead to larger variations in the multipliers, as the issuance of government debt generates larger shifts in excess liquidity. Figure 6b in Appendix F.1 shows a similar when we fix $\rho_{b^G} = 0$ and vary the persistence of spending, η, instead.

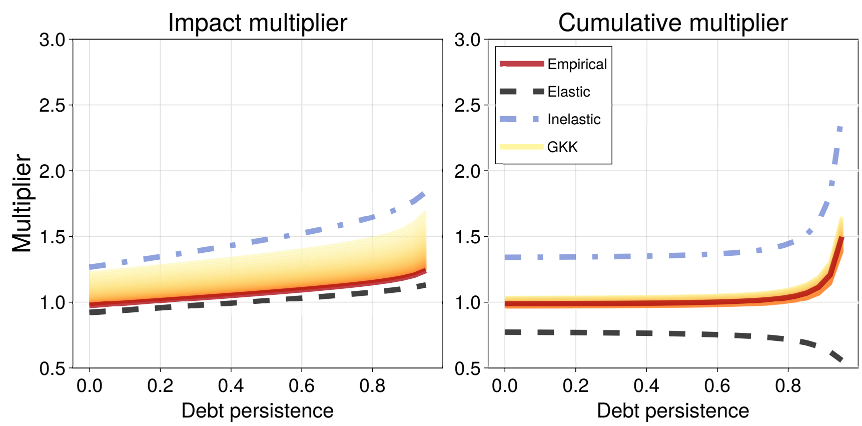

Figure 6: Government spending multipliers for $\eta = 0.5$ and $\rho_{b^G} \in [0, 0.95]$. Left panel: impact multiplier (dy_0/dg_0). Right panel: cumulative multiplier ($\sum_{t=0}^{\infty}(1+r^B)^{-t}dy_t/\sum_{t=0}^{\infty}(1+r^B)^{-t}dg_t$). Red: empirical estimate ($\bar{\Theta}_{r^K} = 24$). Black: perfectly elastic liquidity supply ($\bar{\Theta}_{r^K} \to \infty$). Blue: perfectly inelastic liquidity supply ($\mathcal{D}_{r^K} = 0$). Dark to light yellow: high to low $\bar{\Theta}_{r^K}$ (from $\bar{\Theta}_{r^K} = 24$ to $\bar{\Theta}_{r^K} = 12$ in GKK).

6.2 Asset Purchases vs. Tax Cuts

Our second policy question concerns the Wall Street vs. Main Street debate: Can asset purchases stimulate aggregate output more effectively than tax cuts? We compare two alternative policies for the government.

The first policy, given by $\{d\tilde{b}_t^G, d\tilde{g}_t, d\tilde{r}_t^B, d\tilde{a}_t^G, d\tilde{T}_t\}$, features an asset purchase program in which the government issues liquid debt $d\tilde{b}_t^G$ and holds illiquid asset $d\tilde{a}_t^G$ of the same value: $d\tilde{a}_t^G = d\tilde{b}_t^G$. The government keeps $d\tilde{g}_t = d\tilde{r}_t^B = 0$, and adjusts tax revenue $d\tilde{T}_t$ to satisfy its budget constraint. This policy implies net asset purchases:[13]

$$d\Delta_t \coloneqq d\tilde{a}_t^G - (1 + r^A)d\tilde{a}_{t-1}^G.$$

For the second policy, given by $\{d\check{b}_t^G, d\check{g}_t, d\check{r}_t^B, d\check{a}_t^G, d\check{T}_t\}$, the government keeps illiquid asset holdings at $d\check{a}_t^G = 0$ and pays out $d\Delta_t$ as tax cuts: $d\check{T}_t \coloneqq d\tilde{T}_t - d\Delta_t$. It maintains the same path for debt, $d\check{b}_t^G = d\tilde{b}_t^G$, and $d\check{g}_t = d\check{r}_t^B = 0$.

To parameterize the policy paths, we assume the government debt follows

$$d\tilde{b}_t^G = \rho_{b^G}(d\tilde{b}_{t-1}^G + s_t), \quad s_t = \eta^t s_0.$$

As in Section 6.1, we set $\rho_{b^G} = 0.95$ and $\eta = 0.5$, so that government debt peaks at 1% of annual GDP in one year, and we calculate the implied paths for asset purchases and tax

[13]Since $a^G = 0$ in the steady state and net asset purchases are given by $\Delta_t \coloneqq \tilde{a}_t^G - (1 + r_t^A)\tilde{a}_{t-1}^G$, and the formula for $d\Delta_t$ follows from a first-order approximation with $dr_t^A a^G = 0$.

cuts. Figure 7 shows the paths for the two policies for government debt, net asset purchases ($d\Delta_t$ and 0), and tax revenue ($d\tilde{T}_t$ and $d\check{T}_t$).

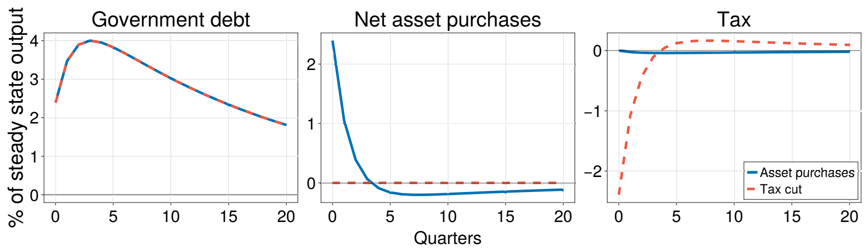

Figure 7: Policy paths for the asset purchase and tax cut experiments. Left panel: government debt. Center panel: net asset purchases. Right panel: taxes. Solid lines: asset purchase program; dashed lines: tax cut program.

Output Responses

Figure 8 compares how different assumptions about the financial sector's liquidity supply elasticities affect output responses to asset purchases and tax cuts.

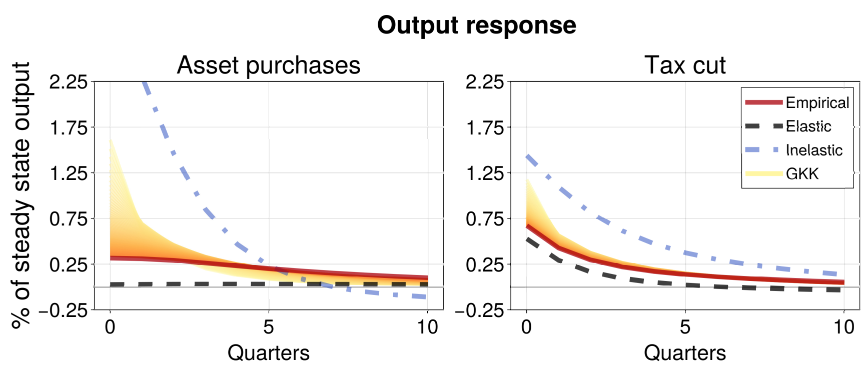

Figure 8: Output responses to asset purchases and tax cuts with $\eta = 0.5, \rho_{bG} = 0.95$. Left panel: asset purchases. Right panel: tax cuts. Red: empirical estimate ($\bar{\Theta}_{r^K} = 24$). Black: perfectly elastic liquidity supply ($\bar{\Theta}_{r^K} \to \infty$). Blue: perfectly inelastic liquidity supply ($\mathcal{D}_{r^K} = 0$). Dark to light yellow: high to low $\bar{\Theta}_{r^K}$ (from $\bar{\Theta}_{r^K} = 24$ to $\bar{\Theta}_{r^K} = 12$ in GKK).

For asset purchases, output responses on impact differ by orders of magnitudes, ranging from 0.03% to more than 3.2% between perfectly elastic and inelastic liquidity supply; for tax cuts, output responses differ almost by a factor of three, ranging from 0.5% to 1.4%. Despite the difference being substantial for both policies, the effects of asset purchases are noticeably more sensitive to assumptions about the financial sector than tax cuts. To understand the contrast between the two policies, we decompose output responses into the three channels using Theorem 1.

Decomposition

Figure 9 shows the decomposition for asset purchases and tax cuts, with each column representing a channel. The left column shows the goods market channel. Asset purchases have little effect through this channel as their impact on tax revenue is limited. Tax cuts, on the contrary, generate substantial responses due to households' high MPCs. The effect of tax cuts emphasizes the importance of household heterogeneity and illiquidity. These features break Ricardian equivalence and generate substantial effects from deficit-financed tax cuts. However, since households' MPCs do not depend on the financial sector, all specifications generate the same outcome through the goods market channel.

By contrast, through the asset market channel, assumptions about the financial sector generate a wide range of outcomes in response to shifts in excess liquidity generated by the two policies. Comparing the two policies, the asset purchase program generates stronger effects through asset markets as it generates larger shifts in excess liquidity. However, as it relies more on the propagation through asset markets, it is more sensitive to the assumption about the financial sector than tax cuts, explaining the wider range of output responses for asset purchases in Figure 8.

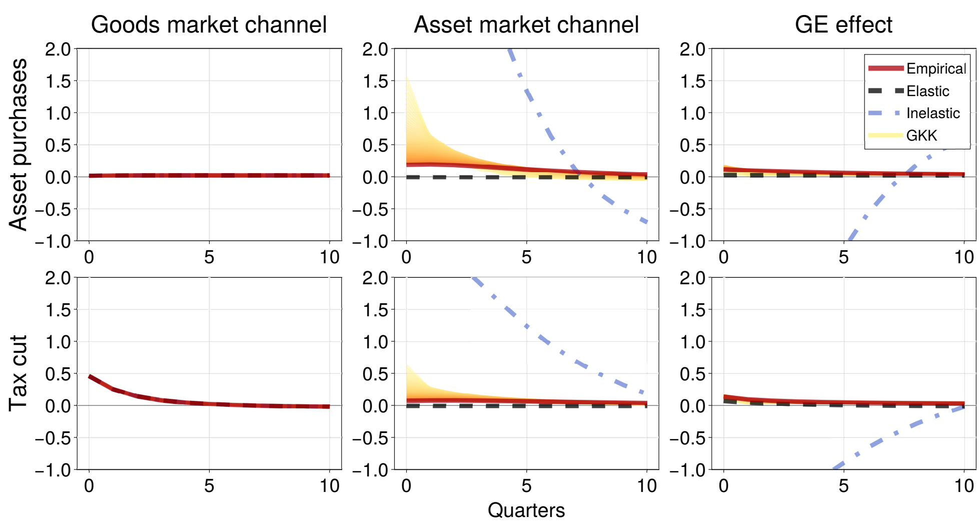

Figure 9: Decomposition of output responses using Theorem 1. First row: asset purchases; second row: tax cuts; both with $\eta = 0.5, \rho_{b^G} = 0.95$. Left column: goods market channel. Center column: asset market channel. Right column: GE effect (the difference between total output responses and the sum of the goods market and asset market channels). Red: empirical estimate ($\bar{\Theta}_{r^K} = 24$). Black: perfectly elastic liquidity supply ($\bar{\Theta}_{r^K} \to \infty$). Blue: perfectly inelastic liquidity supply ($\mathcal{D}_{r^K} = 0$). Dark to light yellow: high to low $\bar{\Theta}_{r^K}$ (from $\bar{\Theta}_{r^K} = 24$ to $\bar{\Theta}_{r^K} = 12$ in GKK).

Relative Effects of Asset Purchases vs. Tax Cuts

To study how our conclusions about the effectiveness of the two policies depend on assumptions about the financial sector, we calculate the difference in output responses to asset purchases and tax cuts:

$$d\boldsymbol{y}^{asset} - d\boldsymbol{y}^{tax} = \underbrace{\left(\mathbf{I} - \boldsymbol{\Psi}_y - \boldsymbol{\Omega}\boldsymbol{\epsilon}_y\right)^{-1}}_{\text{(3) modified Keynesian cross}} \Big(\underbrace{\boldsymbol{\Psi}_T d\boldsymbol{\Delta}}_{\text{(1) goods market}} + \underbrace{\boldsymbol{\Omega}\,\boldsymbol{\epsilon}_T d\boldsymbol{\Delta}}_{\text{(2) asset market}} \Big).$$

The difference is positive when asset purchases have a stronger effect on output and vice versa. As the difference depends only on $d\boldsymbol{\Delta}$, our conclusion about the relative effects does not hinge on our assumption about $d\boldsymbol{g}, d\boldsymbol{r}^B$, as long as both policies feature the same paths. Therefore, we effectively control for responses due to other policy variables by focusing on the difference.

Figure 10 represents the difference in output responses. On one end, when the financial sector features perfectly inelastic liquidity supply, asset purchases have a stronger effect than tax cuts: the difference in output response amounts to 1.8% of steady-state output on impact. A Gertler-Karadi-Kiyotaki type model gives a qualitatively similar prediction: Asset purchases are more effective in stimulating output, with a 0.4% difference. On the other end, when the financial sector features a perfectly elastic liquidity supply, the asset market channel vanishes, the effect of tax cuts dominates through the goods market channel, and the difference is −0.5% on impact. Compared to this benchmark, our empirical elasticities generate a non-negligible response to asset purchases. Still, our estimates of these elasticities are relatively high, which implies modest asset market responses compared to the goods market channel. As a result, they predict that policies targeting households directly, such as tax cuts, can stimulate output more effectively than policies that rely mostly on the asset market channel, such as asset purchases.

The weak effects of the asset purchase program implied by our empirical elasticities can also shed light on similar policies such as Quantitative Easing or Operation Twist—although we do not explicitly model the maturity structure of Treasury debts. To the extent that long-term government bonds share characteristics with illiquid assets in our framework, the effects of QE would similarly depend on the cross-price elasticities through the asset market channel. Large estimates of the cross-price elasticities would suggest more modest effects than standard calibrations predict. Our empirical estimates therefore offers a potential way to reconcile the large effects of QE found in some quantitative models, e.g., Cui and Sterk (2021), with the more mixed empirical evidence (Krishnamurthy and Vissing-Jorgensen, 2011; Weale and Wieladek, 2016).

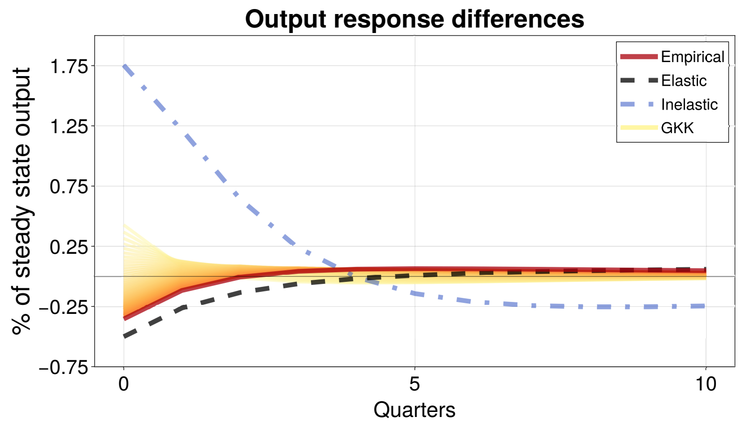

Figure 10: Difference between output response to asset purchases and tax cuts. Positive values mean that responses to asset purchases are larger. Red: empirical estimate ($\bar{\Theta}_{rK} = 24$). Black: perfectly elastic liquidity supply ($\bar{\Theta}_{rK} \to \infty$). Blue: perfectly inelastic liquidity supply ($\mathcal{D}_{rK} = 0$). Dark to light yellow: high to low $\bar{\Theta}_{rK}$ (from $\bar{\Theta}_{rK} = 24$ to $\bar{\Theta}_{rK} = 12$ in GKK).

7 Conclusion

In this paper, we show that, for a large class of macro models with financial frictions, the financial sector's asset supply elasticities provide sufficient information about the underlying financial frictions. By focusing on these elasticities, we can strengthen empirical discipline on these frictions. Such discipline is highly policy-relevant, as we demonstrated in our policy analysis. Moreover, with minimal assumptions on the detailed microfoundation, we can integrate this class of financial intermediation models into state-of-the-art quantiative macro models with rich features such as household heterogeneity and illiquidity. This integration allows us to study a set of important policy questions that standard macro-finance models with a simplified household sector are not suited for.

Generalizing the financial sector's asset supply system to include various types of assets and intermediaries is a natural step for a comprehensive framework to study the transmission of policies and shocks through asset markets. Normative analysis based on such an asset supply system will allow us to understand how characteristics of the financial sector should shape a wide range of government policies, such as open market operations, quantitative easing and tightening, and Operation Twist. We leave these topics for future research.

Data Availability Statement

The data and code underlying this article are available at https://doi.org/10.5281/zenodo.18764264.

References

Aguiar, Mark, Mark Bils, and Corina Boar, "Who Are the Hand-to-Mouth?," *The Review of Economic Studies*, 05 2024, *92* (3), 1293–1340.

Akinci, Ozge and Albert Queralto, "Credit spreads, financial crises, and macroprudential policy," *American Economic Journal: Macroeconomics*, 2022, *14* (2).

Angeletos, George-Marios, Chen Lian, and Christian K. Wolf, "Can Deficits Finance Themselves?," *Econometrica*, 2024, *92* (5), 1351–1390.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, 2021, *89* (5), 2375–2408.

_ , Matthew Rognlie, and Ludwig Straub, "The Intertemporal Keynesian Cross," *Journal of Political Economy*, 2024, *132* (12), 4068–4121.

Bauer, Michael D and Eric T Swanson, "A reassessment of monetary policy surprises and high-frequency identification," *NBER Macroeconomics Annual*, 2023, *37* (1), 87–155.

Baumeister, Christiane and James D Hamilton, "Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks," *American Economic Review*, 2019, *109* (5), 1873–1910.

Bayer, Christian, Benjamin Born, and Ralph Luetticke, "The liquidity channel of fiscal policy," *Journal of Monetary Economics*, 2023, *134*, 86–117.

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.

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

Brunnermeier, Markus K and Yuliy Sannikov, "A macroeconomic model with a financial sector," *American Economic Review*, 2014, *104* (2), 379–421.

Chaudhary, Manav, Zhiyu Fu, and Haonan Zhou, "Anatomy of the Treasury Market: Who Moves Yields?," *Available at SSRN*, 2024.

Chi, Chun-Che, Stephanie Schmitt-Grohé, and Martín Uribe, "Optimal Bank Reserve Remuneration and Capital Control Policy," Technical Report 1 January 2025.

Chiang, Yu-Ting and Piotr Żoch, "Optimal Asset Market Operations," Technical Report 2025.

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

Cui, Wei and Vincent Sterk, "Quantitative easing with heterogeneous agents," *Journal of Monetary Economics*, 2021, *123.*

Cúrdia, Vasco and Michael Woodford, "Credit frictions and optimal monetary policy," *Journal of Monetary Economics*, 2016, *84.*

den Heuvel, Skander J. Van, "The welfare cost of bank capital requirements," *Journal of Monetary Economics*, 2008, *55* (2).

Dávila, Eduardo and Andreas Schaab, "Optimal Monetary Policy with Heterogeneous Agents: Discretion, Commitment, and Timeless Policy," Working Paper 30961, National Bureau of Economic Research February 2023.

Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuño, "Financial Frictions and the Wealth Distribution," Technical Report 3 2023.

Ferrante, Francesco and Nils Gornemann, "Devaluations, Deposit Dollarization, and Household Heterogeneity," *International Finance Discussion Paper*, 2022, (1336).

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), 2274–2303.

_ , _ , Federico Mainardi, Sangmin Oh, and Motohiro Yogo, "Asset demand of us households," Technical Report, National Bureau of Economic Research 2024.

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

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

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?," Technical Report 5 2021.

Jansen, Kristy AE, Wenhao Li, and Lukas Schmid, "Granular treasury demand with arbitrageurs," Technical Report, National Bureau of Economic Research 2024.

Kaplan, Greg and Giovanni L. Violante, "The Marginal Propensity to Consume in Heterogeneous Agent Models," NBER Working Papers 30013, National Bureau of Economic Research, Inc 2022.

— **, Benjamin Moll, and Giovanni L Violante**, "Monetary policy according to HANK," *American Economic Review*, 2018, *108* (3).

— **, Giovanni L. Violante, and Justin Weidner**, "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 2014, *45.*

Karadi, Peter and Anton Nakov, "Effectiveness and addictiveness of quantitative easing," *Journal of Monetary Economics*, 2021, *117* (C).

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

Krishnamurthy, Arvind and Annette Vissing-Jorgensen, "The effects of quantitative easing on interest rates: channels and implications for policy," Technical Report, National Bureau of Economic Research 2011.

Lee, Donggyu, "Quantitative Easing and Inequality," mimeo, Federal Reserve Bank of New York 2021.

Lee, Seungcheol, Ralph Luetticke, and Morten Ravn, "Financial Frictions: Macro vs Micro Volatility," CEPR Discussion Papers 15133 2020.

McKay, Alisdair and Christian K. Wolf, "Optimal Policy Rules in HANK," Technical Report 2023.

Mendicino, Caterina, Lukas Nord, and Marcel Peruffo, "Distributive Effects of Banking Sector Losses," Technical Report 2021.

Ottonello, Pablo, Diego J Perez, and Paolo Varraso, "Are collateral-constraint models ready for macroprudential policy design?," *Journal of International Economics*, 2022, *139.*

Plagborg-Møller, Mikkel and Christian K. Wolf, "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 2021, *89* (2).

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

Weale, Martin and Tomasz Wieladek, "What are the macroeconomic effects of asset purchases?," *Journal of monetary Economics*, 2016, *79*, 81–93.

Wolf, Christian K., "Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result," NBER Working Papers 29193, National Bureau of Economic Research, Inc August 2021.

_ , "Interest Rate Cuts versus Stimulus Payments: An Equivalence Result," *Journal of Political Economy*, 2025, *133* (4), 1235–1275.

Woodford, Michael, "Fiscal Requirements for Price Stability," *Journal of Money, Credit and Banking*, 2001, *33* (3), 669–728.

A Proofs and Derivations

A.1 Proof of Proposition 1

Proof. Totally differentiating $d_t = (\Theta_t - 1)n_t$ at steady state yields

$$d\mathcal{D}_t = d\Theta_t \bar{n} + (\bar{\Theta} - 1)dn_t,$$

where net worth evolves as

$$dn_t = \bar{G}_\Theta d\Theta_{t-1} + \bar{G}_{r^K} dr_t^K + \bar{G}_{r^B} dr_t^B + \bar{G}_n dn_{t-1}.$$

Iterate backward:

$$dn_t = \bar{G}_\Theta \sum_{u=0}^{t-1} \bar{G}_n^u d\Theta_{t-1-u} + \bar{G}_n^{t-s} \mathbf{1}_{\{s \le t\}} (\bar{G}_{r^K} dr_t^K + \bar{G}_{r^B} dr_t^B).$$

Consider a change in return dr_s^K with $dr_t^K = 0$, $\forall t \neq s$ and $dr_t^B = 0$, $\forall t \geq 0$.

From Lemma 2,

$$\frac{d\Theta_{t-u-1}}{dr_s^K} = \begin{cases} \gamma^{s-t+u} \bar{\Theta}_{r^K}, & t > u \geq t - s, \\ 0, & t \leq u, \; u < t - s. \end{cases}$$

Define $\sigma(s) := \frac{1-(\gamma\bar{G}_n)^s}{1-\gamma\bar{G}_n} \times \mathbf{1}_{\{s\geq 0\}}$, then

$$\sum_{u=0}^{t-1} \bar{G}_n^u d\Theta_{t-1-u} = \begin{cases} \gamma^{s-t}\sigma(t)\bar{\Theta}_{r^K} dr_s^K, & s > t, \\ \bar{G}_n^{t-s}\sigma(s)\bar{\Theta}_{r^K} dr_s^K, & s \leq t. \end{cases}$$

Define $\Sigma(t) := \frac{1}{\bar{n}}\bar{G}_\Theta\sigma(t)$ and use $D_t = (\Theta_t - 1)n_t$, the liquidity supply elasticity is:

$$\frac{\partial \mathcal{D}_t / \partial r_s^K}{\mathcal{D}_t} = \begin{cases} \gamma^{s-t-1}\bar{\Theta}_{r^K}\left(\frac{1}{\Theta-1} + \gamma\Sigma(t)\right), & s > t \\ \left(\frac{1}{\bar{n}}\bar{G}_{r^K} + \bar{\Theta}_{r^K}\Sigma(s)\right)\bar{G}_n^{t-s}, & s \leq t \end{cases}$$

□

A.2 Proof of Lemma 3

Proof. We define aggregate functions for the household, production, and financial blocks.

Households

First, the household problem maps after-tax income and returns, $\left\{y_{i,t} - \mathcal{T}(y_{i,t}), r_t^A, r_t^B\right\}_{s=0}^{\infty}$, to consumption, savings in each type of asset, and the associated adjustment cost.

Second, we link $y_{i,t} - \mathcal{T}(y_{i,t})$ to aggregate output, y_t, and total tax revenue T_t. From the firm's problem, we have $\frac{W_t}{P_t}h_t = (1-\alpha)\, y_t$. Because labor unions are identical, $h_{lt} = h_t$, and the labor demand rule implies $h_{i,t} = l(z_{i,t})h_t$ and $y_{i,t} = z_{i,t}l(z_{i,t})(1-\alpha)y_t$.

Given the tax system, after-tax income for household i is

$$y_{i,t} - \mathcal{T}(y_{i,t}) = (1-\tau_t)\big(z_{i,t}l(z_{i,t})(1-\alpha)y_t\big)^{1-\lambda}.$$

To generate tax revenue T_t, tax rate τ_t satisfies

$$\int y_{i,t}di - T_t = (1-\tau_t)\big((1-\alpha)y_t\big)^{1-\lambda}\int \big(z_{i,t}l(z_{i,t})\big)^{1-\lambda}di.$$

Therefore,

$$1 - \tau_t = \frac{(1-\alpha)y_t - T_t}{\big((1-\alpha)y_t\big)^{1-\lambda}\int \big(z_{i,t}l(z_{i,t})\big)^{1-\lambda}di},$$

and individual after-tax income is:

$$y_{i,t} - \mathcal{T}(y_{i,t}) = \frac{\big(z_{i,t}l(z_{i,t})\big)^{1-\lambda}}{\int \big(z_{i,t}l(z_{i,t})\big)^{1-\lambda}di}\Big((1-\alpha)y_t - T_t\Big).$$

As a result, individual household policy rules can be expressed as functions of the idiosyncratic state $\{z_{i,s}\}$ and $\{y_s, r_s^A, r_s^B, T_s\}_{s=0}^{\infty}$. Given the initial distribution of assets and productivity, aggregation yields the aggregate functions:

$$\mathcal{A}_t\left(\left\{y_s, r_s^A; r_s^B, T_s\right\}_{s=0}^{\infty}\right), \quad \mathcal{B}_t\left(\left\{y_s, r_s^A; r_s^B, T_s\right\}_{s=0}^{\infty}\right), \quad \mathcal{C}_t\left(\left\{y_s, r_s^A; r_s^B, T_s\right\}_{s=0}^{\infty}\right),$$

where the consumption function is defined to include the adjustment cost:

$$\mathcal{C}_t\left(\left\{y_s, r_s^A; r_s^B, T_s\right\}_{s=0}^{\infty}\right) := \int c_{i,t} + \Phi(a_{i,t}, a_{i,t-1}, r_t^A) di.$$

Production

From the law of motion for capital, the investment ratio is

$$\frac{x_t}{k_{t-1}} = \Gamma^{-1}\left(\frac{k_t - (1-\delta)\, k_{t-1}}{k_{t-1}}\right) =: \iota(k_t, k_{t-1}). \tag{13}$$

Substitute it into the first order condition for ι_t, we have

$$q_t = \frac{1}{\Gamma'\left(\iota(k_t, k_{t-1})\right)} =: \hat{q}\left(k_t, k_{t-1}\right). \tag{14}$$

This yields the return on capital:

$$r_{t+1}^K = \frac{\alpha \frac{y_{t+1}}{k_t} + \hat{q}\left(k_{t+1}, k_t\right)\left(\frac{k_{t+1}}{k_t}\right) - \iota(k_{t+1}, k_t)}{\hat{q}\left(k_t, k_{t-1}\right)} - 1,$$

which determines capital in each period as a function $\mathcal{K}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right)$, given k_{-1}.

Use it in Equation (13) and (14), we have the investment and capital price functions: $\mathcal{X}_t(\{y_s, r_s^K\}_{s=0}^{\infty})$ and $\mathcal{Q}_t(\{y_s, r_s^K\}_{s=0}^{\infty})$.

Illiquid Returns

For function $\mathcal{R}_t^A(\cdot)$, we use Equation (8):

$$r_{t+1}^A = \frac{1}{a_t}(r_{t+1}^K \omega_{M,t} + r_{t+1}^N n_t).$$

Return on net worth is given by

$$r_{t+1}^N = \frac{r_{t+1}^K \omega_t - r_{t+1}^B d_t}{n_t}.$$

The balance sheet of the mutual fund implies $a_t = \omega_{M,t} + \omega_t - d_t$, and the capital market clearing requires that $\omega_{M,t} + \omega_t = q_t k_t$. Therefore, we have

$$r_{t+1}^A = \frac{1}{q_t k_t - d_t}(r_{t+1}^K q_t k_t - r_{t+1}^B d_t).$$

Using functions $\mathcal{Q}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right)$, $\mathcal{K}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right)$, and $\mathcal{D}_t$,

$$\mathcal{R}_{t+1}^A\left(\left\{r_s^K, r_s^B, y_s; \mathcal{D}_t\right\}_{s=0}^{\infty}\right) := \frac{r_{t+1}^K \mathcal{Q}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right) \mathcal{K}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right) - r_{t+1}^B \mathcal{D}_t}{\mathcal{Q}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right) \mathcal{K}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right) - \mathcal{D}_t}. \tag{15}$$

Market Clearing

From the definition of the aggregate functions, the goods market clearing and liquid asset market clearing conditions are given by

$$\mathcal{C}_t(\{y_s, r_s^A; r_s^B, T_s\}_{s=0}^{\infty}) + \mathcal{X}_t(\{y_s, r_s^K\}_{s=0}^{\infty}) + g_t = y_t, \tag{16}$$

$$\mathcal{B}_t\left(\left\{y_s, r_s^A; r_s^B, T_s\right\}_{s=0}^{\infty}\right) = \mathcal{D}_t\left(\left\{r_s^K, r_s^B\right\}_{s=0}^{\infty}\right) + b_t^G. \tag{17}$$

Given $\left\{g_s, T_s, r_s^B, b_s^G\right\}_{s=0}^{\infty}$, let $\{y_s, r_s^K\}_{s=0}^{\infty}$ be a sequence that satisfies Equations 15, 16, and 17. We solve a_s^G from Equation (9), so the government budget constraint is satisfied. Because the aggregate functions for households are derived under household budget constraints, by the Walras law, the illiquid asset market clears

$$\mathcal{A}_t\left(\left\{y_s, r_s^A, r_s^B; T_s\right\}\right) = \mathcal{Q}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right) \mathcal{K}_t\left(\left\{y_s, r_s^K\right\}_{s=0}^{\infty}\right) - \mathcal{D}_t\left(\left\{r_s^K, r_s^B\right\}_{s=0}^{\infty}\right) - a_t^G.$$

Therefore, the sequence $\{y_s, r_s^K\}_{s=0}^{\infty}$ constitutes an equilibrium. □

A.3 Time 0 Returns

We express r_0^K as a function of output and expected returns:

$$r_0^K = \frac{\alpha \frac{y_0}{k_{-1}} + \hat{q}\left(k_0, k_{-1}\right)\left(\frac{k_0}{k_{-1}}\right) - \iota(k_0, k_{-1})}{\hat{q}\left(k_{-1}, k_{-2}\right)} - 1,$$

where only y_0 and k_0 are not predetermined. From the proof of Lemma 3, we have $k_0 = \mathcal{K}_0\left(\left\{y_s, r_{s+1}^K\right\}_{s=0}^{\infty}\right)$. This allows us to write r_0^K as a function of $\left\{y_s, r_s^K\right\}_{s=0}^{\infty}$.

A.4 Proof of Proposition 2.

Proof. Recall the definition of excess liquidity supply

$$\mathcal{E}_t(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{T}, \boldsymbol{b}^G) := \mathcal{D}_t(r_0^K(\boldsymbol{y}, \boldsymbol{r}^K), \boldsymbol{r}^K, \boldsymbol{r}^B) + b_t^G - \mathcal{B}_t(\boldsymbol{y}, r^A(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B), \boldsymbol{r}^B, \boldsymbol{T}).$$

Market clearing requires: $\mathcal{E}_t(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{T}, \boldsymbol{b}^G) = 0$. Totally differentiating this condition in every period yields

$$\boldsymbol{\epsilon}_{r^K} d\boldsymbol{r^K} + \boldsymbol{\epsilon}_y d\boldsymbol{y} + \boldsymbol{\epsilon}_T d\boldsymbol{T} + d\boldsymbol{b^G} + \boldsymbol{\epsilon}_{r^B} d\boldsymbol{r^B} = \boldsymbol{0},$$

where $\boldsymbol{\epsilon}_{r^K} := \mathbf{D}_{r^K} - \tilde{\mathbf{B}}_{r^K},\ \boldsymbol{\epsilon}_{r^B} := \mathbf{D}_{r^B} - \tilde{\mathbf{B}}_{r^B},\ \boldsymbol{\epsilon}_y := \mathbf{D}_y - \tilde{\mathbf{B}}_y,\ \boldsymbol{\epsilon}_T := -\tilde{\mathbf{B}}_T$, with the matrices defined in Appendix B.3. Rearrange and left-multiply by the inverse of $-\boldsymbol{\epsilon}_{r^K}$ to obtain Equation (10).

For the second part of Proposition 2, note that if $\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B} \to \infty$ and $\bar{\Theta}_{r^B}/\bar{\Theta}_{r^K} \to \varsigma$, Proposition 1 implies

$$\frac{\partial \mathcal{D}_t}{\partial r_s^K} \frac{1}{\bar{\Theta}_{r^K}} \to \begin{cases} \Sigma(s) G_n^{t-s}(\bar{\Theta} - 1)n, & s \leq t, \\ \gamma^{s-t-1}\Big(1 + (\bar{\Theta} - 1)\gamma\Sigma(t)\Big)n, & s > t, \end{cases}$$

$$\frac{\partial \mathcal{D}_t}{\partial r_s^B} \frac{1}{\bar{\Theta}_{r^K}} \to \begin{cases} -\varsigma\Sigma(s) \bar{G}_n^{t-s}(\bar{\Theta} - 1)n, & s \leq t, \\ -\varsigma\gamma^{s-t-1}\Big(1 + (\bar{\Theta} - 1)\gamma\Sigma(t)\Big)n, & s > t. \end{cases}$$

Thus, $\mathbf{D}_{r^K} \frac{1}{\bar{\Theta}_{r^K}} \to \mathbf{D}_{\infty,r},\ \mathbf{D}_{r^B} \frac{1}{\bar{\Theta}_{r^K}} \to -\varsigma \mathbf{D}_{\infty,r}$, where

$$\mathbf{D}_{\infty,r} := \begin{cases} \Sigma(s) \bar{G}_n^{t-s}(\bar{\Theta} - 1)n, & s \leq t \\ \gamma^{s-t-1}\Big(1 + (\bar{\Theta} - 1)\gamma\Sigma(t)\Big)n, & s > t. \end{cases}$$

Assume that the derivatives of $\mathcal{B}_t$ are bounded. Divide the linearized liquid asset market clearing condition by $\bar{\Theta}_{r^K}$. As $\bar{\Theta}_{r^K}, \bar{\Theta}_{r^B} \to \infty$ with $\bar{\Theta}_{r^B}/\bar{\Theta}_{r^K} \to \varsigma$, for all bounded sequences $\{d\boldsymbol{y}, d\boldsymbol{r}^K, d\boldsymbol{r}^B, d\boldsymbol{b}^G\}$, the limit of the liquid asset market clearing condition is

$$\left(\mathbf{I} - \mathbf{B}_{r^A}(r^K - r^B)\frac{qk}{(qk-d)^2}\right)\mathbf{D}_r^\infty \left(d\boldsymbol{r}^K - \varsigma d\boldsymbol{r}^B\right) = \boldsymbol{0}.$$

The condition is satisfied for $d\boldsymbol{r}^K = \varsigma d\boldsymbol{r}^B$.

□

A.5 Proof of Theorem 1.

Proof. The aggregate demand is defined as

$$\Psi_t(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{T}, \boldsymbol{g}) \coloneqq \mathcal{C}_t(\boldsymbol{y}, r^A(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B), \boldsymbol{r}^B, \boldsymbol{T}) + \mathcal{X}_t(\boldsymbol{y}, \boldsymbol{r}^K) + g_t.$$

Goods market clears if $\Psi_t(\boldsymbol{y}, \boldsymbol{r}^K, \boldsymbol{r}^B, \boldsymbol{T}, \boldsymbol{g}) = y_t$. By totally differentiating this condition in every period we have

$$\boldsymbol{\Psi}_{r^K} d\boldsymbol{r}^{\boldsymbol{K}} + \boldsymbol{\Psi}_y d\boldsymbol{y} + \boldsymbol{\Psi}_T d\boldsymbol{T} + d\boldsymbol{b}^{\boldsymbol{G}} + \boldsymbol{\Psi}_{r^B} d\boldsymbol{r}^{\boldsymbol{B}} + d\boldsymbol{g} = d\boldsymbol{y}$$

where $\boldsymbol{\Psi}_{r^K} \coloneqq \tilde{\mathbf{C}}_{r^K} + \mathbf{X}_{r^K}$, $\boldsymbol{\Psi}_{r^B} \coloneqq \tilde{\mathbf{C}}_{r^B}$, $\boldsymbol{\Psi}_y \coloneqq \tilde{\mathbf{C}}_y + \mathbf{X}_y$, $\boldsymbol{\Psi}_T \coloneqq \tilde{\mathbf{C}}_T$, and the matrices are defined in Appendix B.3.

Let $\boldsymbol{\Omega} \coloneqq \boldsymbol{\Psi}_{r^K}(-\boldsymbol{\epsilon}_{r^K}^{-1})$, and use Proposition 2 to write

$$\boldsymbol{\Omega}\left(\boldsymbol{\epsilon}_y d\boldsymbol{y} + \boldsymbol{\epsilon}_T d\boldsymbol{T} + d\boldsymbol{b}^{\boldsymbol{G}} + \boldsymbol{\epsilon}_{r^B} d\boldsymbol{r}^{\boldsymbol{B}}\right) + \boldsymbol{\Psi}_y d\boldsymbol{y} + \boldsymbol{\Psi}_T d\boldsymbol{T} + d\boldsymbol{b}^{\boldsymbol{G}} + \boldsymbol{\Psi}_{r^B} d\boldsymbol{r}^{\boldsymbol{B}} + d\boldsymbol{g} = d\boldsymbol{y}.$$

Finally, rearrange it as

$$d\boldsymbol{y} = \left(\mathbf{I} - \boldsymbol{\Psi}_y - \boldsymbol{\Omega}\ \boldsymbol{\epsilon}_y\right)^{-1} \times \Big(\ d\boldsymbol{g} + \boldsymbol{\Psi}_T d\boldsymbol{T} + \boldsymbol{\Psi}_{r^B} d\boldsymbol{r}^B + \boldsymbol{\Omega}\left(d\boldsymbol{b}^G + \boldsymbol{\epsilon}_T d\boldsymbol{T} + \boldsymbol{\epsilon}_{r^B} d\boldsymbol{r}^{\boldsymbol{B}}\right)\Big),$$

which is the formula in Theorem 1. □