Firm Heterogeneity, Capital Misallocation and Optimal Monetary Policy*

Beatriz González — BdE, CEMFI

Galo Nuño — BdE, CEMFI

Dominik Thaler — ECB

First version: October 2020. This version: January 2026

Abstract

This paper analyzes the link between monetary policy and capital misallocation in a New Keynesian model with heterogeneous firms and financial frictions. In the model, firms with a high return to capital increase their investment more strongly in response to a monetary policy expansion, thus reducing misallocation. This feature creates a new time-inconsistent incentive for the central bank to engineer an unexpected monetary expansion to temporarily reduce misallocation. Nevertheless, near price stability remains the optimal timeless response to time-preference and TFP shocks. Capital misallocation does alter the optimal response to a cost-push shock: the planner tolerates higher inflation to stabilize both output and TFP.

Keywords: Monetary policy, firm heterogeneity, financial frictions, capital misallocation.

JEL classification: E12, E22, E43, E52, L11.

*Beatriz González (Banco de España, CEMFI): beatrizgonzalez@bde.es. Galo Nuño (Banco de España, CEMFI): galo.nuno@bde.es. Dominik Thaler (European Central Bank): dominik.thaler@eui.eu.

Two versions of this paper were previously circulated as "Monetary Policy and Capital Misallocation" and "Optimal Monetary Policy with Heterogeneous Firms". We are especially grateful to Andrew Atkeson, Anmol Bhandari, Sergi Basco, Saki Bigio, Eduardo Davila, Jesús Fernández-Villaverde, Dmitry Khametshin and Andreas Schaab for their input and suggestions. Silvia Albrizio contributed to a previous version of the manuscript that included an empirical analysis, which was removed in this version. We are deeply grateful to her. We thank Klaus Adam, Adrien Auclert, Bence Bardóczy, Pierpaolo Benigno, Paco Buera, Ricardo Caballero, Matias Covarrubias, Jim Costain, Wei Cui, Ester Faia, Luis Franjo, Veronica Guerrieri, Masashige Hamano, Greg Kaplan, Agnieszka Markiewicz, Ben Moll, Pierluca Pannella, Morten Ravn, Ricardo Reis, Kjetil Storesletten, Harald Uhlig, Gianluca Violante, Thomas Winberry, Chris Wolf, Tack Yun, and Francesco Zanetti for excellent comments, as well as participants at several seminars and conferences. The opinions and analysis are the responsibility of the authors, and therefore, do not necessarily coincide with those of the ECB, Banco de España or the Eurosystem. All errors are our own.

1 Introduction

An inefficient allocation of capital across firms may significantly reduce aggregate total factor productivity (TFP). In a market economy, capital is allocated according to the investment decisions of individual firms. Monetary policy is an important driver of these investment decisions, as it directly affects firms' funding costs and indirectly influences firms' revenues and other costs such as wages. If firm investment responds heterogeneously to changes in monetary policy, then monetary policy can affect capital misallocation and thus aggregate TFP.

Traditional analyses of monetary policy design abstract from this channel. In the workhorse model of monetary policy—the New Keynesian model—the central bank faces a short-run trade-off between stabilizing inflation and closing the output gap, taking aggregate productivity as exogenous. Yet if monetary policy itself affects productivity through its impact on the efficiency of capital allocation, then it also influences the economy's medium-run output potential. Should the central bank take this into account when designing optimal policy, and if so, in what way?

The objective of this paper is to shed light on the interaction between monetary policy and capital misallocation and its implications for optimal monetary policy. To this end, we develop a tractable framework that combines the workhorse New Keynesian model with a model of firm heterogeneity, based on Moll (2014), in which capital misallocation arises from financial frictions. The economy is populated by a continuum of firms owned by entrepreneurs, who have access to a constant-returns-to-scale technology. Entrepreneurs are heterogeneous in their net worth and receive idiosyncratic productivity shocks. They face financial frictions, as their borrowing cannot exceed a multiple of their net worth. This results in endogenous capital misallocation: entrepreneurs with productivity above a certain threshold are constrained and therefore borrow as much capital as possible, since their marginal revenue product of capital (MRPK) is higher than their cost of capital. Entrepreneurs below the threshold are unconstrained: their optimal size is zero and they choose to lend their net worth to other entrepreneurs.[1] This economy allows for an aggregate representation akin to that in the standard New Keynesian model, but with endogenous productivity dynamics driven by the evolution of the distribution of capital across heterogeneous entrepreneurs.

[1] This is the tractable limit case of an economy with decreasing returns to scale at the firm level, in which unconstrained firms are optimally very small and the bulk of production is carried out by productive and constrained firms.

Misallocation channel of monetary policy The tractability of the model allows us to analyze monetary policy from both a positive and a normative perspective. Starting with the positive analysis, we first analyze how monetary policy shocks affect capital misallocation. Expansionary monetary policy relaxes the financial constraints disproportionately for firms with high MRPK. This allows them to increase their investment relative to the low-MRPK firms, reducing capital misallocation and increasing measured aggregate TFP.[2] We call this the *capital misallocation channel* of monetary policy.

The fact that monetary policy expansions reduce misallocation might seem in conflict with previous research, such as Reis (2013), Gopinath et al. (2017), or Asriyan et al. (2025), who argue that a decline in real interest rates may fuel capital misallocation in real economies. We show that there is no such conflict: our model also delivers a decline in TFP in response to a fall in real rates due to a negative demand shock when prices are flexible. The difference in the behavior of misallocation compared to a monetary policy shock is due to the different natural rate dynamics: though in response to both shocks the *real* rate drops, the *natural* rate falls only for the demand shock, remaining constant for the monetary policy shock.[3] Just observing the dynamics of real interest rates is not sufficient to infer whether misallocation will improve or worsen. In the presence of nominal rigidities, it is the joint dynamics of real and natural rates that matter.

Optimal monetary policy Next, we investigate the implications of this capital misallocation channel for the optimal conduct of monetary policy by analyzing the Ramsey problem of a benevolent central bank. This is a technical challenge, as the state space of the model includes the distribution of net worth across firms, an infinite-dimensional object. We overcome it by introducing a new algorithm to compute optimal policy problems in the presence of non-trivial heterogeneity.

We first study optimal policy in the absence of shocks. The steady state of the Ramsey problem features zero-inflation, as in the standard complete-markets New Keynesian

[2] The fact that TFP increases after a positive monetary policy shock has been previously documented by Evans (1992), Christiano et al. (2005), Garga and Singh (2021), Jordà et al. (2020), Moran and Queralto (2018), Meier and Reinelt (2024) or Baqaee et al. (2024), among others. While these authors propose complementary mechanisms such as R&D, hysteresis effects, or markup heterogeneity to account for it, which we abstract from, our findings suggest that capital misallocation also plays a significant role.

[3] The natural rate is defined as the real interest rate in the counterfactual flexible price economy.

model, which is nested in our framework.[4] However, capital misallocation renders the problem *time-inconsistent,* as the central bank is tempted to exploit the capital misallocation channel. This time inconsistency manifests itself as a time-0 problem: when starting from the Ramsey zero-inflation steady state, the central bank has an incentive to engineer a temporary monetary expansion in the short run while still committing to price stability in the long run. This strategy allows the central bank to temporarily improve capital allocation, and to increase TFP, even if it means tolerating positive inflation during a certain period of time. We find that this time-0 problem is quantitatively more important than the standard one in the New Keynesian literature—namely, the temptation to exploit the short-run trade-off between inflation and the output gap generated by steady-state markups.

Next, we analyze optimal monetary policy from a "timeless perspective" (Woodford, 2003), in which the central bank respects the commitments that it has optimally made at a date far in the past. This allows us to study systematic changes in monetary policy in response to shocks. To eliminate the time-0 problem and restore time consistency, we augment the Ramsey problem by adding a timeless penalty to the central bank's objective—a term that enforces the central bank's steady-state pre-commitments (Marcet and Marimon, 2019; Dávila and Schaab, 2023).

We start with the analysis of time-preference and TFP shocks. The optimal policy response to these shocks generates only small deviations from price stability. In a simplified version of the model featuring i.i.d. shocks as in Itskhoki and Moll (2019), we can furthermore show analytically that exact price stability is optimal. There is thus no meaningful trade-off between price stability and managing misallocation for these shocks, just as the standard New Keynesian model features no trade-off between price stability and aggregate demand management (this is commonly known as the "divine coincidence", Blanchard and Gali, 2007). This is so because the central bank has no incentive to systematically exploit the capital misallocation channel in response to TFP or time-preference shocks. This logic is familiar from the New Keynesian model with a distorted steady state (Benigno and Woodford, 2005). This implies that, even if the central bank can affect TFP through its monetary policy, it finds optimal *not* to do so in a systematic way, and rather sticks to price stability.

Finally, we analyze the optimal timeless response to cost-push shocks. Under com-

[4] Our model nests the complete-markets model as a particular case in which the borrowing constraint is arbitrarily loose, or in which entrepreneurs' productivity levels are arbitrarily similar.

plete markets, cost-push shocks break the divine coincidence and create an inflation–output trade-off, prompting the central bank to contract output to curb inflation. However, in our baseline with capital misallocation, monetary policy also affects TFP, leading the central bank to tolerate higher inflation to stabilize both output and productivity in the short run. Nevertheless, in both cases long-run price level targeting is optimal, that is the central bank eventually brings the price level back to its original level.

Related literature This paper contributes to three strands of the literature. First, it relates to the extensive literature on capital misallocation, and the different channels that may affect it, such as Hsieh and Klenow (2009) or Midrigan and Xu (2014)—see Restuccia and Rogerson (2017) for a review on this literature. Our framework builds on Moll (2014), whose tractable heterogeneous producer model we augment with a New Keynesian monetary block to understand how monetary policy affects misallocation.[5]

By incorporating monetary policy, we also connect to a second strand of the literature that studies how firm heterogeneity shapes the transmission of monetary policy. Within this literature, a growing body of work shows that financial frictions generate heterogeneous investment responses to monetary policy. In particular, firms' investment is found to be more responsive to monetary policy shocks when their default risk is low (Ottonello and Winberry, 2020), when they have high leverage or fewer liquid assets (Jeenas, 2023), when they are young and do not pay dividends (Cloyne et al., 2023), when their excess bond premia are low (Ferreira et al., 2023), or when a larger share of debt matures (Jungherr et al., 2024). This paper contributes to these studies in two ways. First, we highlight that, in the presence of financial frictions, the investment response to monetary policy varies across firms along the productivity distribution.[6] Second, we move beyond the predominantly positive and empirical focus on firm-level behavior and highlight the normative implications of this heterogeneous response. In particular, we examine how monetary policy affects capital misallocation and thus aggregate productivity, and how this channel in turn influences the optimal conduct of monetary policy.[7] Related work emphasizes complementary mechanisms

[5] Buera and Nicolini (2020) employ a discrete-time version of Moll (2014) with cash-in-advance constraints to analyze the impact of different monetary and fiscal policies after a credit crunch.

[6] In complementary empirical work, Albrizio et al. (2025) analyze the impact of monetary policy on capital misallocation in more detail, both from an intensive and extensive margin. They find that the intensive margin is the main reallocation channel and that firms' investment sensitivity to monetary policy is driven by their MRPK rather than their age, leverage, or cash.

[7] Monetary policy generally also affects capital misallocation in the models cited above, although

through which firm heterogeneity matters for monetary policy and the allocation of resources more generally, including heterogeneity in markups and entry-exit (e.g. Meier and Reinelt, 2024, Bilbiie et al., 2014, Baqaee et al., 2024, or Hamano and Zanetti, 2022), risk-taking (David and Zeke, 2021), firm-level productivity trends (Adam and Weber, 2019), or the importance of the price elasticity of investment (Koby and Wolf, 2020).[8]

Finally, this paper contributes to the literature analyzing optimal monetary policy problems in heterogeneous-agent economies. A number of recent papers, such as Bhandari et al. (2021), Acharya et al. (2023), Bilbiie and Ragot (2020), Le Grand et al. (2025), Mckay and Wolf (2022), or Auclert et al. (2024) propose different approaches to solve these problems. Similar to Nuño and Thomas (2022), Smirnov (2022), or Dávila and Schaab (2023), we set up the problem as one of infinite-dimensional optimal control. We propose a new, simple and broadly applicable algorithm to solve these kinds of problems, which leverages the computational advantages of continuous time. We also differ in focus: while these papers analyze heterogeneous households, to the best of our knowledge this is the first paper to analyze optimal monetary policy in a model with heterogeneous firms.[9]

The structure of the paper is as follows. Section 2 presents the model, which we calibrate in Section 3. Section 4 analyzes the drivers and dynamics of misallocation. Section 5 studies optimal monetary policy. Finally, Section 6 concludes.

2 Model

We develop a closed-economy New Keynesian model with financial frictions and heterogeneous firms. Time is continuous and there is no aggregate uncertainty. The economy is populated by six types of agents: households, the central bank, entrepreneurs that operate input-good firms, capital producer, retailers and final good producers. En-

this is not the focus of their work.

[8] Although some of these studies also link monetary policy to misallocation, they emphasize different transmission channels. For example, Hamano and Zanetti (2022) show that contractionary monetary policy discourages firm entry and protects incumbents from competition, thereby reducing aggregate productivity. Our framework abstracts from the entry–exit (extensive) margin and instead focuses on the expansion and contraction of incumbent firms (intensive margin). The two mechanisms can be, however, complementary.

[9] Martin-Baillon (2021) studies Ramsey policy with heterogeneous firms, although her focus is on fiscal policy rather than monetary policy.

trepreneurs are heterogeneous in their net worth and productivity. They combine capital and labor to produce the input good. This input is sold to imperfectly competitive retail firms, which produce differentiated goods and face price rigidities. A perfectly competitive final-goods producer aggregates the differentiated retail varieties into a homogeneous final good. The retail and final-goods sectors follow the standard structure of New Keynesian models. There is a fixed aggregate capital stock. Although there is no aggregate uncertainty, we consider four types of MIT shocks—that is, one-time surprise shocks no one expects, with a known and commonly understood path from the time of impact onward. Specifically, we analyze: (i) a monetary policy shock, i.e., a shock to the nominal interest rate i_t; (ii) an exogenous TFP shock, denoted by ς_t; (iii) a time-preference shock, that is, a disturbance to the household's discount rate ρ_t^h; and (iv) a cost-push shock, modeled as a shock to the subsidy on the input good τ_t.

2.1 Heterogeneous input-good firms

The heterogeneous-firm block is based on Moll (2014). There is a continuum of entrepreneurs. Each entrepreneur owns some net worth, which they hold in units of capital. They can use this capital for production in their own input-good producing firm—firm for short—or lend it to other entrepreneurs. Similar to Gertler and Karadi (2011), we assume that entrepreneurs are atomistic members of the representative household, to whom they may transfer dividends.[10]

Entrepreneurs are heterogeneous in two dimensions: their net worth a_t and their idiosyncratic productivity z_t.[11] Each entrepreneur owns a constant returns to scale (CRS) technology which uses capital k_t and labor l_t to produce the homogeneous input good y_t:

$$y_t = f_t(z_t, k_t, l_t) = (z_t \varsigma_t k_t)^\alpha (l_t)^{1-\alpha}, \tag{1}$$

where ς_t is the (exogenous) common component of productivity and $\alpha \in (0,1)$ the

[10]This assumption is the only relevant difference between the real side of our model and the model of Moll (2014). We consider it to avoid having to deal with redistributive issues between households and entrepreneurs when analyzing optimal monetary policy. Both models produce linear dividend policies, so they are equivalent from a positive perspective.

[11]For notational simplicity, we use x_t instead of $x(t)$ for the variables depending on time. Furthermore, we suppress the input-good firm's index.

capital share. Idiosyncratic productivity z_t follows a diffusion process,

$$dz_t = \mu(z_t)dt + \sigma(z_t)dW_t, \tag{2}$$

where $\mu(z)$ is the drift and $\sigma(z)$ the diffusion of the process.[12]

Entrepreneurs can use their net worth to produce in their firm with their own technology, or lend it to firms owned by other entrepreneurs. Firms employ labor l_t, which they hire at the real wage w_t, and capital k_t, which is the sum of the entrepreneur's net worth (a_t) and what the firm borrows ($b_t = k_t - a_t$) at the real cost of capital R_t. Capital is borrowed from the agents which save, i.e. both households and lending entrepreneurs.[13]

Firms sell the input good at the real price $m_t = p_t^y/P_t$, which is the inverse of the gross markup of retail products over input goods, being p_t^y the nominal price of the input good and P_t the price of the final good, i.e. the numeraire. Entrepreneurs' income flow consists of their firms' profits and the return on their capital (net worth). Entrepreneurs use these funds to distribute (non-negative) dividends d_t to the household and to invest in additional capital at the real price q_t. Capital depreciates at rate δ. An entrepreneur's flow budget constraint can be expressed as follows

$$\underbrace{q_t \dot{a}_t}_{\text{Net investment}} + \underbrace{d_t}_{\text{Dividends}} = \left[\underbrace{m_t f_t(z_t, k_t, l_t) - w_t l_t - R_t k_t}_{\text{Firm's profits}} + \underbrace{(R_t/q_t - \delta)}_{\text{Return on net worth}} q_t a_t \right]. \tag{3}$$

Firms face a collateral constraint, such that the value of capital used in production cannot exceed $\gamma > 1$ of their net worth,[14]

$$q_t k_t \leq \gamma q_t a_t. \tag{4}$$

[12] The process is bounded with reflective barriers.

[13] Since debt contracts are instantaneous and in units of capital, firms' balance sheets are not exposed to Fisherian debt deflation or financial accelerator effects (Bernanke et al., 1999). Asriyan et al. (2025) include the latter. This keeps the model tractable. Allowing for such effects would amplify the effect of monetary policy shocks discussed later.

[14] Assuming alternatively that firms' borrowing is constrained to a multiple of their earnings (Lian and Ma, 2020) would amplify the effect of monetary policy shocks discussed later on. The increase in high-productivity firms' profits, that (as we explain below) drives the positive impact of expansionary monetary policy on TFP, would relax the constraint of high-productivity firms and improve the allocation of capital further.

Entrepreneurs retire and return to the household according to an exogenous Poisson process with arrival rate η. Upon retirement they pay all their net worth, valued $q_t a_t$, to the household, and they are replaced by a new entrepreneur with the same productivity level. Entrepreneurs maximize the discounted flow of dividends, which is given by

$$V_0(z,a) = \max_{k_t,l_t,d_t} \int_0^{\infty} e^{-\eta t} \Lambda_{0,t} \left(\underbrace{d_t}_{\text{Dividends}} + \eta \underbrace{q_t a_t}_{\text{Terminal value}} \right) dt, \tag{5}$$

subject to the budget constraint (3), the collateral constraint (4), and the productivity process (2).[15] Future profits are discounted by the household's stochastic discount factor $\Lambda_{0,t}$. Below we show that $\Lambda_{0,t} = e^{-\int_0^t r_s ds}$, where r_t is the real interest rate.

We can split the entrepreneurs' problem into two parts: a static profit maximization problem and a dynamic dividend-distribution problem. First, entrepreneurs maximize firm profits given their productivity and net worth,

$$\max_{k_t,l_t} \left\{ m_t f_t(z_t,k_t,l_t) - w_t l_t - R_t k_t \right\}, \tag{6}$$

subject to the collateral constraint (4).

Since the production function has constant returns to scale, entrepreneurs find it optimal to operate their firm at the maximum scale defined by the collateral constraint whenever their idiosyncratic productivity is high enough, that is whenever z exceeds a certain threshold level z_t^*. Else the optimal size of the firm is $k^*(z) = 0$, because they cannot run a profitable firm given their low productivity. In that case the borrowing constraint does not bind and the entrepreneur rents out her capital. From now on, we refer to the two groups of entrepreneurs as 'constrained' and 'unconstrained'. That is, firms' demand for capital and labor is:

$$k_t(z_t,a_t) = \begin{cases} \gamma a_t, & \text{if } z_t \geq z_t^*, \\ 0, & \text{if } z_t < z_t^*, \end{cases} \tag{7}$$

[15]In reality, other sources of misallocation other than financial frictions—such as uninsurable risk (Boar et al., 2025)—may be equally or even more important. We view this paper as a first step toward understanding how firm heterogeneity affects the optimal conduct of monetary policy. Our aim is for this framework to serve as a useful benchmark that can be extended in future work to incorporate additional sources of heterogeneity and frictions, to further explore how firm heterogeneity shapes the design and transmission of optimal monetary policy.

$$l_t(z_t, a_t) = \left(\frac{(1-\alpha)m_t}{w_t}\right)^{1/\alpha} \varsigma_t z_t k_t(z_t, a_t). \tag{8}$$

Firm's profits are then given by

$$\Phi_t(z_t, a_t) = \max\{z_t \varsigma_t \varphi_t - R_t, 0\}\, \gamma a_t, \quad \text{where} \quad \varphi_t = \alpha \left(\frac{1-\alpha}{w_t}\right)^{\frac{1-\alpha}{\alpha}} m_t^{\frac{1}{\alpha}}. \tag{9}$$

Note that the term $z_t \varsigma_t \varphi_t$ is simply the marginal revenue product of capital (MRPK) of the firm with idiosyncratic productivity z_t. The productivity cutoff, above which firms are profitable, is given by

$$z_t^* \varsigma_t \varphi_t = R_t, \tag{10}$$

This expression tells us that the MRPK of the marginal firm equals the marginal cost of capital. Any firm with $z > z^*$ thus makes profits over and above the cost of capital. These profits arise despite perfect competition, because the borrowing constraint binds.

Note that factor demands and profits are linear in net worth. This is a consequence of the CRS technology and makes the model significantly more tractable. As discussed by Moll (2014), the assumption of CRS in firms' production function (1) can be seen as the limiting case of decreasing returns to scale (DRS), $y_t = \left[(z_t \varsigma_t k_t)^\alpha (l_t)^{1-\alpha}\right]^\nu$, $\nu < 1$, when $\nu \to 1$. In the case with DRS, there is a threshold $z^*(a)$ which depends on net worth, such that the firms with $z \leq z^*(a)$ are unconstrained and produce at their optimal level $(k^*(z))$, whereas those with $z > z^*(a)$ are constrained. When this threshold increases, previously marginally constrained firms become unconstrained and reduce their capital stock below the maximum implied by the constraint. When $\nu \to 1$, the optimal size of low-productivity firms, and hence their production, are very small, $k^*(z), y^*(z) \to 0$. Therefore our model should be understood as the tractable limit of the more realistic DRS case.[16] This highlights a crucial point regarding the interpretation of the model: the threshold z^* does not capture entry/exit, but rather the limiting case of expansions and contractions of the optimal size of active firms. Entry and exit of firms in this model is exogenous, and given by the exogenous retirement rate η.

Second, entrepreneurs choose the dividends d_t that they pay to the household. The solution to this problem is derived in Appendix B.1. There we show how entrepreneurs never distribute dividends ($d_t = 0$) until retirement, when they bring all their net worth

[16] Ferreira et al. (2023) find that financially constrained firms are found across the entire firm-size distribution.

home to the household. The intuition is the following. The return on one unit of capital in the hands of the entrepreneur is *at least* $(R_t - \delta q_t + \dot{q}_t)/q_t$, while for the household the return on this unit of capital is *exactly* $(R_t - \delta q_t + \dot{q}_t)/q_t$. It is thus always worthwhile for entrepreneurs to keep their funds in the firm. To keep things simple, as in Gertler and Karadi (2011) we assume the representative household uses a fraction $\psi \in (0,1)$ of these dividends to finance the net worth of the new entrepreneurs, so net dividends are $(1-\psi)$ of the net worth of retiring entrepreneurs.

Using (9) and (3), the law of motion of an entrepreneur's net worth can thus be rewritten as

$$\dot{a}_t = \frac{a_t}{q_t}\left[\gamma \max\left\{z_t \varsigma_t \varphi_t - R_t, 0\right\} + R_t - \delta q_t\right]. \tag{11}$$

2.2 Households

There is a representative household, composed of workers and entrepreneurs, that saves in capital D_t and in nominal instantaneous bonds B_t^N. Nominal bonds are in zero net supply. Workers supply labor L_t. The household maximizes

$$W_t = \max_{C_t, L_t, B_t^N, D_t} \int_0^\infty e^{-\int_0^t \rho_s^h ds} u(C_t, L_t) dt \tag{12}$$

$$\begin{aligned} s.t. \quad \dot{D}_t q_t &= \left(R_t - \delta q_t\right) D_t - S_t^N - C_t + w_t L_t + T_t, \\ \dot{B}_t^N &= \left(i_t - \pi_t\right) B_t^N + S_t^N, \end{aligned} \tag{13}$$

where S_t^N is the investment into nominal bonds and T_t denotes the profits received by the household lump sum. The profits are the sum of net dividends received from entrepreneurs ($(1-\psi)\eta q_t A_t$) and the profits of the capital and retail-goods producers (discussed below). The discount rate ρ_t^h is potentially time-varying.

We assume separable CRRA utility, i.e., $u(C_t, L_t) = \frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon \frac{L_t^{1+\vartheta}}{1+\vartheta}$. Solving this problem (see Appendix B.2 for details), we obtain the Euler equation,

$$\frac{\dot{C}_t}{C_t} = \frac{r_t - \rho_t^h}{\zeta}, \tag{14}$$

the labor supply condition

$$w_t = \frac{\Upsilon L_t^{\vartheta}}{C_t^{-\zeta}}, \tag{15}$$

and the Fisher equation

$$r_t = i_t - \pi_t, \tag{16}$$

where, for convenience, we have used the following definition of the real interest rate

$$r_t \equiv \frac{R_t - \delta q_t + \dot{q}_t}{q_t}, \tag{17}$$

which equals the real return on capital adjusted by capital gains and depreciation. Integrating the Euler equation (14), and using the Definition (17), we can verify that the stochastic discount factor $\Lambda_{0,t}$ is

$$\Lambda_{0,t} \equiv e^{-\int_0^t \rho_t^h ds} \frac{u_C\left(C_t\right)}{u_C\left(C_0\right)} = e^{-\int_0^t r_s ds}.$$

2.3 Final good producers

As usual in New Keynesian models, a competitive representative final good producer aggregates the output produced by a continuum of retailers indexed $j \in [0,1]$,

$$Y_t = \left(\int_0^1 y_{j,t}^{\frac{\varepsilon-1}{\varepsilon}} dj\right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{18}$$

where $\varepsilon > 1$ is the elasticity of substitution across goods. Cost minimization implies

$$y_{j,t}\left(p_{j,t}\right) = \left(\frac{p_{j,t}}{P_t}\right)^{-\varepsilon} Y_t, \text{ where } P_t = \left(\int_0^1 p_{j,t}^{1-\varepsilon} dj\right)^{\frac{1}{1-\varepsilon}}.$$

2.4 Retailers

We differentiate between heterogeneous input-good firms and retailers.[17] Retailers purchase input goods from the input-good firms, differentiate them and sell them to final

[17] Distinguishing between heterogeneous input-good firms and retailers is standard practice in previous New Keynesian models featuring firm heterogeneity and nominal rigidities, such as Ottonello and Winberry (2020) or Jeenas (2023). Besides greater tractability, it avoids the possibly implausible countercyclical behavior of New Keynesian markups interfering with our mechanism, which we see as an advantage. It also does justice to the fact that retail prices are significantly more sticky than intermediate goods prices (for Europe see Alvarez et al., 2006, Gautier et al., 2023).

good producers. Retailers operate under monopolistic competition. Each retailer j chooses the nominal sales price $p_{j,t}$ to maximize real profits subject to price adjustment costs as in Rotemberg (1982), taking as given the demand curve $y_{j,t}(p_{j,t})$ and the price of input goods, p_t^y. The adjustment costs are quadratic in the rate of price change $\dot{p}_{j,t}/p_{j,t}$ and expressed as a fraction of output Y_t,

$$\Theta_t\left(\frac{\dot{p}_{j,t}}{p_{j,t}}\right) = \frac{\theta}{2}\left(\frac{\dot{p}_{j,t}}{p_{j,t}}\right)^2 Y_t, \tag{19}$$

where $\theta > 0$. We assume the government pays a proportional subsidy τ_t on the input good, so that the net real price for the retailer is $\tilde{m}_t = m_t(1-\tau_t)$. This subsidy is financed by a lump-sum tax on the retailers Ψ_t, and it is set so as to exactly offset the distortions caused by imperfect competition in steady state, $\tau = \frac{1}{\varepsilon}$, as it is common in the optimal policy literature. We allow for temporary shocks to the subsidy. These shocks operate like cost-push shocks, as we discuss in Section 5. Suppressing notational dependence on j, each retailer chooses $\{p_t\}_{t\geq 0}$ to maximize the expected profit stream, discounted at the stochastic discount factor of the household,

$$\int_0^\infty \Lambda_{0,t}\left[\left(\frac{p_t}{P_t} - \tilde{m}_t\right)\left(\frac{p_t}{P_t}\right)^{-\varepsilon} Y_t - \Psi_t - \Theta_t\left(\frac{\dot{p}_t}{p_t}\right)\right] dt. \tag{20}$$

The symmetric solution to the pricing problem yields the New Keynesian Phillips curve (see Appendix B.3), which is given by

$$\left(r_t - \frac{\dot{Y}_t}{Y_t}\right)\pi_t = \frac{\varepsilon}{\theta}\left(\tilde{m}_t - m^*\right) + \dot{\pi}_t, \qquad m^* = \frac{\varepsilon - 1}{\varepsilon}. \tag{21}$$

where π_t denotes the inflation rate $\pi_t = \dot{P}_t/P_t$.

2.5 Capital producer

Capital is produced by a representative capital producer owned by the household. The capital producer inelastically supplies δK units of capital, where K is the total amount of capital and δ its depreciation rate, and sells them at price q_t. As input, it uses δK units of the final good. This assumption is to be understood as the analytically convenient limit of a capital producer with a convex cost function $\Xi(\iota)$, where $\Xi'(\iota) \to$

∞ for any investment rate $\iota \neq \delta$, and it implies that aggregate capital stock is fixed.[18]

2.6 Distribution

As previously explained, we assume that, for each entrepreneur retiring to the household, a new entrepreneur enters operating the same technology, that is, with the same productivity level. This new entrepreneur receives a startup capital stock from the household in a lump-sum fashion, equal to a fraction $\psi < 1$ of the net worth of the entrepreneur she replaces. Let $G_t(z,a)$ be the joint distribution of net worth and productivity. The evolution of its density $g_t(z,a)$ is given by the Kolmogorov Forward equation

$$\frac{\partial g_t(z,a)}{\partial t} = \underbrace{-\frac{\partial}{\partial a}[g_t(z,a)s_t(z)a]}_{\text{Retained earnings}} \underbrace{-\frac{\partial}{\partial z}[g_t(z,a)\mu(z)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[g_t(z,a)\sigma^2(z)]}_{\text{Productivity changing randomly}} \underbrace{-\eta g_t(z,a)}_{\text{Entrepreneurs retiring}} \underbrace{+\frac{\eta}{\psi} g_t(z,\frac{a}{\psi})}_{\text{Entrepreneurs entering}}, \quad (22)$$

where $s_t(z)$ is the entrepreneurs' investment rate from (11)

$$s_t(z) \equiv \frac{\dot{a}_t}{a_t} = \frac{1}{q_t}(\underbrace{\gamma \max\{z\varsigma_t\varphi_t - R_t, 0\}}_{\text{Profit rate from operating the firm}} + R_t - \delta q_t), \quad (23)$$

and $1/\psi g_t(z, a/\psi)$ is the density of new entrepreneurs entering.

Using this two-dimensional distribution we can define the one-dimensional distribution of *net-worth shares* as $\omega_t(z) \equiv \frac{1}{A_t}\int_0^\infty a g_t(z,a)da$, where we have defined aggregate net worth as $A_t = \int a dG_t(z,a)$. This distribution measures the share of net worth held by entrepreneurs with productivity z. Due to the linearity of the firms choices in a, it contains all the relevant information in a more compact form. Given this definition and the structure of the problem, net-worth shares are non-negative, continuous, once differentiable everywhere and they integrate up to 1. The law of motion of net worth

[18] Endogenous capital accumulation by itself introduces departures from standard New Keynesian dynamics—see Faia (2008). Given that we focus exclusively on the allocation of capital, we abstract from endogenous aggregate capital dynamics.

shares is given by (see Appendix B.4)

$$\frac{\partial \omega_t(z)}{\partial t} = \left[s_t(z) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta \right] \omega_t(z) - \frac{\partial}{\partial z}\left[\mu(z)\omega_t(z)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)\omega_t(z)\right]. \quad (24)$$

2.7 Market Clearing and Aggregation

Market clearing. In equilibrium, the aggregate capital used in production has to equal the amount of capital available: $K = \int k_t(z,a)dG_t(z,a)$. Define aggregate net debt as $B_t = \int b_t(z,a)dG_t(z,a)$. Since the capital borrowed by an individual entrepreneur equals that used in production minus its net worth $b_t = k_t - a_t$, we have that

$$K = A_t + B_t. \quad (25)$$

Asset market clearing requires that net borrowing of entrepreneurs B_t equals net savings of the household D_t,

$$B_t = D_t. \quad (26)$$

Let $\Omega(z)$ be the cumulative distribution of net-worth shares, i.e. $\Omega_t(z) = \int_0^z \omega_t(x)\,dx$. By combining equations (25), (26), aggregating capital used by firms (7), and solving for A_t, we obtain

$$A_t = \frac{D_t}{\gamma(1-\Omega_t(z_t^*)) - 1}. \quad (27)$$

Labor market clearing implies

$$L_t = \int_0^\infty l_t(z,a)dG_t(z,a). \quad (28)$$

Aggregation. Aggregating up, one can express output as a function of aggregate factors and aggregate TFP

$$Y_t = Z_t K^\alpha L_t^{1-\alpha}, \quad (29)$$

where aggregate TFP Z_t is an endogenous variable given by

$$Z_t = \left(\varsigma_t \mathbb{E}_{\omega_t(\cdot)}\left[z \mid z > z_t^*\right]\right)^\alpha = \left(\varsigma_t \frac{\int_{z_t^*}^\infty x\omega_t(x)\,dx}{1-\Omega_t(z_t^*)}\right)^\alpha. \quad (30)$$

This highlights that, in terms of output, the model is isomorphic to a standard representative-agent New Keynesian model with constant capital and TFP Z_t. TFP is endogenous and evolves over time and, as we discuss below, its evolution depends on factor prices.

Note that TFP Z_t serves as a measure of misallocation. The financial frictions faced by entrepreneurs imply that capital is not optimally allocated. The entrepreneur operating the most productive firm does not have enough net worth to operate the whole capital stock, hence less productive firms operate as well, which is suboptimal and reduces TFP. Thus the more misallocated capital is, the lower is TFP.

Factor prices are

$$w_t = (1-\alpha) m_t Z_t K^{\alpha} L_t{}^{-\alpha}, \tag{31}$$

$$R_t = \alpha m_t Z_t K^{\alpha-1} L_t{}^{1-\alpha} \frac{z_t^*}{\mathbb{E}_{\omega_t(\cdot)}\left[z \mid z > z_t^*\right]}. \tag{32}$$

Finally, the law of motion of the aggregate net worth of entrepreneurs is given by

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}\left[\gamma\left(1-\Omega_t(z_t^*)\right)\left(\alpha m_t Z_t K^{\alpha-1} L_t{}^{1-\alpha} - R_t\right) + R_t - \delta q_t - q_t(1-\psi)\eta\right]. \tag{33}$$

Appendix B.5 derives these aggregate formulae step by step.

2.8 Central Bank

The central bank controls the nominal interest rate i_t on nominal bonds held by households. For the positive analysis in Section 4, we assume that the central bank sets the nominal rate according to a standard Taylor rule of the form

$$di = -\upsilon\left(i_t - \left(\rho^h + \phi\left(\pi_t - \bar{\pi}\right) + \bar{\pi}\right)\right) dt, \tag{34}$$

where $\bar{\pi}$ is the inflation target, ϕ is the sensitivity to inflation deviations and υ determines the persistence of the policy rule. For the normative analysis in Section 5 we assume that the central bank follows the Ramsey-optimal policy.

2.9 Competitive equilibrium

Next, we define the competitive equilibrium. Appendix B.6 reports the corresponding set of equilibrium conditions.

Definition. Given initial conditions A_0, D_0 and $\omega_0(z)$, an interest rate path $\{i_t\}_{t\geq0}$ and exogenous paths $\{\rho_t^h, \tau_t, \varsigma_t\}_{t\geq0}$, a competitive equilibrium is a sequence of prices $\{q_t, w_t, r_t, R_t, m_t, \pi_t\}_{t\geq0}$, aggregates $\{C_t, Y_t, L_t, A_t, D_t, Z_t\}_{t\geq0}$, cutoff values $\{z_t^*\}_{t\geq0}$ and net-worth shares $\{\omega_t(z)\}_{t\geq0}$ such that: (i) Households, input firms, retailers and the final good producer solve their respective problems; (ii) Net worth shares evolve according to the Kolmogorov forward equation stated in equation (24); (iii) Markets clear as stated in Section 2.7.

3 Numerical solution and calibration

Numerical algorithm. We solve the model numerically using a new method, described in Appendix F. It combines a discretization of the model using an upwind finite-difference method similar to the one in Achdou et al. (2021) with a Newton algorithm that computes non-linear transitional dynamics in a single loop.[19]

Calibration. Table 1 summarizes our calibration. We calibrate the parameters of the heterogeneous firms block to match data on Spanish firms. The entrepreneur's exit rate (η) is set to 10 percent, in line with the average exit rate 2007-2020 in the data from the Spanish Statistical Institute (INE). The other parameters of the heterogeneous-firm block are disciplined using detailed firm-level panel data from the quasi-universe of Spanish firms between 2000 and 2016 from the *Central de Balances Integrada*.[20] The fraction of assets of exiting entrepreneurs reinvested (ψ) is 0.1, so that new entrants account for 1 percent of the total capital stock, in line with the data. The borrowing constraint parameter γ is 1.56, implying that entrepreneurs can borrow up to 56% of their net worth, or 36% of their total assets, which targets the aggregate debt to total assets ratio in the data. We assume that individual productivity z follows an Ornstein-

[19] Our method differs from the linear approximation approaches in Reiter (2009), Winberry (2018) or Ahn et al. (2018) since we compute the *nonlinear* responses to MIT shocks. However, the two approaches coincide for small shocks (Boppart et al., 2018). Our method solves for the same solution as the nonlinear version of the sequence space Jacobian approach by Auclert et al. (2021) in a single loop.

[20] This data set includes not only large firms with access to stock and bond markets, but also medium and small firms more reliant on bank credit and internal financing. This contrasts with most papers in this literature, which use data from publicly traded firms (e.g. Compustat). Those firms can potentially behave very differently from the rest of firms in the economy, as documented for example by Caglio et al. (2021). Appendix A.1 details the data definition. Our key variable of interest is firms' MRPK, which we proxy by value added over capital following the literature (see for instance Bau and Matray, 2023).

Uhlenbeck process in logs,[21] with a reflective lower (upper) barrier at some value close to 0 (very high value).[22]

$$d\log(z) = -\varpi_z \log(z)dt + \sigma_z dW_t. \tag{35}$$

We estimate this process using our firm panel data set, as explained in Appendix A.3. The estimate for ϖ_z corresponds to an annual persistence of 0.83, and the annual volatility of the shock is estimated to be 0.73.

Table 1: **Calibration**

	Parameter	**Value**	**Source/target**
η	Firms' death rate	0.1	Average exit rate
ψ	Fraction firms' assets at entry	0.1	Capital of firms younger than 1 year / All firms capital
γ	Borrowing constraint parameter	1.56	Debt / Total Assets firms
ϖ_z	Mean reverting parameter	0.19	Estimate based on firm level data
σ_z	Volatility of the shock	0.73	Estimate based on firm level data
ρ^h	Household's discount rate	0.01	1%
α	Capital share in production function	0.35	Gopinath et al. (2017)
δ	Capital depreciation rate	0.06	Gopinath et al. (2017)
ζ	Inverse intertemporal elasticity of substitution	1	Log utility in consumption
ϑ	Inverse Frisch Elasticity	1	Kaplan et al. (2018)
Υ	Constant in disutility of labor	0.71	Normalization $L = 1$
K	Capital stock	15	Normalization $q = 1$
ϵ	Elasticity of substitution retail goods	10	Mark-up of 11%
θ	Price adjustment costs	100	Slope of Phillips curve of 0.1 as in Kaplan et al. (2018)
$\bar{\pi}$	Inflation target	0	Standard
ϕ	Slope Taylor rule	1.5	Standard
υ	Persistence Taylor rule	0.2	Standard

The conventional macro parameters are set to standard values. The rate of time-preference of the household ρ^h is 0.01, which targets an average real rate of return of 1 percent. The capital share α is set at 0.35 and the capital depreciation rate δ at 0.06, as in Gopinath et al. (2017). We assume log-utility in consumption ($\zeta = 1$) and the inverse Frisch elasticity ϑ is also set to 1, standard values in the literature. We set the constant multiplying the disutility of labor Υ and the quantity of capital K such that aggregate labor supply and the price of capital in steady state are normalized to one.

Regarding the New Keynesian block, the elasticity of substitution of retail goods ϵ is set to 10, so that the steady-state mark-up is $1/(\epsilon - 1) = 0.11$. The Rotemberg cost

[21] By Ito's lemma, this implies that z in levels follows the diffusion process $dz = \mu(z)dt + \sigma(z)dW_t$, where $\mu(z) = z\left(-\varpi_z \log z + \frac{\sigma_z^2}{2}\right)$ and $\sigma(z) = \sigma_z z$.

[22] We truncate the process for z at 48. This corresponds to truncating the MRPK distribution at the same level as in the data.

parameter θ is set to 100, so that the slope of the Phillips curve is $\epsilon/\theta = 0.1$ as in Kaplan et al. (2018). The Taylor rule parameters take the following standard values: $\bar{\pi} = 0$, $\phi = 1.5$ and $\upsilon = 0.2$. Appendix A.2 shows that the model generates a steady-state MRPK distribution in line with the data.

Besides the idiosyncratic shocks, the economy is also affected by aggregate MIT shocks.[23] These can be shocks to time-preference (ρ_t^h), aggregate technology (ς_t), cost-push shocks (τ_t), or monetary policy (i_t). The first three shocks follow a mean-reverting Ornstein-Uhlenbeck process with an annual persistence of 0.8. Only the monetary policy shock has no persistence, because the policy rule itself is persistent.

4 Monetary policy and capital misallocation

In this Section, we analyze the links between monetary policy and capital misallocation. To this end, we first delve into the theoretical mechanisms through which changes in equilibrium prices affect misallocation, to then analyze the response of misallocation in general equilibrium to a monetary policy and a time-preference shock. Understanding these mechanisms is key to understanding the optimal monetary policy explained in the following Section 5.

4.1 The capital misallocation channel of monetary policy

As Section 2 highlighted, misallocation and thus TFP is endogenous and evolves over time. Misallocation is driven by the investment dynamics within the heterogeneous input-goods firms block of the model, which in turn depend on the factor prices determined in general equilibrium.[24] One can thus think about the heterogeneous firm sector as a model block that translates (sequences of) prices into (a sequence of) TFP.

As discussed above, by equation (30) and abstracting for now from the exogenous component of TFP ς_t, TFP depends on the allocation of capital across entrepreneurs:

$$Z_t = \left(\mathbb{E}_{\omega_t(\cdot)}\left[z \mid z > z_t^*\right]\right)^{\alpha}. \tag{36}$$

[23] An MIT shock is defined as a zero-probability event, that is, an unexpected shock that happens once, and after arrival, its dynamics are perfectly anticipated.

[24] Of course, this channel is not specific to the precise model outlined in this paper. Any model featuring frictions that generate heterogeneous investment responses to monetary policy will imply changes in capital allocation across firms and, as a result, endogenous movements in aggregate TFP.

That is, TFP is the capital-weighted average of firms' idiosyncratic productivity. TFP thus depends on the mass of the net-worth distribution, $\omega_t(\cdot)$, above the productivity threshold, z_t^* (the shaded area in Figure 1). Entrepreneurs below the cutoff z^* are unconstrained, operate at their optimal size $k(z) = 0$, and lend their net worth to constrained entrepreneurs above the cutoff. Equation (36) allows us to identify how changes in equilibrium prices affect aggregate TFP in this economy (i) by changing the *net-worth distribution*, $\omega_t(\cdot)$; and (ii) by changing the *productivity-threshold* z_t^*. We now explore these two margins in isolation.

We start by analyzing the case in which the dynamics of TFP are driven purely by changes in the net-worth distribution, which happens when the cutoff z_t^* is constant. In this case, the *excess investment rate* is key for the dynamics of TFP. We define the excess investment rate as the ratio of profits over net worth

$$\tilde{\Phi}_t(z) \equiv \frac{\Phi_t(z, a_t)}{q_t a_t} \quad = \quad \max\left\{\frac{\gamma\varphi_t}{q_t}\left(z - z_t^*\right), 0\right\}, \tag{37}$$

where $\tilde{\Phi}_t(z)$ is the return on net worth that a firm with MRPK $\varphi_t z$ makes *in excess of* the cost of capital R_t. Since entrepreneurs reinvest all profits, $\tilde{\Phi}(z)$ also describes the speed at which the net worth of an entrepreneur with productivity z grows *in excess of* the growth rate of the unconstrained entrepreneurs with productivity $z \leq z_t^*$.[25]

Proposition 1. ***(TFP response to the slope)***. *Conditional on a constant cutoff z^*, the dynamics of Z_t are fully determined by the slope of the excess investment rate, $\frac{\gamma\varphi_t}{q_t}$. An increase in $\frac{\gamma\varphi_t}{q_t}$ leads to an increase in the growth rate of TFP through changes in the net-worth distribution:*

$$\frac{\partial}{\partial\left(\frac{\gamma\varphi_t}{q_t}\right)} \left.\frac{d\log Z_t}{dt}\right|_{z^*} = \frac{\int_{z^*}^{\infty} z^2\omega_t(z)}{\int_{z^*}^{\infty} z\omega_t(z)\,dz} - \frac{\int_{z^*}^{\infty} z\omega_t(z)}{\int_{z^*}^{\infty} \omega_t(z)\,dz} > 0.$$

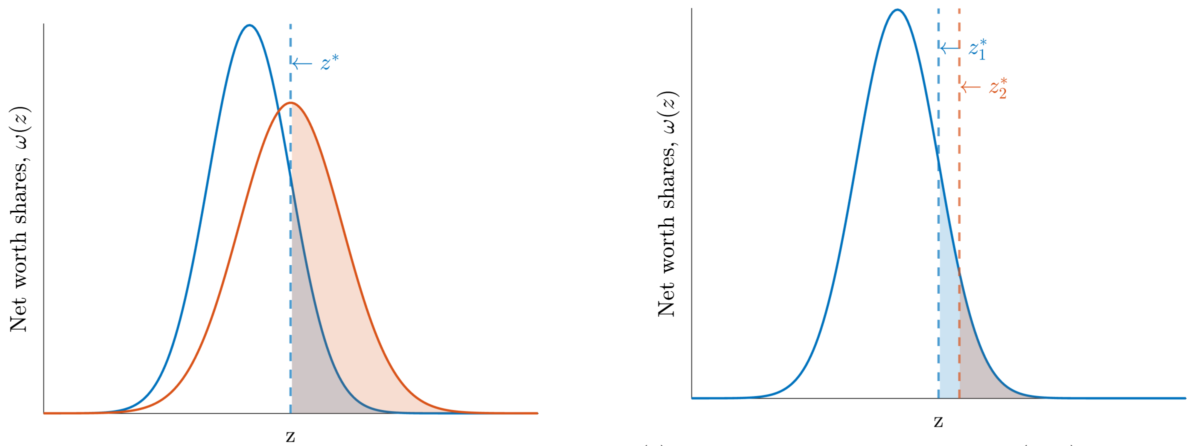
Proof. See Appendix *B.7.1*.

This Proposition states that the slope $\frac{\gamma\varphi_t}{q_t}$ determines how, conditional on a constant cutoff z^*, the net-worth share distribution moves, and hence in which direction TFP

[25] $\tilde{\Phi}(z)$ is also a measure of how constrained a firm is, since $\tilde{\Phi}(z)$ is the Lagrange multiplier of the borrowing constraint in the firm's maximization problem. From the first-order condition of the firm for k, we get that $MRPK_t = R_t + q_t\lambda^{BC}$, where λ^{BC} is the multiplier on the borrowing constraint. Hence, $\gamma\lambda^{BC} = \tilde{\Phi}_t(z)$.

Figure 1: The capital misallocation channel.

(a) Change in the net-worth distribution. (b) Shift in the productivity threshold.



Notes: The figure illustrates the net-worth share distribution $\omega(z)$ and the productivity-threshold z^* (blue). The light blue area is the initial mass of constrained firms. Panel (a) shows the impact of a change in the net-worth distribution. Panel (b) shows the impact of an increase in the threshold (from blue dashed line to orange dashed line). The new mass of constrained firms after the change is depicted by the shaded orange area in both panels.

evolves. $\frac{\gamma\varphi_t}{q_t}$ is a function of prices. If the slope increases, then high-productivity firms' profitability advantage widens, such that they grow faster than low-productivity firms. Thus the net-worth distribution shifts to the right, the allocation of capital improves, and TFP increases, as represented in Panel (a) of Figure 1. Note that, in the model, high-productivity firms have a high MRPK, which is given by $\varphi_t z$. So we can equivalently say that an increase in the relative growth rate of high-MRPK firms improves TFP.

We turn next to the case when the net-worth distribution remains constant and only the cutoff changes in response to price changes.[26] In this case the response of TFP growth depends exclusively on the changes in the growth rate of the cutoff, according to the following Proposition.

Proposition 2. ***(TFP response to the cutoff)***. *Conditional on a constant density $\omega(\cdot)$, the dynamics of Z_t are fully determined by the threshold z_t^*. The partial derivative of TFP growth with respect to the growth rate of the threshold $\frac{dz_t^*}{dt}$ is positive:*

[26]This happens in the limit of iid shocks, that is, the limit as $\varsigma_z \to \infty$, as discussed in Itskhoki and Moll (2019).

$$\frac{\partial}{\partial\left(\frac{dz_t^*}{dt}\right)}\left.\frac{d\log Z_t}{dt}\right|_{\omega(\cdot)} = \frac{\alpha\omega(z_t^*)\int_{z_t^*}^{\infty}(z-z_t^*)\,\omega(z)\,dz}{\int_{z_t^*}^{\infty}\omega(z)\,dz\int_{z_t^*}^{\infty}z\omega(z)\,dz} > 0.$$

Proof. See Appendix B.7.2.

This Proposition implies that a change in prices that raises the threshold $z_t^* = R_t/\varphi_t$ increases TFP.[27] Panel (b) in Figure 1 illustrates how an increase in the threshold decreases the share of constrained firms by crowding out low-productivity entrepreneurs. The intuition is simple: low-MRPK constrained firms that were close to the threshold become unconstrained and reduce their capital optimally to 0, which implies that these entrepreneurs stop using their net worth for production, and instead they lend it to more productive firms, decreasing misallocation. Changes in the productivity-threshold thus capture changes in the share of constrained versus unconstrained firms. This mechanism is different from the extensive margin: it is not meant to capture firm entry and exit, which in our model is exogenously given by the probability of retiring η. Rather, it captures the idea that previously constrained firms become unconstrained and vice versa.

So far we have considered how, in partial equilibrium, price movements affect TFP independently through the two margins. Together, these two margins simultaneously determine the evolution of TFP. However, rather than being two independent mechanisms, they are tightly connected through general equilibrium forces, which tie the price movements together. In particular, inspecting the market clearing condition for capital allows us to conclude that, in general equilibrium, prices must move in such a way that the two channels always work in the same direction.

Corollary 1. ***(Co-movement of slope and cutoff).*** *An increase in the slope of the excess investment rate, $\frac{\gamma\varphi_t}{q_t}$, is associated to an increase in the growth rate of the threshold $\frac{dz_t^*}{dt}$*

$$\frac{\partial\left(\frac{dz_t^*}{dt}\right)}{\partial\left(\frac{\gamma\varphi_t}{q_t}\right)} > 0.$$

Therefore, when $\frac{\gamma\varphi_t}{q_t}$ increases TFP increases.

[27]We consider the effect of the *growth rate* of the threshold $\frac{dz_t^*}{dt}$ on the *growth rate* of TFP $\frac{d\log Z_t}{dt}$, and not that of the *level* z_t^* on the *level* of TFP Z_t. We focus on growth rates because, in general equilibrium, equation (27) implies that neither z_t^* nor Z_t can jump discretely in response to monetary policy induced price changes; instead, they adjust only gradually over time.

Proof. See Appendix *B.7.3*.

This is intuitive: an increase in the slope of the excess investment function $\frac{\gamma\varphi_t}{q_t}$ implies a shift of the net-worth share distribution to the right. *Ceteris paribus*, this increases the demand for capital and reduces its supply. Thus, the threshold z_t^* has to increase to clear the capital market. Thus TFP increases through both margins.

This corollary allows us to characterize the capital misallocation channel of monetary policy, that is, the response of misallocation and TFP to monetary policy. If expansionary monetary policy shocks increase the slope $\frac{\gamma\varphi_t}{q_t}$—which is the case in a wide range of New Keynesian models—then misallocation decreases and TFP increases. To see if this is indeed the case, we now turn to general equilibrium analysis.

4.2 Misallocation in general equilibrium

Households' time-preference shock without nominal rigidities. While our main focus in this Section is on monetary policy shocks, it is instructive to begin with the response to a time-preference shock—specifically, a temporary decline in the household's discount rate, ρ^h, abstracting for the moment from nominal rigidities, or equivalently assuming strict inflation targeting (dashed orange line in Figure 2).[28] This shock triggers a decrease of real interest rates and a hump-shaped decline of TFP. The drop in TFP results from a deterioration of the capital allocation due to the two interconnected effects discussed above.

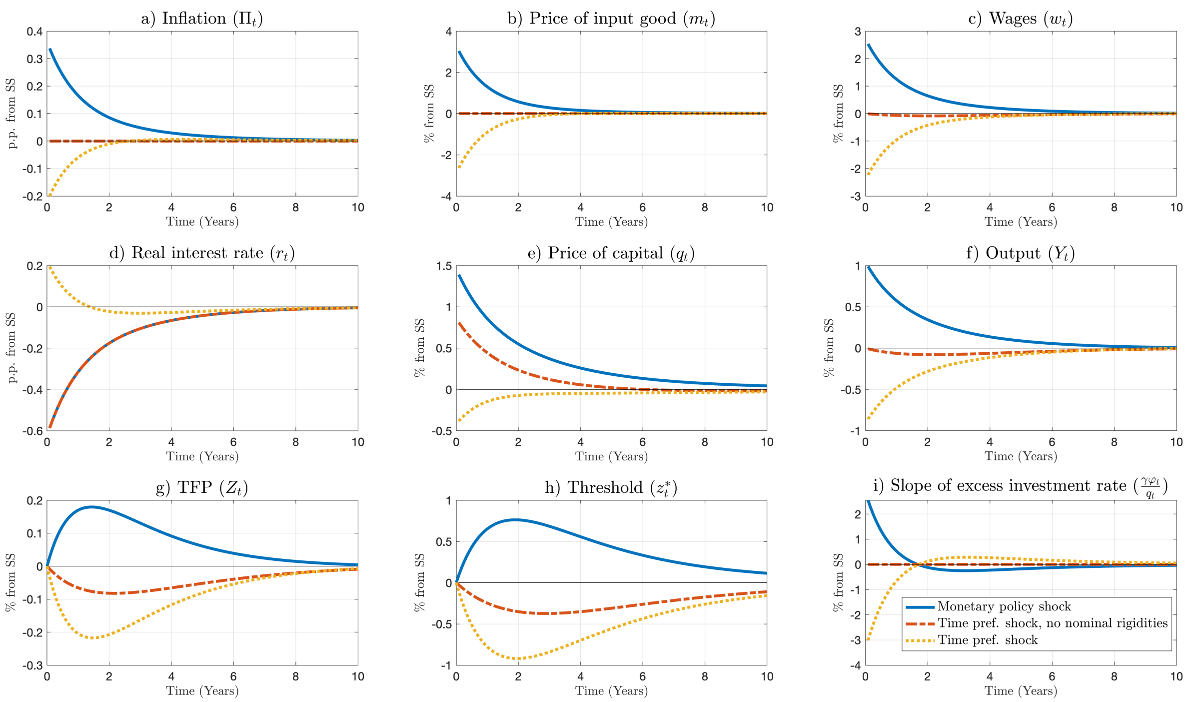
First, the share of net worth held by high-MRPK firms declines, because their profitability advantage over low-MRPK firms narrows, lowering their relative growth rate. Equivalently, the slope of the excess-investment $\frac{\gamma\varphi_t}{q_t}$ falls (panel i). This drop is almost entirely due to the rise in capital prices q_t (panel e), since $\varphi_t = \alpha\left(\frac{1-\alpha}{w_t}\right)^{\frac{1-\alpha}{\alpha}} m_t^{\frac{1}{\alpha}}$ barely moves because wages and input-good prices remain essentially unchanged (panels b and c). These price movements closely resemble those in the complete-markets limit in which capital allocation is efficient and TFP is exogenous—i.e., the representative-agent New Keynesian model (Figure 9 in Appendix C.1). Thus the price dynamics shaping the net-worth distribution are standard.[29]

[28]For comparability, we pick the time path for ρ^h such that the implied real rate coincides with that from the monetary policy shock analyzed next. This implies a drop in ρ^h of 24.6 b.p., and a gradual return to the steady state level.

[29]The complete-markets economy is the standard representative agent New Keynesian model with

Second, some previously unconstrained low-MRPK firms expand and become constrained. This is reflected in a decline in the cutoff $z_t^* = \frac{r_t q_t + \delta q_t - \dot{q}_t}{\varphi_t}$ (eq. 10). The fall in z_t^* results from the drop in the real interest rate r_t, only partially offset by the increase in q_t. Intuitively, the flatter slope of the excess-investment function shifts the net-worth distribution leftward, reducing demand from firms above the original cutoff. The threshold must therefore fall to clear the capital market. Panel (a) of Figure 3 decomposes the TFP response into contributions from individual price movements. Input-good prices m_t remain constant and wages w_t matter little. Rising capital prices q_t raise TFP, but the decline in real rates r_t reduces it, and this negative effect dominates, producing the overall decline in TFP.

Figure 2: Impulse responses.

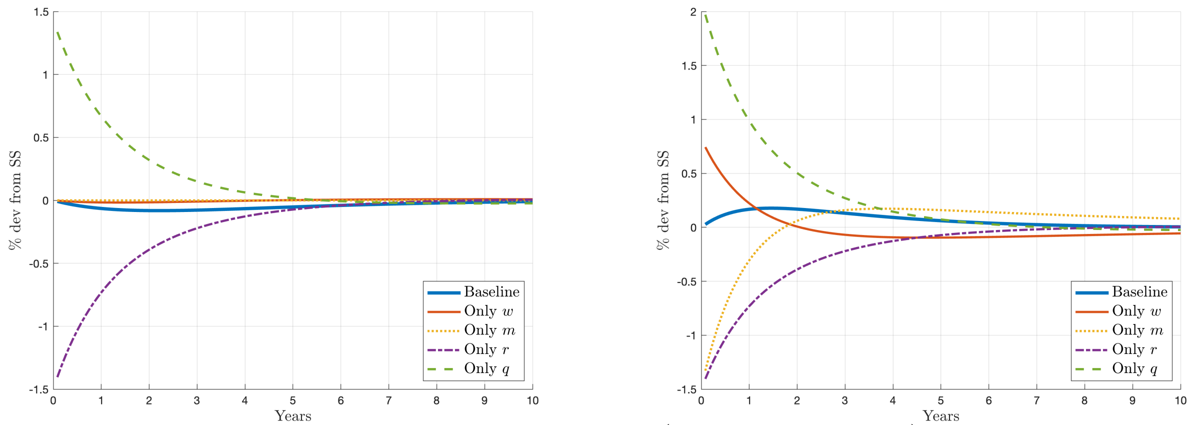


Notes: The figure shows the deviations from steady state of the economy. The solid blue line is the response of the baseline economy to an expansionary monetary policy shock of 25 basis points. The dashed orange line is the response of the economy to a time-preference shock in the absence of nominal rigidities, where the path for ρ^h is chosen so as to reproduce the path of the real rate for the monetary policy shock. The yellow dotted line is the response to the same shock with sticky prices.

fixed capital. It represents a limit case of the baseline economy where either the borrowing constraint is infinitely loose, so that the net-worth distribution becomes irrelevant and only the most productive entrepreneur operates, or where the variance in entrepreneurial productivity z is 0. In this case, capital allocation is efficient (no misallocation) and TFP is exogenous. Appendix B.8 compares the baseline and complete-markets models.

Figure 3: Decomposing the effect of a monetary policy shock on TFP.

(a) Time-preference shock (b) Monetary policy shock



Notes: The figure decomposes the effect of a time-preference shock (absent nominal rigidities) and monetary policy shock on TFP (bold blue line) into the effect of the individual factor price changes. This is done by computing how TFP would have evolved if all prices except one had remained at steady state.

Monetary policy shock. Now we turn to the baseline model *with* nominal rigidities and analyze a monetary policy shock, that is, a 25 b.p. reduction in the nominal rate (blue lines in Figure 2) when the central bank follows the Taylor rule in equation (34). The natural rate, defined as the real interest rate in the counterfactual flexible price economy, remains constant as monetary policy does not affect real variables under flexible prices. The decline in nominal rates causes an increase in output and inflation and a drop in the real rate (panels a, d, f) through the standard New Keynesian transmission mechanism.[30]

Furthermore, TFP now increases (panel g). This is so, first, because the distribution of net worth shifts towards more productive firms as the slope of the excess investment rate $\frac{\gamma\varphi_t}{q_t}$ increases (panel i). This increase happens because now the changes in input-good prices m_t and wages w_t cause φ_t to rise by more than the price of capital, q_t (panel e).[31] Second, the cutoff $z_t^* = \frac{r_t q_t + \delta q_t - \dot{q}_t}{\varphi_t}$ also increases (panel h), because the increase in the price of capital q_t overcompensates the decrease in the real rate r_t and

[30] As before, factor prices respond largely as in the complete-markets representative agent model (see Figure 9 in the appendix).

[31] Input-good prices m_t and wages w_t affect φ_t in opposite directions, as the higher prices (panel b) increase excess profits whereas higher wages (panel c) reduce them. However, the elasticity with respect to both variables is different, being larger ($\frac{1}{\alpha}$) for prices than for wages ($\frac{1-\alpha}{\alpha}$). As the increase of wages and input-good prices is roughly of the same magnitude, the different elasticities explain why φ_t increases.

the increase in φ_t. In contrast to the time-preference shock, the monetary policy shock raises the relative profitability of high-MRPK firms. As a result, the net-worth share of constrained high-MRPK firms expands. Consequently, the threshold z_t^* must rise to equilibrate the capital market (see the corollary above), implying that constrained low-MRPK firms contract their scale and become unconstrained. We refer to the impact of monetary policy on misallocation as the *capital misallocation channel of monetary policy*.

Panel b of Figure 3 illustrates how each factor price contributes to TFP dynamics. The rise in input-good prices m_t, which increases φ_t, raises the excess investment rate but lowers the threshold z_t^*. In the short run, the latter effect dominates, resulting in a net decline in TFP, whereas in the long run, the former effect prevails, yielding a net positive impact. The increase in wages w_t has the opposite—though quantitatively smaller—pattern. As under the time-preference shock, the decline in the real interest rate r_t exerts a negative effect on TFP, while the rise in capital prices q_t contributes positively, with the latter ultimately dominating.

Households' time-preference shock with nominal rigidities. For completeness we also report the response to the time-preference shock with nominal rigidities under a Taylor rule. The response (dotted yellow lines in Figure 2) is a combination of the response to the time-preference shock under flexible prices—or equivalently under strict inflation targeting—and the response to a negative monetary policy shock. To see this, note that the central bank lowers the nominal rate more gradually than under strict inflation targeting, causing the real interest rate to remain above the natural rate (compare the dotted yellow and dashed orange lines in panel d). The Taylor rule therefore implies a temporary, contractionary deviation from the strict inflation-targeting benchmark, leading TFP to decline even more than in the flexible-price case.[32]

Robustness. Note that the responses of capital misallocation and TFP discussed above depend directly on the relative movements of factor prices, which can only be determined numerically in general equilibrium. We show in Appendix C.7 that the positive response of TFP to monetary policy shocks is robust to a wide range of parameters. The key question, therefore, is whether the data indeed reveal a positive relationship between expansionary monetary policy shocks, a decrease in capital misallocation and an increase in TFP.

Empirical support. It has been widely documented that expansionary monetary

[32]The responses to a permanent time-preference shock are qualitatively similar - see Appendix C.2.

policy shocks indeed raise TFP (see Evans, 1992; Christiano et al., 2005; Garga and Singh, 2021; Jordà et al., 2020; Moran and Queralto, 2018; Meier and Reinelt, 2024 or Baqaee et al., 2024). The peak effect on TFP predicted by our model (0.72 p.p. increase in TFP to a 1 p.p. shock) falls within the range 0.4-1.7 p.p. of medium-run peak responses of TFP to monetary policy shocks estimated by those papers. Furthermore, their estimated responses are also hump-shaped over the medium-run. The model is thus consistent with this evidence.

The mechanism proposed in this paper to explain the positive TFP response—namely, that firms with higher marginal revenue product of capital (MRPK) exhibit a stronger response to monetary policy shocks—is also supported by empirical evidence. Albrizio et al. (2025), using firm-level data from the Spanish Central de Balances, provide evidence consistent with this channel. They show that firms with a MRPK above their industry average increase their capital investment more than the average firm in response to monetary policy easing, thereby contributing to a more efficient allocation of capital in the aggregate, which would be associated to an increase in TFP.[33] As such, the capital misallocation channel can complement alternative explanations for the increase of TFP in response to expansionary monetary policy shocks put forward in the literature, including mechanisms related to research and development (R&D), hysteresis effects, and markup heterogeneity.

Relation to the misallocation literature. Our results help reconcile the apparent tension between the literature showing how monetary policy shocks affect TFP on the one hand, and the literature finding that low real rates may fuel misallocation on the other hand (e.g. Gopinath et al., 2017 or Asriyan et al., 2025).[34] We show that there is no such conflict: our model delivers both an increase in TFP in response to an expansionary monetary policy shock and a decline in TFP in response to a negative demand shock when prices are flexible. The difference in the behavior of misallocation is due to the different natural rate dynamics: though in response to both shocks the *real* rate drops, the *natural* rate falls only for the demand shock, remaining constant for the monetary policy shock.

[33] Using the same data, the working paper version of this paper (Gonzalez et al., 2023) shows that the investment response at the firm level across different MRPK levels is also quantitatively similar to that predicted by the model.

[34] They find that misallocation increases and TFP decreases in response to a decline in real interest rates in flexible-price economies.

5 Optimal monetary policy

Having examined the interaction between monetary policy and capital misallocation, we now turn to the analysis of optimal monetary policy.

5.1 The Ramsey problem

We assume that the central bank sets its policy instrument—the nominal interest rate i_t—such as to maximize household utility under full commitment. That is, the central bank solves the following Ramsey problem: it chooses time paths for all equilibrium variables, which include the distribution $\omega(z)_t$, so as to maximize household welfare

$$\max_{\{q_t,w_t,r_t,R_t,m_t,\pi_t,i_t,C_t,Y_t,L_t,A_t,D_t,Z_t,z_t^*,\omega_t(z)\}_{t\geq 0}} \mathbb{E}_0 \int_0^\infty e^{-\int_0^t \rho_s^h ds} u(C_t, L_t) dt \tag{38}$$

subject to all the private equilibrium conditions defined above, the initial conditions on state variables $\{\omega_0(z), D_0, A_0\}$, and the transversality conditions of forward-looking variables $\lim_{t\to\infty} \{q_t, \pi_t, C_t, Y_t\}$. In the Appendix D.1 we present the full problem, spelling out all constraints (i.e. implementability conditions) and we derive the optimality conditions.[35]

Relative to the standard New Keynesian framework—namely, the complete-markets version of our model—the central bank's problem gains an additional dimension: monetary policy now influences TFP through the capital misallocation channel, and the policymaker must take this effect into account.[36] This additional layer of complexity makes the problem substantially more demanding computationally: the central bank's control variables now include the distribution of net-worth shares $\omega_t(z)$. Solving for the optimal policy thus requires computing first-order conditions with respect to an infinite-dimensional object. To address this challenge, we develop a new algorithm that enables us to solve a broad class of continuous-time Ramsey problems with heterogeneous agents nonlinearly under perfect foresight in a largely automated manner. Further details are provided in Appendix F.

Simplified model. To offer additional analytical insights about the model, in Appendix E we propose a simplified version of the model. Borrowing from Itskhoki

[35] We focus on policies with stationary Lagrange multipliers.

[36] Appendix B.8 shows the equilibrium conditions of the complete-markets model and how they compare to our baseline economy.

and Moll (2019), we assume that firm-level productivity shocks follow an i.i.d. power law distribution with parameter ξ: $G(z_t^*) = 1 - (z_t^*)^{-\xi}$ if $z_t^* > 1$ (and zero elsewhere). In this case, aggregate endogenous TFP only depends on entrepreneurs' aggregate net wealth A_t, so we no longer need to keep track of the distribution of wealth among entrepreneurs. To ensure stability of the equilibrium, we require risk aversion to be sufficiently low, specifically $\zeta < \frac{Y-\delta K}{Y}$.

5.2 Optimal Ramsey policy

5.2.1 Steady state.

Let us focus first on the steady state of the Ramsey problem.

Proposition 3. *In the baseline model, if the Ramsey problem has a steady state, it features zero inflation as long as the household's budget constraint is binding for the central bank.*

Proof. See Appendix *D.2*.

The optimality of zero inflation in the long run—a hallmark feature of the standard New Keynesian model—thus extends to our setting. The intuition is exactly the same as in the standard New Keynesian model (Woodford, 2003; Benigno and Woodford, 2005), and does not depend on the misallocation channel.[37] On the one hand, the central bank considers the direct costs of inflation, due to the Rotemberg adjustment costs, which are minimized when inflation is zero. On the other hand, the central bank trades off the benefits and costs of inflation through the Phillips curve. The benefit of inflation is a *contemporaneous* reduction in markups. The cost is that firms, in anticipation of future inflation, increase markups. As the proof illustrates, in steady state the benefit and cost from the Phillips curve exactly offset each other. Thus, only the direct costs of inflation matter, and zero inflation is optimal.

The condition that the household budget constraint is binding is mild, and holds in the case of our calibrated economy. For the simplified version of the model, we can furthermore prove the existence of the zero-inflation Ramsey steady state analytically—see Appendix E.2.2.

[37]Key to this result is the assumption that misallocation and nominal rigidities do not interact at the level of an individual firm: misallocation occurs among the heterogeneous input-good producers, while price stickiness affects the retail firms.

5.2.2 Time inconsistency and the time-0 problem

Ramsey policy plans, which are devised under full commitment, are not necessarily time consistent. A policy is time inconsistent if it is optimal when first planned, but the policymaker would like to deviate from it further in the future. Intuitively, such inconsistency arises because the policymaker's incentives change over time: after private agents have formed expectations based on past commitments, the policymaker may wish to deviate from them to achieve short-term gains. This dynamic tension between commitment and discretion is at the heart of the time-inconsistency problem, as described by Kydland and Prescott (1977).

Mathematically, time inconsistency arises when the Ramsey problem features *binding forward-looking* constraints—conditions linking current policy choices or prices to agents' expectations of the future. The Lagrange multipliers associated with these constraints represent the shadow value for the central bank of deviating from previous commitments. If these multipliers are zero, the constraints are slack, and the policymaker has no incentive to deviate from its past commitments—implying time consistency. Conversely, when the multipliers are nonzero, the policymaker gains from unexpectedly relaxing these constraints, meaning that the Ramsey plan is time inconsistent and the central bank would prefer to deviate from its earlier announced plan.

Time inconsistency is a common feature in monetary macroeconomics, and often affects the steady state of the Ramsey plan. Consider the standard New Keynesian model. Here the only forward-looking constraint is the Phillips curve.[38] If an appropriately chosen subsidy makes the steady-state allocation efficient, that constraint is slack and the Ramsey plan is time consistent in steady state. Otherwise, output is inefficient in the Ramsey steady state and the central bank has an incentive to exploit the Phillips curve in order to close the output gap. The Phillips curve thus binds and the Ramsey plan is time inconsistent (Benigno and Woodford, 2005).

In our model with capital misallocation, there are two forward-looking conditions: the Phillips curve and the Euler equation for capital prices. Since the zero-inflation steady state is distorted by capital misallocation, those two constraints bind, no matter whether the subsidy is in place or not. Therefore, the Ramsey plan is time inconsistent in steady state. We prove this analytically for the simplified version of the model,

[38] The Euler equation does not play any role in the central bank's problem, except to back out the path of real rates consistent with the optimal inflation and output gap. It is thus slack.

by showing that the Lagrange multipliers associated with these two constraints are non-zero in steady state:

Proposition 4. *In the simplified model, the optimal policy is time inconsistent.*

Proof. See Appendix *E.2.2*

In the baseline model, we can only evaluate the time inconsistency numerically: we find the respective multipliers to be non-zero in steady state.

Time inconsistency manifests in the form of what is often referred to as time-0 problem: if the economy starts at the zero-inflation steady state, the solution of the central bank's Ramsey problem involves a temporary deviation from zero inflation, even though zero inflation is optimal in the long run, i.e. in the Ramsey steady state.

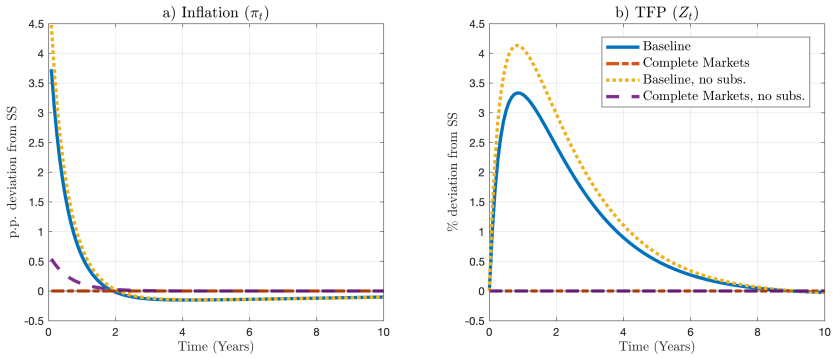
Figure 4 illustrates the time-0 problem in our baseline model (solid blue lines) and the complete-markets case (dashed orange line), which coincides with the standard New Keynesian model.[39] Thanks to the subsidy offsetting the markup, the Ramsey plan in the model with complete markets is time-consistent. Hence, inflation and the rest of the variables remain constant at their steady state values, and there is no time-0 problem. This contrasts with the baseline model, where the central bank engineers a sizable surprise monetary expansion, increasing inflation (panel a). The resulting dynamics are very similar to those caused by an expansionary monetary policy shock, which were described in detail in Section 4.2: the share of capital held by high-MRPK entrepreneurs grows such that TFP increases (panel b). The central bank thus engineers a monetary expansion, tolerating a temporary increase in inflation, in order to achieve a persistent rise in TFP, brought about by a more efficient allocation of capital.[40] Appendix C.4 shows the evolution of factor prices and the resulting increases in both the slope of the excess investment function, $\frac{\gamma\varphi_t}{q_t}$, and the cutoff z^*, which drive this increase in TFP. It also shows how the net-worth share distribution shifts to the right in response to the shock.

To assess the severity of the time-0 problem in our framework, we compare the optimal policy described above with the one obtained in the absence of a subsidy that

[39]The set of equilibrium conditions of the complete-markets economy is presented in Appendix B.8. In Appendix E.2.4 we prove that the Ramsey plan is time-consistent with complete markets.

[40]The desire of the central bank to redistribute resources to high-MRPK entrepreneurs is reminiscent of the case with optimal fiscal policy analyzed by Itskhoki and Moll (2019). They find that optimal fiscal policy in economies starting at below steady-state net-worth levels initially redistributes from households towards entrepreneurs in order to speed up net-worth accumulation, and thus increase TFP growth.

Figure 4: Time-0 problem.



Notes: The figure shows the deviations from steady state when the planner solves the Ramsey problem without pre-commitments and in the absence of shocks. The baseline economy is the solid blue line, and the complete-markets economy the dashed orange line. The dotted yellow line and the purple dashed line repeat the same exercise in the absence of the subsidy that offsets the markup distortion.

offsets the markup distortion. In this case, the complete-markets representative agent model is time-inconsistent as well and thus also exhibits a time-0 problem, as illustrated by the dashed purple lines in Figure 4.[41] Nonetheless, the figure shows that capital misallocation leads to a substantially more pronounced time-0 problem: the optimal inflation path in the model with misallocation (dotted yellow line) is roughly six times higher than in the complete-markets representative agent model, in which the markup is the only relevant steady state distortion (purple line). We therefore conclude that the time-0 problem is not only sizable in absolute terms—with average inflation reaching about 3% in the first year—but also far exceeds the one arising from markups in the standard New Keynesian model.

5.3 Optimal policy from a timeless perspective

Next we analyze the optimal policy response to unexpected shocks hitting an economy initially at its zero-inflation steady state. We consider policies that are optimal from a "timeless perspective" (Woodford, 2003). The interpretation is standard: under timeless optimality, the central bank honors its prior commitments and refrains from exploiting short-term incentives to deviate from them. These pre-commitments are the ones consistent with an optimal plan chosen at a time far in the past. Timelessly optimal policy thus allows us to study systematic monetary policy responses to shocks under the ex-ante optimal, time-invariant policy rule. Building on the argument by

[41] Appendix C.4 presents the same exercise displaying the dynamics of a broader set of variables.

Boppart et al. (2018) one can reinterpret the timeless response to MIT shocks as a first-order approximation to the optimal response under uncertainty. We operationalize timeless optimality following the timeless Ramsey approach by Dávila and Schaab (2023), applying the recursive Lagrangian logic of Marcet and Marimon (2019) to our continuous-time setting. That is, we augment the central bank's objective function by introducing a *timeless penalty*. This penalty adjusts the Ramsey problem so as to account for the central bank's steady-state pre-commitments. It renders the problem recursive and ensures that its solution in steady state is time consistent.[42]

The next Proposition characterizes the *timeless penalty*:

Proposition 5. *The penalty that makes the optimal policy time-consistent at the stationary Ramsey plan, and the associated timeless objective function are given by*

$$\overbrace{\int_0^\infty e^{-\int_0^t \rho_s^h ds}\left\{\frac{C_t^{1-\zeta}}{1-\zeta}-\Upsilon\frac{L_t^{1+\vartheta}}{1+\vartheta}\right\}dt\underbrace{-\overbrace{\lambda_{PC}C_0^{-\zeta}\pi_0 Y_0}^{\text{Phillips curve penalty}}-\overbrace{\lambda_{EE}C_0^{-\zeta}q_0}^{\text{Euler equation penalty}}}_{\text{Timeless penalty}}}^{\text{Timeless objective}}$$

where λ_{PC} and λ_{EE} are the steady-state values of the multipliers associated to the Phillips curve and the Euler equation, respectively.

Proof. See Appendix *D.3*

The first component of the penalty (*Phillips curve penalty*), $\lambda_{PC}C_0^{-\zeta}\pi_0 Y_0$, represents the central bank's commitment not to deviate from the firms' expectations embedded in the Phillips curve. The second component (*Euler equation penalty*), $\lambda_{EE}C_0^{-\zeta}q_0$, represents the commitment not to surprise the households relative to their expectations embedded in the Euler equation. The penalty modifies the central bank's objective so that the marginal benefit of deviating from zero inflation is exactly offset by the marginal cost imposed by the penalty. In this way, the penalty induces the central bank optimizing today to adopt a timeless perspective—namely, to internalize how its current actions would have affected the past through the anticipation of those actions by households and firms.

[42] In the Ramsey plan, these pre-commitments are encoded in the initial values of the Lagrange multipliers associated to forward-looking equations. Therefore the penalty contains one term per forward-looking equation, weighted by the steady state multiplier of that equation.

Figure 14 in Appendix C.5 illustrates these costs of anticipation. It shows how a marginal deviation from zero inflation is no longer socially desirable when it *is* anticipated, even though it is desirable when it *is not* anticipated. This is because inflation no longer leads to the same increase in TFP, once it is anticipated. The penalties encode these costs of anticipation at the margin for the steady state.

The entrepreneurs' savings decision does not imply a forward-looking optimality condition, so unlike in Dávila and Schaab (2023), there is no distributional penalty in our setting. The intuition for this is straightforward. Entrepreneurs face both a static profit-maximization problem and a dynamic saving decision. However, the dynamic problem yields a corner solution—entrepreneurs always fully invest their profits—so the optimal choice does not depend on the expectation of future variables, and there is no temptation to manipulate such expectations.

While the proof in the Appendix D.3 focuses on the baseline model, the same penalty applies also to the simplified model and to the complete-markets economy. In the baseline and the simplified model, both steady-state multipliers are non-zero, as we show analytically for the simplified model in the proof of Proposition 4 and numerically for the baseline model. In the complete-markets economy, however, the Euler equation is slack (see Appendix E.2.4), so the second component of the penalty drops out.

5.3.1 Time-preference and TFP shocks

We focus first on time-preference and technology shocks. For those shocks, the "*divine coincidence*" holds in the complete-markets economy (Blanchard and Gali, 2007). This implies that, in the standard New Keynesian model, it is optimal for the central bank to keep inflation at zero at all times in response to these shocks: once constrained by the pre-commitments, the central bank cannot improve the flexible-price allocation (see Appendix E.2.4).

Once we extend the model to include firm heterogeneity, inflation may also be used to modify the capital allocation. This might potentially imply that the divine coincidence does not hold anymore in our model. We start by considering the simplified model. As the next Proposition states, the divine coincidence extends to this case.

Proposition 6. *In the simplified model with $\delta = 0$, the "divine coincidence" holds, that is, $\pi_t = 0$ is optimal $\forall t$ in response to shocks to the discount rate (ρ_t^h) or to the exogenous component of TFP (ς_t).*

Proof. See Appendix *E.2.3*.

The proof establishes that the Lagrange multipliers associated with the forward-looking constraints are constant. This implies that the tightness with which the forward-looking constraints bind is invariant to the realization of time-preference or TFP shocks. Consequently, the penalty term disciplining the timeless policy offsets the incentive to stimulate demand not only in steady state, but also upon the arrival of these shocks. The intuition mirrors that in the standard New Keynesian model with an inefficient steady state (Benigno and Woodford, 2005). In both cases, the timeless planner is tempted to engineer an unexpected short-run stimulus: under complete markets, to push output toward its efficient level; and in our model, to further improve the allocation of capital. However, in both cases the central bank recognizes that, once anticipated, such a policy loses its benefits, and therefore refrains from it in the Ramsey steady state. Furthermore, in both cases the benefit from a surprise stimulus does not get larger or smaller in response to a TFP or a time-preference shock. Therefore, the central bank also does not systematically deviate from zero inflation in response to those shocks. That is, the flex-price allocation is optimal, once taking pre-commitments into account.

While the Proposition above pertains to the fully nonlinear, perfect-foresight model, a second-order approximation to the timeless planner's objective function helps to further sharpen the underlying intuition.

Proposition 7. *In the simplified model with $\delta = 0$, the objective function under timelessly optimal policy can be approximated to second order by (where variables with hats denote log deviations from the zero-inflation steady state)*

$$-\int_0^\infty e^{-\rho t} \frac{Y^{1-\zeta}}{2} \left[\frac{\theta}{1-\zeta} \pi_t^2 - \frac{2}{1-\zeta} \frac{\vartheta + 1}{1-\alpha} \left(\hat{Y}_t - \hat{Z}_t \right) \hat{\tau}_t \right] dt. \tag{39}$$

Proof. See Appendix *E.4*.

In the absence of cost-push shocks ($\hat{\tau}_t = 0$), only the inflation term is relevant. In that case, the central bank faces no trade-off and zero inflation is trivially optimal, reiterating Proposition 6. How come there is no output gap term? As we show in the proof, in the absence of pre-commitments, the second-order approximation of the objective around the zero-inflation steady state is the sum of four terms: the volatility of inflation π_t; the volatility of output gap $\tilde{Y}_t$;[43] the volatility of TFP, $\hat{Z}_t$; and a linear term in TFP deviations, $\hat{Z}_t$ (see Appendix E.3):

$$-\int_0^\infty e^{-\rho t}\frac{C^{-\zeta}Y}{2}\left[\theta\pi_t^2+\left(\zeta+\frac{\vartheta+\alpha}{1-\alpha}\right)\tilde{Y}_t^2+\frac{\vartheta+1}{1-\alpha}\frac{\zeta-1}{\zeta+\frac{\vartheta+\alpha}{1-\alpha}}\hat{Z}_t^2-2\left(1-\int_0^t\hat{\rho}_s ds\right)\hat{Z}_t\right]dt, \tag{40}$$

These four terms reflect the three potential inefficiencies of our economy. The first two terms—the inflation and the output gap term—are standard in New Keynesian models and capture the welfare losses from price adjustment costs and from monopolistic markups.[44] In our model, capital misallocation constitutes an additional inefficiency, which is reflected in the TFP terms, which are now endogenous to policy. Note that the linear TFP term points at the time inconsistency of the zero inflation Ramsey steady state: the central bank has an incentive to increase TFP. However, when taking pre-commitments into account, the last three of those four terms cancel out. That is, any concerns for the volatilities of output gap and TFP, or the desire to increase TFP, that the time-0 Ramsey planner trades off, are exactly canceled out by the pre-commitments from a timeless perspective. By contrast, when TFP is exogenous as in the standard New Keynesian model, the terms in Z_t are independent of policy and drop out, such that the timeless objective can be approximated by the standard two quadratic terms in π_t and $\hat{Y}_t$.[45] This comparison highlights that the familiar timeless incentive to manipulate the output gap is exactly offset by the timeless incentive to manipulate endogenous TFP in our simplified model.

The baseline framework includes two deviations from the simplified model of Proposition 7, both of which may potentially invalidate the divine coincidence. First, capital depreciates ($\delta > 0$), so that under zero inflation the aggregate resource constraint becomes $C_t + \delta K = Y_t$, breaking the identity $C_t = Y_t$. The introduction of depreciation

[43] The output gap is defined relative to the flexible-price zero-markup allocation with the same TFP level $\tilde{Y}_t \equiv \hat{Y}_t - \frac{1}{\zeta+\frac{\vartheta+\alpha}{1-\alpha}}\frac{(\vartheta+1)}{(1-\alpha)}\hat{Z}_t$.

[44] As standard in New Keynesian models, a constant subsidy eliminates the effect of the markup in steady state. However, following shocks markups vary over time due to nominal rigidities. Consequently, the subsidy no longer fully offsets markups, giving rise to a wedge between the marginal rate of substitution and the marginal product of labor (the labor wedge). The output gap term reflects this distortion.

[45] When the markup distortion is not corrected for by a subsidy in steady state, the timeless planner's objective function contains the same two terms in standard representative-agent model, but the weights are different. See appendix E.4.3.

with constant capital is equivalent to having government consumption in the aggregate resource constraint. Benigno and Woodford (2005) show how adding government consumption breaks the divine coincidence in the standard New Keynesian model when the steady state is distorted. While in that model the steady state is distorted by markups, in our environment the steady state distortion is caused by capital misallocation, but the logic is the same. Second, the presence of persistent firm-level productivity shocks implies that the aggregate capital allocation depends not only on the total net worth of firms, but also on its distribution across firms. This heterogeneity may further disrupt the divine coincidence.

However, there is a special case in which the result survives in the baseline model:

Proposition 8. *In the baseline model, with $\delta = 0$ and logarithmic utility, the divine coincidence holds, that is, $\pi_t = 0\ \forall t$ in response to shocks to the exogenous component of TFP (ς_t).*

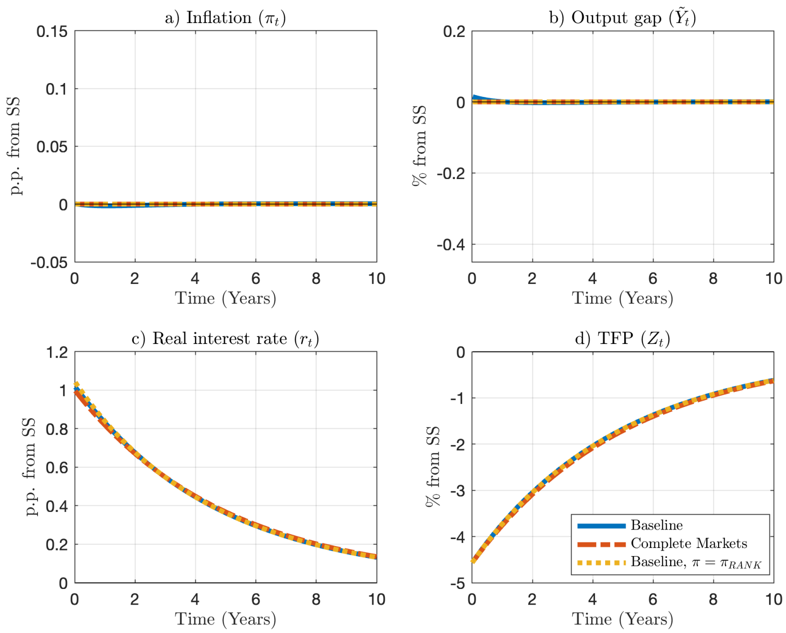
Proof. See Appendix *D.4*.

The intuition underlying the Proposition is straightforward. Suppose labor and inflation are held constant. In that case, consumption, the real interest rate, the price of capital and wages move one-for-one with exogenous productivity, leaving firms' net worth—and hence endogenous aggregate productivity—unchanged. Logarithmic utility ensures that income and substitution effects cancel, making it optimal for the household to keep labor supply constant. It follows that if zero inflation is optimal in steady state, it remains optimal in the presence of productivity shocks.

The preceding propositions characterize optimal timeless policy under restrictions introduced for analytical tractability. Once these restrictions are relaxed, timeless welfare can no longer be expressed in the compact form given by (39), although the linear-quadratic approximation to welfare in (40) remains valid. In this case, the central bank may optimally trade off deviations from strict inflation targeting—which closes the output gap—against reductions in timeless welfare losses from capital misallocation. Nevertheless, we show below that for empirically plausible parameterizations, including our baseline calibration, optimal policy remains close to the divine coincidence.

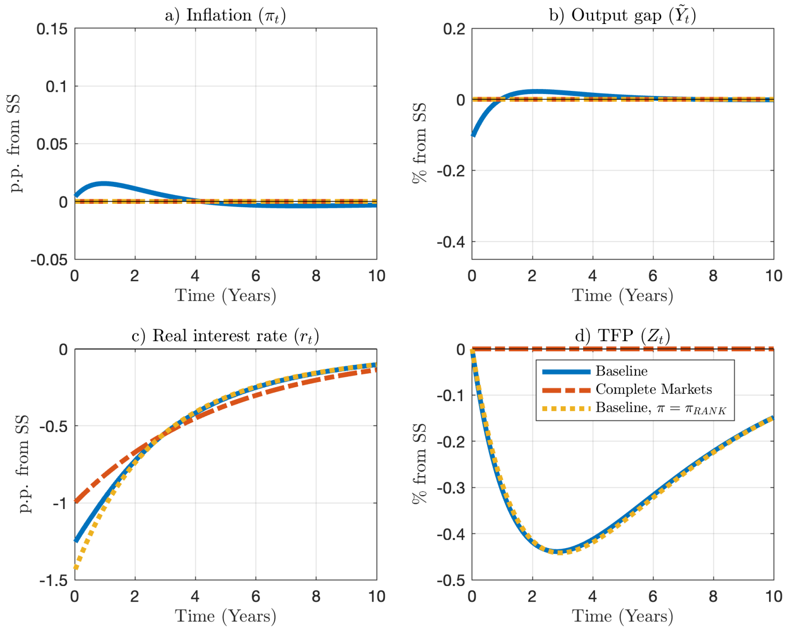
TFP shock. Figure 5 shows the optimal policy response to a TFP shock. The shock is scaled so that its impact effect on the real interest rate matches that of the time-preference shock in absolute value. In the complete-markets economy (orange dashed line), the central bank stabilizes inflation and the output gap at their steady-state

Figure 5: Optimal monetary policy response to an exogenous TFP shock.



Notes: The figure shows the optimal response from a timeless perspective (in deviations from steady state) to a 12.5% decrease in the exogenous component of TFP (ς_t), which is equivalent to a 4.4% drop in TFP (due to the exponent of ς_t). The shock is mean reverting with a yearly persistence of 0.8. The baseline economy is the solid blue line, and the complete-markets economy is the dashed orange line. The figure also shows the paths of the variables under strict inflation targeting (yellow lines).

Figure 6: Optimal monetary policy response to a time-preference shock.



Notes: The figure shows the optimal response from a timeless perspective (in deviations from steady state) to a 1 p.p. decrease in the household's time-preference rate, ρ_t^h, which is mean reverting with a yearly persistence of 0.8. The baseline economy is the solid blue line, and the complete-markets economy is the dashed orange line. The figure also shows the paths of the variables under strict inflation targeting (dotted yellow line).

value of zero, delivering the exact divine coincidence. In the baseline economy with misallocation (blue solid line), the optimal inflation response remains quantitatively negligible. Moreover, outcomes under strict inflation targeting (yellow dashed line) are virtually indistinguishable from those under the optimal policy. Hence, the baseline model exhibits an "approximate divine coincidence". This result is consistent with the fact that the model departs from the environment of Proposition 8 only through a positive but small rate of capital depreciation.

Households' time-preference shock. A similar result applies for time-preference shocks. As shown in Figure 6 the optimal response is again very close to strict inflation targeting, though deviations are somewhat larger than under a TFP shock. Recall from Proposition 6 that the divine coincidence holds exactly in the simplified environment with i.i.d. firm-level shocks and no depreciation. By contrast, the baseline model features persistent idiosyncratic shocks and positive depreciation. These departures naturally generate somewhat larger deviations from the divine coincidence, though they remain quantitatively modest.

Notice that the policy rate path that the optimal policy prescribes, however, differs with respect to the complete-markets case. Under incomplete markets, a negative demand shock leads to an endogenous fall in both the productivity of the marginal firm z_t^* and in aggregate TFP Z_t through the increase in capital misallocation. This, in turn, amplifies the reduction of the natural rate brought about by the demand shock itself, such that the natural rate drops more than in the case with complete markets. As the interest rate mimics the natural rate, it declines more, and more persistently, than in the standard New Keynesian model.[46]

These results carry important implications. In principle, monetary policy could be deployed to mitigate the impact of shocks by exploiting the capital misallocation channel. However, under a timeless-perspective policy, the strength of these incentives is quantitatively weak. As a consequence, strict inflation targeting delivers allocations that are nearly indistinguishable from those implied by the fully optimal policy. Put differently, even in the presence of misallocation, inflation targeting emerges as a robust approximation to optimal monetary policy.[47]

5.3.2 Cost-push shocks

Cost-push shocks create a trade-off between inflation and output gap stabilization in the complete-markets model. We model cost-push shocks as temporary changes in the subsidy τ_t on the input good. We calibrate the shock so that, in the complete-markets economy, the initial decline in the real interest rate matches the impact response under the time-preference shock.

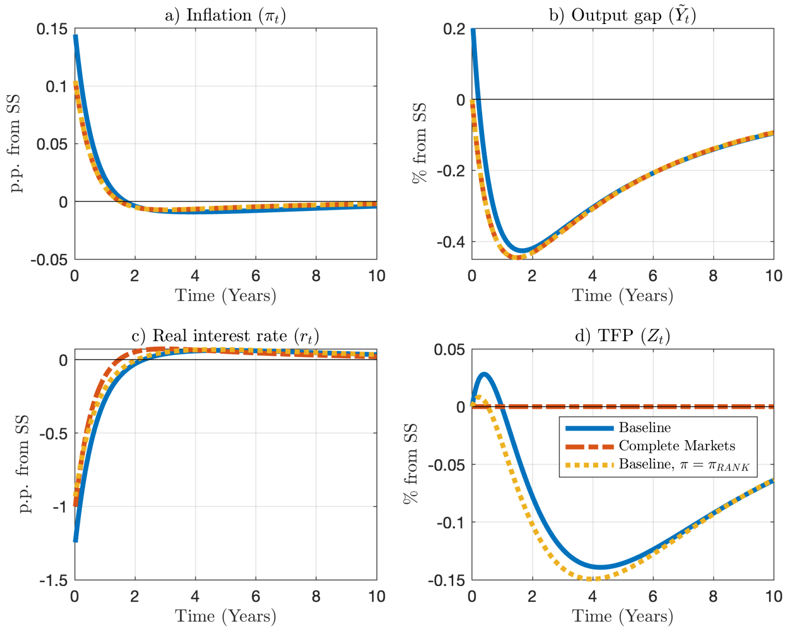
The dashed orange line in Figure 7 shows the timeless Ramsey plan in the complete-markets benchmark. The reduction in the subsidy increases costs, causing firms to produce less. This gives rise to a trade-off between inflation and output stabilization: the central bank lets inflation increase (panel a) to partially cushion the fall in the output gap (panel b), a strategy known as "*leaning against the wind*".

With capital misallocation, the optimal prescription differs qualitatively from the complete-markets benchmark, even though the resulting quantitative deviations are relatively small. Here, in addition to the short-run inflation–output gap trade-off, mon-

[46] In the working paper version of this paper (Gonzalez et al., 2023), we showed that, because of this, the zero lower bound becomes more binding in the baseline model.

[47] Appendix C.7 reports the sum of absolute changes in inflation following TFP and time-preference shocks under the timeless perspective, across alternative parameterizations relative to the baseline. The results show that, for reasonable parameter values, deviations from the divine coincidence remain small.

Figure 7: Optimal monetary policy response to a cost-push shock.



Notes: The figure shows the optimal response from a timeless perspective (in deviations from steady state) to a 195 b.p. decrease in τ, i.e. a *ceteris paribus* increase in costs of around 2%, that is mean reverting with a yearly persistence of 0.8. The baseline economy is the solid blue line, and the complete-markets economy is the dashed orange line. The figure also shows the paths of the variables in the basline model under strict inflation targeting (yellow line).

etary policy also affects resource misallocation and hence medium-run TFP. This additional channel motivates the central bank to attenuate the policy response and tolerate a larger increase in inflation (blue solid line, panel a in Figure 7). This causes the output gap to increase on impact, rather than dropping as in the complete-markets economy (panel b). This stance not only stabilizes the output gap more effectively than would the counterfactual policy of replicating the complete-markets response (yellow dashed line), but it also dampens fluctuations in endogenous productivity (panel d).

In the long run, the central bank exactly makes up for the initial inflationary burst, both in the complete-markets and the baseline economy, so that the price level reverts to its pre-shock value.[48] That is, the long-run price level targeting feature of the Ramsey policy (Gali, 2008) survives in the presence of capital misallocation.

Those two features—the temporarily positive output response and the long-run price level targeting—are robust to parameter variations.[49] For the simplified model with an

[48]Figure 15 in the Appendix shows this feature numerically in the baseline and the complete-markets economy.

[49]Appendix C.7 reports the at-impact output response under alternative parameterizations. In all cases, the response is positive. Furthermore, it shows that the price level targeting feature is present

additional analytically convenient assumption regarding capital depreciation costs, we can prove these two features analytically using a linear-quadratic approximation to the timeless problem.

Proposition 9. *Assume that capital depreciation costs are paid lump sum by the household. Up to first order, the timeless optimal response to a cost-push shock starting from the Ramsey steady state in the simplified model has the following features:*

1. *The output gap is strictly positive at some time t, while it is weakly negative in the complete-markets economy $\forall t$.*

2. *The price level converges back to its initial level in the long run, as in the complete-markets economy.*

Proof. See Appendix *E.4*.

Using the timeless objective from Proposition 5, we quantify the welfare differences between different policies. The time-0 optimal policy increases time-0 welfare by an amount equivalent to raising steady-state consumption by 5% for one year (annual consumption equivalent variation – ACEV). From a timeless perspective, however, this policy is suboptimal: it reduces *timeless* welfare by 250% ACEV.[50] In the complete-markets model, these numbers are one order of magnitude smaller, stressing again that the capital misallocation channel magnifies the time-0 problem.

By contrast, the welfare differences between the timelessly optimal policy and strict inflation targeting in response to shocks are small. For TFP and time-preference shocks, the gap is on the order of 10^{-6}% and 10^{-4}% ACEV, respectively. Strict inflation targeting is therefore close to optimal, especially for TFP shocks. For a cost-push shock, the gains from "*leaning against the wind*" relative to strict inflation targeting are modest—about 10^{-3}% ACEV—but twice as large as in the complete-markets benchmark. Our conclusion is that, despite the severity of the time-0 problem, the timelessly optimal response to shocks is only modestly affected by the capital misallocation channel of monetary policy.

across alternative parameterizations.

[50] We refer to *time-0* welfare as the discounted sum of household utility, and to *timeless* welfare as the augmented welfare measure defined in Proposition 5—that is, the discounted utility sum plus the timeless penalty term.

6 Conclusions

This paper introduces a tractable model with heterogeneous firms, financial frictions, and nominal rigidities in order to understand the link between monetary policy and capital misallocation, and its policy implications. We establish that monetary policy easing endogenously improves capital allocation by disproportionately relaxing financial constraints for productive firms, thereby raising TFP.

This new *capital misallocation channel* generates a time-0 problem: starting from the zero-inflation steady state, the central bank would like to engineer a one-off expansion to boost TFP. Yet, once commitment is imposed (*timeless perspective*), we find that the optimal policy is close to price stability when facing TFP and time-preference shocks, similar to the complete-markets benchmark. However, the optimal policy entails more aggressive and persistent interest-rate adjustments because the endogenous response of TFP amplifies the fluctuations of the natural rate. When facing negative cost-push shocks, the central bank tolerates a larger short-run rise in inflation and a temporarily positive output response. This more expansionary stance stabilizes real activity more effectively than the complete-markets benchmark and mitigates fluctuations in endogenous productivity.

The model presented in this paper abstracts from several relevant mechanisms driving firm dynamics, such as endogenous default, size-varying capital constraints, labor market frictions, uninsurable risk or decreasing returns to scale, among many others. This allows us to provide a clear understanding of the forces linking monetary policy with capital misallocation, as well as highlighting the similarities and differences with the standard New Keynesian model. A natural extension would be to add more of these features to study their impact on the optimal conduct of monetary policy.

References

Acharya, S., E. Challe, and K. Dogra (2023). Optimal monetary policy according to hank. *American Economic Review 113*(7), 1741–1782. 1

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2021, 04). Income and wealth distribution in macroeconomics: A continuous-time approach. *The Review of Economic Studies 89*(1), 45–86. 3, F.1, F.2, F.5, 67

Adam, K. and H. Weber (2019). Optimal trend inflation. *American Economic Review 109*(2), 702–37. 1

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018). When inequality matters for macro and macro matters for inequality. *NBER macroeconomics annual 32*(1), 1–75. 19, F.4

Albrizio, S., B. Gonzalez, and D. Khametshin (2025). A tale of two margins: Monetary policy and capital misallocation. *European Economic Review, accepted*. 6, 4.2

Almunia, M., D. Lopez Rodriguez, and E. Moral-Benito (2018). Evaluating the macro-representativeness of a firm-level database: an application for the spanish economy. *Banco de España Ocassional Paper* (1802). A.1

Alvarez, L. J., E. Dhyne, M. Hoeberichts, C. Kwapil, H. L. Bihan, P. Luennemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen, and J. Vilmunen (2006). Sticky prices in the euro area: A summary of new micro-evidence. *Journal of the European Economic Association 4*(2/3), 575–584. 17

Asriyan, V., L. Laeven, A. Martin, A. Van der Ghote, and V. Vanasco (2025). Falling interest rates and credit reallocation: lessons from general equilibrium. *Review of Economic Studies 92*(4), 2197–2227. 1, 13, 4.2

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica 89*(5), 2375–2408. 19, F.2, F.4

Auclert, A., M. Cai, M. Rognlie, and L. Straub (2024). Optimal long-run fiscal policy with heterogeneous agents. Technical report. 1

Baqaee, D. R., E. Farhi, and K. Sangani (2024). The supply-side effects of monetary policy. *Journal of Political Economy 132*(4), 1065–1112. 2, 1, 4.2

Bau, N. and A. Matray (2023). Misallocation and capital market integration: Evidence from india. *Econometrica 91*(1), 67–106. 20, A.2

Benigno, P. and M. Woodford (2005, December). Inflation Stabilization And Welfare: The Case Of A Distorted Steady State. *Journal of the European Economic Association 3*(6), 1185–1236. 1, 5.2.1, 5.2.2, 5.3.1, 5.3.1, E.2.3, E.4.1, E.4.3

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999, December). The financial accelerator in a quantitative business cycle framework. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1 of *Handbook of Macroeconomics*, Chapter 21, pp. 1341–1393. Elsevier. 13

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2021). Inequality, business cycles, and monetary-fiscal policy. *Econometrica 89*(6), 2559–2599. 1, F.1, F.4

Bigio, S. and Y. Sannikov (2021). A model of credit, money, interest, and prices. Technical report, National Bureau of Economic Research. F.2

Bilbiie, F. O., I. Fujiwara, and F. Ghironi (2014). Optimal monetary policy with endogenous entry and product variety. *Journal of Monetary Economics 64*, 1–20. 1

Bilbiie, F. O. and X. Ragot (2020). Optimal monetary policy and liquidity with heterogeneous households. *Review of Economic Dynamics*. 1

Blanchard, O. and J. Gali (2007). Real wage rigidities and the new keynesian model. *Journal of Money, Credit and Banking 39*, 35–65. 1, 5.3.1

Boar, C., D. Gorea, and V. Midrigan (2025). The aggregate costs of uninsurable business risk. Technical report, Working paper. 15

Boppart, T., P. Krusell, and K. Mitman (2018). Exploiting mit shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control 89*, 68–92. 19, 5.3, F.2

Buera, F. J. and J. P. Nicolini (2020). Liquidity traps and monetary policy: Managing a credit crunch. *American Economic Journal: Macroeconomics 12*(3), 110–38. 5

Caglio, C. R., R. M. Darst, and S. Kalemli-Özcan (2021). Risk-taking and monetary policy transmission: Evidence from loans to smes and large firms. Technical report, National Bureau of Economic Research. 20

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005, February). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy 113*(1), 1–45. 2, 4.2

Cloyne, J., C. Ferreira, M. Froemel, and P. Surico (2023). Monetary policy, corporate finance, and investment. *Journal of the European Economic Association 21*(6), 2586–2634. 1

David, J. M. and D. Zeke (2021). Risk-taking, capital allocation and optimal monetary policy. Technical report. 1

Dávila, E. and A. Schaab (2023). Optimal monetary policy with heterogeneous agents: Discretion, commitment, and timeless policy. *NBER Working Paper N. 30961*. 1, 1, 5.3, 5.3, F.1, F.2, F.4

Debortoli, D. and R. Nunes (2006, July). On Linear Quadratic Approximations. MPRA Paper 544, University Library of Munich, Germany. E.4.1

Evans, C. L. (1992, April). Productivity shocks and real business cycles. *Journal of Monetary Economics 29*(2), 191–208. 2, 4.2

Faia, E. (2008). Ramsey monetary policy with capital accumulation and nominal rigidities. *Macroeconomic Dynamics 12*(S1), 90–99. 18

Ferreira, M. H., T. Haber, and C. Rorig (2023). Financial constraints and firm size: Micro-evidence and aggregate implications. 16

Ferreira, T., D. Ostry, and J. H. Rogers (2023). Firm financial conditions and the transmission of monetary policy. 1

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review 98*(1), 394–425. A.1

Gali, J. (2008). Monetary policy, inflation, and the business cycle: An introduction to the new keynesian framework. *Princeton University Press*. 5.3.2

Garga, V. and S. R. Singh (2021). Output hysteresis and optimal monetary policy. *Journal of Monetary Economics 117*, 871–886. 2, 4.2

Gautier, E., P. Karadi, C. Conflitti, B. Fabo, L. Fadejeva, C. Fuss, T. Kosma, V. Jouvanceau, F. Martins, J.-O. Menz, and T. Messner (2023, July). Price adjustment in the euro area in the low-inflation period: evidence from consumer and producer micro price data. Occasional Paper Series 319, European Central Bank. 17

Gertler, M. and P. Karadi (2011). A model of unconventional monetary policy. *Journal of monetary Economics 58*(1), 17–34. 2.1, 2.1

Gonzalez, B., G. Nuno, D. Thaler, and S. Albrizio (2023, Dec). Firm heterogeneity, capital misallocation and optimal monetary policy. CEPR Discussion Papers 18695, C.E.P.R. Discussion Papers. 33, 46

Gopinath, G., Ş. Kalemli-Özcan, L. Karabarbounis, and C. Villegas-Sanchez (2017). Capital allocation and productivity in south europe. *The Quarterly Journal of Economics 132*(4), 1915–1967. 1, 3, 4.2, A.2

Hamano, M. and F. Zanetti (2022). Monetary policy, firm heterogeneity, and product variety. *European Economic Review 144*, 104089. 1, 8

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics 124*(4), 1403–1448. 1, A.2

Itskhoki, O. and B. Moll (2019). Optimal development policies with financial frictions. *Econometrica 87*(1), 139–173. 1, 26, 5.1, 40, B.7.2, E.1

Jeenas, P. (2023). Firm balance sheet liquidity, monetary policy shocks, and investment dynamics. Technical report, Universidad Pompeu Fabra. 1, 17

Jordà, Ò., S. R. Singh, and A. M. Taylor (2020). The long-run effects of monetary policy. Technical report, National Bureau of Economic Research. 2, 4.2

Juillard, M., D. Laxton, P. McAdam, and H. Pioro (1998). An algorithm competition: First-order iterations versus newton-based techniques. *Journal of Economic Dynamics and Control 22*(8-9), 1291–1318. F.2, F.4

Jungherr, J., M. Meier, T. Reinelt, and I. Schott (2024, August). Corporate debt maturity matters for monetary policy. Working Paper Series 2024-30, Federal Reserve Bank of San Francisco. 1

Kaplan, G., B. Moll, and G. L. Violante (2018). Monetary policy according to hank. *American Economic Review 108*(3), 697–743. 3, B.3

Koby, Y. and C. Wolf (2020). Aggregation in heterogeneous-firm models: Theory and measurement. Technical report. 1

Kydland, F. E. and E. C. Prescott (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of political economy 85*(3), 473–491. 5.2.2

Le Grand, F., A. Martin-Baillon, and X. Ragot (2025). Should monetary policy care about redistribution? Optimal monetary and fiscal policy with heterogeneous agents. *Review of Economic Studies, accepted*. 1, F.1, F.4

Lian, C. and Y. Ma (2020, 09). Anatomy of corporate borrowing constraints. *The Quarterly Journal of Economics 136*(1), 229–291. 14

Marcet, A. and R. Marimon (2019, September). Recursive Contracts. *Econometrica 87*(5), 1589–1631. 1, 5.3

Martin-Baillon, A. (2021). When should we tax firms? optimal corporate taxation with firm heterogeneity. Technical report. 9

Mckay, A. and C. Wolf (2022). Optimal Policy Rules in HANK. Mimeo. 1

Meier, M. and T. Reinelt (2024). Monetary policy, markup dispersion, and aggregate tfp. *Review of Economics and Statistics 106*(4), 1012–1027. 2, 1, 4.2

Midrigan, V. and D. Y. Xu (2014, February). Finance and misallocation: Evidence from plant-level data. *American Economic Review 104*(2), 422–58. 1

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review 104*(10), 3186–3221. 1, 1, 5, 2.1, 10, 2.1

Moran, P. and A. Queralto (2018). Innovation, productivity, and monetary policy. *Journal of Monetary Economics 93*, 24–41. 2, 4.2

Nuño, G. and C. Thomas (2022). Optimal Redistributive Inflation. *Annals of Economics and Statistics* (146), 3–63. 1, F.1, F.2, F.4, F.7

Ottonello, P. and T. Winberry (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica 88*(6), 2473–2502. 1, 17

Ragot, X. (2019). Managing inequality over the business cycles: Optimal policies with heterogeneous agents and aggregate shocks. 2019 Meeting Papers 1090, Society for Economic Dynamics. F.1

Reis, R. (2013). The portuguese slump and crash and the euro crisis. *Brookings Papers on Economic Activity*, 143–193. 1

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control 33*(3), 649–665. 19, F.4

Restuccia, D. and R. Rogerson (2017). The causes and costs of misallocation. *Journal of Economic Perspectives 31*(3), 151–74. 1

Rotemberg, J. J. (1982). Sticky prices in the united states. *Journal of Political Economy 90*(6), 1187–1211. 2.4

Smirnov, D. (2022). Optimal monetary policy in hank. Technical report. 1, F.1, F.2, F.4

Trimborn, T., K.-J. Koch, and T. Steger (2008). Multidimensional transitional dynamics: A simple numerical procedure. *Macroeconomic Dynamics 12*(3), 301Ã¢ÂÂ319. F.2, F.4

Winberry, T. (2018). A method for solving and estimating heterogeneous agent macro models. *Quantitative Economics 9*(3), 1123–1151. 19, F.4

Woodford, M. (2003). Interest and prices: Foundations of a theory of monetary policy. *Princeton University Press*. 1, 5.2.1, 5.3

Online Appendix

Appendix Contents

A Calibration

A.1 Firm level data

For calibration we use annual firm-level balance-sheet data from *Central de Balances Integrada* (CBI) of Banco de España, covering 1999–2016. The CBI provides detailed administrative information for the quasi-universe of Spanish firms (see Almunia et al., 2018). We retain only high-quality observations as defined by the CBI. The marginal revenue product of capital (MRPK) is proxied by the log of value added over tangible capital, restricting the sample to firms with positive value added. We drop the top 5 percent of the capital-weighted MRPK distribution to remove outliers. All nominal variables are deflated using industry price indices, yielding a revenue-based MRPK measure (Foster et al., 2008). In the final sample, mean MRPK is 0.77 in levels (-0.87 in logs) with a standard deviation of 0.66 (1.39 in logs). The capital-weighted MRPK distribution is shown in Figure 8. We use this MRPK series to compute the process of idiosyncratic shocks - see AppendixA.3. We use aggregate total debt (short- plus long-term) and aggregate total assets from the CBI to calibrate the borrowing-constraint parameter. The entry size is computed as the ratio of the aggregate capital of entrants to the aggregate capital of all active firms. Both statistics are computed annually and then averaged over the sample period.

For the calibration of the exit rate η we use data from the *Directorio Central de Empresas*, which is a data set maintained by INE, and it contains aggregate data on all firms operating in Spain, and its status (incumbent, entrant or exiter). The data set can be accessed here.

A.2 MRPK and its steady state distribution

A firm's MRPK is no directly observable. However, in a Cobb-Douglas production framework, such as the one presented in Section 2, a firm's MRPK is proportional to its average revenue product of capital (ARPK):

$$MRPK_t \equiv \alpha k_t^{\alpha-1} l_t^{1-a} \propto ARPK_t \equiv y_t/k_t = k_t^{\alpha-1} l_t^{1-a}$$

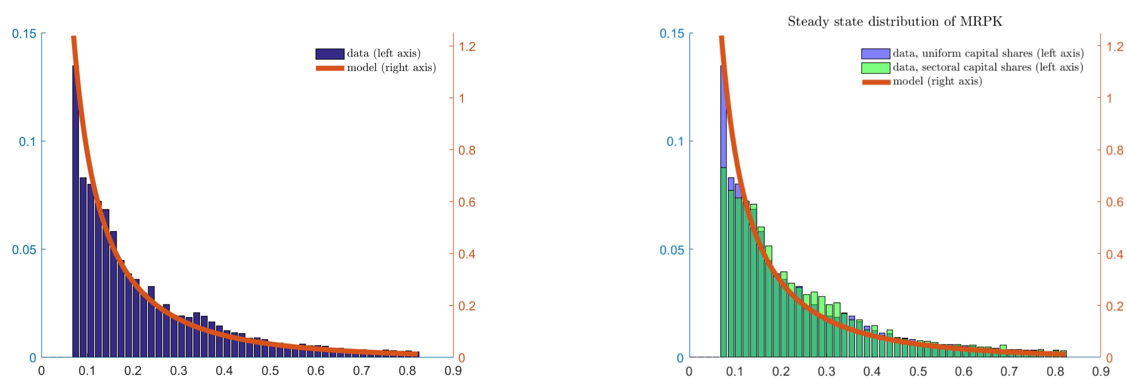
We thus use the easily measurable ARPK as an empirical measure for the unobservable MRPK, following the literature (see for instance Bau and Matray, 2023). To account for the use of intermediate inputs, which we don't model explicitly, we use value added (sales minus intermediate inputs) instead of sales.

However, capital shares α may differ across sectors. This would imply that the ARPK is no longer a valid proxy for the MRPK in cross sectoral comparisons. In the following we explain that all our results are robust to this concern.

Steady state MRPK distribution Figure 8 shows that the model predicts a steady state MRPK distribution that is in line with the MRPK distribution implied by the data, assuming a uniform capital share for all sectors (panel a). In Figure 8 panel b, we relax this assumption and allow for sectoral difference in the capital shares. Following Hsieh and Klenow (2009) and Gopinath et al. (2017), we take the sectoral capital shares of a relatively undistorted economy such as the United States. The fit of the model worsens only slightly in the direction of underpredicting the measured misallocation. The baseline calibration can hence be considered conservative.

Figure 8: MRPK distribution

(a) Uniform capital shares (b) Sectoral capital shares



Notes: Panel (a) shows the steady state distribution of firms MRPK in the model (orange solid line) and compares it to the data (histogram with blue bars). Panel (b) shows the robustness to differences in sectoral capital shares. We drop observations above an MRPK of 0.82, which implies dropping firms in the 5% upper tail of the capital-weighted MRPK distribution. Note that, by construction, the model cannot explain firms with an MRPK below the cost of capital ($R = r + \delta = 0.07$ in steady state), which we also drop in this figure.

A.3 Estimating the process for idiosyncratic productivity shocks

We assume that individual productivity z in logs follows an Ornstein-Uhlenbeck process

$$d\log(z) = -\varpi_z \log(z)dt + \sigma_z dW_t.$$

To estimate this continuous time process on discrete data, we approximate it by an AR(1) process using an Euler-Maruyama approximation

$$\log(z_t^j) = \rho_z \log(z_{t-1}^j) + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_z \sqrt{\Delta t}),$$

where $\rho_z \approx 1 - \varpi_z \Delta t \approx \exp(-\varpi_z \Delta t)$.

In the model, firm level productivity z is proportional to firm level MRPK

$$MRPK_t(z) = z\varphi_t$$

Using this we can rewrite the discretized process for z as

$$\log(MRPK_t(z_t^j)/\varphi_t) = \rho_z \log(MRPK_{t-1}(z_{t-1}^j)/\varphi_{t-1}) + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_z)$$

$$\log(MRPK_t(z_t^j)) = \rho_z \log(MRPK_{t-1}(z_{t-1}^j)) + f(\varphi_t, \varphi_{t-1}) + \varepsilon_t,$$

We estimate this equation using OLS on our panel data specified above, capturing the term $f(\varphi_t, \varphi_{t-1})$ by using year fixed effects. We find $\rho_z = 0.83$, and the standard deviation of the shock is $\sigma = 0.73$. This estimate is robust to including sector fixed effects to account for sectoral differences in capital shares (see Appendix A.2). As the data frequency is annual, $\Delta t = 1$, we back out the implied to continuous time parameter $\varpi_z = -\log(\rho_z) = 0.189$.

B Further details on the model

B.1 Entrepreneur's intertemporal problem

The Hamilton-Jacobi-Bellman (HJB) equation of the entrepreneur is given by

$$r_t V_t(z,a) = \max_{d_t \geq 0} d_t + s_t^a(z,a,d)\frac{\partial V}{\partial a} + \mu(z)\frac{\partial V}{\partial z} + \frac{\sigma^2(z)}{2}\frac{\partial^2 V}{\partial z^2} + \eta\left(q_t a_t - V_t(z,a)\right) + \frac{\partial V}{\partial t}.$$

We guess and verify a value function of the form $V_t(z,a) = \kappa_t(z)\, q_t a$. The first-order condition is

$$\kappa_t(z) - 1 = \lambda_d \text{ and } \min\{\lambda_d, d_t\} = 0,$$

where $\lambda_d = 0$ if $\kappa_t(z) = 1$. If $\kappa_t(z) > 1\ \forall z, t$, then $d_t = 0$ and the firm does not pay dividends until it closes down. If this is the case, then the value of $\kappa_t(z)$ can be obtained from

$$\begin{aligned}
&(r_t + \eta)\,\kappa_t(z)\, q_t = \\
&\eta q_t + (\gamma \max\{\varsigma_t z_t \varphi_t - R_t, 0\} + R_t - \delta q_t)\kappa_t(z) + \mu(z) q_t \frac{\partial \kappa_t}{\partial z} + \frac{\sigma^2(z)}{2} q_t \frac{\partial^2 \kappa_t}{\partial z^2} + \frac{\partial\left(q_t \kappa_t\right)}{\partial t}.
\end{aligned} \tag{41}$$

Lemma. $\kappa_t(z) > 1\ \forall z, t$

Proof. The drift of the entrepreneur's capital holdings is

$$s_t^a = \frac{1}{q_t}\left[(\gamma \max\{\varsigma_t z_t \varphi_t - R_t, 0\} + R_t - \delta q_t\right] \geq \frac{R_t - \delta q_t}{q_t}$$

which is expected to hold with strict inequality eventually if $\exists\ \mathbb{P}\left(z_t \geq z_t^*\right) > 0$ (which is satisfied in equilibrium since z is unbounded), and hence

$$\mathbb{E}_0 a_t = \mathbb{E}_0 a_0 e^{\int_0^t s_u^a du} > a_0 e^{\int_0^t \frac{R_s - \delta q_s}{q_s} ds}. \tag{42}$$

The value function is then

$$\begin{aligned}
\kappa_{t_0}(z)\, q_{t_0} a_{t_0} &= V_{t_0}(z, a_{t_0}) \\
&= \mathbb{E}_{t_0} \int_0^\infty e^{-\int_{t_0}^t (r_s + \eta) ds}\left(d_t + \eta q_t a_t\right) dt
\end{aligned}$$

$$
\begin{aligned}
&\geq \mathbb{E}_{t_0}\int_0^\infty e^{-\int_{t_0}^t (r_s+\eta)ds}\eta q_t a_t dt = \mathbb{E}_{t_0}\int_0^\infty e^{-\int_{t_0}^t\left(\frac{\overbrace{R_s-\delta q_s+\dot{q}_s}^{r_s}}{q_s}+\eta\right)ds}\eta q_t a_t dt \\
&= \mathbb{E}_{t_0}\int_0^\infty e^{-\int_{t_0}^t\left(\frac{R_s-\delta q_s}{q_s}+\eta\right)ds-\log\frac{q_t}{q_{t_0}}}\eta q_t a_t dt = \mathbb{E}_{t_0}\int_0^\infty e^{-\int_{t_0}^t\left(\frac{R_s-\delta q_s}{q_s}+\eta\right)ds}\eta q_{t_0} a_t dt \\
&> \mathbb{E}_{t_0}\int_0^\infty e^{-\int_{t_0}^t\left(\frac{R_s-\delta q_s}{q_s}+\eta\right)ds}\eta q_{t_0} a_{t_0} e^{\int_0^t\frac{R_s-\delta q_s}{q_s}ds}dt = \int_0^\infty e^{-\eta t}\eta q_{t_0} a_{t_0} dt = q_{t_0} a_{t_0},
\end{aligned}
$$

where in the first equality we have employed the linear expression of the value function, in the second equation (5), in the third the fact that dividends are non-negative, in the fourth the definition of the real rate (17) and in the last line the inequality (42). Hence $\kappa_{t_0}(z) > 1$ for any t_0.

B.2 Household's problem

We can rewrite the household's problem as

$$
W_t = \max_{C_t, L_t, D_t, B_t^N, S_t^N} \mathbb{E}_0 \int_0^\infty e^{-\rho_t^h t}\left(\frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon\frac{L_t^{1+\vartheta}}{1+\vartheta}\right)dt. \tag{43}
$$

$$
\begin{aligned}
s.t. \quad \dot{D}_t &= \left[\left(R_t - \delta q_t\right)D_t + w_t L_t - C_t - S_t^N + \Pi_t\right]/q_t, & (44)\\
\dot{B_t^N} &= S_t^N + \left(i_t - \pi_t\right)B_t^N, & (45)
\end{aligned}
$$

where S_t^N is the investment into nominal bonds. The Hamiltonian is

$$
\begin{aligned}
&H = \left(\frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon\frac{L_t^{1+\vartheta}}{1+\vartheta}\right) \\
&+\chi_{H,t}\left[\left(\left(R_t-\delta q_t\right)D_t + w_t L_t - C_t - S_t^N + \left(q_t\iota_t - \iota_t - \Phi\left(\iota_t\right)\right)K_t + \Pi_t\right)/q_t\right] \\
&+\eta_t\left[S_t^N + \left(i_t-\pi_t\right)B_t^N\right]
\end{aligned}
$$

The first-order conditions are

$$
C_t^{-\zeta} - \chi_{H,t}/q_t = 0 \tag{46}
$$

$$-\Upsilon L_t^{\vartheta}+\chi_{H,t}w_t/q_t=0 \tag{47}$$

$$-\chi_{H,t}/q_t+\eta_t=0 \tag{48}$$

$$\dot{\chi}_{H,t}=\rho_t^h\chi_{H,t}-\chi_{H,t}\left(R_t-\delta q_t\right)/q_t \tag{49}$$

$$\dot{\eta}_t=\rho_t^h\eta_t-\eta_t\left[\left(i_t-\pi_t\right)\right] \tag{50}$$

(46) and (47) combine to the optimality condition for labor

$$w_t=\frac{L_t^{\vartheta}}{C_t^{-\zeta}}.$$

(46) can be rewritten as

$$\chi_{H,t}=C_t^{-\zeta}q_t,$$

Now take derivative with respect to time

$$\dot{\chi}_{H,t}=C_t^{-\zeta}\,\dot{q}_t-\zeta C_t^{-\zeta-1}\,\dot{C}_tq_t,$$

and plug this into (49)

$$C_t^{-\zeta}\,\dot{q}_t-\zeta C_t^{-\zeta-1}\,\dot{C}_tq_t=\rho_t^hC_t^{-\zeta}q_t-C_t^{-\zeta}q_t\left(R_t-\delta q_t\right)/q_t.$$

Written this way, the Euler equation is linear in a single time derivative, that of the costate $\chi_{H,t}=C_t^{-\zeta}q_t$. It is in this form that the Euler equation enters the central bank's problem. If we divide by $C_t^{-\zeta}q_t\zeta$ and rearrange, we get the Euler equation in the form we are used to have

$$\frac{\dot{C}_t}{C_t}=\frac{\frac{R_t-\delta q_t+\dot{q}_t}{q_t}-\rho_t^h}{\zeta}.$$

(48) can be rewritten as

$$\eta_t=\chi_{H,t}/q_t$$

Now take derivative with respect to time

$$\dot{\eta}_t = \frac{\dot{\chi}_{H,t} q_t - \chi_{H,t} \dot{q}}{q_t^2}$$

Use these two expressions and the definition of $\dot{\chi}_{H,t}$ in (50) to get the second Euler equation

$$\frac{\dot{C}_t}{C_t} = \frac{(i_t - \pi_t) - \rho_t^h}{\zeta}.$$

Combining the two Euler equations, we get the Fisher equation

$$\frac{R_t - \delta q_t + \dot{q}_t}{q_t} = (i_t - \pi_t)$$

Finally using the definition of $r_t \equiv \frac{R_t - \delta q_t + \dot{q}_t}{q_t}$ we can rewrite the first Euler equation and the Fisher equation as in the main text.

B.3 New Keynesian Phillips curve

The proof is equivalent to that of Lemma 1 in Kaplan et al. (2018), though we follow the Hamiltonian approach. We assume the government pays a proportional subsidy τ on the input good, so that the net real price for the retailer is $\tilde{m}_t = m_t(1 - \tau)$. The retailer's problem is

$$\max_{\{\pi, p_t\}_{t \geq 0}} \int_0^\infty e^{\int_0^t -r_s ds} \left[\left(\frac{p_t}{P_t} - \tilde{m}_t \right) \left(\frac{p_t}{P_t} \right)^{-\varepsilon} Y_t - \frac{\theta}{2} \pi_t^2 Y_t \right] dt$$

subject to

$$\dot{p}_t = \pi_t p_t.$$

Note that

$$e^{\int_0^t -r_s ds} = e^{\int_0^t -\frac{\dot{C}_s}{C_s}\zeta - \rho_s ds} = C_0{}^\zeta C_t^{-\zeta} e^{\int_0^t -\rho_s ds}$$

Plug this into the objective, dropping the constant C_0 we can define the Hamiltonian $\mathbb{H}$, where $\chi_{F,t}$ is the multiplier on the constraint:

$$\mathbb{H} = C_t^{-\zeta}\left[\left(\frac{p_t}{P_t} - \tilde{m}_t\right)\left(\frac{p_t}{P_t}\right)^{-\varepsilon} Y_t - \frac{\theta}{2}\pi_t^2 Y_t\right] + \chi_{F,t}\pi_t p_t.$$

The first-order condition (FOC) with respect to p_t is

$$C_t^{-\zeta} Y_t\left[\frac{1}{P_t}\left(\frac{p_t}{P_t}\right)^{-\varepsilon} - \varepsilon\left(\frac{p_t}{P_t} - \tilde{m}_t\right) p_t^{-\varepsilon-1} P_t^{\varepsilon}\right] - \rho_t\chi_{F,t} + \dot{\chi}_{F,t} + \pi_t\chi_{F,t} = 0.$$

The first-order condition with respect to π_t is

$$C_t^{-\zeta}\theta\pi_t Y_t = \chi_{F,t} p_t.$$

Solving the latter FOC for $\chi_{F,t}$ and differentiating with respect to time, we get

$$\dot{\chi}_{F,t} = \frac{\theta\left(C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t\dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}_t\pi_t Y_t\right) p_t - \dot{p}_t C_t^{-\zeta}\theta\pi_t Y_t}{p_t^2}.$$

We multiply the first FOC by p_t and use the latter two equations to eliminate $\chi_{F,t}$

$$\begin{aligned} 0 \ &= \ C_t^{-\zeta} Y_t\left[\frac{p_t}{P_t}\left(\frac{p_t}{P_t}\right)^{-\varepsilon} - \varepsilon\left(\frac{p_t}{P_t} - \tilde{m}_t\right) p_t^{-\varepsilon} P_t^{\varepsilon}\right] - \rho_t C_t^{-\zeta}\theta\pi_t Y_t \\ &+ \ \theta\left(C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t\dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}\pi_t Y_t\right) \\ & \quad \ldots - \frac{\dot{p}_t C_t^{-\zeta}\theta\pi_t Y_t}{p_t} + C_t^{-\zeta}\theta Y_t\pi_t^2. \end{aligned}$$

Since $\frac{\dot{p}_t}{p_t} = \pi_t$, the last two terms cancel. Now use symmetry $P_t = p_t$ and divide by θ to get

$$\frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t\left(\frac{1-\varepsilon}{\varepsilon} + \tilde{m}_t\right) - \rho_t C_t^{-\zeta}\pi_t Y_t + \left(C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t\dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}\pi_t Y_t\right) = 0$$

Notice that the last term in parenthesis is equal to $\frac{dC_t^{-\zeta}\pi_t Y_t}{dt}$. Thus, written this way, the Phillips curve is linear in a single time derivative, related to the firm's costate $\chi_{F,t}$. We exploit this feature when we derive the timeless penalty, and it is in this form the

the Phillips curve enters the Ramsey problem.

If we divide by $C_t^{-\zeta} Y_t$ we obtain the New Keynesian Phillips curve

$$\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon}+\tilde{m}_t\right)+\dot{\pi}_t=\left(\zeta\frac{\dot{C}}{C_t}+\rho_t-\frac{\dot{Y}_t}{Y_t}\right)\pi_t$$

The total profit of retailers, net of the lump-sum tax, which is transferred to the households lump sum, is

$$\Pi_t=\left(1-m_t\right)Y_t-\frac{\theta}{2}\pi_t^2 Y_t. \tag{51}$$

B.4 Distribution

The joint distribution of net worth and productivity is given by the Kolmogorov Forward equation

$$\frac{\partial g_t(z,a)}{\partial t}=-\frac{\partial}{\partial a}[g_t(z,a)s_t(z)a]-\frac{\partial}{\partial z}[g_t(z,a)\mu(z)]+\frac{1}{2}\frac{\partial^2}{\partial z^2}[g_t(z,a)\sigma^2(z)]-\eta g_t(z,a)+\eta/\psi g_t(z,a/\psi), \tag{52}$$

where $1/\psi g_t(z,a/\psi)$ is the distribution of entry firms.

To characterize the law of motion of net-worth shares, defined as $\omega_t(z)=\frac{1}{A_t}\int_0^\infty a g_t(z,a)da$, first we take the derivative of $\omega_t(z)$ wrt time

$$\frac{\partial\omega_t(z)}{\partial t}=-\frac{\dot{A}_t}{A_t^2}\int_0^\infty a g_t(z,a)da+\frac{1}{A_t}\int_0^\infty a\frac{\partial g_t(z,a)}{\partial t}da. \tag{53}$$

Next, we plug in the derivative of $g_t(z,a)$ wrt time from equation(52) into equation (53),

$$\begin{aligned}\frac{\partial\omega_t(z)}{\partial t}=&-\frac{\dot{A}_t}{A_t^2}\int_0^\infty a g_t(z,a)da+\frac{1}{A_t}\int_0^\infty a\left(-\frac{\partial}{\partial a}[g_t(z,a)s_t(z)a]\right)da\\&-\frac{\partial}{\partial z}\mu(z)\frac{1}{A_t}\int_0^\infty a g_t(z,a)da+\frac{1}{2}\frac{\partial^2}{\partial z^2}\sigma^2(z)\frac{1}{A_t}\int_0^\infty a g_t(z,a)da\\&-\frac{1}{A_t}\int_0^\infty \eta a g_t(z,a)da+\frac{1}{A_t}\int_0^\infty \eta a/\psi g_t(z,a/\psi)da.\end{aligned}$$

Using integration by parts and the definition of net worth shares, we obtain the second order partial differential equation that characterizes the law of motion of net-worth shares,

$$\frac{\partial \omega_t(z)}{\partial t} = \left[s_t(z) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta \right] \omega_t(z) - \frac{\partial}{\partial z}\mu(z)\omega_t(z) + \frac{1}{2}\frac{\partial^2}{\partial z^2}\sigma^2(z)\omega_t(z). \tag{54}$$

The stationary distribution is therefore given by the following second order partial differential equation,

$$0 = (s(z) - (1-\psi)\eta)\,\omega(z) - \frac{\partial}{\partial z}\mu(z)\omega(z) + \frac{1}{2}\frac{\partial^2}{\partial z^2}\sigma^2(z)\omega(z). \tag{55}$$

B.5 Market clearing and aggregation

Define the cumulative function of net-worth shares as

$$\Omega_t(z) = \int_0^z \omega_t(z)dz. \tag{56}$$

Using the optimal choice for k_t from equation (7), we obtain

$$K_t = \int k_t(z,a)dg_t(z,a) = \int_{z_t^*}^{\infty} \left[\int \gamma a \frac{1}{A_t} g_t(z,a)da \right] dz A_t = \gamma(1-\Omega(z_t^*))A_t. \tag{57}$$

By combining equations (25), (26) and (57), and solving for A_t,we obtain

$$A_t = \frac{D_t}{\gamma(1-\Omega(z_t^*)) - 1}, \tag{58}$$

Labor market clearing implies

$$L_t = \int_0^{\infty} l_t(z,a)dg_t(z,a). \tag{59}$$

Define the following auxiliary variable,

$$X_t \equiv \int_{z_t^*}^{\infty} z\omega_t(z)dz = \mathbb{E}\left[z \mid z > z_t^*\right](1-\Omega(z_t^*)). \tag{60}$$

Using labor demand from (8) , X_t and using the definition of φ_t, we obtain

$$L_t = \int_{z_t^*}^{\infty} \int \left(\frac{\varphi_t}{\alpha m_t}\right)^{\frac{1}{1-\alpha}} \varsigma_t z \gamma a_t g_t(z,a) da dz = \left(\frac{\varphi_t}{\alpha m_t}\right)^{\frac{1}{1-\alpha}} \gamma A_t \varsigma_t X_t. \tag{61}$$

Plugging in (8) into production function (1), and using again the definition of shares, we obtain

$$Y_t = \int_{z_t^*}^{\infty} \int \underbrace{\frac{\varsigma_t z \varphi_t}{\alpha m_t} \gamma a}_{y_t(z,a)} g_t(z,a) da dz = \varsigma_t \frac{\varphi_t}{\alpha m_t} X_t \gamma A_t = (\gamma \varsigma_t X_t A_t)^{\alpha} L_t^{1-\alpha}. \tag{62}$$

If we employ equation (57), we get

$$Y_t = \left(\varsigma_t \frac{X_t}{(1-\Omega(z_t^*)).}\right)^{\alpha} K_t^{\alpha} L_t^{1-\alpha} = Z_t K_t^{\alpha} L_t{}^{1-\alpha}, \tag{63}$$

where the TFP term Z_t is defined as

$$Z_t = \left(\varsigma_t \frac{X_t}{(1-\Omega(z_t^*))}\right)^{\alpha} = \left(\varsigma_t \mathbb{E}\left[z \mid z > z_t^*\right]\right)^{\alpha}. \tag{64}$$

Aggregate profits of retailers are given by

$$\Phi_t^{Agg} = \int \gamma \max\left\{\varsigma_t z \varphi_t - R_t, 0\right\} a dG_t(z,a) = \left[\varsigma_t \varphi_t X_t - R_t \left(1 - \Omega(z^*)\right)\right] \gamma A_t. \tag{65}$$

Aggregating the budget constraint of all input good firms, using the linearity of savings policy (11) and using (58), we obtain

$$\begin{aligned} \dot{A}_t &= \int \dot{a} dG(z,a,t) - \eta \int (1-\psi) a dG(z,a,t) \; = \\ &= \int_0^{\infty} \frac{1}{q_t} \left[\gamma \max\left\{\varsigma_t z \varphi_t - R_t, 0\right\} + R_t - \delta q_t - q_t(1-\psi)\eta\right] a dG(z,a), \end{aligned} \tag{66}$$

Dividing by A_t both sides of this equation, using the definition of net worth shares and the fact that these integrate up to one, we obtain

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}(\gamma \varphi_t \varsigma_t X_t - R_t \gamma (1-\Omega(z_t^*)) + R_t - \delta q_t - q_t(1-\psi)\eta). \tag{67}$$

Using the definition of Z_t, and substituting φ_t using equation (61), we can simplify equation (67) as

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}(1-\Omega(z_t^*))\gamma(\alpha m_t Z_t {L_t}^{1-\alpha}\left((1-\Omega(z_t^*))\gamma A_t\right)^{\alpha-1} - R_t) + R_t - \delta q_t - q_t(1-\psi)\eta). \tag{68}$$

Using (57) and (58) we can replace $(1-\Omega(z_t^*))\gamma A_t$ by K_t, which delivers equation (33).

Finally, we can obtain factor prices

$$w_t = (1-\alpha) m_t Z_t {K_t}^{\alpha} {L_t}^{-\alpha} \tag{69}$$

$$R_t = \alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} \frac{z_t^*}{\gamma X_t} \tag{70}$$

where wages come from substituting the definition of φ_t into equation (61); and interest rates come from plugging in the wage expression (69) into the cutoff rule (10) and using equation (58). We could equivalently write equation (70) in terms of real rate of return r_t :

$$r_t = \frac{1}{q_t}\left(\alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} \frac{z_t^*}{\gamma X_t}\right) - \delta + \frac{\dot{q}}{q_t} \tag{71}$$

We can easily get these equations in terms of aggregate net worth instead of capital by simply using equation (57), i.e. $A_t\gamma(1-\Omega(z_t^*)) = K_t$, and using that $\varsigma_t \mathbb{E}\left[z \mid z > z_t^*\right] = \varsigma_t \frac{X_t}{(1-\Omega(z_t^*))} = \frac{\int_{z_t^*}^{\infty} \varsigma_t z \omega_t(z) dz}{(1-\Omega(z_t^*))}$ (see equation 64).

B.6 Full set of equations

Conditional on some monetary policy, the competitive equilibrium is described by the following 14 equations, for the 15 variables $\{\omega(z), w, r, q, R, A, L, C, D, Z, z^*, \pi, m, i, Y\}$. Remember that $\mu(z) = z\left(-\varpi_z \log z + \frac{\sigma^2}{2}\right)$ and $\sigma(z) = \sigma_z z$, and that government bonds are in zero net supply ($B_t^N = 0$, hence $S_t^N = 0$). Note that there are more variables here than in the equilibrium definition, because we have introduced several auxiliary variables.

To close the model we need to define a particular policy. This can be a simple Taylor rule, such as the one we assume for the positive part of the paper:

$$di = -\upsilon\left(i_t - \left(\rho_t^h + \phi\left(\pi_t - \bar{\pi}\right) + \bar{\pi}\right)\right)dt.$$

or it can be Ramsey policy.When we consider Ramsey policy, these 15 equations (in appropriate transformations) are the constraints of the Ramsey problem.

- Household block:

$$\frac{\dot{C}_t}{C_t} = \frac{r_t - \rho_t^h}{\zeta}, \tag{72}$$

$$r_t = i_t - \pi_t, \tag{73}$$

$$r_t = \frac{R_t - \delta q_t + \dot{q}_t}{q_t}, \tag{74}$$

$$w_t = \frac{\Upsilon L_t^{\vartheta}}{C_t^{-\zeta}}, \tag{75}$$

$$\dot{D}_t = \left[\left(R_t - \delta q_t\right) D_t + w_t L_t - C_t + T_t\right]/q_t, \tag{76}$$

$$\text{where } T_t = \left(1 - m_t\right) Y_t - \frac{\theta}{2}\pi_t^2 Y_t + q_t(1-\psi)\eta A_t + \left(q_t - 1\right)\delta.$$

- Production block:

$$\frac{\partial \omega_t\left(z\right)}{\partial t} = \left(s_t(z) - (1-\psi)\eta - \frac{\dot{A}_t}{A_t}\right)\omega_t\left(z\right) \tag{77}$$

$$-\frac{\partial}{\partial z}\left[\mu(z)\omega_t\left(z\right)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)\omega_t\left(z\right)\right], \tag{78}$$

$$\text{where } s_t(z) \equiv \frac{1}{q_t}(\gamma \max\left\{z_t \varsigma_t \varphi_t - R_t, 0\right\} + R_t - \delta q_t)$$

$$\text{and } \varphi_t \equiv \alpha\left(\frac{(1-\alpha)}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}},$$

$$w_t = (1-\alpha)m_t Z_t K^{\alpha} {L_t}^{-\alpha}, \tag{79}$$

$$R_t = \alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha}\frac{z_t^*}{\mathbb{E}\left[z \mid z > z_t^*\right]}, \tag{80}$$

$$Z_t = \left(\varsigma_t \mathbb{E}\left[z \mid z > z_t^*\right]\right)^{\alpha} \tag{81}$$

$$\text{where } \mathbb{E}\left[z \mid z > z_t^*\right] \equiv \frac{\int_{z_t^*}^{\infty} z\omega_t(z)dz}{(1-\Omega(z_t^*))}$$

$$\text{and } \Omega_t(z^*) \equiv \int_0^{z^*} \omega_t\left(z\right)\mathrm{dz},$$

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}\left[\gamma(1-\Omega(z_t^*))\left(\alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} - R_t\right) + R_t - \delta q_t - q_t(1-\psi)\eta)\right], \tag{82}$$

$$\left(r_t - \frac{\dot{Y}_t}{Y_t}\right)\pi_t = \frac{\varepsilon}{\theta}\left(m_t(1-\tau_t) - m^*\right) + \dot{\pi}_t, \quad m^* = \frac{\varepsilon - 1}{\varepsilon}, \tag{83}$$

$$Y_t = Z_t K_t^{\alpha} {L_t}^{1-\alpha}. \tag{84}$$

- Market clearing:

$$K = A_t + D_t, \tag{85}$$

$$A_t = \frac{D_t}{\gamma(1-\Omega(z_t^*)) - 1}. \tag{86}$$

B.7 Proofs of subsection 4.2

To isolate the role of endogenous TFP, we consider the case without aggregate exogenous TFP shocks, i.e. $\varsigma_t = 1 \ \ \forall t$. In this case, total factor productivity is given by

$$Z_t = \left(\frac{\int_{z_t^*}^{\infty} z\omega_t\left(z\right)dz}{\int_{z_t^*}^{\infty} \omega_t\left(z\right)dz} \right)^{\alpha}.$$

We compute the growth rate of TFP

$$\begin{aligned}
\frac{1}{Z_t}\frac{dZ_t}{dt} &= \frac{d\log Z_t}{dt} = \alpha\left[\frac{d}{dt}\left(\log\int_{z_t^*}^{\infty} z\omega_t\left(z\right)dz\right) - \frac{d}{dt}\left(\log\int_{z_t^*}^{\infty}\omega_t\left(z\right)dz\right)\right] \\
&= \alpha \quad \left[\frac{\int_{z_t^*}^{\infty} z\frac{\partial\omega_t(z)}{\partial t}dz - z_t^*\omega_t(z_t^*)\frac{dz_t^*}{dt}}{\int_{z_t^*}^{\infty} z\omega_t\left(z\right)dz} + \frac{-\int_{z_t^*}^{\infty}\frac{\partial\omega_t(z)}{\partial t}dz + \omega_t(z_t^*)\frac{dz_t^*}{dt}}{\int_{z_t^*}^{\infty}\omega_t\left(z\right)dz}\right],
\end{aligned}$$

where the dynamics of the density are

$$\begin{aligned}
\frac{\partial\omega_t\left(z\right)}{\partial t} &= \left[\underbrace{\frac{\gamma\varphi_t}{q_t}\max\left\{\left(z-z^*\right),0\right\}}_{\equiv\tilde{\Phi}_t(z)} + \underbrace{\frac{R_t-\delta q_t}{q_t} - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta}_{\equiv\tilde{\Xi}_t}\right]\omega_t(z) \\
&+ \varsigma_z\frac{\partial}{\partial z}\left(\log(z)\omega_t(z)\right) + \frac{\sigma_z^2}{2}\frac{\partial^2}{\partial z^2}\omega_t(z).
\end{aligned} \tag{87}$$

From there we can analyze two limit cases.

B.7.1 Constant cutoff z^*

First, we analyze the case in which the cutoff remains approximately constant. In this case, the growth rate is

$$\left.\frac{1}{Z_t}\frac{dZ_t}{dt}\right|_{z^*} = \frac{\int_{z^*}^{\infty} z\frac{\partial\omega_t(z)}{\partial t}dz}{\int_{z^*}^{\infty} z\omega_t\left(z\right)dz} - \frac{\int_{z^*}^{\infty}\frac{\partial\omega_t(z)}{\partial t}dz}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}.$$

We now show that, in this case, (i) prices only influence TFP though changes in the slope of the excess investment rate, $\frac{\gamma\varphi_t}{q_t}$; and (ii) that this response is positive. The derivative of the TFP growth rate with respect to a price or a function of prices x_t is

$$\frac{\partial}{\partial x_t}\left.\frac{d\log Z_t}{dt}\right|_{z^*}=\frac{\int_{z^*}^{\infty}z\frac{\partial\dot{\omega}_t(z)}{\partial x_t}dz}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}\frac{\partial\dot{\omega}_t(z)}{\partial x_t}dz}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz},$$

where

$$\left.\frac{\partial\dot{\omega}_t\left(z\right)}{\partial x_t}\right|_{z^*}=\left.\frac{\partial}{\partial x_t}\left(\tilde{\Phi}_t(z)+\tilde{\Xi}_t\right)\right|_{z^*}\omega\left(z\right),$$

given the definitions of $\tilde{\Phi}_t(z)$ and $\tilde{\Xi}_t$ above. Then we have:

$$\begin{aligned}\frac{\partial}{\partial x_t}\left.\frac{d\log Z_t}{dt}\right|_{z^*}&=\frac{\int_{z^*}^{\infty}z\frac{\partial\tilde{\Phi}(z)}{\partial x_t}\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}\frac{\partial\tilde{\Phi}(z)}{\partial x_t}\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}\\&+\frac{\partial\tilde{\Xi}_t}{\partial x_t}\underbrace{\left(\frac{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}\right)}_{0}.\end{aligned}$$

This expression shows how only the excess investment rate $\tilde{\Phi}(z)$ matters to understand the impact of changes in prices on the growth rate of TFP. Conditional on z^*, price changes affect the excess investment rate by affecting its slope $\frac{\gamma\varphi_t}{q_t}$. So the effect of a shock on TFP growth is determined by its effect on $\frac{\gamma\varphi_t}{q_t}$. This proves claim (i).

To prove that an increase in the slope $\frac{\gamma\varphi_t}{q_t}$ increases TFP growth, we compute

$$\begin{aligned}\frac{\partial}{\partial\left(\frac{\gamma\varphi_t}{q_t}\right)}\left.\frac{d\log Z_t}{dt}\right|_{z^*}&=\frac{\int_{z^*}^{\infty}z\frac{\partial\tilde{\Phi}_t(z)}{\partial\left(\frac{\gamma\varphi_t}{q_t}\right)}\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}\frac{\partial\tilde{\Phi}_t(z)}{\partial\left(\frac{\gamma\varphi_t}{q_t}\right)}\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}\\&=\frac{\int_{z^*}^{\infty}z\left(z-z^*\right)\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}z\left(z-z^*\right)\omega_t\left(z\right)dz}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz},\end{aligned}$$

To uncover the sign, we analyze the term

$$\frac{\int_{z^*}^{\infty}\left(z-z^*\right)z\omega_t\left(z\right)}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}\left(z-z^*\right)\omega_t\left(z\right)}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}=\frac{\int_{z^*}^{\infty}z^2\omega_t\left(z\right)}{\int_{z^*}^{\infty}z\omega_t\left(z\right)dz}-\frac{\int_{z^*}^{\infty}z\omega_t\left(z\right)}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}.\tag{88}$$

We define $\bar{\omega}_t(z)\equiv\frac{\omega_t(z)}{\int_{z^*}^{\infty}\omega_t(z)dz}\mathbb{I}_{z>z^*}$ and $\tilde{\omega}_t(z)\equiv\frac{z\omega_t(z)}{\int_{z^*}^{\infty}\omega_t(z)zdz}\mathbb{I}_{z>z^*}$. These are continuous probability density functions over the domain $[z^*,\infty)$, as they are non-negative and sum up to 1. They satisfy the monotone likelihood ratio condition as

$$I\left(z\right)=\frac{\tilde{\omega}_{t}(z)}{\bar{\omega}_{t}(z)}=z\frac{\int_{z^{*}}^{\infty}z\omega_{t}\left(z\right)dz}{\int_{z^{*}}^{\infty}\omega_{t}\left(z\right)dz}$$

is non decreasing. This implies that function $\tilde{\omega}_t(z)$ dominates $\bar{\omega}_t(z)$ first-order stochastically. Hence

$$\frac{\int_{z^{*}}^{\infty}z\omega_{t}\left(z\right)}{\int_{z_{t}^{*}}^{\infty}\omega_{t}\left(z\right)dz}=\mathbb{E}_{\bar{\omega}_{t}(z)}\left[z\right]=\int_{z^{*}}^{\infty}z\bar{\omega}_{t}(z)zdz\quad<\quad\int_{z_{t}^{*}}^{\infty}z\tilde{\omega}_{t}\left(z\right)dz=\mathbb{E}_{\tilde{\omega}_{t}(z)}\left[z\right]=\frac{\int_{z^{*}}^{\infty}z^{2}\omega_{t}\left(z\right)}{\int_{z_{t}^{*}}^{\infty}z\omega_{t}\left(z\right)dz}.$$

Therefore, equation (88) is positive. An increase in the slope of the excess investment rate, $\frac{\gamma\varphi_t}{q_t}$, thus increases TFP growth, which proves claim (ii):

$$\frac{\partial}{\partial\left(\frac{\gamma\varphi_{t}}{q_{t}}\right).}\left.\frac{d\log Z_{t}}{dt}\right|_{z^{*}}=\frac{\int_{z^{*}}^{\infty}z^{2}\omega_{t}\left(z\right)}{\int_{z_{t}^{*}}^{\infty}z\omega_{t}\left(z\right)dz}-\frac{\int_{z^{*}}^{\infty}z\omega_{t}\left(z\right)}{\int_{z_{t}^{*}}^{\infty}\omega_{t}\left(z\right)dz}>0.$$

B.7.2 Constant $\omega(z)$: the limit with iid shocks

Next, we consider the limit of iid shocks, that is, the limit as $\varsigma_z\to\infty$. In this case, as discussed in Itskhoki and Moll (2019), the distribution $\omega\left(z\right)$ is constant and the growth rate of TFP simplifies to

$$\frac{1}{Z_{t}}\frac{dZ_{t}}{dt}=\alpha\omega(z_{t}^{*})\frac{\int_{z_{t}^{*}}^{\infty}\left(z-z_{t}^{*}\right)\omega\left(z\right)dz}{\int_{z_{t}^{*}}^{\infty}\omega\left(z\right)dz\int_{z_{t}^{*}}^{\infty}z\omega\left(z\right)dz}\frac{dz_{t}^{*}}{dt}.$$

Notice that $\alpha\omega(z_t^*)\frac{\int_{z_t^*}^{\infty}(z-z_t^*)\omega(z)dz}{\int_{z_t^*}^{\infty}\omega(z)dz\int_{z_t^*}^{\infty}z\omega(z)dz}>0$ for any value of the cutoff. In this case, the growth rate of TFP depends linearly with the growth rate of the cutoff: if the later increases, so does the former.

B.7.3 Corollary 1

The market clearing condition for capital ((27)) can be rewritten as

$$K\quad=\quad A_{t}\gamma\int_{z_{t}^{*}}^{\infty}\omega_{t}\left(x\right)dx. \tag{89}$$

Take the time derivative on both sides, taking into account that aggregate capital is constant and manipulating the right-hand side,

$$
\begin{aligned}
\dot{K}_t &= \dot{A}_t \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx + A_t \frac{\partial \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx}{\partial t}, \\
0 &= \dot{A}_t \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx + A_t \gamma \left(\int_{z_t^*}^{\infty} \frac{\partial \omega_t(z)}{\partial t} dz - \omega_t(z_t^*) \frac{\partial z_t^*}{\partial t} \right).
\end{aligned}
$$

Now take the derivative with respect to the slope $\frac{\gamma \varphi_t}{q_t}$,

$$
\begin{aligned}
0 &= \frac{\partial \left\{ \dot{A}_t \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx + A_t \gamma \left(\int_{z_t^*}^{\infty} \frac{\partial \omega_t(z)}{\partial t} dz - \omega_t(z_t^*) \frac{\partial z_t^*}{\partial t} \right) \right\}}{\partial \frac{\gamma \varphi_t}{q_t}} && (90) \\
&= \frac{\partial \left\{ \dot{A}_t \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx \right\}}{\partial \frac{\gamma \varphi_t}{q_t}} + \frac{\partial \left\{ A_t \gamma \left(\int_{z_t^*}^{\infty} \frac{\partial \omega_t(z)}{\partial t} dz - \omega_t(z_t^*) \frac{\partial z_t^*}{\partial t} \right) \right\}}{\partial \frac{\gamma \varphi_t}{q_t}} && (91) \\
&= \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx \frac{\partial \dot{A}_t}{\partial \frac{\gamma \varphi_t}{q_t}} + A_t \gamma \frac{\partial \left(\int_{z_t^*}^{\infty} \frac{\partial \omega_t(z)}{\partial t} dz - \omega_t(z_t^*) \frac{\partial z_t^*}{\partial t} \right)}{\partial \frac{\gamma \varphi_t}{q_t}} && (92)
\end{aligned}
$$

Note that here we have used the fact that $\omega_t(x)$ and A_t do not depend on current prices because they are states, but $\dot{A}_t$ does. Furthermore, note that z_t^* also does not depend on current prices, by the capital market clearing condition (2.7).

By (87) we have that

$$
\frac{\frac{\partial \omega_t(z)}{\partial t}}{\partial \frac{\gamma \varphi_t}{q_t}} = \frac{\partial \tilde{\Phi}_t(z)}{\partial \frac{\gamma \varphi_t}{q_t}} = (z - z_t^*)\, \omega_t(z).
$$

By (66), and the definition of z_t^* ((10)), we have that

$$
\begin{aligned}
\dot{A}_t &= \int_0^{\infty} \left[\frac{\gamma \varphi_t}{q_t} \max\{z - z_t^*, 0\} + \frac{R_t}{q_t} - \delta - (1-\psi)\eta \right] \omega_t(z) A_t dz, \\
&= \int_0^{\infty} \left[\frac{\gamma \varphi_t}{q_t} \max\{z - z_t^*, 0\} + \frac{z_t^*}{\gamma} \frac{\gamma \varphi_t}{q_t} - \delta - (1-\psi)\eta \right] \omega_t(z) A_t dz, \\
\Rightarrow \frac{\partial \dot{A}_t}{\partial \frac{\gamma \varphi_t}{q_t}} &= \int_{z_t^*}^{\infty} (z - z_t^*)\, \omega_t(z) A_t dz + \frac{z_t^*}{\gamma} \omega_t(z) A_t.
\end{aligned}
$$

Using these two equations in (90), we get the following

$$
\begin{aligned}
0 &= \gamma \int_{z_t^*}^{\infty} \omega_t\left(x\right) dx \left(\int_{z_t^*}^{\infty} \left(z - z_t^*\right) \omega_t(z) A_t dz + \frac{z_t^*}{\gamma} \omega_t(z) A_t \right) \\
&+ A_t \gamma \int_{z_t^*}^{\infty} \left(z - z_t^*\right) \omega_t\left(z\right) dz - A_t \gamma \omega_t(z_t^*) \frac{\partial \frac{\partial z_t^*}{\partial t}}{\partial \frac{\gamma \varphi_t}{q_t}}.
\end{aligned}
$$

Thus

$$
\frac{\partial \frac{\partial z_t^*}{\partial t}}{\partial \frac{\gamma \varphi_t}{q_t}} = \frac{\left(\int_{z_t^*}^{\infty} \left(z - z_t^*\right) \omega_t(z) dz \gamma + z_t^* \omega_t(z) \right) \int_{z_t^*}^{\infty} \omega_t\left(x\right) dx + \int_{z_t^*}^{\infty} \left(z - z_t^*\right) \omega_t\left(z\right) dz}{\omega_t(z_t^*)} > 0
$$

This expression is positive. Thus we have shown that an increase in the slope of the excess investment function ($\frac{\gamma \varphi_t}{q_t}$) is associated to an increase in the growth rate of the threshold ($\frac{\partial z_t^*}{\partial t}$). Combining this with propositions 1 and 2, we can conclude that an increase in $\frac{\gamma \varphi_t}{q_t}$ is associated to and increase in TFP.

B.8 Baseline vs complete markets

In this Appendix we want to highlight the differences between the model presented in this paper and the standard representative agent New Keynesian model with fixed capital (complete markets). Note first that the baseline economy collapses to the standard complete-markets economy if the collateral constraint is made infinitely slack (assuming that the support of entrepreneurs productivity distribution is bounded above). In that case entrepreneurial net worth becomes irrelevant and only the entrepreneur with the highest level of productivity z_t produces, since she can frictionlessly lend all the capital in the economy. Her productivity determines aggregate productivity $Z_t = (z_t^{max})^{\alpha}$.In contrast, in the baseline model with incomplete markets, entrepreneurs' firms can only use capital up to a multiple γ of their net worth , i.e. $\gamma a_t \leq k_t$. Thus entrepreneurs need to accumulate net worth (in units of capital) to alleviate these financial frictions. Hence, in the baseline model, the distribution of aggregate capital across entrepreneurs and the representative household matters and aggregate productivity depends on the expected productivity of constrained firms, $Z_t = \left(\varsigma_t \mathbb{E}\left[z \mid z > z_t^*\right]\right)^{\alpha}$. The rest of the agents (retailers, final good producers, capital producer) are identical in both economies.

Below we report the equilibrium conditions for the complete-markets economy.

Comparing them with those of the baseline economy shows that they are identical except for the treatment of productivity: in the baseline, TFP is endogenous, $Z_t = (\varsigma_t \mathbb{E}\left[z \mid z > z_t^*\right])^\alpha$, whereas under complete markets it is exogenous, $Z_t = (\varsigma_t)^\alpha$. In addition, the condition equating the cost of capital R_t to the marginal return on capital differs only by the presence of an endogenous adjustment term in the baseline model.

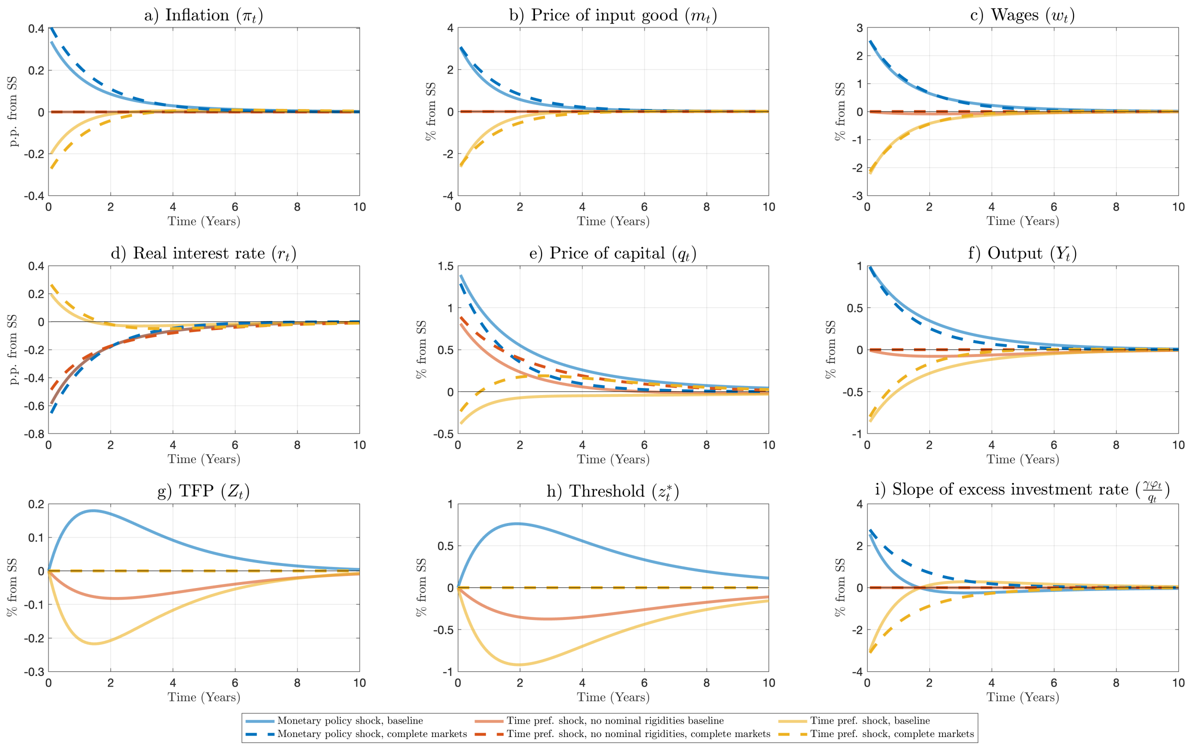
The competitive equilibrium of the complete-markets model with capital consists of the following 13 equations for the 13 variables $\{w, r, R, q, L, C, D, \pi, m, \tilde{m}, i, Y, T\}$:

$$
\begin{aligned}
\tilde{m}_t &= m_t(1-\tau_t) \\
w_t &= (1-\alpha) m_t Z_t K^\alpha L_t{}^{-\alpha} \\
R_t &= \alpha m_t Z_t K^{\alpha-1} L^{1-\alpha} \\
K &= D_t \\
\frac{\dot{C}_t}{C_t} &= \frac{r_t - \rho_t^h}{\zeta} \\
w_t &= \frac{\Upsilon L_t^\vartheta}{C_t^{-\zeta}} \\
\dot{D}_t &= \left[\left(R_t - \delta q_t\right) D_t + w_t L_t - C_t + T_t\right] / q_t \\
r_t &= i_t - \pi_t \\
r_t &= \frac{R_t - \delta q_t + \dot{q}_t}{q_t} \\
\left(r_t - \frac{\dot{Y}_t}{Y_t}\right) \pi_t &= \frac{\varepsilon}{\theta}\left(\tilde{m}_t - m^*\right) + \dot{\pi}_t, \qquad m^* = \frac{\varepsilon - 1}{\varepsilon} \\
Y_t &= Z_t K^\alpha L_t{}^{1-\alpha} \\
T_t &= \left(1 - m_t\right) Y_t - \frac{\theta}{2} \pi_t^2 Y_t + \left[q_t - 1\right] \delta K \\
di &= -\upsilon\left(i_t - \left(\rho_t^h + \phi\left(\pi_t - \bar{\pi}\right) + \bar{\pi}\right)\right) dt.
\end{aligned}
$$

C Further numerical results

C.1 Baseline vs complete markets

Figure 9: Impulse responses.

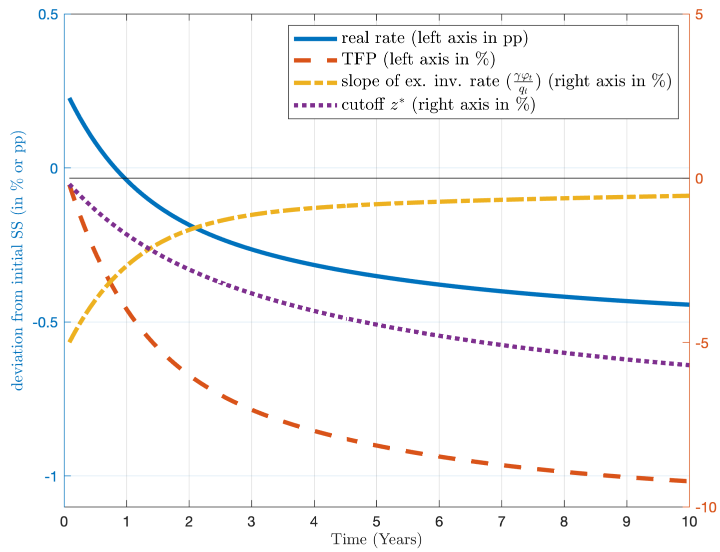


Notes: Deviations from steady state. The semi- transparent blue line is the response in the baseline model to a 25bp monetary policy shock. The semi-transparent orange solid line is the response of the baseline model *with flexible prices* to a time-preference shock whose path is chosen such that the path of real interest rates coincides with that of the monetary policy shock, as in Figure 2. Note that the two semi-transparent lines for the real interest rates are on top of each other, appearing in a brown color. The semi-transparent yellow line is the response to the same time-preference shock, but in the baseline model *with nominal rigidities*. Those three lines coincide with those in figure 2. The dashed lines correspond to the complete-markets counterpart. The blue dashed line is the response in the complete-markets representative agent New Keynesian model to a 25bp monetary policy shock. The dashed orange and dashed yellow lines are the responses to the same time-preference shock as in the baseline model, but in the complete-markets economy *without nominal rigidities* (dashed orange) and *with nominal rigidities* (dashed yellow).

C.2 Dynamics after a permanent real interest rate decline

Figure 10 displays the transition dynamics toward a new steady state following a permanent decline in the household discount rate, ρ^h, from 1% to 0.5%. It shows that the decline in real rates (solid blue line) is accompanied by a decline in TFP (dashed orange line). This is both a consequence of the decline in the threshold (dotted purple line) and the lower slope of the excess investment rate (dashed-dotted yellow line), which increase the share of low-MRPK firms in production. The initial increase in real rates is a consequence of the nominal rigidities and the Taylor rule: as nominal rates do not decrease as fast as the natural rate on impact, it initially produces a fall in inflation that mechanically increases the real rate. As nominal rates progressively adjust, this effect disappears after one year.

Figure 10: Transition to a low-real-rate steady state.

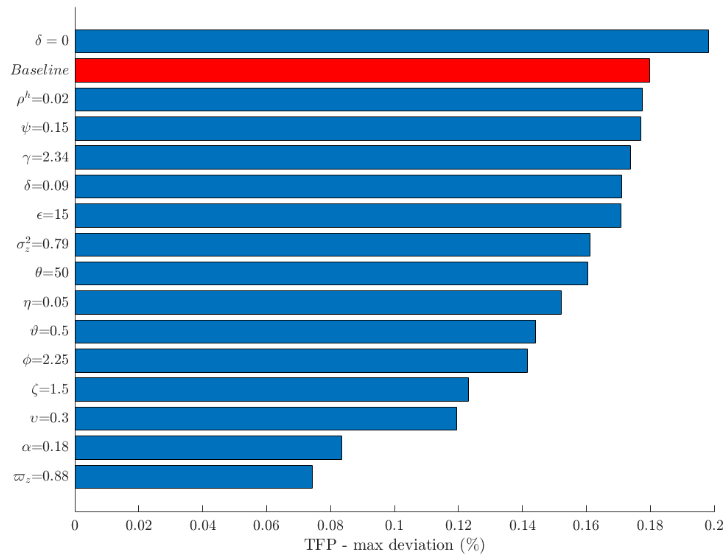


Notes: The figure shows the paths after an unexpected and permanent decline in the household's discount rate from 1% to 0.5% expressed in deviations from the initial steady state. The lines depict real rates r (solid blue), TFP Z (dashed orange), the slope of the excess investment function (dashed-dotted yellow) and the threshold z^* (dotted purple line).

C.3 Robustness of misallocation channel

We perform a sensitivity analysis of the capital misallocation channel, that is, the increase in TFP following an expansionary monetary policy shock, when the Central Bank follows a standard and exogenous Taylor rule. We vary, one at a time, the key structural parameters $\zeta, \alpha, \eta, \vartheta, \delta, \epsilon, \varpi_z, \sigma_z^2, \gamma, \rho_h, \theta, \psi, \tau, \upsilon, \phi$ by $\pm 50\%$ and recompute the general equilibrium response to the monetary expansionary shock of -25bps. A few parameters require restricted changes: for γ, the 50% decrease would violate the lower bound, so we set $\gamma = 1$; for α, we increase it by 25% to remain within plausible values; for risk aversion, we reduce ζ by 25%; for the steady state subsidy, we set it to zero, $\tau = 0$. For each parameter, *we only show the results of the variation for which the peak of TFP is the lowest.* We also solve the model under $\delta = 0$ and show it for

Figure 11: Sensitivity of TFP response to monetary policy expansion.



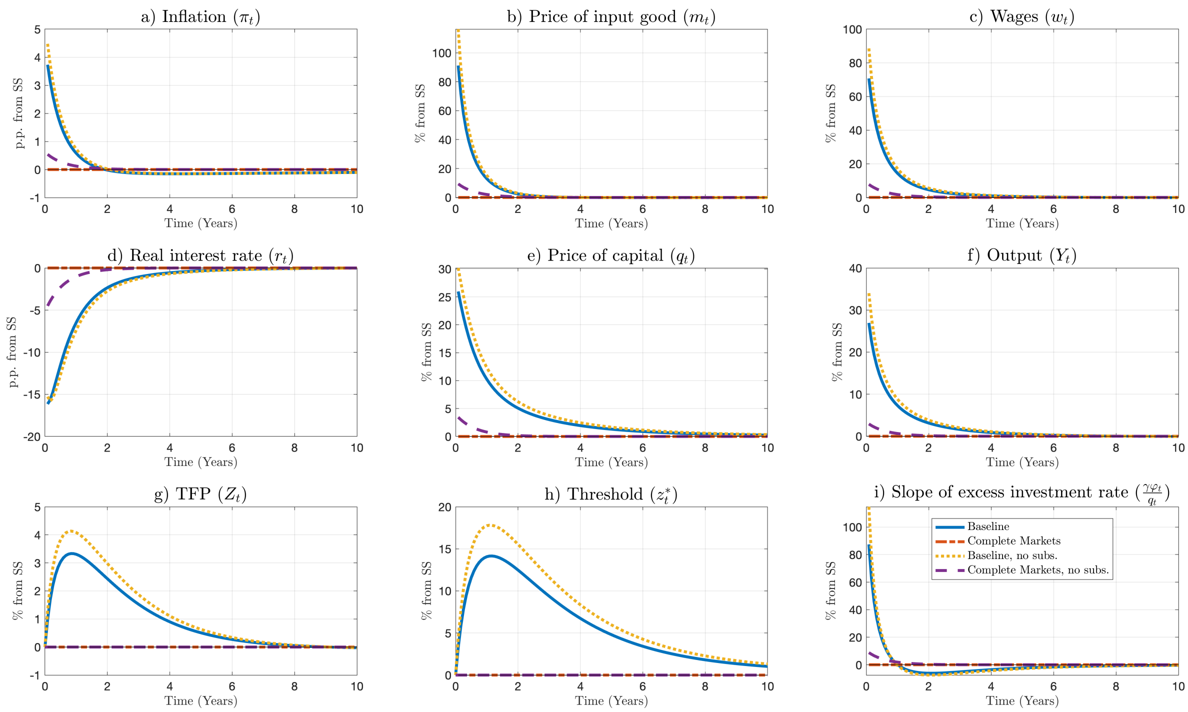
Notes: This figure reports the sensitivity analysis of the response of TFP after an expansionary monetary policy shock ok -25bps. The Figure shows the peak response of TFP, in % deviations from steady state.

completeness. As Figure 11 shows, there is always a positive response of TFP following an expansionary monetary policy shock.

C.4 Time-0 problem

We complement Figure 4 by reporting the full set of nine equilibrium variables under the Ramsey problem without pre-commitment and in the absence of shocks. The solid blue line corresponds to the baseline economy, while the dashed orange line depicts the complete-markets benchmark. The dotted yellow and purple dashed lines repeat the same experiment but without the subsidy that offsets the steady-state markup distortion. As already emphasized in the main text, these additional panels make clear that the time-0 problem is far more pronounced in the baseline incomplete-markets economy than in the standard complete-markets model.

Figure 12: Time-0 problem

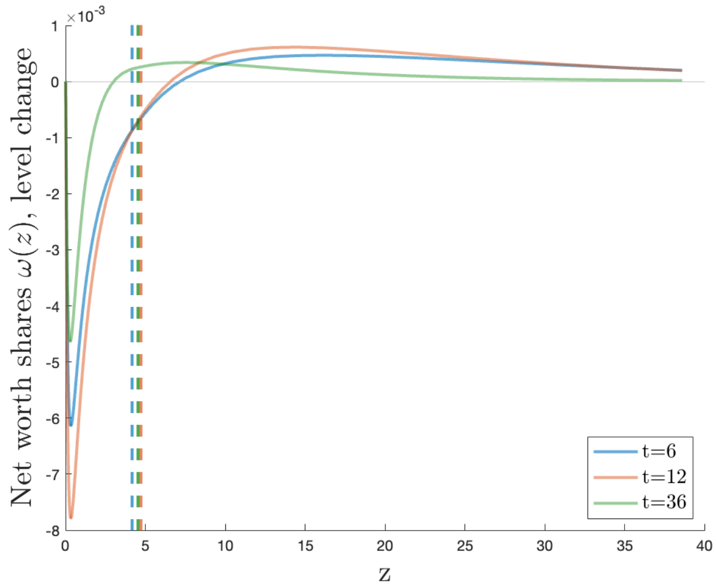


Notes: The figure shows the deviations from steady state of the economy when the planner solves the Ramsey problem without pre-commitments and in the absence of shocks. The baseline economy is the solid blue line, and the complete-markets economy the dashed orange line. The dotted yellow line and the purple dashed line repeat the same exercise in the absence of the subsidy that undoes the markup distortion.

Figure 13 illustrates the distributional effects associated with the time-0 problem in the baseline economy. The solid lines show the deviations of the net-worth share distribution from its steady state at three horizons: 6 months (blue), 12 months (orange), and 36 months (green). The corresponding vertical dashed lines, in matching colors, display the movements in the productivity threshold at each horizon.

The figure shows a clear reallocation of net worth toward more productive entrepreneurs: their wealth shares rise, while those of low-productivity entrepreneurs fall. Consistent with the previous figure, the threshold shifts to the right, implying that the marginal entrepreneur shrinks their optimal size to 0. This movement in the distribution is the mechanism through which the time-0 policy intervention affects capital allocation and endogenous TFP.

Figure 13: Time-0 problem: distribution



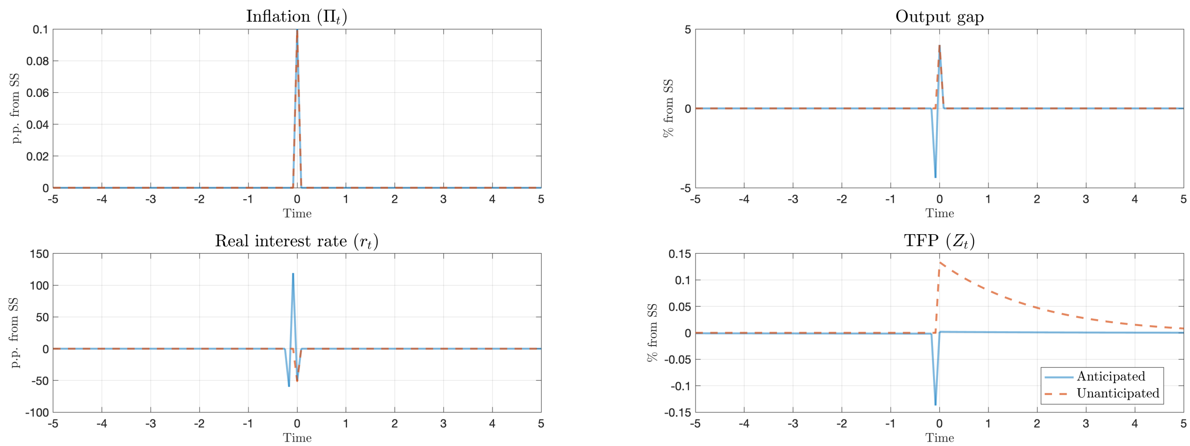
Notes: The figure shows the changes in net-worth share distribution from SS at $t = 6$ (blue), after 12 months (orange), and after 36 months (green), together with the threshold z^* at each of these points in time in an horizontal dashed line.

C.5 Inflation from the perspective of time-0 and from the timeless perspective

Figure 14 illustrates the intuition behind why the Ramsey steady state features zero inflation (Proposition 3), and why the penalty makes the Ramsey steady state optimal (5). It shows the effect of a positive deviation from the zero inflation policy for one month in the absence of any fundamental shocks, starting at the zero-inflation steady state . This can be seen as a particularly short-lived monetary policy shock, or as the partial derivative of the equilibrium dynamics with respect to inflation at one point in time. We compare two cases. One where the deviation is not anticipated, that is it happens in $t = 0$; and another where it is anticipated very far in advance (it was known since time $t = -\infty$). In both cases we center the impulse responses around the time of the policy shock.

The unanticipated shock is qualitatively similar to an expansionary monetary policy shock. Naturally, it has no effect before it happens. The anticipated shock however has fundamentally different effects as there is a strong negative anticipatory effect on TFP.

Figure 14: Anticipated vs unanticipated shocks



Notes: IRFs of a 1 month inflation shock, which is either unanticipated or anticipated from the far away future.

Using the linear quadratic approximation to the welfare function discussed in Section 5.3.1 and derived in Section E.3

$$-\int_0^\infty e^{-\int_0^t \rho_s^h ds} \frac{C^{-\zeta} Y}{2} \left(\theta \pi_t^2 + \left(\zeta + \frac{\vartheta + \alpha}{1 - \alpha} \right) \tilde{Y}_t^2 + \frac{\vartheta + 1}{(1 - \alpha)} \frac{\zeta - 1}{\zeta + \frac{\vartheta + \alpha}{1 - \alpha}} \hat{Z}_t^2 - 2 \hat{Z}_t \right), \ dt$$

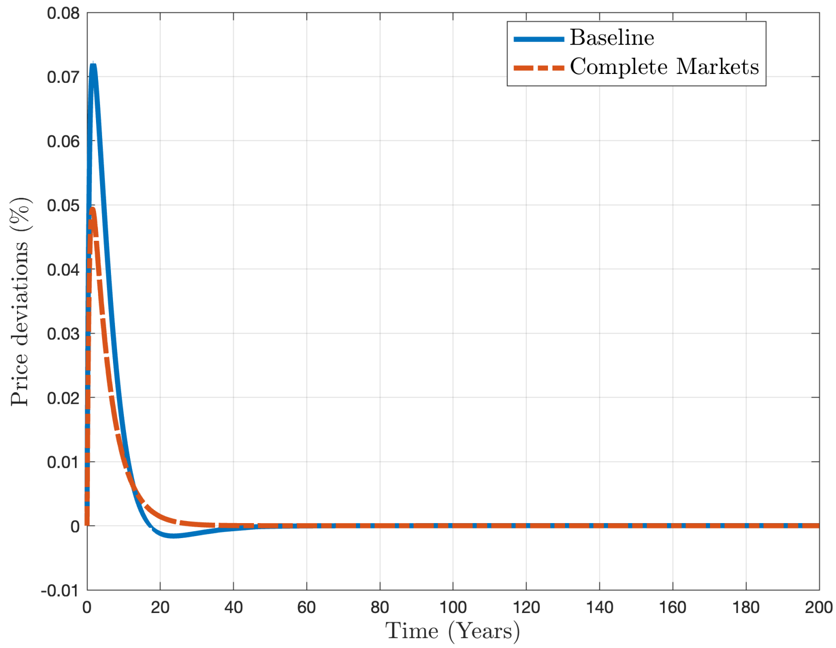
we can compare those two policies in terms of welfare. The unanticipated policy shock increases TFP $\hat{Z}_t$ at the expense of some volatility in output gap $\tilde{Y}_t^2$, inflation π_t^2 and TFP $\hat{Z}^2$. The effect of the linear term dominates and this policy is thus improving welfare (from the time-0 perspective) relative to the steady state.

The anticipated shock also causes costly volatility in all three variables. However, the effect on average TFP is much less benign. Just before the implementation of the temporary inflation policy, it drops dramatically, as firms increase their markups (which causes a decline in the output gap) in anticipation of future inflation. Upon implementation, TFP raises again, undoing the previous decline. Thus the benefits from the linear term no longer outweigh the costs resulting from the quadratic terms, and this policy is welfare reducing. Symetrically, a temporary reduction in inflation would have the opposite effects and also be welfare reducing. Zero inflation is thus optimal in steady state.

C.6 Price-level targeting after a cost-push shock.

Figure 15 shows that, following a cost-push shock, the optimal timeless response drives the price level back to its pre-shock value (blue solid line). Thus, the well-known long-run price-level-targeting property of the Ramsey policy in the standard RANK model (orange dashed line) persists even in the presence of capital misallocation.

Figure 15: Price-level targeting after a cost-push shock.



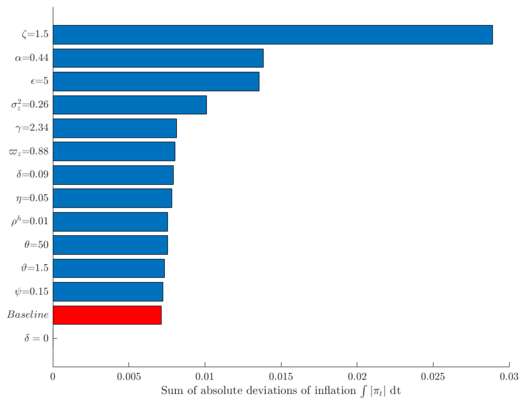
Notes: The figure shows the optimal response from a timeless perspective (in deviations from steady state) to a 195 b.p. decrease in τ, i.e. a *ceteris paribus* increase of costs of around 2%, that is mean reverting with a yearly persistence of 0.8. The baseline economy is the solid blue line, and the complete-markets economy the dashed orange line. Percentage deviations from the steady state price level.

C.7 Sensitivity of optimal monetary policy results with respect to model parameters

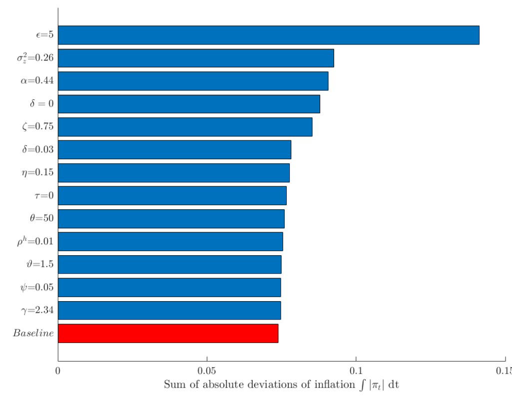
As in the positive Section, we perform a sensitivity analysis of the optimal timeless policy results. In particular, we examine whether the quasi-divine coincidence under the timeless Ramsey policy holds in response to TFP and preference shocks, and whether, following a cost-push shock, the initial output response remains positive, suggesting a less powerful "leaning against the wind". Specifically, we assess how the optimal policy results change under alternative parameterizations of key structural parameters: $\zeta, \alpha, \eta, \vartheta, \delta, \epsilon, \varpi_z, \sigma_z^2, \gamma, \rho_h, \theta, \psi, \tau$. For each parameter, we vary its baseline value by

Figure 16: Sensitivity of optimal monetary policy from a timeless perspective with respect to parameters: TFP and time-preference shocks.

(a) TFP shock.



(b) Time-preference shock.



Notes: This figure reports the sensitivity analysis of the optimal timeless policy results, following a shock of the same size as the one in the main text. Panels (a)–(b) show the absolute cumulative deviation of inflation.

±50%—increasing and decreasing it one at a time—and recompute the optimal timeless Ramsey allocation for each shocks.[51]

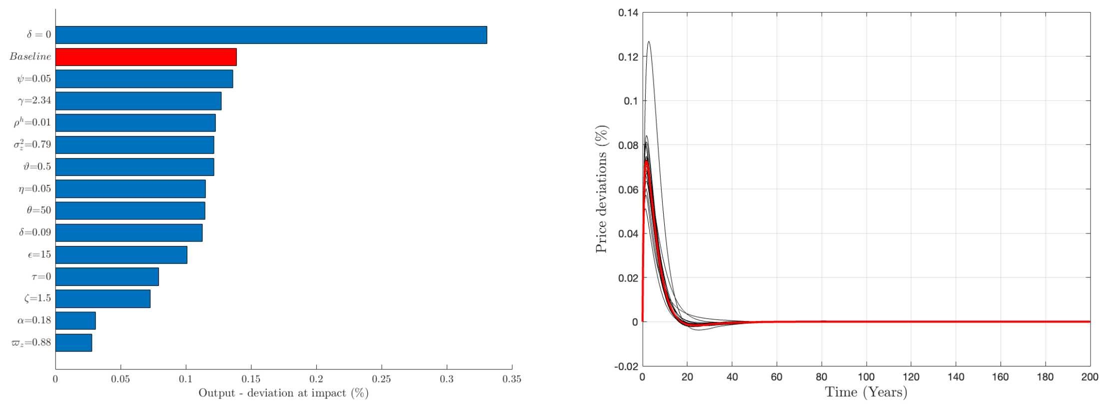
When analyzing TFP and time-preference shocks—reported in panels (a) and (b) of Figure 16, respectively—we compute the absolute cumulative deviation of inflation. Within each parameter sensitivity, we keep only the simulation exhibiting the largest deviation compared to the benchmark. The figure includes only those cases in which this deviation exceeds that of the baseline calibration, with the exception of the case $\delta = 0$, which we always include for completeness. Across all alternative parameterizations, we find that the quasi-divine coincidence continues to hold, as deviations are quantitatively small. Note that when we have $\delta = 0$, we verify numerically the results of Proposition 8: the divine coincidence exactly holds for this case.

When analyzing cost-push shocks (Figure 17), panel (a) shows the percentage change

[51] There are a few exceptions where we cannot increase or decrease a parameter by the full 50%. For example, for (γ), we apply a 50% increase ($1.56 \times 1.5 = 2.34$), but since the lower bound for this parameter is 1, we set ($\gamma = 1$) in the alternative case. Other instances where we deviate from the ±50% rule include: increasing the persistence of idiosyncratic shocks so that the correlation rises from 0.8 (baseline) to 0.9; raising the capital share (α) by only 25% (0.35×1.25) to keep it within reasonable bounds; and reducing risk aversion to ($\zeta = 1 \times 0.75$). For the steady state subsidy, we only set it to zero, $\tau = 0$. In addition, we also run the model with ($\delta = 0$), in combination with the ±50% parameter variations.

Figure 17: Sensitivity of optimal monetary policy from a timeless perspective with respect to parameters: cost-push shocks.

(a) Output deviation at impact. (b) Price level targeting.



Notes: This figure reports the sensitivity analysis of the optimal timeless policy results, following a shock of the same size as the one in the main text. Panel (a) displays the output change (%) at impact following a cost-push shock. Panel (b) shows the price level deviations in percentage terms for all the alternative parametrizations (in black) and for the baseline (red).

in output at impact, and panel (b) shows the percentage price deviations for all simulations. Within each parameter sensitivity, we keep only the simulation exhibiting the lowest output deviation at impact. In panel (a) we report only those cases where the deviation is below that of the baseline calibration, with the exception of the case $\delta = 0$ which we include for completeness. In the complete-markets benchmark, the at-impact output deviation would be zero. Across all alternative parameterizations in the baseline model, this deviation remains positive, indicating that the central bank is willing to tolerate greater inflationary pressures so that output temporarily rises on impact. In panel (b), we report the path of prices for all the simulations performed in black, and the baseline simulation in red. It shows that the price level targeting feature survives in the baseline model across all alternative parametrizations.

D Analytical characterization of optimal policy in the baseline model

D.1 The central bank's problem

The Ramsey problem is given by

$$\max_{\{\pi_t, A_t, D_t, Y_t, L_t, R_t, w_t, q_t, m_t, C_t, z_t^*, \omega_t(z), Z_t\}_{t\geq 0}} \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon \frac{L_t^{1+\vartheta}}{1+\vartheta} \right\} dt$$

subject to the following equilibrium conditions for all $t \geq 0$

$$\begin{aligned}
\frac{\partial \omega_t(z)}{\partial t} &= \left(\frac{1}{q_t} (\gamma \max\left\{ z_t \varsigma_t \alpha \left(\frac{(1-\alpha)}{w_t} \right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - R_t, 0 \right\} + R_t - \delta q_t) - (1-\psi)\eta - \frac{\dot{A}_t}{A_t} \right) \omega_t(z) \\
&\quad - \frac{\partial}{\partial z} \left[\mu(z) \omega_t(z) \right] + \frac{1}{2} \frac{\partial^2}{\partial z^2} \left[\sigma^2(z) \omega_t(z) \right] \\
w_t &= (1-\alpha) m_t Z_t K^\alpha {L_t}^{-\alpha} \\
R_t &= \alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} z_t^* Z_t^{-1/\alpha} \varsigma_t \\
\dot{A}_t &= \frac{A_t}{q_t} \left[\gamma \left(1 - \Omega(z_t^*)\right) \left(\alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} - R_t \right) + R_t - \delta q_t - q_t (1-\psi) \eta) \right] \\
K &= A_t + D_t \\
Y_t &= Z_t K^\alpha {L_t}^{1-\alpha} \\
D_t &= A_t \left(\gamma \left(1 - \Omega(z_t^*)\right) - 1 \right) \\
Z_t &= \left(\varsigma_t \frac{\Gamma(z_t^*)}{1 - \Omega(z_t^*)} \right)^\alpha \\
\dot{D}_t &= \frac{1}{q_t} \left(\left(R_t - \delta q_t\right) D_t + w_t L_t - C_t + (1 - m_t) Y_t \right) \\
&\quad + \frac{1}{q_t} \left(-\frac{\theta}{2} \pi_t^2 Y_t + q_t (1-\psi) \eta A_t + \delta (q_t - 1) K \right) \\
w_t &= \frac{\Upsilon L_t^\vartheta}{C_t^{-\zeta}} \\
C_t^{-\zeta} q_t \rho_t^h &= \left(R_t - \delta q_t\right) C_t^{-\zeta} + \left(\dot{q}_t C_t^{-\zeta} - \dot{C}_t \zeta q_t C_t^{-\zeta-1} \right) \\
\rho_t^h C_t^{-\zeta} \pi_t Y_t &= \left(C_t^{-\zeta} \dot{\pi}_t Y_t + C_t^{-\zeta} \pi_t \dot{Y}_t - \zeta C_t^{-\zeta-1} \dot{C} \pi_t Y_t \right) + \frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t \left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau) m_t \right)
\end{aligned}$$

where $\Gamma_t(z_t^*) = \int_{z_t^*}^\infty z \omega_t(z) \, dz$ and $\Omega_t(z_t^*) = \int_0^{z_t^*} \omega_t(z) \, dz$ and subject to the initial conditions $\{\omega_0(z), D_0, A_0\}$ and the transversality conditions for $\{C_\infty, Y_\infty, \pi_\infty, q_\infty\}$.

Note that we have written the dynamic equations in a specific way such that the time derivatives can be summarized by *one single time derivative* which appears *linearly.*[52] Also note that we eliminated the auxiliary variables r_t and removed the redundant equation for i_t.

The Lagrangian is:

$$
\begin{aligned}
\mathbb{L} \quad = \quad & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \frac{C_t^{1-\zeta}}{1-\zeta} - \frac{\Upsilon L_t^{1+\vartheta}}{1+\vartheta} \right. \\
+ \quad & \int \phi_t(z) \left[\begin{array}{c} \left(\frac{1}{q_t} (\gamma \max \left\{ z\varsigma_t \alpha \left(\frac{(1-\alpha)}{w_t} \right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - R_t, 0 \right\} + R_t - \delta q_t) - (1-\psi)\eta - \frac{\dot{A}_t}{A_t} \right) \omega_t(z) \\ - \frac{\partial \omega_t(z)}{\partial t} - \frac{\partial}{\partial z} \left[\mu(z) \omega_t(z) \right] + \frac{1}{2} \frac{\partial^2}{\partial z^2} \left[\sigma^2(z) \omega_t(z) \right] \end{array} \right] dz \\
+ \quad & \lambda_{1,t} \left[-w_t + (1-\alpha) m_t Z_t K^\alpha {L_t}^{-\alpha} \right] \\
+ \quad & \lambda_{2,t} \left[-R_t + \alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} z_t^* Z_t^{-1/\alpha} \varsigma_t \right] \\
+ \quad & \lambda_{3,t} \left[-\dot{A}_t + A_t \frac{\gamma \left(1 - \Omega(z_t^*)\right) \left(\alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} - R_t \right) + R_t - \delta q_t - q_t (1-\psi)\eta)}{q_t} \right] \\
+ \quad & \lambda_{4,t} \left[-K + A_t + D_t \right] \\
& \lambda_{5,t} \left[-Y_t + Z_t K^\alpha {L_t}^{1-\alpha} \right] \\
+ \quad & \lambda_{6,t} \left[-D_t + A_t \left(\gamma \left(1 - \Omega(z_t^*)\right) - 1 \right) \right] \\
+ \quad & \lambda_{7,t} \left[-Z_t \left(1 - \Omega(z_t^*)\right)^\alpha + \left(\varsigma_t \Gamma(z_t^*) \right)^\alpha \right] \\
+ \quad & \lambda_{8,t} \left[\begin{array}{c} -\dot{D}_t + \frac{1}{q_t} \left((R_t - \delta q_t) D_t + w_t L_t - C_t + (1-m_t) Y_t \right) \\ + \frac{1}{q_t} \left(-\frac{\theta}{2} \pi_t^2 Y_t + q_t (1-\psi) \eta A_t + \delta (q_t - 1) K \right) \end{array} \right] \\
+ \quad & \lambda_{9,t} \left[-w_t + \Upsilon L_t^\vartheta C_t^\zeta \right] \\
+ \quad & \lambda_{EE,t} \left[\rho_t^h C_t^{-\zeta} q_t - (R_t - \delta q_t) C_t^{-\zeta} - \left(\dot{q}_t C_t^{-\zeta} - \dot{C}_t \zeta q_t C_t^{-\zeta-1} \right) \right] \\
+ \quad & \lambda_{PC,t} \left[\rho_t^h C_t^{-\zeta} \pi_t Y_t - \frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t \left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau) m_t \right) - \left(C_t^{-\zeta} \dot{\pi}_t Y_t + C_t^{-\zeta} \pi_t \dot{Y}_t - \zeta C_t^{-\zeta-1} \dot{C} \pi_t Y_t \right) \right] dt \left. \vphantom{\frac{C}{C}} \right\}
\end{aligned}
$$

To derive the first-order conditions for the planner's problem, we apply calculus of variations. Whenever the Lagrangian contains time derivatives, we need to apply integration by parts. The following general rule is useful: Consider the following general term (which may form part of the Lagrangian)

$$
-\int_0^\infty e^{-\rho_t^h t} \lambda_t \dot{x}_t dt,
$$

where x is a variable and λ is a Lagrange multiplier. We integrate this term by parts:

[52]Note that we have already plugged in the conditions $\chi_{H,t} = C_t^{-\zeta} q_t$ and $\chi_{F,t} = C_t^{-\zeta} \pi_t Y_t$. In the Euler equation the single time derivative is: $\frac{\partial \chi_{H,t}}{\partial t} = \frac{\partial C_t^{-\zeta} q_t}{\partial t} = \dot{q}_t C_t^{-\zeta} - \dot{C}_t \zeta q_t C_t^{-\zeta-1}$. In the Phillips curve: $\frac{\partial \chi_{F,t}}{\partial t} = \frac{\partial C_t^{-\zeta} \pi_t Y_t}{\partial t} = C_t^{-\zeta} \dot{\pi}_t Y_t + C_t^{-\zeta} \pi_t \dot{Y}_t - \zeta C_t^{-\zeta-1} \dot{C}_t \pi_t Y_t$

$$-\int_0^\infty e^{-\rho_t^h t}\lambda_t \dot{x}_t dt = \lambda_0 x_0 - \lim_{t\to\infty} e^{-\rho_t^h t}\lambda_t x_t + \int_0^\infty e^{-\rho_t^h t} x_t \left(\dot{\lambda}_t - \rho_t^h \lambda_t\right) dt.$$

We take the Gateaux derivative of this general term

$$\frac{d}{d\kappa}\left[\lambda_0\left(x_0+\kappa h_0\right) - \lim_{t\to\infty} e^{-\rho_t^h t}\lambda_t\left(x_t+\kappa h_t\right) + \int_0^\infty e^{-\rho_t^h t}\left(x_t+\kappa h_t\right)\left(\dot{\lambda}_t - \rho_t^h\lambda_t\right)dt\right]$$
$$= \lambda_0 h_0 - \lim_{t\to\infty} e^{-\rho_t^h t}\lambda_t h_t + \int_0^\infty e^{-\rho_t^h t} h_t\left(\dot{\lambda}_t - \rho_t^h\lambda_t\right)dt$$

First-order optimality requires the Gateaux derivative of the Lagrangian to be zero for any function h. Thus, the first-order condition will contain the term $\dot{\lambda}_t - \rho_t^h\lambda_t$. Furthermore, optimality will demand a initial condition or a terminal condition for the multiplier λ, depending on whether x is a forward-looking variable or a predetermined state variable. In the former case, we get an initial condition for λ_0 and in the latter case a terminal condition for λ_∞.

Applying this rule to the law of motion of wealth and household savings is straightforward. To make it easier to understand the first-order conditions resulting from the application of this rule, we report the Lagrangian after integrating by parts the parts related to the Euler equation, Phillips curve *and the law of motion of the distribution.* This part of the Lagrangian can be rewritten as follows

$$\begin{aligned}
\mathbb{L} \;=\; & ... + \int_0^\infty e^{-\int_0^t \rho_s^h ds}\lambda_{PC,t}\left\{\rho_t^h C_t^{-\zeta}\pi_t Y_t - \frac{\varepsilon}{\theta}C_t^{-\zeta}Y_t\left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau)m_t\right)\right\}dt \\
- & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\lambda_{PC,t}\left\{\frac{dC_t^{-\zeta}\pi_t Y_t}{dt}\right\}dt \\
+ & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\lambda_{EE,t}\left\{\rho_t^h C_t^{-\zeta}q_t - \left(R_t - \delta q_t\right)C_t^{-\zeta}\right\}dt \\
- & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\lambda_{EE,t}\left\{\frac{dC_t^{-\zeta}q_t}{dt}\right\}dt + ... \\
- & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\int \phi_t(z)\left[\frac{d}{dz}\left[\mu(z)\omega_t(z)\right] + \frac{1}{2}\frac{d^2}{dz^2}\left[\sigma^2(z)\omega_t(z)\right]\right]dzdt, \\
- & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\int \phi_t(z)\left[\frac{d\omega_t(z)}{dt}\right]dzdt,
\end{aligned}$$

where we have simply used the fact that $\frac{dC_t^{-\zeta}q_t}{dt} = \dot{q}_t C_t^{-\zeta} - \dot{C}_t \zeta q_t C_t^{-\zeta-1}$ and that $\frac{dC_t^{-\zeta}\pi_t Y_t}{dt} = C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t \dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}\pi_t Y_t$. This is convenient, because now we can apply the above rule, (with these two terms corresponding to the x). That is, we integrate the second, fourth and sixth line of the (trimmed) Lagrangian by parts with respect to time, and the fifth line with respect to state z

$$\begin{aligned}
\mathbb{L} &= ... + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC,t} \left\{ \rho_t^h C_t^{-\zeta} \pi_t Y_t - \frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t \left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau) m_t \right) \right\} dt \\
&+ \lambda_{PC,0} \left\{ C_0^{-\zeta} \pi_0 Y_0 \right\} - \lambda_{PC,\infty} \left\{ C_\infty^{-\zeta} \pi_\infty Y_\infty \right\} \\
&+ \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(\dot{\lambda}_{PC,t} - \rho_t^h \lambda_{PC,t} \right) C_t^{-\zeta} \pi_t Y_t dt \\
&+ \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{EE,t} \left\{ C_t^{-\zeta} q_t \rho_t^h - (R_t - \delta q_t) C_t^{-\zeta} \right\} dt \\
&+ \lambda_{EE,0} \left\{ q_0 C_0^{-\zeta} \right\} - \lambda_{EE,\infty} \left\{ q_\infty C_\infty^{-\zeta} \right\} \\
&+ \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(\dot{\lambda}_{EE,t} - \rho_t^h \lambda_{EE,t} \right) q_t C_t^{-\zeta} dt + ... \\
&+ \int_0^\infty e^{-\rho^h t} \int \omega_t(z) \left[\left(\dot{\phi}_t(z) - \rho_t^h \phi_t(z) \right) + \mu(z) \frac{d}{dz} \left[\phi_t(z) \right] + \frac{\sigma^2(z)}{2} \frac{d^2}{dz^2} \left[\phi_t(z) \right] \right] dz dt \\
&+ \int \left[\phi_0(z) \omega_0(z) - \lim_{T \to \infty} \phi_T(z) \omega_T(z) \right] dz.
\end{aligned}$$

Note that the terms $\lambda_{PC,\infty} \left\{ C_\infty^{-\zeta} \pi_\infty Y_\infty \right\}$ and $\lambda_{EE,\infty} \left\{ q_\infty C_\infty^{-\zeta} \right\}$ are zero by the transversality condition. Taking the Gateaux derivative directly leads to the below first-order conditions (FOCs), which include initial conditions. The first-order conditions are:

- **Inflation**. The FOC with respect to inflation is

$$\begin{aligned}
0 &= \lambda_{PC,t} \left\{ \rho_t^h C_t^{-\zeta} Y_t \right\} + \left(\dot{\lambda}_{PC,t} - \rho_t^h \lambda_{PC,t} \right) C_t^{-\zeta} Y_t - \lambda_{8,t} \theta \pi_t Y_t, \\
0 &= \lambda_{PC,0} C_0^{-\zeta} Y_0.
\end{aligned}$$

- **Wages**. The FOC with respect to wages is

$$\begin{aligned}
0 = &- \int \phi_t(z) \frac{\gamma z \varsigma_t (1-\alpha)}{w_t q_t} \left(\frac{(1-\alpha)}{w_t} \right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} \omega_t(z) \mathbf{1}_{\{z \geq z_t^*\}} dz \\
&- \lambda_{1,t} - \lambda_{9,t} + \lambda_{8,t} L_t.
\end{aligned}$$

- **Rental rates**. The FOC with respect to rental rates is

$$
\begin{aligned}
0 &= \int \phi_t(z) \frac{\gamma}{q_t} \left(-\gamma \mathbf{1}_{\{z \geq z_t^*\}} + 1\right) \omega_t(z)\, dz - \lambda_{2,t} \\
&+ \lambda_{3,t} A_t \frac{1-\gamma\left(1-\Omega(z_t^*)\right)}{q_t} - \lambda_{EE,t} C_t^{-\zeta} + \lambda_{8,t} \frac{D_t}{q_t}.
\end{aligned}
$$

- **Capital prices**. The FOC with respect to capital prices is

$$
\begin{aligned}
0 &= -\int \phi_t(z) \frac{\gamma\left(z\varsigma_t \alpha \left(\frac{(1-\alpha)}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - R_t\right) \mathbf{1}_{\{z \geq z_t^*\}} + R_t}{q_t^2} \omega_t(z)\, dz \\
& -\lambda_{3,t} A_t \frac{\gamma\left(1-\Omega(z_t^*)\right)\left(\alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} - R_t\right) + R_t}{q_t^2} \\
& +\lambda_{EE,t}\left(C_t^{-\zeta}\rho_t^h + \delta C_t^{-\zeta}\right) + \left(\dot{\lambda}_{EE,t} - \rho_t^h \lambda_{EE,t}\right) C_t^{-\zeta} \\
& -\lambda_{8,t} \frac{\left(\left(R_t - \delta q_t\right) D_t + w_t L_t - C_t + \left(1-m_t\right) Y_t\right)}{q_t^2} \\
& -\lambda_{8,t} + \frac{\left(-\frac{\theta}{2}\pi_t^2 Y_t + \left(-\iota_t - \frac{\phi^k}{2}(\iota_t - \delta)^2\right) K_t\right)}{q_t^2} \\
0 &= \lambda_{EE,0} C_0^{-\zeta}.
\end{aligned}
$$

- **Intermediate-good prices**. The FOC with respect to intermediate-good prices is

$$
\begin{aligned}
0 &= \int \phi_t(z) \frac{1}{q_t} \gamma z \varsigma_t \left(\frac{(1-\alpha)}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}-1} \mathbf{1}_{\{z \geq z_t^*\}} \omega_t(z)\, dz \\
&+ \lambda_{1,t}(1-\alpha) Z_t {K_t}^{\alpha} {L_t}^{-\alpha} + \lambda_{2,t} \alpha Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} z_t^* Z_t^{-1/\alpha} \varsigma_t \\
&+ \lambda_{3,t} \frac{A_t \gamma\left(1-\Omega(z_t^*)\right) \alpha Z_t K_t^{\alpha-1} {L_t}^{1-\alpha}}{q_t} - \lambda_{PC,t} \frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t (1-\tau)
\end{aligned}
$$

- **Cutoff**. The derivative with respect to the cutoff is

$$
\begin{aligned}
0 &= \lambda_{2,t} \alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} Z_t^{-1/\alpha} \varsigma_t - \lambda_{3,t} \frac{A_t \gamma \Omega'(z_t^*)\left(\alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} - R_t\right)}{q_t} \\
&- \lambda_{6,t} A_t \gamma \Omega'(z_t^*) + \lambda_{7,t}\left(\alpha Z_t \left(1-\Omega(z_t^*)\right)^{\alpha-1} \Omega'(z_t^*) + \varsigma_t^{\alpha} \alpha \left(\Gamma(z_t^*)\right)^{\alpha-1} \Gamma_t'(z^*)\right),
\end{aligned}
$$

where $\Gamma_t'(z_t^*) = -z_t^* \omega_t(z_t^*)$ and $\Omega_t'(z_t^*) = \omega_t(z_t^*)$.

- **Density**. When taking the Gateaux derivative with respect to $\omega_t(z)$, the result is

$$\begin{aligned}
0 &= \phi_t(z)\left(\frac{1}{q_t}(\gamma\max\left\{z\varsigma_t\alpha\left(\frac{(1-\alpha)}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - R_t, 0\right\} + R_t - \delta q_t) - (1-\psi)\eta - \frac{\dot{A}_t}{A_t}\right) \\
&+ \left(\frac{\partial \phi_t(z)}{\partial t} - \rho_t^h \phi_t(z)\right) + \mu(z)\frac{\partial}{\partial z}\left[\phi_t(z)\right] + \frac{\sigma^2(z)}{2}\frac{\partial^2}{\partial z^2}\left[\phi_t(z)\right] \\
&- \lambda_{3,t}\left[A_t \frac{\gamma\left(\alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} - R_t\right)}{q_t}\right]\mathbf{1}_{z<z_t^*} \\
&- \lambda_{6,t}A_t\gamma\mathbf{1}_{z<z_t^*} + \lambda_{7,t}Z_t\alpha\left(1-\Omega(z_t^*)\right)^{\alpha-1}\mathbf{1}_{z<z_t^*} + \lambda_{7,t}\alpha\left(\varsigma_t\Gamma(z_t^*)\right)^{\alpha-1} z\mathbf{1}_{z>z_t^*}
\end{aligned}$$

  and the transversality condition is $\lim_{T\to\infty}\phi_T(z) = 0$. Notice that no derivative is taken with respect to $\omega_0(z)$ as it is predetermined. We have applied the fact that the Gateaux derivative of $\Omega(z_t^*) = \int \omega_t(z)\mathbf{1}_{z<z_t^*}dz$ with respect to $\omega_t(z)$ is $\mathbf{1}_{z<z_t^*}$.

- **Consumption**. The FOC with respect to consumption is

$$\begin{aligned}
0 &= C_t^{-\zeta} - \zeta\lambda_{EE,t}\left\{C_t^{-\zeta-1}q_t\rho_t^h - \left(R_t - \delta q_t\right)C_t^{-\zeta-1}\right\} - \zeta\left(\dot{\lambda}_{EE,t} - \rho_t^h\lambda_{EE,t}\right)q_t C_t^{-\zeta-1} \\
&\quad + \lambda_{9,t}\Upsilon\zeta L_t^{\vartheta}C_t^{\zeta-1} - \lambda_{8,t} + \lambda_{PC,t}\zeta C_t^{-\zeta-1}Y_t\left\{\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon} + \tilde{m}_t\right) + \rho_t^h\pi_t\right\} \\
&\quad + \left(\dot{\lambda}_{PC,t} - \rho_t^h\lambda_{PC,t}\right)\zeta C_t^{-\zeta-1}\pi_t Y_t, \\
0 &= -\lambda_{PC,0}\zeta C_0^{-\zeta-1}\pi_0 Y_0 - \lambda_{EE,0}\zeta q_0 C_0^{-\zeta-1}.
\end{aligned}$$

- **Labor**. The FOC with respect to labor is

$$\begin{aligned}
0 &= -\Upsilon L_t^{\vartheta} - \lambda_{1,t}\alpha(1-\alpha)m_t Z_t K^{\alpha}{L_t}^{-\alpha-1} + \lambda_{2,t}(1-\alpha)\alpha m_t Z_t K^{\alpha-1}{L_t}^{-\alpha}z_t^* Z_t^{-1/\alpha}\varsigma_t \\
&+ \lambda_{3,t}\frac{A_t\gamma\left(1-\Omega(z_t^*)\right)\alpha(1-\alpha)m_t Z_t K^{\alpha-1}{L_t}^{-\alpha}}{q_t} + \lambda_{9,t}\vartheta\Upsilon L_t^{\vartheta-1}C_t^{\zeta} + \lambda_{8,t}w_t \\
&+ \lambda_{5,t}(1-\alpha)Z_t K^{\alpha}{L_t}^{-\alpha}.
\end{aligned}$$

- **TFP**. The FOC with respect to TFP is

$$\begin{aligned}
0 &= \lambda_{1,t}(1-\alpha)m_t K^{\alpha}{L_t}^{-\alpha} - \lambda_{2,t}(\alpha-1)m_t K^{\alpha-1}{L_t}^{1-\alpha}z_t^* Z_t^{-\frac{1}{\alpha}}\varsigma_t \\
&+ \lambda_{3,t}\frac{A_t\gamma\left(1-\Omega(z_t^*)\right)\alpha m_t K^{\alpha-1}{L_t}^{1-\alpha}}{q_t} \\
&- \lambda_{7,t}\left(1-\Omega(z_t^*)\right)^{\alpha} + \lambda_{5,t}K^{\alpha}{L_t}^{1-\alpha}.
\end{aligned}$$

- **Entrepreneur's net worth**. The FOC with respect to net worth is

$$\begin{aligned} 0 &= \dot{\lambda}_{3,t} + \lambda_{3,t}\left[\frac{\gamma\left(1-\Omega(z_t^*)\right)\left(\alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} - R_t\right)}{q_t} - \rho_t^h\right] \\ &+ \lambda_{4,t} + \lambda_{6,t}\left(\gamma\left(1-\Omega(z_t^*)\right)-1\right) + \lambda_{8,t}(1-\psi)\eta. \end{aligned}$$

- **Household's wealth**. The FOC with respect to wealth is

$$0 = \lambda_{4,t} - \lambda_{6,t} + \dot{\lambda}_{8,t} + \lambda_{8,t}\left(\frac{R_t - \delta q_t}{q_t} - \rho_t^h\right).$$

- **Output**. The FOC with respect to output Y_t is

$$\begin{aligned} 0 &= \lambda_{8,t}\frac{1}{q_t}\left(1 - m_t - \frac{\theta}{2}\pi_t^2\right) - \lambda_{5,t} \\ &+ \lambda_{PC,t}\left\{-\frac{\varepsilon}{\theta}C_t^{-\zeta}\left(\frac{1-\varepsilon}{\varepsilon} + \tilde{m}_t\right) + \rho_t^h C_t^{-\zeta}\pi_t\right\} + \left(\dot{\lambda}_{PC,t} - \rho_t^h\lambda_{PC,t}\right)C_t^{-\zeta}\pi_t, \\ 0 &= \lambda_{PC,0}C_0^{-\zeta}\pi_0. \end{aligned}$$

D.2 Proof of Proposition 3

As shown above in Appendix (D.1), the FOC with respect to. π_t is

$$\frac{\partial \mathbb{L}}{\partial \pi_t}: \quad 0 = \underbrace{\lambda_{PC,t}\left\{\rho_t C_t^{-\zeta}Y_t\right\}}_{\text{Marg. benefit of inflation}} + \underbrace{\left(\dot{\lambda}_{PC,t} - \rho_t\lambda_{PC,t}\right)C_t^{-\zeta}Y_t}_{\text{Marg. cost of inflation anticipation}} - \underbrace{\lambda_{8,t}\theta\pi_t Y_t}_{\text{Marg. price adjustement cost}},$$

Consider the sources of those three terms in the Lagrangian above. The first term in the FOC results from the π_t term in the Phillips curve: it reflects the contemporaneous benign effect of inflation. The second term results from the growth rate $\dot{\pi}_t$: it reflects the negative effect of the anticipation of inflation. The third term comes from the total resource constraint: it reflects the Rotemberg resource cost of price adjustments.

We can simplify this to

$$0 = \dot{\lambda}_{PC,t}C_t^{-\zeta}Y_t + \lambda_{8,t}\theta\pi_t Y_t.$$

Notice that this simplification is possible because the firm and the central bank use the same discount rate: the ρ_t-term in the marginal benefit results from the integration by parts of the firm's Lagrangian and thus reflects the firms' discounting, while the ρ_t-term in the marginal cost of inflation results from the integration by parts of the planner's Lagrangian.

In steady state ($\dot{\lambda}_{PC,t} = 0$), this simplifies further

$$0 = \lambda_8 \theta \pi Y.$$

If the multiplier on the budget constraint λ_8 is strictly positive in steady state and such a steady state exists, π must be 0 in steady state. This proves Proposition 3.

In Section C.5, we illustrate this result numerically.

D.3 Proof of Proposition 5

As shown in above in Appendix (D.1), the FOCs with respect to q_0, C_0, π_0 and Y_0, yield the following initial conditions:

$$\begin{aligned}
\frac{\partial \mathbb{L}}{\partial q_0}: \quad 0 &= \lambda_{EE,0} C_0^{-\zeta}, \\
\frac{\partial \mathbb{L}}{\partial C_0}: \quad 0 &= -\lambda_{EE,0} \zeta q_0 C_0^{-\zeta-1} - \lambda_{PC,0} \zeta C_0^{-\zeta-1} \pi_0 Y_0, \\
\frac{\partial \mathbb{L}}{\partial \pi_0}: \quad 0 &= \lambda_{PC,0} C_0^{-\zeta} Y_0, \\
\frac{\partial \mathbb{L}}{\partial Y_0}: \quad 0 &= \lambda_{PC,0} C_0^{-\zeta} \pi_0.
\end{aligned}$$

Since $C_0, Y_0, q_0 > 0$ these four conditions boil down to the two initial conditions $\lambda_{PC,0} = 0$ and $\lambda_{EE,0} = 0$.

We derive the penalty term using a guess-and-verify approach. We guess that timeless penalty is $\left(-\lambda_{EE} C_0^{-\zeta} q_0 - \lambda_{PC} C_0^{-\zeta} \pi_0 Y_0\right)$. We therefore add this penalty to the objective function, and the respective FOCs become

$$\begin{aligned}
\frac{\partial \mathbb{L}}{\partial q_0}: \quad 0 &= \left(\lambda_{EE,0} - \lambda_{EE}\right) C_0^{-\zeta}, \\
\frac{\partial \mathbb{L}}{\partial C_0}: \quad 0 &= -\left(\lambda_{EE,0} - \lambda_{EE}\right) \zeta q_0 C_0^{-\zeta-1} - \left(\lambda_{PC,0} - \lambda_{PC}\right) \zeta C_0^{-\zeta-1} \pi_0 Y_0,
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \mathbb{L}}{\partial \pi_0}: \quad 0 &= \left(\lambda_{PC,0} - \lambda_{PC}\right) C_0^{-\zeta} Y_0, \\
\frac{\partial \mathbb{L}}{\partial Y_0}: \quad 0 &= \left(\lambda_{PC,0} - \lambda_{PC}\right) C_0^{-\zeta} \pi_0.
\end{aligned}
$$

Since C_0, π_0 and Y_0 are all strictly positive, these conditions are equivalent to the condition that the two Lagrange multipliers on the forward looking constraints are equal to their steady state values.

$$
\begin{aligned}
\lambda_{EE,0} &= \lambda_{EE} \\
\lambda_{PC,0} &= \lambda_{PC}
\end{aligned}
$$

Thus, with the timeless penalty, the central bank's problem is now time-consistent in steady state. This verifies that our guess was correct, and proves Proposition 5.

D.4 Proof of divine coincidence for log utility and $\delta = 0$

The set of constraints of the planner's problem from Section D.1, after eliminating Z_t and breaking each of the forward looking constraints into two

$$
\begin{aligned}
\frac{\partial \omega_t(z)}{\partial t} &= \left(\frac{1}{q_t}\left[\gamma \max\left\{\varsigma_t z \alpha \left(\frac{(1-\alpha)}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - R_t, 0\right\} + R_t\right] - (1-\psi)\eta - \frac{\dot{A}_t}{A_t}\right)\omega_t(z) \\
&\quad - \frac{\partial}{\partial z}\left[\mu(z)\omega_t(z)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)\omega_t(z)\right], \\
w_t &= (1-\alpha) m_t Y_t {L_t}^{-1}, \\
R_t &= \alpha m_t Y_t K^{-1} z_t^* \frac{1-\Omega(z_t^*)}{\Gamma(z_t^*)}, \\
\frac{\dot{A}_t}{A_t} &= \frac{1}{q_t}\left[\gamma\left(1-\Omega(z_t^*)\right)\left(\alpha m_t \left(\varsigma_t \frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha} K^{\alpha-1} {L_t}^{1-\alpha} - R_t\right) + R_t - q_t(1-\psi)\eta\right)\right], \\
K &= A_t + D_t, \\
Y_t &= \left(\varsigma_t \frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha} K^{\alpha} {L_t}^{1-\alpha}, \\
D_t &= A_t\left(\gamma\left(1-\Omega(z_t^*)\right) - 1\right), \\
Z_t &= \left(\varsigma_t \frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha}, \\
q_t \dot{D}_t &= R_t D_t + w_t L_t - C_t + (1-m_t) Y_t \\
&\quad - \frac{\theta}{2}\pi_t^2 Y_t + q_t(1-\psi)\eta A_t,
\end{aligned}
$$

$$
\begin{aligned}
w_t &= \frac{\Upsilon L_t^{\vartheta}}{C_t^{-\zeta}},\\
\chi_{H,t}\rho_t^h &= \left(R_t - \delta q_t\right)/q_t \chi_{H,t} + \dot{\chi}_{H,t},\\
\chi_{H,t} &= C_t^{-\zeta} q_t,\\
\rho_t^h \chi_{F,t} &= \dot{\chi}_{F,t} + \frac{\varepsilon}{\theta}\frac{\chi_{F,t}}{\pi_t}\left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau)m_t\right),\\
\chi_{F,t} &= C_t^{-\zeta}\pi_t Y_t.
\end{aligned}
$$

Now we divide the 2nd, 3rd, 6th, 9th, 10th equation by ς_t^{α}. Furthermore we multiply and divide the RHS of the 1st, 4th, 11th and 14th equation by ς_t^{α} and use $\zeta = 1$

$$
\begin{aligned}
\frac{\partial \omega_t\left(z\right)}{\partial t} &= \left(\frac{\varsigma_t^{\alpha}}{q_t}\left[\gamma \max\left\{z\alpha\left((1-\alpha)\frac{\varsigma_t^{\alpha}}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - \frac{R_t}{\varsigma_t^{\alpha}}, 0\right\} + \frac{R_t}{\varsigma_t^{\alpha}}\right] - (1-\psi)\eta - \frac{\dot{A}_t}{A_t}\right)\omega_t\left(z\right)\\
&\quad -\frac{\partial}{\partial z}\left[\mu(z)\omega_t\left(z\right)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)\omega_t\left(z\right)\right],\\
\frac{w_t}{\varsigma_t^{\alpha}} &= (1-\alpha)m_t\frac{Y_t}{\varsigma_t^{\alpha}}{L_t}^{-1},\\
\frac{R_t}{\varsigma_t^{\alpha}} &= \alpha m_t \frac{Y_t}{\varsigma_t^{\alpha}} K^{-1} z_t^* \frac{1-\Omega(z_t^*)}{\Gamma(z_t^*)},\\
\frac{\dot{A}_t}{A_t} &= \frac{\varsigma_t^{\alpha}}{q_t}\left[\gamma\left(1-\Omega(z_t^*)\right)\left(\alpha m_t\left(\frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha} K^{\alpha-1}{L_t}^{1-\alpha} - \frac{R_t}{\varsigma_t^{\alpha}}\right) + \frac{R_t}{\varsigma_t^{\alpha}} - \frac{q_t}{\varsigma_t^{\alpha}}(1-\psi)\eta)\right],\\
K &= A_t + D_t,\\
\frac{Y_t}{\varsigma_t^{\alpha}} &= \left(\frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha} K^{\alpha}{L_t}^{1-\alpha},\\
D_t &= A_t\left(\gamma\left(1-\Omega(z_t^*)\right) - 1\right),\\
Z_t &= \left(\varsigma_t \frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha},\\
\frac{q_t}{\varsigma_t^{\alpha}}\dot{D}_t &= \frac{R_t}{\varsigma_t^{\alpha}}D_t + \frac{w_t}{\varsigma_t^{\alpha}}L_t - \frac{C_t}{\varsigma_t^{\alpha}} + \left(1-m_t\right)\frac{Y_t}{\varsigma_t^{\alpha}},\\
&\quad -\frac{\theta}{2}\pi_t^2\frac{Y_t}{\varsigma_t^{\alpha}} + \frac{q_t}{\varsigma_t^{\alpha}}(1-\psi)\eta A_t,\\
\frac{w_t}{\varsigma_t^{\alpha}} &= \Upsilon L_t^{\vartheta}\frac{C_t}{\varsigma_t^{\alpha}},\\
\chi_{H,t}\rho_t^h &= \left(\frac{R_t}{\varsigma_t^{\alpha}}\right)/\frac{q_t}{\varsigma_t^{\alpha}}\chi_{H,t} + \dot{\chi}_{H,t},\\
\chi_{H,t} &= \left(\frac{C_t}{\varsigma_t^{\alpha}}\right)^{-1}\frac{q_t}{\varsigma_t^{\alpha}},\\
\rho_t^h \chi_{F,t} &= \dot{\chi}_{F,t} + \frac{\varepsilon}{\theta}\frac{\chi_{F,t}}{\pi_t}\left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau)m_t\right),\\
\chi_{F,t} &= \left(\frac{C_t}{\varsigma_t^{\alpha}}\right)^{-1}\pi_t\frac{Y_t}{\varsigma_t^{\alpha}}.
\end{aligned}
$$

Now we define $\tilde{X}_t \equiv \frac{X_t}{\varsigma_t^{\alpha}}$ for any arbitrary variable and use this definition, and drop the equation defining Z_t, which is now redundant

$$
\begin{aligned}
\frac{\partial \omega_t(z)}{\partial t} &= \left(\frac{1}{\tilde{q}_t}(\gamma \max\left\{z\alpha\left(\frac{(1-\alpha)}{\tilde{w}_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} - \tilde{R}_t, 0\right\} + \tilde{R}_t) - (1-\psi)\eta - \frac{\dot{A}_t}{A_t}\right)\omega_t(z) \\
&\quad -\frac{\partial}{\partial z}\left[\mu(z)\omega_t(z)\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma^2(z)\omega_t(z)\right], \\
\tilde{w}_t &= (1-\alpha)m_t\tilde{Y}_t {L_t}^{-1}, \\
\tilde{R}_t &= \alpha m_t \tilde{Y}_t K^{-1} z_t^* \frac{1-\Omega(z_t^*)}{\Gamma(z_t^*)}, \\
\frac{\dot{A}_t}{A_t} &= \frac{1}{\tilde{q}_t}\left[\gamma\left(1-\Omega(z_t^*)\right)\left(\alpha m_t\left(\frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha} K^{\alpha-1}{L_t}^{1-\alpha} - \tilde{R}_t\right) + \tilde{R}_t - \tilde{q}_t(1-\psi)\eta)\right], \\
K &= A_t + D_t, \\
\tilde{Y}_t &= \left(\frac{\Gamma(z_t^*)}{1-\Omega(z_t^*)}\right)^{\alpha} K^{\alpha}{L_t}^{1-\alpha}, \\
D_t &= A_t\left(\gamma\left(1-\Omega(z_t^*)\right)-1\right), \\
\tilde{q}_t\dot{D}_t &= \left(\tilde{R}_t - \delta\tilde{q}_t\right)D_t + \tilde{w}_t L_t - \tilde{C}_t + (1-m_t)\tilde{Y}_t, \\
&\quad -\frac{\theta}{2}\pi_t^2\tilde{Y}_t + \tilde{q}_t(1-\psi)\eta A_t \\
\tilde{w}_t &= \Upsilon L_t^{\vartheta}\tilde{C}_t, \\
\chi_{H,t}\rho_t^h &= \tilde{R}_t/\tilde{q}_t\chi_{H,t} + \dot{\chi}_{H,t}, \\
\chi_{H,t} &= \left(\tilde{C}_t\right)^{-1}\tilde{q}_t, \\
\rho_t^h\chi_{F,t} &= \dot{\chi}_{F,t} + \frac{\varepsilon}{\theta}\pi_t\chi_{F,t}\left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau)m_t\right), \\
\chi_{F,t} &= \left(\tilde{C}_t\right)^{-1}\pi_t\tilde{Y}_t.
\end{aligned}
$$

The central bank's objective function including the timeless penalty is given by the following expression, and can be re-expressed in therms of normalized variables as follows

$$
\begin{aligned}
&\log(C_t) - \frac{\Upsilon L_t^{1+\vartheta}}{1+\vartheta} - \left(\lambda_{EE}C_0^{-1}q_0 + \lambda_{PC}C_0^{-1}\pi_0 Y_0\right) \\
&= \log(\tilde{C}_t\varsigma_t^{\alpha}) - \frac{\Upsilon L_t^{1+\vartheta}}{1+\vartheta} - \left(\lambda_{EE}\tilde{C}_0^{-1}\tilde{q}_0 + \lambda_{PC}\tilde{C}_0^{-1}\pi_0\tilde{Y}_0\right) \\
&= \log(\tilde{C}_t) + \alpha\log(\varsigma_t) - \frac{\Upsilon L_t^{1+\vartheta}}{1+\vartheta} - \left(\lambda_{EE}\tilde{C}_0^{-1}\tilde{q}_0 + \lambda_{PC}\tilde{C}_0^{-1}\pi_0\tilde{Y}_0\right)
\end{aligned}
$$

We have thus rewritten the original planner's problem in normalized form. The normalized problem is to choose $\pi_t, A_t, D_t, \tilde{Y}_t, L_t, \tilde{R}_t, \tilde{w}_t, \tilde{q}_t, m_t, \tilde{C}_t, z_t^*, \omega_t(z)$ so as to maximize the normalized objective subject to the normalized constraints. This normalized problem is equivalent to the original problem. Given a solution to the normalized problem, the original variables can be backed out inverting the definition $\tilde{X}_t \equiv \frac{X_t}{\varsigma_t^\alpha}$.

Note that the exogenous component of TFP ς_t does no longer appear in the constraints of the planner. Instead, it now appears in the objective. But since it appears additively there, it does not affect the planner's first-order conditions. Thus, the planner's choices of $\pi_t, A_t, D_t, \tilde{Y}_t, L_t, \tilde{R}_t, \tilde{w}_t, \tilde{q}_t, m_t, \tilde{C}_t, z_t^*, \omega_t\,(z)$are independent of ς_t.

Thus, the inflation rate does not respond to shocks to the exogenous component of TFP ς_t. In particular, if the Ramsey steady state features zero inflation (as we have shown before) zero inflation is also optimal after a shock to the exogenous component of TFP under timeless policy. This is the divine coincidence.

E Analytical characterization of optimal policy in the simplified model and complete-markets economy

In this Appendix we present a simplified version of the model that is analytically tractable. This formulation allows us to derive several key results that, in the baseline model, can only be obtained numerically, and to provide a clearer comparison with the complete-markets (RANK) benchmark.

E.1 Overview of the simplified model

We consider now a simpler version of the model. The only difference, borrowed from Itskhoki and Moll (2019) is to assume a different process for firm-level productivity. Instead of following a persistent OU-process, we now assume that firm-level productivity shocks are iid and follow a power law distribution with parameter $\xi > 1$:

$$G(z_t^*) = 1 - (z_t^*)^{-\xi}.$$

Such that $\Omega(z_t^*) = 1 - (z_t^*)^{-\xi}$and $\mathbb{E}[z \mid z > z_t^*] = \frac{\xi}{\xi - 1} z_t^*$. This leaves us with the following system of equilibrium conditions:

$$\begin{aligned}
w_t &= (1-\alpha) m_t Z_t K^{\alpha} {L_t}^{-\alpha}, && (93)\\
R_t &= \alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} z_t^* Z_t^{-1/\alpha} \varsigma_t, && (94)\\
\frac{\dot{A}_t}{A_t} &= \frac{1}{q_t}\left[\gamma (z_t^*)^{-\xi}\left(\alpha m_t Z_t K^{\alpha-1} {L_t}^{1-\alpha} - R_t\right) + (R_t - \delta q_t) - q_t(1-\psi)\eta\right], && (95)\\
K &= A_t + D_t, && (96)\\
Y_t &= Z_t K^{\alpha} {L_t}^{1-\alpha}, && (97)\\
D_t &= A_t\left(\gamma (z_t^*)^{-\xi} - 1\right), && (98)\\
Z_t &= \left(\varsigma_t \frac{\xi}{\xi-1} z_t^*\right)^{\alpha}, && (99)\\
q_t \dot{D}_t &= (R_t - \delta q_t) D_t + w_t L_t - C_t + (1-m_t) Y_t && (100)\\
&\quad -\frac{\theta}{2}\pi_t^2 Y_t + q_t(1-\psi)\eta A_t + \delta (q_t - 1) K, &&\\
w_t &= \frac{\Upsilon L_t^{\vartheta}}{C_t^{-\zeta}}, && (101)\\
C_t^{-\zeta} q_t \rho_t^h &= (R_t - \delta q_t) C_t^{-\zeta} + \left(\dot{q}_t C_t^{-\zeta} - \dot{C}_t \zeta q_t C_t^{-\zeta-1}\right), && (102)
\end{aligned}$$

$$\rho_t^h C_t^{-\zeta}\pi_t Y_t = \left(C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t \dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}\pi_t Y_t\right) + \frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t \left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau) m_t\right). \quad (103)$$

Reduction in the number of variables. These two assumptions not simplify the system of equilibrium conditions and allow us to reduce the system substantially. First, combining (96) and (98) to eliminate D_t we get

$$z_t^* = \left(\frac{A_t \gamma}{K}\right)^{1/\xi}, \quad (104)$$

Plugging this into (99), we get that aggregate TFP only depends on entrepreneurs' net wealth A_t and exogenous TFP shocks ς_t

$$Z_t = \left(\varsigma_t \frac{\xi}{\xi - 1}\right)^{\alpha} \left(\frac{A_t \gamma}{K}\right)^{\alpha/\xi}, \quad (105)$$

or, equivalently,

$$A_t = \frac{K}{\gamma} Z_t^{\xi/\alpha} \left(\frac{\xi - 1}{\xi \varsigma_t}\right)^{\xi}. \quad (106)$$

Second, combining (93), (97) and (101) to eliminate L and w we get

$$m_t = \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}. \quad (107)$$

Using (107) to eliminate m in (103) we have

$$\rho_t^h C_t^{-\zeta}\pi_t Y_t = \left(C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t \dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}\pi_t Y_t\right) \quad (108)$$

$$+\frac{\varepsilon}{\theta} C_t^{-\zeta} Y_t \left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right), \quad (109)$$

Third, using (94), (104) and (106) we get

$$R_t = \alpha m_t K^{-1} Y_t \frac{\xi - 1}{\xi},$$

and using (107) we have

$$R_t = \frac{\xi - 1}{\xi} \frac{\alpha \Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha} - 1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha} + 1} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}. \quad (110)$$

Fourth, using equations (97) and (104) in (95) to eliminate $z^*$and L we get

$$\dot{A}_t = \frac{A_t}{q_t}\left[\frac{K}{A_t}\left(\alpha m_t Y_t K^{-1} - R_t\right) + (R_t - \delta q_t) - q_t(1-\psi)\eta)\right].$$

Using (110) and (107) to eliminate R and m we get

$$\dot{A}_t = \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right) - A_t\left((1-\psi)\eta+\delta\right) \tag{111}$$

Fifth, we can replace (100) by the aggregate resource constraint

$$Y_t\left(1-\frac{\theta}{2}\pi_t^2\right) = C_t + \delta. \tag{112}$$

Sixths, we plug (110) into (102) to eliminate R

$$\left(\rho_t^h+\delta\right)C_t^{-\zeta}q_t = \frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} + \left(\dot{q}_t C_t^{-\zeta} - \dot{C}_t\zeta q_t C_t^{-\zeta-1}\right) \tag{113}$$

This leaves us with a nonlinear system of 5 equations (102), (112), (108), (111), (113) in 6 unknowns $(C_t, Y_t, \pi_t, Z_t, q_t, A)$:

$$\begin{aligned}
\dot{A}_t &= \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right) - A_t\left((1-\psi)\eta+\delta\right) && (114)\\
Z_t &= \left(\varsigma_t\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A_t\gamma}{K}\right)^{\alpha/\xi}, && (115)\\
C_t+\delta K &= Y_t\left(1-\frac{\theta}{2}\pi_t^2\right), && (116)\\
\left(\rho_t^h+\delta\right)C_t^{-\zeta}q_t &= \frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} + \left(\dot{q}_t C_t^{-\zeta} - \dot{C}_t\zeta q_t C_t^{-\zeta-1}\right), && (117)\\
\rho_t^h C_t^{-\zeta}\pi_t Y_t &= \left(C_t^{-\zeta}\dot{\pi}_t Y_t + C_t^{-\zeta}\pi_t\dot{Y}_t - \zeta C_t^{-\zeta-1}\dot{C}\pi_t Y_t\right) && (118)\\
&\quad +\frac{\varepsilon}{\theta}C_t^{-\zeta}Y_t\left(\frac{1-\varepsilon}{\varepsilon} + (1-\tau)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right). && (119)
\end{aligned}$$

The first equation is the law of motion of entrepreneurial net worth, which determines TFP via the second equation. The third is the total resource constraint. The fourth equation is the Euler equation for capital. The last equation is the Phillips curve. The last three equations also hold in the standard representative agent New Keynesian model with fixed capital. However, the fourth equation is block recursive and thus usually ignored, while the third and the fifths together yield the Phillips curve in its usual formulation.

E.1.1 Zero-inflation steady state

In the zero-inflation steady state we have

$$
\begin{aligned}
0 &= \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}C^{\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)-A\left((1-\psi)\eta+\delta\right),\\
Z &= \left(\varsigma\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A\gamma}{K}\right)^{\alpha/\xi},\\
C+\delta K &= Y,\\
\left(\rho^h+\delta\right)C^{-\zeta}q &= \frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}},\\
\frac{(\varepsilon-1)}{\varepsilon(1-\tau)}C^{-\zeta} &= \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}}.
\end{aligned}
$$

We can solve for A and Z in closed form

$$
\begin{aligned}
A &= \frac{\left(\rho^h+\delta\right)\frac{K}{\xi-1}}{(1-\psi)\eta-\rho^h},\\
Z &= \left(\varsigma\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A\gamma}{K}\right)^{\alpha/\xi}.
\end{aligned}
$$

The total resource constraint and the Phillips curve can be combined to

$$
\begin{aligned}
\frac{(\varepsilon-1)}{\varepsilon(1-\tau)}C^{-\zeta} &= \left(\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}(C+\delta K)^{\frac{\vartheta+\alpha}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\right),\\
Y &= C+\delta K,\\
q &= \frac{\xi-1}{\xi\left(\rho^h+\delta\right)}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}}C^{\zeta}.
\end{aligned}
$$

The first equation implicitly defines the steady state value C, conditional on the Z above. Since the left hand side is decreasing in C and takes any value in $(0,\infty)$ for positive C, and the RHS increasing in C and positive, such a value exists. The second equation then defines Y, and the third q. Thus, it exists a steady state of the private equilibrium withe zero inflation.

E.2 Nonlinear Ramsey problem

Now we set up the Ramsey problem and characterize its solution, starting with the steady state, then turning to the divine coincidence. Finally we compare the case of the complete-markets textbook economy.

E.2.1 Optimal monetary policy of simple model: first-order conditions

We now set up the Ramsey problem and determine its first-order conditions.

Central bank objective. The central bank objective is

$$\frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon \frac{L_t^{1+\vartheta}}{1+\vartheta} = \frac{\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta\right)^{1-\zeta}}{1-\zeta} - \Upsilon \frac{Y_t^{\frac{1+\vartheta}{1-\alpha}} Z_t^{\frac{-(1+\vartheta)}{1-\alpha}} K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}.$$

Lagragian. The Lagragian is

$$\begin{aligned}
&\int_0^\infty e^{-\int_0^t \rho_s^h ds}\left[\frac{\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta\right)^{1-\zeta}}{1-\zeta} - \Upsilon \frac{Y_t^{\frac{1+\vartheta}{1-\alpha}} Z_t^{\frac{-(1+\vartheta)}{1-\alpha}} K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{A,t}\left[-\dot{A}_t + \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} C_t^{\zeta} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right) - A_t\left((1-\psi)\eta+\delta\right)\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{Z,t}\left[-Z_t + \left(\varsigma_t\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A_t\gamma}{K}\right)^{\alpha/\xi}\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{TR,t}\left[-C_t + Y_t\left(1-\frac{\theta}{2}\pi_t^2\right) - \delta K\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{EE,t}\left[\rho_t^h\chi_{H,t} + \delta C_t^{-\zeta} q_t - \chi_{\dot{H},t} - \frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{PC,t}\left[\rho_t^h\chi_{F,t} - \chi_{\dot{F},t}\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{PC,t}\left[-\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon}\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta\right)^{-\zeta} Y_t + (1-\tau_t)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right)\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{H,t}\left[C_t^{-\zeta} q_t - \chi_{H,t}\right] \\
&+e^{-\int_0^t \rho_s^h ds}\lambda_{F,t}\left[C_t^{-\zeta}\pi_t Y_t - \chi_{F,t}\right] dt
\end{aligned}$$

We integrate by parts as above:

$$\int_0^\infty e^{-\int_0^t \rho_s^h ds}\left[\frac{\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta\right)^{1-\zeta}}{1-\zeta}-\Upsilon\frac{Y_t^{\frac{1+\vartheta}{1-\alpha}}Z_t^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{A,t}\left[\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta\right)^{\zeta}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right)\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{A,t}\left[-A_t\left((1-\psi)\eta+\delta\right)\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\left(\dot{\lambda}_{A,t}-\rho_t^h\lambda_{A,t}\right)A_t$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{Z,t}\left[-Z_t+\left(\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A_t\gamma}{K}\right)^{\alpha/\xi}\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{TR,t}\left[-C_t+Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{EE,t}\left[\rho_t^h\chi_{H,t}+\delta\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta}q_t-\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\left(\dot{\lambda}_{EE,t}-\rho_t^h\lambda_{EE,t}\right)\chi_{H,t}$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{PC,t}\left[-\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon}\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta}Y_t+\frac{(1-\tau_t)\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right)\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\left[\lambda_{PC,t}\rho_t^h\chi_{F,t}+\left(\dot{\lambda}_{PC,t}-\rho_t^h\lambda_{PC,t}\right)\chi_{F,t}\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{H,t}\left[C_t^{-\zeta}q_t-\chi_{H,t}\right]$$
$$+e^{-\int_0^t \rho_s^h ds}\lambda_{F,t}\left[C_t^{-\zeta}\pi_tY_t-\chi_{F,t}\right]dt$$
$$+\lambda_{PC,0}\left\{\chi_{F,0}\right\}+\lambda_{EE,0}\left\{\chi_{H,0}\right\}-\lambda_{A,\infty}\left\{A_\infty\right\}$$

The first-order conditions with respect to $C_t, Y_t, \pi_t, Z_t, q_t, A, \chi_{F,t}, \chi_{H,t}$ are

$$\frac{dL}{dC_t}:0 \quad = \quad -\lambda_{TR,t}-\lambda_{H,t\zeta}C_t^{-\zeta-1}q_t-\lambda_{F,t}\zeta C_t^{-\zeta-1}\pi_tY_t$$

$$\begin{aligned}\frac{dL}{dY_t}:0 \quad = \quad & \left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta}\left(1-\frac{\theta}{2}\pi_t^2\right)-\Upsilon\frac{1+\vartheta}{1-\alpha}\frac{Y_t^{\frac{1+\vartheta}{1-\alpha}-1}Z_t^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\\ & +\lambda_{A,t}\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right)Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\\ & \times\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{\zeta}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)Y_t^{-1}\\ & +\lambda_{A,t}\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right)Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\end{aligned}$$

$$
\begin{aligned}
&\times\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{\zeta}\zeta\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-1}\left(1-\frac{\theta}{2}\pi_t^2\right)\\
&+\lambda_{TR,t}\\
&-\lambda_{EE,t}\left[\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right]\\
&-\lambda_{EE,t}\left[\zeta\delta\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta-1}\left(1-\frac{\theta}{2}\pi_t^2\right)q_t\right]\\
&-\lambda_{PC,t}\frac{\varepsilon}{\theta}\left(\left[(Y_t-\delta K)^{-\zeta}-\zeta Y_t\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta-1}\left(1-\frac{\theta}{2}\pi_t^2\right)\right]\frac{1-\varepsilon}{\varepsilon}\right)\\
&-\lambda_{PC,t}\frac{\varepsilon}{\theta}\left(\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)(1-\tau_t)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{d\pi_t}:0 \quad = \quad & \lambda_{TR,t}\left[Y_t\theta\pi_t\right]+\lambda_{F,t}C_t^{-\zeta}Y_t\\
&-\lambda_{PC,t}\left[\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon}\zeta\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta-1}\theta\pi_tY_t^2\right)\right],\\
&-\lambda_{A,t}\left[\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\theta\pi_t\zeta\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{\zeta-1}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{dZ_t}:0 \quad = \quad & \left[\frac{(1+\vartheta)}{1-\alpha}\Upsilon\frac{Y_t^{\frac{1+\vartheta}{1-\alpha}}Z_t^{\frac{-(1+\vartheta)}{1-\alpha}-1}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\right]\\
&+\lambda_{A,t}\left[\frac{-(\vartheta+1)}{1-\alpha}\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{\zeta}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}-1}\frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right)\right]\\
&+\lambda_{EE,t}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}-1}\right]-\lambda_{Z,t}\\
&+\lambda_{PC,t}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau_t)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}-1}\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{dq_t}:0 \quad = \quad & +\lambda_{EE,t}\left[\delta\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{-\zeta}\right]+\lambda_{H,t}C_t^{-\zeta}\\
&-\lambda_{A,t}\left[\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t^2(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{\zeta}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A_t}{K}(\xi-1)\right)\right],
\end{aligned}
$$

$$
\frac{dL}{dA_t}:0 \quad = \quad +\lambda_{A,t}\left[\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1}\left(Y_t\left(1-\frac{\theta}{2}\pi_t^2\right)-\delta K\right)^{\zeta}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(\frac{1}{K}(\xi-1)\right)\right]
$$

$$
\begin{aligned}
&+\left(\dot{\lambda}_{A,t}-\rho_t^h\lambda_{A,t}\right)+\lambda_{A,t}\left(-(1-\psi)\eta-\delta\right)\\
&+\lambda_{Z,t}\left[\alpha/\xi\left(\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A_t\gamma}{K}\right)^{\alpha/\xi}/A_t\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{d\chi_{F,t}}:0 &= \dot{\lambda}_{PC,t}+\lambda_{PC,t}\rho_t^h-\lambda_{PC,t}\rho_t^h-\lambda_{F,t},\\
\frac{dL}{d\chi_{H,t}}:0 &= \dot{\lambda}_{EE,t}+\rho_t^h\lambda_{EE,t}-\rho_t^h\lambda_{EE,t}-\lambda_{H,t}.
\end{aligned}
$$

The boundary conditions are

$$
\begin{aligned}
\lambda_{PC,0} &= \lambda_{EE,0}=0,\\
\lim_{t\to\infty}e^{-\int_0^t\rho_s^h ds}\lambda_{A,t} &= 0.
\end{aligned}
$$

E.2.2 The Ramsey SS and time inconsistency

Consider the SS. We guess that $\pi=0$, $\lambda_H=\lambda_F=\lambda_{TR}=0$. The planner's FOCs in steady state now read

$$
\frac{dL}{dC}:0 = 0
$$

$$
\begin{aligned}
\frac{dL}{dY}:0 = &\ (Y-\delta K)^{-\zeta}-\Upsilon\frac{1+\vartheta}{1-\alpha}\frac{Y^{\frac{1+\vartheta}{1-\alpha}-1}Z^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\\
&+\lambda_A\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}(Y-\delta K)^{\zeta}\\
&\times\left\{\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)Y^{-1}+\zeta(Y-\delta K)^{-1}\right\}\\
&-\lambda_{EE}\left[\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}}+\zeta\delta(Y-\delta K)^{-\zeta-1}q\right]\\
&-\lambda_{PC}\frac{\varepsilon}{\theta}\left(\begin{array}{c}\left[(Y-\delta K)^{-\zeta}-\zeta Y(Y-\delta K)^{-\zeta-1}\right]\frac{1-\varepsilon}{\varepsilon}\\ +\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)(1-\tau)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\end{array}\right),
\end{aligned}
$$

$$
\frac{dL}{d\pi}:0 = 0,
$$

$$
\frac{dL}{dZ}:0 = \left[\frac{(1+\vartheta)}{1-\alpha}\Upsilon\frac{Y^{\frac{1+\vartheta}{1-\alpha}}Z^{\frac{-(1+\vartheta)}{1-\alpha}-1}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\right]
$$

$$
\begin{aligned}
&+\lambda_A\left[\frac{-(\vartheta+1)}{1-\alpha}\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}(Y-\delta K)^{\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}-1}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\right]-\lambda_Z,\\
&+\lambda_{EE}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}-1}\right]\\
&+\lambda_{PC}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}-1}\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{dq}:0 \quad = \quad & +\lambda_{EE}\left[\delta(Y-\delta K)^{-\zeta}\right]\\
& -\lambda_A\left[\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q^2(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}(Y-\delta K)^{\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{dA}:0 \quad = \quad & +\lambda_A\left[\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}(Y-\delta K)^{\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(\frac{1}{K}(\xi-1)\right)-(1-\psi)\eta-\delta-\rho^h\right]\\
& +\lambda_Z\left[\alpha/\xi\left(\varsigma\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A\gamma}{K}\right)^{\alpha/\xi}/A\right],
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{d\chi_F}:0 &= 0,\\
\frac{dL}{d\chi_H}:0 &= 0.
\end{aligned}
$$

Note that 4 conditions are immediately satisfied. Now we need to show that values of λ_{EE} and λ_{PC} exist that satisfy the other 2 conditions.

Using the zero-inflation steady-state versions of the private equilibrium conditions from Section E.1.1, we can simplify a bit

$$
\begin{aligned}
\frac{dL}{dY}:0 \quad = \quad & C^{-\zeta}-\frac{\varepsilon-1}{\varepsilon(1-\tau)}C^{-\zeta}\\
& +\lambda_A\left(\rho_t^h+\delta\right)\frac{1}{\xi-1}K\left(1+\frac{A_t}{K}(\xi-1)\right)\left\{\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)Y^{-1}+\zeta C^{-1}\right\}\\
& +\lambda_{EE}\left[\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\frac{\alpha}{K}\frac{1-\varepsilon}{\varepsilon(1-\tau)}C^{-\zeta}-\zeta\delta C^{-\zeta-1}q\right]\\
& -\lambda_{PC}\frac{\varepsilon}{\theta}\left(\left[C^{-\zeta}-\zeta Y C^{-\zeta-1}\right]\frac{1-\varepsilon}{\varepsilon}-\frac{1-\varepsilon}{\varepsilon}C^{-\zeta}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\right),
\end{aligned}
$$

$$\frac{dL}{dZ}:0 = 1-\lambda_A\left[\frac{(\vartheta+1)}{1-\alpha}\alpha C^{\zeta}\frac{1}{q\xi}\left(1+\frac{A}{K}(\xi-1)\right)\right]-\lambda_Z\left[\frac{(\varepsilon-1)}{\varepsilon(1-\tau)}C^{-\zeta}YZ^{-1}\right]^{-1}$$
$$+\lambda_{EE}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\alpha K^{-1}\right]+\lambda_{PC}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\right],$$

-

$$\frac{dL}{dq}:0 = \lambda_{EE}\delta C^{-\zeta}-\lambda_A\frac{\left(\rho_t^h+\delta\right)\left(\frac{K}{\xi-1}+A\right)}{q},$$

$$\frac{dL}{dA}:0 = -\lambda_A(1-\psi)\eta+\lambda_Z\left[\alpha/\xi\left(\varsigma\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A\gamma}{K}\right)^{\alpha/\xi}/A\right].$$

Now we solve

$$\lambda_A = \lambda_{EE}\frac{\delta q C^{-\zeta}}{\left(\rho_t^h+\delta\right)\left(\frac{K}{\xi-1}+A\right)},$$

$$\lambda_Z = \lambda_{EE}\frac{\delta q C^{-\zeta}}{\left(\rho_t^h+\delta\right)\left(\frac{K}{A(\xi-1)}+1\right)}\frac{(1-\psi)\eta}{\alpha/\xi\left(\varsigma\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A\gamma}{K}\right)^{\alpha/\xi}},$$

$$\frac{dL}{dY}:0 = 1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}$$
$$+\lambda_{EE}\left\{\delta q\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)Y^{-1}+\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\frac{\alpha}{K}\frac{1-\varepsilon}{\varepsilon(1-\tau)}\right\}$$
$$-\lambda_{PC}\frac{\varepsilon}{\theta}\left(\left[1-\zeta YC^{-1}\right]\frac{1-\varepsilon}{\varepsilon}-\frac{1-\varepsilon}{\varepsilon}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\right),$$

$$\frac{dL}{dZ}:0 = 1+\lambda_{EE}\frac{\xi-1}{\xi\left(\rho^h+\delta\right)}\left[\rho^h\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\alpha}{K}\right]-\frac{\rho\delta}{[\rho+\delta]}\frac{\xi}{K}\right]$$
$$+\lambda_{PC}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\right].$$

Thus

$$\lambda_{EE} = \frac{1+\lambda_{PC}\left[\frac{\vartheta+1}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\right]}{\frac{\xi-1}{\xi\left(\rho^h+\delta\right)}\left[\frac{\rho\delta}{\rho+\delta}\frac{\xi}{K}-\rho^h\left[\frac{\vartheta+1}{1-\alpha}\frac{\alpha}{K}\right]\right]}$$

Plug this in the FOC for Y to solve for the multipliers:

$$\lambda_{PC} = \frac{\theta}{1-\varepsilon}\frac{\rho^h\frac{\vartheta+1}{1-\alpha}\frac{\alpha}{\xi}+\left(\frac{\varepsilon-1}{\varepsilon(1-\tau)}-1\right)\frac{\rho\delta}{\rho+\delta}}{\frac{\vartheta+1}{1-\alpha}\frac{\rho\delta}{\rho+\delta}+\left(1-\zeta\frac{Y}{C}\right)\left(\frac{\vartheta+1}{1-\alpha}\rho^h\frac{\alpha}{\xi}-\frac{\rho\delta}{\rho+\delta}\right)},$$

$$\lambda_{EE} = \frac{-\left(1-\zeta\frac{Y}{C}+\frac{\varepsilon(1-\tau)}{1-\varepsilon}\frac{(\vartheta+1)}{1-\alpha}\right)}{\frac{\xi-1}{\left(\rho^h+\delta\right)K}\left(\frac{\vartheta+1}{1-\alpha}\frac{\rho\delta}{\rho+\delta}+\left(1-\zeta\frac{Y}{C}\right)\left(\frac{\vartheta+1}{1-\alpha}\rho^h\frac{\alpha}{\xi}-\frac{\rho\delta}{\rho+\delta}\right)\right)},$$

$$\lambda_{A} = \frac{-\left(1-\zeta\frac{Y}{C}+\frac{\varepsilon(1-\tau)}{1-\varepsilon}\frac{(\vartheta+1)}{1-\alpha}\right)}{\left(\frac{\vartheta+1}{1-\alpha}\frac{\delta\rho}{\rho+\delta}+\left(1-\zeta\frac{Y}{C}\right)\left(\frac{\vartheta+1}{1-\alpha}\rho^h\frac{\alpha}{\xi}-\frac{\delta\rho}{\rho+\delta}\right)\right)}\frac{\delta qC^{-\zeta}}{1+(\xi-1)\frac{A}{K}},$$

$$\lambda_{Z} = \frac{-\left(1-\zeta\frac{Y}{C}+\frac{\varepsilon(1-\tau)}{1-\varepsilon}\frac{(\vartheta+1)}{1-\alpha}\right)}{\left(\frac{\vartheta+1}{1-\alpha}\frac{\delta}{\rho+\delta}+\left(1-\zeta\frac{Y}{C}\right)\left(\frac{\vartheta+1}{1-\alpha}\frac{\rho^h}{\rho}\frac{\alpha}{\xi}-\frac{\delta}{\rho+\delta}\right)\right)}\frac{\delta qC^{-\zeta}}{\left(\frac{1}{A}+\frac{\xi-1}{K}\right)}\frac{1}{\alpha/\xi\left(\varsigma\frac{\xi}{\xi-1}\right)^{\alpha}\left(\frac{A\gamma}{K}\right)^{\alpha/\xi}}.$$

This verifies our initial guess and specifies the Lagrange multipliers.

Together with the existence of a private equilibrium steady state with zero inflation shown above, this proves that a **Ramsey steady state** with 0 inflation exists.

Time inconsistency. If the steady state is time consistent, the central bank has no incentives to deviate from the initial allocation, when the latter coincides with the steady state of the optimal Ramsey problem. In this case, an economy starting at the zero-inflation steady state remains there indefinitely in the absence of exogenous shocks. To analyze whether the Ramsey plan is time consistent in steady state, we need to check whether the (backward-looking) Lagrange multipliers associated with forward-looking equations are zero. This is because in the standard Ramsey problem, the initial value of these multipliers is zero, and thus the problem can only be time-invariant if the initial and steady-state values of those multipliers are zero. In our problem, the Lagrange multiplier on the Phillips curve λ_{PC}, and the Lagrange multiplier on the capital Euler equation, λ_{EE}, are non-zero, and therefore the Ramsey plan is time inconsistent in steady state. Since any Ramsey plan we consider converges to the steady state, the optimal Ramsey policy is time inconsistent.

E.2.3 Divine coincidence for $\delta = 0$

Consider now the special case of no depreciation $\delta = 0$. In this case the same steps as in the previous Section, but outside the steady state, lead to the conclusion that the divine coincidence holds for TFP and time-preference shocks. Consider that there are no cost-push shocks, that is $\tau_t = \tau$.

We guess that $\pi_t = 0$, $\lambda_{Z,t} = \lambda_{A,t} = \lambda_{H,t} = \lambda_{F,t} = \lambda_{TR,t} = 0$ and that $\lambda_{PC,t} = \lambda_{PC}$ and $\lambda_{EE,t} = \lambda_{EE}$, The planner's dynamic FOCs now read

$$\frac{dL}{dC_t} : 0 \quad = \quad 0,$$

$$\begin{aligned}
\frac{dL}{dY_t} : 0 \quad = \quad & Y_t^{-\zeta} - \Upsilon \frac{1+\vartheta}{1-\alpha} \frac{Y_t^{\frac{\vartheta+\alpha}{1-\alpha}} Z^{\frac{-(1+\vartheta)}{1-\alpha}} K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta} \\
& -\lambda_{EE} \left(\frac{\vartheta+\alpha}{1-\alpha} + 1 \right) \frac{\xi - 1}{\xi} \frac{\alpha \Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha} - 1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \\
& -\lambda_{PC} \left[\frac{\varepsilon}{\theta} \left((1-\zeta) Y_t^{-\zeta} \frac{1-\varepsilon}{\varepsilon} + \left(\frac{\vartheta+\alpha}{1-\alpha} + 1 \right) (1-\tau) \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \right) \right],
\end{aligned}$$

$$\frac{dL}{d\pi_t} : 0 \quad = \quad 0,$$

$$\begin{aligned}
\frac{dL}{dZ_t} : 0 \quad = \quad & \left[\frac{\Upsilon}{1-\alpha} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(1+\vartheta)}{1-\alpha}-1} K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}} \right] \\
& +\lambda_{EE} \left[\frac{(\vartheta+1)}{1-\alpha} \frac{\xi-1}{\xi} \frac{\alpha\Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}-1} \right] \\
& +\lambda_{PC} \left[\frac{(\vartheta+1)}{1-\alpha} \frac{\varepsilon}{\theta} (1-\tau) \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}-1} \right],
\end{aligned}$$

$$\frac{dL}{dq_t}:0 = 0,$$

$$\frac{dL}{dA_t}:0 = 0.$$

$$\begin{aligned}\frac{dL}{d\chi_{F,t}}:0 &= 0,\\ \frac{dL}{d\chi_{H,t}}:0 &= 0.\end{aligned}$$

Note that six conditions are immediately satisfied. Now we need to show that values of λ_{EE} and λ_{PC} exist that satisfy the other 2 conditions.

The Phillips curve, given our guess $\pi_t = 0$, is $-\frac{1-\varepsilon}{\varepsilon(1-\tau)}C_t^{-\zeta} = \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}$. Also $Y_t = C_t$. We can use this to simplify the previous condition, and rearrange

$$\begin{aligned}\frac{dL}{dY_t}:0 &= Y_t^{-\zeta} + \frac{1-\varepsilon}{\varepsilon(1-\tau_t)}Y_t^{-\zeta}\\ &\quad +\lambda_{EE}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\alpha K^{-1}\frac{1-\varepsilon}{\varepsilon(1-\tau)}Y_t^{-\zeta}\\ &\quad -\lambda_{PC}\left[\frac{\varepsilon}{\theta}\left((1-\zeta)Y_t^{-\zeta}\frac{1-\varepsilon}{\varepsilon}-\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{1-\varepsilon}{\varepsilon}Y_t^{-\zeta}\right)\right],\\ 0 &= 1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}+\lambda_{EE}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\alpha K^{-1}\frac{1-\varepsilon}{\varepsilon(1-\tau)}\\ &\quad -\lambda_{PC}\left[\frac{1-\varepsilon}{\theta}\left((1-\zeta)-\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\right)\right].\end{aligned}$$

We can rearrange the FOC for Z_t

$$\begin{aligned}\frac{dL}{dZ_t}:0 &= \frac{\Upsilon}{1-\alpha}Y_t^{\frac{1+\vartheta}{1-\alpha}}Z_t^{\frac{-(1+\vartheta)}{1-\alpha}-1}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}\\ &\quad\times\left\{1+\lambda_{EE}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\alpha K^{-1}\right]+\lambda_{PC}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\right]\right\}, &(120)\\ 0 &= 1+\lambda_{EE}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\alpha K^{-1}\right]+\lambda_{PC}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\right]. &(121)\end{aligned}$$

Thus

$$\lambda_{EE} = -\frac{1+\lambda_{PC}\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\varepsilon}{\theta}(1-\tau)\right]}{\left[\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\alpha K^{-1}\right]}.$$

Plug this in the FOC for Y_t to solve for the two multipliers:

$$\lambda_{PC} = \frac{\theta}{(1-\varepsilon)(1-\zeta)},$$

$$\lambda_{EE} = -\frac{1+\frac{\varepsilon}{(1-\varepsilon)(1-\zeta)}\left[\frac{(\vartheta+1)}{1-\alpha}(1-\tau)\right]}{\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\alpha K^{-1}}.$$

This verifies our initial guess and specifies the Lagrange multipliers. Note that these expressions do not depend on exogenous TFP shocks ς_t or on time-preference shocks ρ_t^h. Thus our guess is valid not only in the Ramsey steady state, but even in the presence of these shocks: The **divine coincidence** holds for TFP and time-preference shocks. Note that λ_{EE} contains τ, which we have assumed to be constant. Thus the divine coincidence does not hold for cost-push shocks.

While we do not formally prove this, the divine coincidence generally fails for the case that $\delta > 0$. The role played by depreciation is similar to that played by government spending in Benigno and Woodford, 2005:f it is zero, such that $C = Y$ in the zero-inflation steady state, the divine coincidence holds despite an inefficient steady state, if it is nonzero it fails.

E.2.4 Comparison to the case with complete-markets economy

We briefly compare these results to the well known counterparts in the textbook complete-markets economy.

The complete-markets model can be seen as a particular case of our model, where $\lambda_{Z,t} = \lambda_{A,t} = \lambda_{EE,t} = 0$, and Z_t and A_t are not choice variables. For simplicity we focus on the case of $\delta = 0$. In this case there are 5 FOCs:

$$\frac{dL}{dC_t} : 0 = -\lambda_{TR,t} - \lambda_{H,t}\zeta q_t C_t^{-\zeta-1} - \lambda_{F,t}\zeta C_t^{-\zeta-1}\pi_t Y_t,$$

$$
\begin{aligned}
\frac{dL}{dY_t}:0 &= Y_t^{-\zeta}\left(1-\frac{\theta}{2}\pi_t^2\right)^{1-\zeta} - \Upsilon\frac{1+\vartheta}{1-\alpha}\frac{Y_t^{\frac{1+\vartheta}{1-\alpha}-1}Z_t^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta} \\
&+\lambda_{TR,t}\left(1-\frac{\theta}{2}\pi_t^2\right) \\
&-\lambda_{EE,t}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \\
&-\lambda_{PC,t}\frac{\varepsilon}{\theta}(1-\zeta)Y_t^{-\zeta}\left(1-\frac{\theta}{2}\pi_t^2\right)^{-\zeta}\frac{1-\varepsilon}{\varepsilon} \\
&-\lambda_{PC,t}\frac{\varepsilon}{\theta}\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)(1-\tau)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} + \lambda_{F,t}C_t^{-\zeta}\pi_t, \\
\frac{dL}{d\pi_t}:0 &= Y_t^{1-\zeta}\left(1-\frac{\theta}{2}\pi_t^2\right)^{-\zeta}\theta\pi_t - \lambda_{TR,t}\left[Y_t\theta\pi_t\right] \\
&+\lambda_{F,t}C_t^{-\zeta}Y_t - \lambda_{PC,t}\left[\zeta\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon}Y_t^{1-\zeta}\left(1-\frac{\theta}{2}\pi_t^2\right)^{-\zeta-1}\theta\pi_t\right)\right], \\
\frac{dL}{d\chi_{F,t}}:0 &= \dot{\lambda}_{PC,t}+\lambda_{PC,t}\rho_t^h-\lambda_{PC,t}\rho_t^h-\lambda_{F,t}, \\
\frac{dL}{d\chi_{H,t}}:0 &= \dot{\lambda}_{EE,t}+\rho_t^h\lambda_{EE,t}-\rho_t^h\lambda_{EE,t}-\lambda_{H,t}.
\end{aligned}
$$

Now we guess that $\pi_t = 0$, $\lambda_{H,t} = \lambda_{F,t} = \lambda_{TR,t} = \lambda_{EE,t} = 0$ and that $\lambda_{PC,t} = \lambda_{PC}$. Notice that the Lagrange multiplier on the Euler equation is now slack. The planner's FOCs now read

$$
\begin{aligned}
\frac{dL}{dC_t}:0 &= 0, \\
\frac{dL}{dY_t}:0 &= Y_t^{-\zeta} - \Upsilon\frac{1+\vartheta}{1-\alpha}\frac{Y_t^{\frac{1+\vartheta}{1-\alpha}-1}Z_t^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta} \\
&-\lambda_{PC}\left[\frac{\varepsilon}{\theta}\left((1-\zeta)Y_t^{-\zeta}\frac{1-\varepsilon}{\varepsilon}+\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)(1-\tau)\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y_t^{\frac{\vartheta+\alpha}{1-\alpha}}Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}\right)\right], \\
\frac{dL}{d\pi_t}:0 &= 0, \\
\frac{dL}{d\chi_{F,t}}:0 &= 0, \\
\frac{dL}{d\chi_{H,t}}:0 &= 0.
\end{aligned}
$$

Using again the Phillips curve, the FOC for Y_t can be solved to find the multiplier λ_{PC}:

$$
\begin{aligned}
0 &= Y_t^{-\zeta} + \frac{1-\varepsilon}{\varepsilon(1-\tau)} Y_t^{-\zeta} - \lambda_{PC}\left[\frac{\varepsilon}{\theta}\left((1-\zeta)Y_t^{-\zeta}\frac{1-\varepsilon}{\varepsilon} - \left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{1-\varepsilon}{\varepsilon}Y_t^{-\zeta}\right)\right], \\
0 &= 1 + \frac{1-\varepsilon}{\varepsilon(1-\tau)} - \lambda_{PC}\frac{\varepsilon}{\theta}\left((1-\zeta)\frac{1-\varepsilon}{\varepsilon} - \left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\frac{1-\varepsilon}{\varepsilon}\right), \\
\lambda_{PC} &= \frac{-1+\frac{\varepsilon(1-\tau)}{\varepsilon-1}}{(1-\tau)\frac{\varepsilon}{\theta}\left(\zeta - 1 + \frac{\vartheta+1}{1-\alpha}\right)}.
\end{aligned}
$$

This verifies our guess and shows the well known results that in the complete-markets model the Ramsey steady state feature 0 inflation and that the divine coincidence holds for TFP and time-preference shocks. Notice, when the optimal subsidy is in place $-1+\frac{\varepsilon(1-\tau)}{\varepsilon-1} = -1+\frac{\varepsilon(1-1/\varepsilon)}{\varepsilon-1} = 0$,so that $\lambda_{PC} = 0$ and the problem is time consistent. Otherwise, the Ramsey steady state is time inconsistent.

E.3 A linear-quadratic approximation of the objective function

In this and the next two sections we employ local approximations to derive further insights about the Ramsey problem. First we derive a second order linear-quadratic objective function. Second, in Section E.4 we then show how we can express the objective as a purely quadratic function, for both the simplified model and the complete-markets model. Using those results, in Section we E.5 we then derive the linear FOCs to the planner's linear quadratic problem. Finally, we use the resulting system of optimality conditions to characterize optimal policy in response to cost-push shocks τ_t.

A linear quadratic approximation of welfare. We now approximate the welfare function up to second order around the zero-inflation steady state. We approximate all variables, endogenous or exogenous, except the time-preference shock, which we keep nonlinear except indicated otherwise. We denote a log-linearized variable $\hat{x}_t \equiv \log \frac{x_t}{x}$. In the case of inflation, we just linearize around the steady state $\hat{x}_t \equiv x_t - x$.

Using the total resource constraint and the production function, the period utility is

$$
\frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon\frac{L_t^{1+\vartheta}}{1+\vartheta} = \frac{\left(Y_t\left(1-\frac{\theta}{2}\hat{\pi}_t^2\right) - \delta K\right)^{1-\zeta}}{1-\zeta} - \Upsilon\frac{Y_t^{\frac{1+\vartheta}{1-\alpha}} Z_t^{\frac{-(1+\vartheta)}{1-\alpha}} K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta},
$$

which can be expressed as

$$\frac{\left(Ye^{\hat{Y}_t}\left(1-\frac{\theta}{2}\hat{\pi}_t^2\right)-\delta K\right)^{1-\zeta}}{1-\zeta}-\Upsilon\frac{e^{\frac{1+\vartheta}{1-\alpha}\hat{Y}_t-\frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t}Y^{\frac{1+\vartheta}{1-\alpha}}Z^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}$$

The first-order approximation is

$$(Y-\delta)^{-\zeta}Y\hat{Y}_t-\Upsilon\frac{Y^{\frac{1+\vartheta}{1-\alpha}}Z^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\frac{1+\vartheta}{1-\alpha}\hat{Y}_t+\Upsilon\frac{Y^{\frac{1+\vartheta}{1-\alpha}}Z^{\frac{-(1+\vartheta)}{1-\alpha}}K^{\frac{-\alpha(1+\vartheta)}{1-\alpha}}}{1+\vartheta}\frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t.$$

Taking into account the steady-state relationship $YC^{-\zeta}\frac{\varepsilon-1}{\varepsilon(1-\tau)}=\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y^{\frac{1+\vartheta}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}}$, this simplifies to

$$\hat{Y}_t\left(Y-\delta K\right)^{-\zeta}Y\left[1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right]+\left(Y-\delta\right)^{-\zeta}Y\frac{\varepsilon-1}{\varepsilon(1-\tau)}\hat{Z}_t.$$

The second-order approximation is

$$C^{-\zeta}Y\left\{\hat{Y}_t\left[1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right]+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\hat{Z}_t+\frac{1}{2}\left[\left[1-\zeta\frac{Y}{Y-\delta K}\right]-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\right]\hat{Y}_t^{\,2}\right\} \tag{122}$$

$$+C^{-\zeta}Y\left\{-\frac{1}{2}\theta\hat{\pi}_t^2-\frac{1}{2}\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\hat{Z}_t^{\,2}+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\frac{(\vartheta+1)}{1-\alpha}\hat{Y}_t\hat{Z}_t\right\}. \tag{123}$$

A second order approximation to welfare is thus:

$$\begin{aligned}
&-\int_0^\infty e^{-\int_0^t\rho_s^h ds}\left(1-\int_0^t\hat{\rho}_s ds\right)C^{-\zeta}Y\left(\hat{Y}_t\left[1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right]+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\hat{Z}_t\right)\,dt\\
&-\int_0^\infty e^{-\int_0^t\rho_s^h ds}C^{-\zeta}Y\left(\frac{1}{2}\left[\left[1-\zeta\frac{Y}{Y-\delta K}\right]-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\right]\hat{Y}_t^{\,2}\right)\,dt\\
&-\int_0^\infty e^{-\int_0^t\rho_s^h ds}C^{-\zeta}Y\left(-\frac{1}{2}\theta\hat{\pi}_t^2-\frac{1}{2}\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\hat{Z}_t^{\,2}\right)\,dt\\
&-\int_0^\infty e^{-\int_0^t\rho_s^h ds}C^{-\zeta}Y\left(\frac{\varepsilon-1}{\varepsilon(1-\tau)}\frac{(\vartheta+1)}{1-\alpha}\hat{Y}_t\hat{Z}_t\right)\,dt,
\end{aligned}$$

With the appropriate subsidy such that $\frac{\varepsilon-1}{\varepsilon(1-\tau)}=1$ and with $\delta=0$ and noting that $\pi_t=\hat{\pi}_t$ by definition, we get the expression reported in the main text:

$$-\int_0^\infty e^{-\int_0^t\rho_s^h ds}\frac{C^{-\zeta}Y}{2}\left(\theta\pi_t^2+\left(\zeta+\frac{\vartheta+\alpha}{1-\alpha}\right)\tilde{Y}_t^2+\frac{\vartheta+1}{(1-\alpha)}\frac{\zeta-1}{\zeta+\frac{\vartheta+\alpha}{1-\alpha}}\hat{Z}_t^2-2\hat{Z}_t\right)\,dt,$$

where $\tilde{Y}_t$ is the output gap relative to the flexible-price zero-markup allocation with the same TFP $\tilde{Y}_t \equiv \hat{Y}_t - \frac{1}{\zeta + \frac{\vartheta+\alpha}{1-\alpha}} \frac{(\vartheta+1)}{(1-\alpha)} \hat{Z}_t$. If we also consider a second-order approximation to the time-preference shock, the expression becomes

$$-\int_0^\infty e^{-\rho^h t} \frac{C^{-\zeta} Y}{2} \left(\theta \pi_t^2 + \left(\zeta + \frac{\vartheta + \alpha}{1-\alpha} \right) \tilde{Y}_t^2 + \frac{\vartheta + 1}{(1-\alpha)} \frac{\zeta - 1}{\zeta + \frac{\vartheta+\alpha}{1-\alpha}} \hat{Z}_t^2 - 2\hat{Z}_t \left(1 - \int_0^t \hat{\rho}_s ds \right) \right) dt.$$

Notice that this approximation also holds in all versions of the model we consider: in the baseline model, in the simplified model with iid shocks and in the complete-markets textbook economy, where $\hat{Z}_t$ is a term independent of policy.

E.4 A purely quadratic approximation of the objective function

The results in this Section are derived in a version of the model that deviates slightly from the simplified model used elsewhere in the paper. Specifically, we assume that capital depreciation costs, δK, are not borne by the owners of capital per unit (households and entrepreneurs) but are instead covered by the household in a lump-sum manner. This assumption removes the δ term from the law of motion for entrepreneurial capital and from the Euler equation, although it remains present in the household's budget constraint:

$$\begin{aligned}
\dot{A}_t &= \alpha \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q_t(1-\alpha)} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} C_t^\zeta Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \frac{1}{\xi} \left(1 + \frac{A_t}{K} (\xi - 1) \right) - A_t \left((1-\psi)\eta \right), && (124) \\
\rho_t^h C_t^{-\zeta} q_t &= \frac{\xi - 1}{\xi} \frac{\alpha \Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1} Y_t^{\frac{\vartheta+\alpha}{1-\alpha}+1} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} + \left(\dot{q}_t C_t^{-\zeta} - \dot{C}_t \zeta q_t C_t^{-\zeta-1} \right), && (125) \\
C_t + \delta K &= Y_t \left(1 - \frac{\theta}{2} \pi_t^2 \right). && (126)
\end{aligned}$$

Consistent with the earlier sections of this appendix, we consider the case $\delta = 0$ where these costs are zero when analyzing TFP and time-preference shocks. For the cost-push shock, however, we assume δK to be strictly positive (though possibly very small).

This assumption is intended to balance tractability and closeness to the main model. Why not simply set $\delta = 0$? Retaining some form of depreciation—or, equivalently, any other exogenous expenditure—is necessary to ensure the existence of a stable equilibrium under cost-push shocks, even if δ is arbitrarily small. This existence issue is specific to the simplified model; in the baseline model, an equilibrium exists even when

$\delta = 0$. Why not model depreciation as we do elsewhere (see Section E.1)? Doing so would render the analytical treatment of the cost-push shock intractable. [53] However, we have solved the simplified model numerically using the depreciation specification from Section E.1, and verified that the qualitative results established for the cost-push shock remain robust. As shown numerically in the main text, these results also extend to the baseline model.

Repeating the same steps as in Section E.2.2 under this modified assumption yields the following steady-state values for the multipliers on the Phillips curve and the Euler equation, while all other multipliers remain zero in steady state:

$$\lambda_{PC} = \frac{1}{\frac{1-\varepsilon}{\theta}\left(1-\zeta\frac{Y}{C}\right)}, \tag{127}$$

$$\lambda_{EE} = -\frac{1+\frac{1}{1-\zeta\frac{Y}{C}}\frac{\vartheta+1}{1-\alpha}\frac{\varepsilon(1-\tau)}{1-\varepsilon}}{\frac{(\vartheta+1)}{1-\alpha}\frac{\xi-1}{\xi}\alpha K^{-1}}.$$

For $\delta = 0$ we have that $C = Y$ such that these multipliers collapse to those in Section E.2.2.

E.4.1 General approach to finding the quadratic objective objective

The second-order approximation of the objective (122) contains linear terms in $\hat{Y}_t$ and $\hat{Z}_t$. These terms are a result of the steady state inefficiencies (capital misallocation, markups) and are indicative of the time inconsistency of the problem. They also imply that a naive linear quadratic approach, which maximizes this quadratic objective subject to linear approximations of the private equilibrium conditions is not appropriate, as shown by Benigno and Woodford (2005).

In this Section we provide a general approach to find a purely quadratic objective that approximates the timeless planner's objective up to second order. This procedure is the continuous time counterpart of the procedure suggested by Benigno and Woodford (2005) and detailed in Debortoli and Nunes (2006). The main idea is that a purely quadratic approximation to the planner's objective under the timeless perspective is

[53] Adding the term $\delta C_t^{-\zeta} q_t$ to the Euler equation would require solving a system of 7 ODEs, for which no analytical solution exists.

always feasible, even when the steady state is distorted, as it is the case of our baseline model. This requires adopting the timeless perspective assumption, since then (and only then) the linear terms in the second-order Taylor expansion of welfare can be eliminated, making policy analysis around distorted steady states as tractable as in efficient ones.

We now explain this procedure in general terms, before we then apply it to our problem in the next Section.

Take as given an objective $u(x_t)$ where x_t is an $n-$dimensional vector of variables. There are also $n-1$ constraints $g_i\left(x_t\right)+f_i\dot{x}_t=0$,where g_i is a nonlinear function, f_i is a constant vector, and $i=1,..,n-1$. In our problem, $x_t=\left(\hat{\pi}_t,\hat{Y}_t,\hat{Z}_t,\hat{C}_t,\hat{\chi}_{H,t},\hat{\chi}_{F,t}\right)$.

1.- Linearize the nonlinear FOCs. The nonlinear problem is[54]

$$\max\int_0^\infty e^{-\rho t}u(x_t)dt$$

subject to $g_i\left(x_t\right)+h_i\frac{dx_t}{dt}=0$. The FOCs are

$$\frac{du}{dx_{j,t}}+\sum_{i=1}^{n-1}\left[\lambda_{i,t}\left(\frac{dg_i}{dx_{j,t}}+\rho_t f_{i,j}\right)-\frac{d\lambda_{i,t}}{dt}\right]=0, \tag{128}$$

where the last term comes from[55]

$$\frac{d}{d\alpha}\int_0^\infty\lambda_{i,t}e^{-\rho t}\left[g_i\left(x_t\right)+f_{i,j}\dot{x}_t+\alpha f_{i,j}\dot{h_t}\right]dt\bigg|_{\alpha=0}=\int_0^\infty\lambda_{i,t}e^{-\int_0^t\rho_s ds}f_{i,j}\dot{h}_tdt,$$

and integrating by parts-

$$\lim_{T\to\infty}\lambda_{i,T}e^{-\int_0^T\rho_s ds}f_{i,ji}h_T-\lambda_{i,0}f_{i,j}h_0-\int_0^\infty f_{i,j}\frac{d}{dt}\left(\lambda_{i,t}e^{-\int_0^t\rho_s ds}\right)h_tdt.$$

As the FOC equals zero for any h_t this implies (128).

In steady state, (128) implies

$$\frac{du}{dx_{j,ss}}+\sum_{j=i}^{n-1}\lambda_{i,ss}\left(\frac{dg_i}{dx_{j,ss}}+\rho f_{i,j}\right)=0, \tag{129}$$

[54]For expositional purposes, we limit ourselves to the case with a constant discount ρ.

[55]$f_{i,j}$ is the $j-$th element of vector f_i.

2.- Derive a linear-quadratic approximation. Now, consider a second-order approximation to the utility function

$$u(x_t) \approx u_{ss} + \sum_j \frac{du}{dx_{j,ss}} (x_{j,t} - x_{j,ss}) + s.o.t \equiv \tilde{u}(x_t).$$

where $s.o.t.$ are second order-terms. Similarly, the constraints can be approximated as

$$g_i (x_t, \dot{x}_t) \approx \sum_j \frac{dg_i}{dx_{j,ss}} (x_{j,t} - x_{j,ss}) + \sum_j f_{i,j} \dot{x}_{j,t} + s.o.t \equiv \tilde{g}_i(x_t, \dot{x}_t),$$

In equilibrium, up to second order we have that $\tilde{g}_i(x_t, \dot{x}_t) = 0$. We can thus add the term $\sum_{i=1}^{n-1} \lambda_{i,ss} \tilde{g}_i(x_t, \dot{x}_t)$ to our second order objective, without changing the welfare measure:

$$\int_0^\infty e^{-\rho t} \left(\tilde{u}(x_t) + \sum_i \lambda_{i,ss} \tilde{g}_i(x_t, \dot{x}_t) \right) dt, \tag{130}$$

If we just focus on the linear terms, we get

$$\int_0^\infty e^{-\rho t} \left(\sum_j \frac{du}{dx_{j,ss}} (x_t - x_{ss}) + \sum_i \lambda_{i,ss} \left[\sum_j \frac{dg_i}{dx_{j,ss}} (x_{j,t} - x_{j,ss}) + \sum_j f_{i,j} \dot{x}_{j,t} \right] \right) dt.$$

Integrating by parts the last term, we get

$$\lambda_{i,ss} \int_0^\infty e^{-\rho t} f_{i,j} \dot{x}_{j,t} dt = \lambda_{i,ss} \int_0^\infty e^{-\rho t} \rho f_{i,j} x_{j,t} dt,$$

plus the boundary conditions. We can then subtract the term independent of policy $\int_0^\infty e^{-\rho t} \rho \lambda_{i,ss} f_{i,j} x_{j,ss}$, so that we get

$$\int_0^\infty e^{-\rho t} \left(\sum_j \frac{du}{dx_{j,ss}} (x_{j,t} - x_{j,ss}) + \sum_i \lambda_{i,ss} \left[\sum_j \frac{dg_i}{dx_{j,ss}} (x_{j,t} - x_{j,ss}) + \sum_j \rho f_{i,j} (x_{j,t} - x_{j,ss}) \right] \right) dt.$$

Grouping all the terms in each variable j, we get

$$\int_0^\infty e^{-\rho t} \sum_j \left[\left(\frac{du}{dx_{j,ss}} + \sum_i \lambda_{i,ss} \left(\frac{dg_i}{dx_{j,ss}} + \rho \frac{dg_i}{d\dot{x}_{j,ss}} \right) \right) (x_{j,t} - x_{j,ss}) \right] dt,$$

As $\frac{du}{dx_{j,ss}} + \sum_{j=i}^{n-1} \lambda_{i,ss} \left(\frac{dg_i}{dx_{j,ss}} + \rho \frac{dg_i}{d\dot{x}_{j,ss}} \right) = 0$ due to equation (129), the first order terms must sum to 0.

Thus we have derived a purely quadratic objective, which is given by

$$\int_0^\infty e^{-\rho t} \left(u_{ss} + \tilde{u}^{2nd}(x_t) + \sum_{j=i}^{n-1} \lambda_{i,ss} \tilde{g_i}^{2nd}(x_t, \dot{x}_t) \right) dt,$$

where $\tilde{u}^{2nd}(x_t)$ and $\tilde{g_i}^{2nd}(x_t, \dot{x}_t)$ denotes the second order terms of the approximation.

E.4.2 The purely quadratic objective in the simplified model

Now we apply this procedure to our model. First, following the aforementioned steps, we need to derive second-order approximations to the equilibrium conditions with nonzero steady state multipliers. Having eliminated consumption in the Phillips curve using the resource constraint, just as we did when we set up the Lagrangian, we thus need to approximate:

$$\begin{aligned} 0 &= \rho_t^h \chi_{H,t} - \frac{d\chi_{H,t}}{dt} - \frac{\xi - 1}{\xi} \frac{\alpha \Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha} - 1} Y_t^{\frac{\vartheta+1}{1-\alpha}} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}}, \\ 0 &= \rho_t^h \chi_{F,t} - \frac{d\chi_{F,t}}{dt} - \frac{\varepsilon}{\theta} \left(\left(Y_t \left(1 - \frac{\theta}{2} \pi_t^2 \right) - \delta K \right)^{-\zeta} Y_t \frac{1-\varepsilon}{\varepsilon} + (1-\tau_t) \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y_t^{\frac{\vartheta+1}{1-\alpha}} Z_t^{\frac{-(\vartheta+1)}{1-\alpha}} \right), \end{aligned}$$

We need second-order approximations. First, we express the quantities as log (or linear) deviations, which we denote by hats.[56] In particular, we consider linear deviations of $\chi_{H,t}$, $\chi_{F,t}$ and π_t from their steady state values (which is zero for χ_F and π) and log deviations otherwise. In the case of the cost-push shock we define $\hat{\tau}_t = \log(1-\tau_t) - \log(1-\tau)$. Thus we can write:

$$\begin{aligned} 0 &= \rho_t^h \left(\hat{\chi}_{H,t} - \chi_H \right) - \frac{d\hat{\chi}_{H,t}}{dt} - \frac{\xi - 1}{\xi} \frac{\alpha \Upsilon}{(1-\alpha)} K^{\frac{-\alpha(\vartheta+1)}{1-\alpha} - 1} Y^{\frac{\vartheta+1}{1-\alpha}} Z^{\frac{-(\vartheta+1)}{1-\alpha}} \exp\left(\frac{\vartheta+1}{1-\alpha} \hat{Y}_t - \frac{\vartheta+1}{1-\alpha} \hat{Z}_t \right), \\ 0 &= \rho_t^h \left(\hat{\chi}_{F,t} \right) - \frac{d\hat{\chi}_{F,t}}{dt} - \left(\left(Y \exp\left(\hat{Y}_t \right) \left(1 - \frac{\theta}{2} \hat{\pi}_t^2 \right) - \delta K \right)^{-\zeta} Y \exp\left(\hat{Y}_t \right) \frac{1-\varepsilon}{\theta} \right) \\ &\quad - \left(\frac{\varepsilon}{\theta} (1-\tau) \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)} Y^{\frac{\vartheta+1}{1-\alpha}} Z^{\frac{-(\vartheta+1)}{1-\alpha}} \exp\left(\frac{\vartheta+1}{1-\alpha} \hat{Y}_t - \frac{\vartheta+1}{1-\alpha} \hat{Z}_t + \hat{\tau}_t \right) \right). \end{aligned}$$

Using the steady-state relations from Section E.1.1, we have that

[56] We do not consider (yet) any approximation to the time-preference shock.

$$
\begin{aligned}
\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+1}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}} &= q\rho^h C^{-\zeta}, \\
\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}}Z^{\frac{-(\vartheta+1)}{1-\alpha}} &= YC^{-\zeta}\frac{\varepsilon-1}{\varepsilon(1-\tau)}.
\end{aligned}
$$

We can thus simplify the system as

$$
\begin{aligned}
0 &= \rho_t^h\hat{\chi}_{H,t} - \frac{d\hat{\chi}_{H,t}}{dt} - q\rho^h C^{-\zeta}\exp\left(\left(\frac{\vartheta+\alpha}{1-\alpha}+1\right)\hat{Y}_t - \frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t\right), \\
0 &= \rho_t^h\hat{\chi}_{F,t} - \frac{d\hat{\chi}_{F,t}}{dt} \\
&\quad -\frac{1-\varepsilon}{\theta}\left(\left(Y\exp\left(\hat{Y}_t\right)\left(1-\frac{\theta}{2}\hat{\pi}_t^2\right)-\delta K\right)^{-\zeta}Y\exp\left(\hat{Y}_t\right)\right) \\
&\quad -\frac{1-\varepsilon}{\theta}\left(-YC^{-\zeta}\exp\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t - \frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)\right).
\end{aligned}
$$

Taking a second-order approximation

$$
\begin{aligned}
0 &= \rho_t^h\hat{\chi}_{H,t} - q\rho^h C^{-\zeta}\left(\left(\frac{\vartheta+1}{1-\alpha}\right)\hat{Y}_t - \frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t\right) - \frac{d\hat{\chi}_{H,t}}{dt} \\
&- q\rho^h C^{-\zeta}\frac{1}{2}\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t - \frac{\vartheta+1}{1-\alpha}\hat{Z}_t\right)^2, \\
0 &= \rho^h\hat{\chi}_{F,t} - \frac{d\hat{\chi}_{F,t}}{dt} - \frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\left(1-\zeta\frac{Y}{C}\right)\hat{Y}_t - \left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t - \frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)\right) \\
&- \frac{1}{2}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\left(\zeta\theta\frac{Y}{C}\hat{\pi}_t{}^2 + \left(1-3\zeta\frac{Y}{C}+(\zeta+1)\zeta\left(\frac{Y}{C}\right)^2\right)\hat{Y}_t{}^2\right)\right) \\
&- \frac{1}{2}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(-\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t - \frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)^2\right).
\end{aligned}
$$

Now, as in equation (130) in the general case, the purely quadratic objective can be obtained by adding the SS multipliers times the constraint to the second-order utility (122):

$$
W \approx \int_0^\infty e^{-\int_0^t \rho_s^h ds}C^{-\zeta}Y
$$

$$
\begin{aligned}
&\times\left\{\hat{Y}_t\left[1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right]+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\hat{Z}_t+\frac{1}{2}\left[\left[1-\zeta\frac{Y}{C}\right]-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\right]\hat{Y}_t^{\,2}\right\}dt\\
&+\int_0^\infty e^{-\int_0^t\rho_s^h ds}C^{-\zeta}Y\left\{-\frac{1}{2}\theta\hat{\pi}_t^2-\frac{1}{2}\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\hat{Z}_t^{\,2}+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\frac{(\vartheta+1)}{1-\alpha}\hat{Y}_t\hat{Z}_t\right\}dt\\
&-\lambda_{PC}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\left(1-\zeta\frac{Y}{C}\right)\hat{Y}_t-\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t-\frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)\right)\\
&-\lambda_{EE}q\rho^hC^{-\zeta}\left(\left(\frac{\vartheta+1}{1-\alpha}\right)\hat{Y}_t-\frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t\right)\\
&-\lambda_{PC}\frac{1}{2}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\zeta\theta\frac{Y}{C}\hat{\pi}_t^{\,2}+\left(1-3\zeta\frac{Y}{C}+(\zeta+1)\zeta\left(\frac{Y}{C}\right)^2\right)\hat{Y}_t^{\,2}\right)\\
&-\lambda_{PC}\frac{1}{2}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(-\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t-\frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)^2\right)\\
&-\lambda_{EE}q\rho^hC^{-\zeta}\frac{1}{2}\left(\frac{\vartheta+1}{1-\alpha}\right)^2\left(\hat{Y}_t-\hat{Z}_t\right)^2dt,
\end{aligned}
$$

where we have already canceled out the terms $\rho_t^h\hat{\chi}_{H,t}-\rho_t^h\hat{\chi}_{H,t}$ and $\rho_t^h\hat{\chi}_{F,t}-\rho_t^h\hat{\chi}_{F,t}$. Rearrange the constraints, one line for each variable as

$$
\begin{aligned}
W\ \approx\ &\int_0^\infty e^{-\int_0^t\rho_s^h ds}C^{-\zeta}Y\\
&\times\left\{\hat{Y}_t\left[1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right]+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\hat{Z}_t+\frac{1}{2}\left[\left[1-\zeta\frac{Y}{C}\right]-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\right]\hat{Y}_t^{\,2}\right\}dt\\
&\int_0^\infty e^{-\int_0^t\rho_s^h ds}C^{-\zeta}Y\left\{-\frac{1}{2}\theta\hat{\pi}_t^2-\frac{1}{2}\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\hat{Z}_t^{\,2}+\frac{\varepsilon-1}{\varepsilon(1-\tau)}\frac{(\vartheta+1)}{1-\alpha}\hat{Y}_t\hat{Z}_t\right\}dt\\
&+\hat{Y}_t\left[\lambda_{PC}\frac{\varepsilon-1}{\theta}C^{-\zeta}Y\left(1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}\right)-\lambda_{EE}q\rho^hC^{-\zeta}\left(\frac{\vartheta+1}{1-\alpha}\right)\right]\\
&+\hat{Z}_t\left[\lambda_{PC}\frac{\varepsilon-1}{\theta}C^{-\zeta}Y\frac{\vartheta+1}{1-\alpha}+\lambda_{EE}q\rho^hC^{-\zeta}\frac{(\vartheta+1)}{1-\alpha}\right]\\
&-\hat{\tau}_t\lambda_{PC}\frac{\varepsilon-1}{\theta}C^{-\zeta}Y\\
&-\lambda_{PC}\frac{1}{2}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\zeta\theta\frac{Y}{C}\hat{\pi}_t^{\,2}+\left(1-3\zeta\frac{Y}{C}+(\zeta+1)\zeta\left(\frac{Y}{C}\right)^2\right)\hat{Y}_t^{\,2}\right)\\
&-\lambda_{PC}\frac{1}{2}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(-\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t-\frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)^2\right)\\
&-\lambda_{EE}q\rho^hC^{-\zeta}\frac{1}{2}\left(\frac{\vartheta+1}{1-\alpha}\right)^2\left(\hat{Y}_t-\hat{Z}_t\right)^2dt,
\end{aligned}
$$

Notice that $\lambda_{PC}C^{-\zeta}Y\frac{1-\varepsilon}{\theta}=\frac{C^{-\zeta}Y}{1-\zeta\frac{Y}{C}}$, and $\lambda_{EE}q\rho^hC^{-\zeta}=\frac{\frac{1}{1-\zeta\frac{Y}{C}}\frac{\vartheta+1}{1-\alpha}-\frac{\varepsilon-1}{\varepsilon(1-\tau)}}{\frac{(\vartheta+1)}{1-\alpha}}YC^{-\zeta}$. Substitute these terms:

$$
\begin{aligned}
W \quad \approx \quad & \int_0^\infty e^{-\int_0^t \rho_s^h ds} C^{-\zeta} Y \\
& \times \left\{ \hat{Y}_t \left[1 - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right] + \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \hat{Z}_t + \frac{1}{2} \left[\left[1 - \zeta \frac{Y}{C} \right] - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y_t}^2 \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} C^{-\zeta} Y \left\{ -\frac{1}{2} \theta \hat{\pi}_t^2 - \frac{1}{2} \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \hat{Z_t}^2 + \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \right\} dt \\
& + \hat{Y}_t C^{-\zeta} Y \left[-1 + \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right] \\
& - \hat{Z}_t C^{-\zeta} Y \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \\
& + \hat{\tau}_t C^{-\zeta} Y \frac{1}{1 - \zeta \frac{Y}{C}} \\
& - \frac{1}{2} \frac{C^{-\zeta} Y}{1 - \zeta \frac{Y}{C}} \left(\left(\zeta \theta \frac{Y}{C} \hat{\pi_t}^2 + \left(1 - 3\zeta \frac{Y}{C} + (\zeta+1)\zeta \left(\frac{Y}{C} \right)^2 \right) \hat{Y_t}^2 \right) - \left(\frac{\vartheta+1}{1-\alpha} \hat{Y}_t - \frac{\vartheta+1}{1-\alpha} \hat{Z}_t + \hat{\tau}_t \right)^2 \right) \\
& - \left(\frac{1}{1 - \zeta \frac{Y}{C}} \frac{\vartheta+1}{1-\alpha} - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right) Y C^{-\zeta} \frac{1}{2} \left(\frac{\vartheta+1}{1-\alpha} \right) \left(\hat{Y}_t - \hat{Z}_t \right)^2 dt,
\end{aligned}
$$

The linear terms in $\hat{Y}_t$ and $\hat{Z}_t$ in the welfare function cancel (and $\hat{\tau}_t$ drops as t.i.p.) and we have a purely quadratic objective

$$
\begin{aligned}
W \quad \approx \quad & \int_0^\infty e^{-\int_0^t \rho_s^h ds} C^{-\zeta} Y \left\{ \frac{1}{2} \left[\left[1 - \zeta \frac{Y}{C} \right] - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y_t}^2 - \frac{1}{2} \theta \hat{\pi}_t^2 \right. \\
& - \frac{1}{2} \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \hat{Z_t}^2 + \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \\
& - \frac{1}{2} \frac{C^{-\zeta} Y}{1 - \zeta \frac{Y}{C}} \left(\left(\zeta \theta \frac{Y}{C} \hat{\pi_t}^2 + \left(1 - 3\zeta \frac{Y}{C} + (\zeta+1)\zeta \left(\frac{Y}{C} \right)^2 \right) \hat{Y_t}^2 \right) \right) \\
& - \frac{1}{2} \frac{C^{-\zeta} Y}{1 - \zeta \frac{Y}{C}} \left(- \left(\frac{\vartheta+1}{1-\alpha} \hat{Y}_t - \frac{\vartheta+1}{1-\alpha} \hat{Z}_t + \hat{\tau}_t \right)^2 \right) \\
& \left. - \left(\frac{1}{1 - \zeta \frac{Y}{C}} \frac{\vartheta+1}{1-\alpha} - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right) C^{-\zeta} Y \frac{1}{2} \left(\frac{\vartheta+1}{1-\alpha} \right) \left(\hat{Y}_t - \hat{Z}_t \right)^2 \right\} dt,
\end{aligned}
$$

$$
\begin{aligned}
W \quad \approx \quad & \int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ \frac{1}{2} \left[\left[1 - \zeta \frac{Y}{C} \right] - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y_t}^2 - \frac{1}{2} \theta \hat{\pi}_t^2 \right\} dt \\
& + \int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ -\frac{1}{2} \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \hat{Z_t}^2 + \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \right\} dt \\
& + \int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ -\frac{1}{2} \frac{1}{1 - \zeta \frac{Y}{C}} \left(\left(\zeta \theta \frac{Y}{C} \hat{\pi_t}^2 + \left(1 - 3\zeta \frac{Y}{C} + (\zeta+1)\zeta \left(\frac{Y}{C} \right)^2 \right) \hat{Y_t}^2 \right) \right) \right\} dt
\end{aligned}
$$

$$+\int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ -\frac{1}{2} \frac{1}{1-\zeta\frac{Y}{C}} \left(-\left(\left(\frac{\vartheta+1}{1-\alpha} \hat{Y}_t \right)^2 + \left(\frac{\vartheta+1}{1-\alpha} \hat{Z}_t \right)^2 \right) \right) \right\} dt$$
$$+\int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ -\frac{1}{2} \frac{1}{1-\zeta\frac{Y}{C}} \left(-\left(-2\left(\frac{\vartheta+1}{1-\alpha} \right)^2 \hat{Y}_t \hat{Z}_t + 2\frac{\vartheta+1}{1-\alpha} \hat{Y}_t \hat{\tau}_t \right) \right) \right\} dt$$
$$+\int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ -\frac{1}{2} \frac{1}{1-\zeta\frac{Y}{C}} \left(2\frac{\vartheta+1}{1-\alpha} \hat{Z}_t \hat{\tau}_t \right) \right\} dt$$
$$+\int_0^\infty e^{-\rho^h t} C^{-\zeta} Y \left\{ -\frac{1}{2} \left(\frac{1}{1-\zeta\frac{Y}{C}} \frac{\vartheta+1}{1-\alpha} - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \right) \left(\frac{\vartheta+1}{1-\alpha} \right) \left(\hat{Y}_t^2 - 2\hat{Y}_t \hat{Z}_t + \hat{Z}_t^2 \right) \right\} dt,$$

Simplifying terms, we get an approximation of welfare containing only second order terms:

$$\begin{aligned} W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} C^{-\zeta} Y \left[\left\{ \left[1-\zeta\frac{Y}{C} \right] - \frac{1}{1-\zeta\frac{Y}{C}} \left(1-3\zeta\frac{Y}{C} + (\zeta+1)\zeta\left(\frac{Y}{C}\right)^2 \right) \right\} \frac{1}{2} \hat{Y}_t^2 \right] dt. \\ & + \int_0^\infty e^{-\int_0^t \rho_s^h ds} C^{-\zeta} Y \left[-\frac{\theta}{2} \hat{\pi}_t^2 \left[1 + \frac{\zeta}{1-\zeta\frac{Y}{C}} \frac{Y}{C} \right] + \frac{1}{1-\zeta\frac{Y}{C}} \frac{\vartheta+1}{1-\alpha} \left(\hat{Y}_t - \hat{Z}_t \right) \hat{\tau}_t \right] dt. \end{aligned}$$

Special case with $\delta = 0$ When there is no depreciation ($\delta = 0$) the expression simplifies and the first term drops out:

$$W \approx \int_0^\infty e^{-\int_0^t \rho_s^h ds} Y^{1-\zeta} \left[-\frac{\theta}{2} \hat{\pi}_t^2 \frac{1}{1-\zeta} + \frac{1}{1-\zeta} \frac{\vartheta+1}{1-\alpha} \left(\hat{Y}_t - \hat{Z}_t \right) \hat{\tau}_t \right] dt.$$

Taking a second-order approximation around the time-preference shock, we get

$$W \approx \int_0^\infty e^{-\rho^h t} Y^{1-\zeta} \left[-\frac{\theta}{2} \hat{\pi}_t^2 \frac{1}{1-\zeta} + \frac{1}{1-\zeta} \frac{\vartheta+1}{1-\alpha} \left(\hat{Y}_t - \hat{Z}_t \right) \hat{\tau}_t \right] dt.$$

That is, welfare depends only on inflation and the cross-term between the cost-push shock and the difference of the log-deviations of output minus TFP. This expression makes evident why, in the absence of cost-push shocks (i.e., with TFP or preference shocks), the optimal policy features zero inflation and divine coincidence holds.

E.4.3 The purely quadratic objective with complete markets

We proceed as above in the case of complete markets, adding the quadratic constraints from above times the steady-state multipliers to the objective, we have

$$
\begin{aligned}
W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(\hat{Y}_t Y^{1-\zeta} \left[1 - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right] + \frac{1}{2} Y^{1-\zeta} \left[(1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y}_t{}^2 \right) dt \\
& \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(-\frac{1}{2} Y^{1-\zeta} \theta \hat{\pi}_t^2 + Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \right) dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} \left((1-\zeta) \hat{Y}_t \frac{1-\varepsilon}{\theta} Y^{1-\zeta} \right) dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} \left(\frac{\varepsilon}{\theta} (1-\tau) Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{\vartheta+1}{1-\alpha} \hat{Y}_t - \frac{\vartheta+1}{1-\alpha} \hat{Z}_t + \hat{\tau}_t \right) \right) dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} \frac{1}{2} \left(\left(\zeta \theta \hat{\pi}_t{}^2 + (1-\zeta)^2 \hat{Y}_t{}^2 \right) Y^{1-\zeta} \frac{1-\varepsilon}{\theta} \right) dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} \frac{1}{2} \left(\frac{\varepsilon}{\theta} (1-\tau) Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left[\left(\frac{\vartheta+1}{1-\alpha} \right)^2 \left(\hat{Y}_t - \hat{Z}_t \right)^2 \right] \right) dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} \frac{1}{2} \left(\frac{\varepsilon}{\theta} (1-\tau) Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left[2 \frac{\vartheta+1}{1-\alpha} \hat{Y}_t \hat{\tau}_t - 2 \frac{\vartheta+1}{1-\alpha} \hat{Z}_t \hat{\tau}_t + \hat{\tau}_t^2 \right] \right) dt.
\end{aligned}
$$

Dropping the t.i.p. terms in $\hat{Z}_t$ and $\hat{\tau}_t$

$$
\begin{aligned}
W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \hat{Y}_t Y^{1-\zeta} \left[1 - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right] + \frac{1}{2} Y^{1-\zeta} \left[(1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y}_t{}^2 \right\} dt \\
& \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ -\frac{1}{2} Y^{1-\zeta} \theta \hat{\pi}_t^2 + Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \hat{Y}_t \lambda_{PC} \frac{\varepsilon-1}{\theta} Y^{1-\zeta} \left(1 - \zeta - \frac{\vartheta+1}{1-\alpha} \right) dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} Y^{1-\zeta} \frac{\varepsilon-1}{\theta} \frac{1}{2} \left(\left(\zeta \theta \hat{\pi}_t{}^2 + (1-\zeta)^2 \hat{Y}_t{}^2 \right) - \left[\left(\frac{\vartheta+1}{1-\alpha} \right)^2 \left(\hat{Y}_t - \hat{Z}_t \right)^2 + 2 \frac{\vartheta+1}{1-\alpha} \hat{Y}_t \hat{\tau}_t \right] \right) dt.
\end{aligned}
$$

Notice again that now $\lambda_{PC} \frac{\varepsilon-1}{\theta} = \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}$.Substituting this term, we get

$$
\begin{aligned}
W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \hat{Y}_t Y^{1-\zeta} \left[1 - \frac{\varepsilon - 1}{\varepsilon(1-\tau)} \right] + \frac{1}{2} Y^{1-\zeta} \left[(1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y}_t{}^2 \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ -\frac{1}{2} Y^{1-\zeta} \theta \hat{\pi}_t^2 + Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \right\} dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} \hat{Y}_t \left[1 - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \right] Y^{1-\zeta} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} Y^{1-\zeta} \frac{1}{2} \left(\zeta \theta \hat{\pi}_t{}^2 + (1-\zeta)^2 \hat{Y}_t{}^2 \right) \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \lambda_{PC} Y^{1-\zeta} \frac{\varepsilon-1}{\theta} \frac{1}{2} \left[\left(\frac{\vartheta+1}{1-\alpha} \right)^2 \left(\hat{Y}_t - \hat{Z}_t \right)^2 + 2 \frac{\vartheta+1}{1-\alpha} \hat{Y}_t \hat{\tau}_t \right] dt.
\end{aligned}
$$

The linear term in $\hat{Y}_t$ in the welfare function cancels and we have a purely quadratic objective:

$$
\begin{aligned}
W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \frac{1}{2} Y^{1-\zeta} \left[(1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y_t}^2 \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ -\frac{1}{2} Y^{1-\zeta} \theta \hat{\pi}_t^2 + Y^{1-\zeta} \frac{\varepsilon-1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} \hat{Y}_t \hat{Z}_t \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} Y^{1-\zeta} \frac{1}{2} \left(\left(\zeta\theta\hat{\pi_t}^2 + (1-\zeta)^2 \hat{Y_t}^2 \right) \right) \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left\{ \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} Y^{1-\zeta} \frac{1}{2} \left(-\left[\left(\frac{\vartheta+1}{1-\alpha} \right)^2 \left(\hat{Y_t}^2 - 2\hat{Y}_t \hat{Z}_t \right) + 2\frac{\vartheta+1}{1-\alpha} \hat{Y}_t \hat{\tau}_t \right] \right) \right\} dt.
\end{aligned}
$$

Next, we collect all the terms

$$
\begin{aligned}
W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(-\frac{1}{2} Y^{1-\zeta} \theta \left[1 - \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \zeta \right] \hat{\pi}_t^2 \right) dt \\
& - \int_0^\infty e^{-\int_0^t \rho_s^h ds} Y^{1-\zeta} \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left(\frac{\vartheta+1}{1-\alpha} \right) \hat{Y}_t \hat{\tau}_t dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \frac{1}{2} Y^{1-\zeta} \left[(1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] \hat{Y_t}^2 dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \frac{1}{2} Y^{1-\zeta} \left[\frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left((1-\zeta)^2 - \left(\frac{\vartheta+1}{1-\alpha} \right)^2 \right) \right] \hat{Y_t}^2 dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(Y^{1-\zeta} \left[\frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{\vartheta+1}{1-\alpha} \right) + \frac{\left(1-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\right)}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left(\frac{\vartheta+1}{1-\alpha} \right)^2 \right] \hat{Y}_t \hat{Z}_t \right) dt.
\end{aligned}
$$

Now divide and multiply the last 2 lines of the expression by the coefficient of $\hat{Y_t}^2$

$$
\begin{aligned}
W \approx & \int_0^\infty e^{-\int_0^t \rho_s^h ds} \left(-\frac{1}{2} Y^{1-\zeta} \theta \hat{\pi_t}^2 \left[1 - \frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \zeta \right] \right) dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} \hat{Y}_t \hat{\tau}_t Y^{1-\zeta} \left\{ \frac{\frac{\varepsilon-1}{\varepsilon(1-\tau)} - 1}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left(\frac{\vartheta+1}{1-\alpha} \right) \right\} dt \\
& + \int_0^\infty e^{-\int_0^t \rho_s^h ds} Y^{1-\zeta} \left\{ (1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) + \frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left((1-\zeta)^2 - \left(\frac{\vartheta+1}{1-\alpha} \right)^2 \right) \right\} \\
& \times \left[\hat{Y_t}^2 \frac{1}{2} + \hat{Y}_t \hat{Z}_t \frac{\frac{\varepsilon-1}{\varepsilon(1-\tau)} \frac{(\vartheta+1)}{1-\alpha} + \frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left(\frac{\vartheta+1}{1-\alpha} \right)^2}{\left[(1-\zeta) - \frac{\varepsilon-1}{\varepsilon(1-\tau)} \left(\frac{1+\vartheta}{1-\alpha} \right) \right] + \frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)} \left((1-\zeta)^2 - \left(\frac{\vartheta+1}{1-\alpha} \right)^2 \right)} \right] dt
\end{aligned}
$$

Consider the large fraction at the end. After some algebra we get

$$\frac{\frac{\varepsilon-1}{\varepsilon(1-\tau)}\frac{(\vartheta+1)}{1-\alpha}+\frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\left(\frac{\vartheta+1}{1-\alpha}\right)^2}{\left[(1-\zeta)-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\right]+\frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\left((1-\zeta)^2-\left(\frac{\vartheta+1}{1-\alpha}\right)^2\right)} \quad = \quad -\frac{\frac{(\vartheta+1)}{1-\alpha}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}.$$

Also the coefficient in the third line can be simplified further to yield:

$$\begin{aligned}&\left[(1-\zeta)-\frac{\varepsilon-1}{\varepsilon(1-\tau)}\left(\frac{1+\vartheta}{1-\alpha}\right)\right]-\frac{-1+\frac{\varepsilon(1-\tau)}{\varepsilon-1}}{(1-\tau)\frac{\varepsilon}{\theta}\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\frac{1-\varepsilon}{\theta}\left((1-\zeta)^2-\left(\frac{\vartheta+1}{1-\alpha}\right)^2\right)\\ =&-(1-\zeta)\frac{1-\varepsilon}{(1-\tau)\varepsilon}-\frac{\vartheta+1}{1-\alpha}.\end{aligned}$$

Thus we can write

$$\begin{aligned}W \quad \approx \quad & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\left(-\frac{1}{2}Y^{1-\zeta}\theta\hat{\pi}_t{}^2\left[1-\frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\zeta\right]\right)dt\\ &+\int_0^\infty e^{-\int_0^t \rho_s^h ds}\hat{Y}_t\hat{\tau}_t Y^{1-\zeta}\left[\frac{\frac{\varepsilon-1}{\varepsilon(1-\tau)}-1}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\left(\frac{\vartheta+1}{1-\alpha}\right)\right]dt\\ &-\int_0^\infty e^{-\int_0^t \rho_s^h ds}\frac{1}{2}Y^{1-\zeta}\left[\frac{\vartheta+1}{1-\alpha}-(1-\zeta)\frac{\varepsilon-1}{(1-\tau)\varepsilon}\right]\left[\hat{Y}_t-\frac{\frac{(\vartheta+1)}{1-\alpha}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\hat{Z}_t\right]^2 dt\\ = \quad & \int_0^\infty e^{-\int_0^t \rho_s^h ds}\left(-\frac{1}{2}Y^{1-\zeta}\theta\hat{\pi}_t{}^2\left[1+\frac{1-\frac{\varepsilon-1}{(1-\tau)\varepsilon}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\zeta\right]\right)dt\\ &+\int_0^\infty e^{-\int_0^t \rho_s^h ds}\left(\hat{Y}_t\hat{\tau}_t Y^{1-\zeta}\left[\frac{\frac{\varepsilon-1}{\varepsilon(1-\tau)}-1}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\left(\frac{\vartheta+1}{1-\alpha}\right)\right]\right)dt\\ &+\int_0^\infty e^{-\int_0^t \rho_s^h ds}\left(-\frac{1}{2}Y^{1-\zeta}\left\{\left[\frac{\vartheta+1}{1-\alpha}-(1-\zeta)\frac{\varepsilon-1}{(1-\tau)\varepsilon}\right]\right\}\tilde{Y}_t^2\right)dt,\end{aligned}$$

where

$$\tilde{Y}_t \equiv \hat{Y}_t-\frac{\frac{(\vartheta+1)}{1-\alpha}}{\left(\zeta-1+\frac{\vartheta+1}{1-\alpha}\right)}\hat{Z}_t.$$

is the output gap relative to the efficient level of output. This coincides with the result in Benigno and Woodford (2005).[57]

Note that, with the markup-offsetting subsidy in place, the welfare function becomes

$$W = \int_0^\infty e^{-\int_0^t \rho_s^h ds} Y^{1-\zeta} \left(-\hat{\pi}_t^2 \frac{\theta}{2} - \frac{1}{2} \left\{ \frac{\vartheta + 1}{1-\alpha} - (1-\zeta) \right\} \tilde{Y}_t^2 \right) dt,$$

and taking a second-order approximation for the time-preference shock

$$W = \int_0^\infty e^{-\rho^h t} Y^{1-\zeta} \left(-\hat{\pi}_t^2 \frac{\theta}{2} - \frac{1}{2} \left\{ \frac{\vartheta + 1}{1-\alpha} - (1-\zeta) \right\} \tilde{Y}_t^2 \right) dt.$$

We can also write the linear Phillips curve in terms of $\tilde{Y}_t$:

$$\rho^h \hat{\pi}_t = \dot{\hat{\pi}}_t + \frac{1-\varepsilon}{\theta} \left(\left[(1-\zeta) - \frac{\vartheta+1}{1-\alpha} \right] \tilde{Y}_t - \hat{\tau}_t \right). \tag{131}$$

E.5 The Linear-quadratic problem

Now we are ready to express and solve the planner's problem as a simple linear quadratic problem. That is, we maximize the purely quadratic objective subject to linearized constraints.

E.5.1 Linear-quadratic problem of the simplified model

We analyze the problem in the presence of only cost-push shocks. The central bank maximizes

$$\begin{aligned}
&\int_0^\infty e^{-\rho^h t} \left[C^{-\zeta} Y \left\{ \left[1 - \zeta \frac{Y}{C} \right] - \frac{1}{1-\zeta\frac{Y}{C}} \left(1 - 3\zeta \frac{Y}{C} + (\zeta+1)\zeta \left(\frac{Y}{C} \right)^2 \right) \right\} \frac{1}{2} \hat{Y}_t^2 \right] dt \\
&+ \int_0^\infty e^{-\rho^h t} \left[C^{-\zeta} Y \left\{ -\frac{\theta}{2} \hat{\pi}_t^2 \left[1 + \frac{\zeta}{1-\zeta\frac{Y}{C}} \frac{Y}{C} \right] + \frac{1}{1-\zeta\frac{Y}{C}} \frac{\vartheta+1}{1-\alpha} \left(\hat{Y}_t - \hat{Z}_t \right) \hat{\tau}_t \right\} \right] dt \\
&+ \int_0^\infty e^{-\rho^h t} \hat{\lambda}_{A,t} \left[-A \frac{d\hat{A}_t}{dt} - \hat{A}_t A (1-\psi)\eta \right] dt \\
&+ \int_0^\infty e^{-\rho^h t} \hat{\lambda}_{A,t} \left[\alpha \frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)} Y^{\frac{\vartheta+1}{1-\alpha}+\zeta} Z^{\frac{-(\vartheta+1)}{1-\alpha}} \frac{1}{\xi} \right] \\
&\times \left[\left(1 + \frac{A}{K} (\xi - 1) \right) \left(\left(\frac{\vartheta+1}{1-\alpha} \right) \hat{Y}_t + \zeta \hat{C}_t - \frac{(\vartheta+1)}{1-\alpha} \hat{Z}_t - \hat{q}_t \right) + \frac{A}{K} (\xi - 1) \hat{A}_t \right] dt \\
&+ \int_0^\infty e^{-\rho^h t} \hat{\lambda}_{TR,t} \left[-\hat{C}_t + \hat{Y}_t \right] dt
\end{aligned}$$

[57] They consider a more general version of the complete-markets model. Our version is nested in theirs.

$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{Z,t}\left[-\hat{Z}_t+\frac{\alpha}{\xi}\hat{A}_t\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{EE,t}\left[-\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\left(\left(\frac{\vartheta+1}{1-\alpha}\right)\hat{Y}_t-\frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t\right)\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{EE,t}\left[\rho^h\hat{\chi}_{H,t}-\frac{d\hat{\chi}_{H,t}}{dt}\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{PC,t}\left[\rho^h\hat{\chi}_{F,t}-\frac{d\hat{\chi}_{F,t}}{dt}-\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\left(1-\zeta\frac{Y}{C}\right)\hat{Y}_t-\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t-\frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)\right)\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{H,t}\left[-\zeta C^{-\zeta}q\hat{C}_t+C^{-\zeta}q\hat{q}_t-\hat{\chi}_{H,t}\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{F,t}\left[C^{-\zeta}Y\hat{\pi}_t-\hat{\chi}_{F,t}\right]dt,$$

with respect to $\hat{\pi}_t,\hat{Y}_t,\hat{C}_t,\hat{Z}_t,\hat{q}_t,\hat{A}_t,\hat{\chi}_{H,t},\hat{\chi}_{F,t}$. If we integrate by parts, we get

$$\int_0^\infty e^{-\rho^h t}\left[C^{-\zeta}Y\left[\left\{\left[1-\zeta\frac{Y}{C}\right]-\frac{1}{1-\zeta\frac{Y}{C}}\left(1-3\zeta\frac{Y}{C}+(\zeta+1)\zeta\left(\frac{Y}{C}\right)^2\right)\right\}\frac{1}{2}\hat{Y}_t^{\,2}\right]\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\left[C^{-\zeta}Y\left[-\frac{\theta}{2}\hat{\pi}_t^2\left[1+\frac{\zeta}{1-\zeta\frac{Y}{C}}\frac{Y}{C}\right]+\frac{1}{1-\zeta\frac{Y}{C}}\frac{\vartheta+1}{1-\alpha}\left(\hat{Y}_t-\hat{Z}_t\right)\hat{\tau}_t\right]\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{A,t}\left[-\rho^h A\hat{A}_t-\hat{A}_tA(1-\psi)\eta\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{A,t}\left[-\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\right]$$
$$\times\left[\left(1+\frac{A}{K}(\xi-1)\right)\left(\left(\frac{\vartheta+1}{1-\alpha}\right)\hat{Y}_t+\zeta\hat{C}_t-\frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t-\hat{q}_t\right)+\frac{A}{K}(\xi-1)\hat{A}_t\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}A\hat{A}_t\frac{d\hat{\lambda}_{A,t}}{dt}dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{TR,t}\left[-\hat{C}_t+\hat{Y}_t\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{Z,t}\left[-\hat{Z}_t+\frac{\alpha}{\xi}\hat{A}_t\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{EE,t}\left[-\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\left(\left(\frac{\vartheta+1}{1-\alpha}\right)\hat{Y}_t-\frac{(\vartheta+1)}{1-\alpha}\hat{Z}_t\right)\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\chi}_{H,t}\frac{d\hat{\lambda}_{EE,t}}{dt}dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{PC,t}\left[-\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(\left(1-\zeta\frac{Y}{C}\right)\hat{Y}_t-\left(\frac{\vartheta+1}{1-\alpha}\hat{Y}_t-\frac{\vartheta+1}{1-\alpha}\hat{Z}_t+\hat{\tau}_t\right)\right)\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\chi}_{F,t}\frac{d\hat{\lambda}_{PC,t}}{dt}dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{H,t}\left[-\zeta C^{-\zeta}q\hat{C}_t+C^{-\zeta}q\hat{q}_t-\hat{\chi}_{H,t}\right]dt$$
$$+\int_0^\infty e^{-\rho^h t}\hat{\lambda}_{F,t}\left[C^{-\zeta}Y\hat{\pi}_t-\hat{\chi}_{F,t}\right]dt,$$

The first-order conditions

$$\frac{dL}{dC_t}:0\;=\;\hat{\lambda}_{A,t}\left[\zeta\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\right]-\hat{\lambda}_{TR,t}-\hat{\lambda}_{H,t}\zeta C^{-\zeta}q,$$

$$
\begin{aligned}
\frac{dL}{dY_t}:0 &= C^{-\zeta}Y\left\{\left[1-\zeta\frac{Y}{C}\right]-\frac{1}{1-\zeta\frac{Y}{C}}\left(1-3\zeta\frac{Y}{C}+(\zeta+1)\zeta\left(\frac{Y}{C}\right)^2\right)\right\}+\frac{1}{1-\zeta\frac{Y}{C}}\frac{\vartheta+1}{1-\alpha}C^{-\zeta}Y\hat{\tau}_t \\
&+ \hat{\lambda}_{A,t}\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\left(\frac{\vartheta+1}{1-\alpha}\right) \\
&+ \hat{\lambda}_{TR,t}-\hat{\lambda}_{EE,t}\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\left(\frac{\vartheta+1}{1-\alpha}\right) \\
&\quad -\hat{\lambda}_{PC,t}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\left(1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}\right),
\end{aligned}
$$

$$
\frac{dL}{d\pi_t}:0 = -\theta\hat{\pi}_t\left[1+\frac{\zeta}{1-\zeta\frac{Y}{C}}\frac{Y}{C}\right]C^{-\zeta}Y+C^{-\zeta}Y\hat{\lambda}_{F,t},
$$

$$
\begin{aligned}
\frac{dL}{dZ_t}:0 &= -\frac{1}{1-\zeta\frac{Y}{C}}\frac{\vartheta+1}{1-\alpha}C^{-\zeta}Y\hat{\tau}_t \\
&- \hat{\lambda}_{A,t}\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\frac{(\vartheta+1)}{1-\alpha} \\
&\quad -\hat{\lambda}_{Z,t}+\hat{\lambda}_{EE,t}\frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{(\vartheta+1)}{1-\alpha} \\
&\quad -\hat{\lambda}_{PC,t}\frac{1-\varepsilon}{\theta}C^{-\zeta}Y\frac{\vartheta+1}{1-\alpha},
\end{aligned}
$$

$$
\frac{dL}{dq_t}:0 = -\hat{\lambda}_{A,t}\alpha\tfrac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\tfrac{1}{\xi}\left(1+\tfrac{A}{K}(\xi-1)\right)+\hat{\lambda}_{H,t}C^{-\zeta}q \quad ,
$$

$$
\begin{aligned}
\frac{dL}{dA_t}:0 &= \hat{\lambda}_{A,t}\left[-\rho^h A+\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\frac{A}{K}(\xi-1)-A(1-\psi)\eta\right] \\
&\quad +A\frac{d\hat{\lambda}_{A,t}}{dt}+\hat{\lambda}_{Z,t}\frac{\alpha}{\xi},
\end{aligned}
$$

$$
\begin{aligned}
\frac{dL}{d\chi_{F,t}}:0 &= \frac{d\hat{\lambda}_{PC,t}}{dt}-\hat{\lambda}_{F,t}, \\
\frac{dL}{d\chi_{H,t}}:0 &= \frac{d\hat{\lambda}_{EE,t}}{dt}-\hat{\lambda}_{H,t}.
\end{aligned}
$$

The boundary conditions are

$$\begin{aligned}\hat{\lambda}_{PC,t} &= \hat{\lambda}_{EE,t} = 0, \\ \lim_{t\to\infty} e^{-\rho^h t}\hat{\lambda}_{A,t} &= 0.\end{aligned}$$

Notice that those first-order conditions (by construction) coincide with the (log)linearized versions of the first-order conditions of the nonlinear Ramsey problem E.2.1. The only difference regards the initial conditions: While there we solved the problem from a time 0 perspective, here the initial conditions reflect the timeless perspective: the require that the multipliers for the Phillips curve and Euler equation start at their steady state values.

E.5.2 Optimal response to cost-push shocks

Simplifying the system of ODEs We can write this system of ODEs in a compressed way, bunching terms depending only on parameters:

$$0 = \Omega_1\hat{\lambda}_{A,t} - \hat{\lambda}_{TR,t} - \Omega_2\hat{\lambda}_{H,t},$$

$$0 = \Omega_{14}\hat{Y}_t - \Omega_3\hat{\tau}_t + \Omega_4\hat{\lambda}_{A,t} + \hat{\lambda}_{TR,t} - \Omega_5\hat{\lambda}_{EE,t} - \Omega_6\hat{\lambda}_{PC,t},$$

$$0 = \Omega_7\hat{\pi}_t + \hat{\lambda}_{F,t},$$

$$0 = \Omega_3\hat{\tau}_t - \Omega_4\hat{\lambda}_{A,t} - \hat{\lambda}_{Z,t} + \Omega_5\hat{\lambda}_{EE,t} + \Omega_8\hat{\lambda}_{PC,t}$$

$$0 = -\Omega_9\hat{\lambda}_{A,t} + \Omega_{10}\hat{\lambda}_{H,t} \quad ,$$

$$0 = \Omega_{11}\hat{\lambda}_{A,t} + \Omega_{12}\frac{d\hat{\lambda}_{A,t}}{dt} + \Omega_{13}\hat{\lambda}_{Z,t},$$

$$\begin{aligned}
0 &= \frac{d\hat{\lambda}_{PC,t}}{dt} - \hat{\lambda}_{F,t}, \\
0 &= \frac{d\hat{\lambda}_{EE,t}}{dt} - \hat{\lambda}_{H,t}.
\end{aligned}$$

where we have used the following definitions

$$\begin{aligned}
\Omega_1 &= \left[\zeta\alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\right], \\
\Omega_2 &= \zeta C^{-\zeta}q, \\
\Omega_3 &= -\frac{1}{1-\zeta\frac{Y}{C}}\frac{\vartheta+1}{1-\alpha}C^{-\zeta}Y \\
\Omega_4 &= \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right)\left(\frac{\vartheta+1}{1-\alpha}\right) \\
\Omega_5 &= \frac{\xi-1}{\xi}\frac{\alpha\Upsilon}{(1-\alpha)}K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}-1}Y^{\frac{\vartheta+\alpha}{1-\alpha}+1}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\left(\frac{\vartheta+1}{1-\alpha}\right), \\
\Omega_6 &= \frac{1-\varepsilon}{\theta}YC^{-\zeta}\left(1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}\right), \\
\Omega_7 &= \theta\left[1+\frac{\zeta}{1-\zeta\frac{Y}{C}}\frac{Y}{C}\right]C^{-\zeta}Y, \\
\Omega_8 &= -\frac{1-\varepsilon}{\theta}YC^{-\zeta}\frac{\vartheta+1}{1-\alpha}, \\
\Omega_9 &= \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\left(1+\frac{A}{K}(\xi-1)\right), \\
\Omega_{10} &= C^{-\zeta}q, \\
\Omega_{11} &= -\rho^h A + \alpha\frac{\Upsilon K^{\frac{-\alpha(\vartheta+1)}{1-\alpha}}}{q(1-\alpha)}Y^{\frac{\vartheta+1}{1-\alpha}+\zeta}Z^{\frac{-(\vartheta+1)}{1-\alpha}}\frac{1}{\xi}\frac{A}{K}(\xi-1) - A(1-\psi)\eta, \\
\Omega_{12} &= A, \\
\Omega_{13} &= \frac{\alpha}{\xi}. \\
\Omega_{14} &= C^{-\zeta}Y\left\{\left[1-\zeta\frac{Y}{C}\right]-\frac{1}{1-\zeta\frac{Y}{C}}\left(1-3\zeta\frac{Y}{C}+(\zeta+1)\zeta\left(\frac{Y}{C}\right)^2\right)\right\}
\end{aligned}$$

The boundary conditions are

$$\begin{aligned}\hat{\lambda}_{PC,t} &= \hat{\lambda}_{EE,t}=0,\\ \lim_{t\to\infty} e^{-\rho^h t}\hat{\lambda}_{A,t} &= 0.\end{aligned}$$

We may simplify the system as

$$0 = \Omega_{14}\hat{Y}_t - \Omega_3\hat{\tau}_t + \left(\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\hat{\lambda}_{A,t} - \Omega_5\hat{\lambda}_{EE,t} - \Omega_6\hat{\lambda}_{PC,t}, \quad (132)$$

$$0 = \Omega_3\hat{\tau}_t + \left(\frac{\Omega_{11}}{\Omega_{13}}-\Omega_4\right)\hat{\lambda}_{A,t} + \frac{\Omega_{12}}{\Omega_{13}}\frac{d\hat{\lambda}_{A,t}}{dt} + \Omega_5\hat{\lambda}_{EE,t} + \Omega_8\hat{\lambda}_{PC,t}$$

$$\begin{aligned}0 &= \frac{d\hat{\lambda}_{PC,t}}{dt} + \Omega_7\hat{\pi}_t,\\ 0 &= \frac{d\hat{\lambda}_{EE,t}}{dt} - \frac{\Omega_9}{\Omega_{10}}\hat{\lambda}_{A,t}.\end{aligned}$$

We integrate the latter two equations

$$\begin{aligned}\int_0^t \frac{d\hat{\lambda}_{PC,t}}{dt} = \hat{\lambda}_{PC,t} &= -\Omega_7\int_0^t \pi_s ds \equiv -\Omega_7\hat{P}_t,\\ \int_0^t \frac{d\hat{\lambda}_{EE,t}}{dt} = \hat{\lambda}_{EE,t} &= \frac{\Omega_9}{\Omega_{10}}\int_0^t \hat{\lambda}_{A,s} ds \equiv \frac{\Omega_9}{\Omega_{10}}\hat{\lambda}_{CA,t}.\end{aligned}$$

where the first equation in each line results from the fact that these two co-states are predetermined and cannot jump, and the last equation defines the price level deviation $\hat{P}_t = \int_0^t \pi_s ds$ and the accumulated Lagrange multiplier on the net worth equation $\hat{\lambda}_{CA,t} = \int_0^t \hat{\lambda}_{A,s} ds$. Notice that $\lim_{t\to\infty}\dot{x}_t = 0$ (because $\lim_{t\to\infty} x_t = 0$) and $x_0 = 0$ for $x = \left\{\hat{\lambda}_{CA,t}, \hat{P}_t\right\}$.

Using this, we now eliminate $\hat{\lambda}_{PC,t}$, $\hat{\lambda}_{EE,t}$ and $\hat{\lambda}_{A,t}$ in the first 2 equations:

$$0 = \Omega_{14}\hat{Y}_t - \Omega_3\hat{\tau}_t + \left(\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\frac{d\hat{\lambda}_{CA,t}}{dt} - \Omega_5\frac{\Omega_9}{\Omega_{10}}\hat{\lambda}_{CA,t} + \Omega_6\Omega_7\hat{P}_t, \quad (133)$$

$$0 = \Omega_3\hat{\tau}_t + \left(\frac{\Omega_{11}}{\Omega_{13}} - \Omega_4\right)\frac{d\hat{\lambda}_{CA,t}}{dt} + \frac{\Omega_{12}}{\Omega_{13}}\frac{d^2\hat{\lambda}_{CA,t}}{dt^2} + \Omega_5\frac{\Omega_9}{\Omega_{10}}\hat{\lambda}_{CA,t} - \Omega_8\Omega_7\hat{P}_t. \quad (134)$$

We multiply the first equation by $\frac{\Omega_8}{\Omega_6}$ and obtain

$$0 = \Omega_{14}\frac{\Omega_8}{\Omega_6}\hat{Y}_t - \Omega_3\frac{\Omega_8}{\Omega_6}\hat{\tau}_t + \frac{\Omega_8}{\Omega_6}\left(\Omega_4 + \Omega_1 - \Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\frac{d\hat{\lambda}_{CA,t}}{dt} - \Omega_5\frac{\Omega_8}{\Omega_6}\frac{\Omega_9}{\Omega_{10}}\hat{\lambda}_{CA,t} + \Omega_8\Omega_7\hat{P}_t,$$

and adding the two equations we get

$$\begin{aligned} 0 = {} & \Omega_{14}\frac{\Omega_8}{\Omega_6}\hat{Y}_t + \left(\Omega_3\left(1 - \frac{\Omega_8}{\Omega_6}\right)\right)\hat{\tau}_t + \left(\frac{\Omega_8}{\Omega_6}\left(\Omega_4 + \Omega_1 - \Omega_2\frac{\Omega_9}{\Omega_{10}}\right) + \frac{\Omega_{11}}{\Omega_{13}} - \Omega_4\right)\frac{d\hat{\lambda}_{CA,t}}{dt} \\ & + \frac{\Omega_{12}}{\Omega_{13}}\frac{d^2\hat{\lambda}_{CA,t}}{dt^2} + \Omega_5\frac{\Omega_9}{\Omega_{10}}\left(1 - \frac{\Omega_8}{\Omega_6}\right)\hat{\lambda}_{CA,t}. \end{aligned}$$

We can re-write it as

$$\frac{d^2\hat{\lambda}_{CA,t}}{dt^2} + \Psi_1\frac{d\hat{\lambda}_{CA,t}}{dt} + \Psi_2\hat{\lambda}_{CA,t} = F_t.$$

where

$$\begin{aligned} \Psi_1 &= \frac{\Omega_{13}}{\Omega_{12}}\left(\frac{\Omega_8}{\Omega_6}\left(\Omega_4 + \Omega_1 - \Omega_2\frac{\Omega_9}{\Omega_{10}}\right) + \frac{\Omega_{11}}{\Omega_{13}} - \Omega_4\right), \\ \Psi_2 &= \Omega_5\frac{\Omega_{13}}{\Omega_{12}}\frac{\Omega_9}{\Omega_{10}}\left(1 - \frac{\Omega_8}{\Omega_6}\right), \\ \Psi_3 &= -\frac{\Omega_{13}}{\Omega_{12}}\left(\Omega_3\left(1 - \frac{\Omega_8}{\Omega_6}\right)\right), & (135) \\ \Psi_4 &= -\Omega_{14}\frac{\Omega_8}{\Omega_6}, & (136) \\ F_t &= \Psi_3\hat{\tau}_t + \Psi_4\hat{Y}_t. \end{aligned}$$

Solving the forward-looking ODE This is a non-homogeneous linear second order ODE. The characteristic polynomial is

$$r^2 + \Psi_1 r + \Psi_2.$$

Its roots are

$$r = \frac{-\Psi_1 \pm \sqrt{\Psi_1^2 - 4\Psi_2}}{2}. \tag{137}$$

We focus on the case where one root is positive and one negative and both are real. The solution has two roots, r_1 and r_2, with $r_2 > r_1$. The first root r_1 will be negative, while the second root r_2 will be positive. A solution to the homogeneous ODE

$$\frac{d^2\hat{\lambda}_{CA,t}}{dt^2} + \Psi_1 \frac{d\hat{\lambda}_{CA,t}}{dt} + \Psi_2 \hat{\lambda}_{CA,t} = 0 \tag{138}$$

is

$$\hat{\lambda}_{CA,t} = c_1 e^{r_1 t} + c_2 e^{r_2 t},$$

where c_1 and c_2 are constants. The boundary condition $\lim_{t\to\infty} \dot{\hat{\lambda}}_{CA,t} = 0$ implies $c_2 = 0$. A particular solution to the ODE can be obtained by variation of coefficients:

$$\hat{\lambda}_{CA,t} = \frac{1}{(r_2 - r_1)} \left(\int_t^\infty e^{-r_1(s-t)} F_s ds - \int_t^\infty e^{-r_2(s-t)} F_s ds \right),$$

A general solution to the equation is thus

$$\hat{\lambda}_{CA,t} = \frac{1}{(r_2 - r_1)} \left(\int_t^\infty e^{-r_1(s-t)} \left(\Psi_3 \hat{\tau}_s + \Psi_4 \hat{Y}_s \right) ds - \int_t^\infty e^{-r_2(s-t)} \left(\Psi_3 \hat{\tau}_s + \Psi_4 \hat{Y}_s \right) ds \right) + c_1 e^{r_1 t}, \tag{139}$$

where the constant c_1 is determined depending the initial conditions.

Assuming an auto-regressive process for the cost-push shock $\hat{\tau}_t = \tau^0 e^{-\tau^1 t}$,we can express $\hat{\lambda}_{CA,t}$ as

$$\begin{aligned}
\hat{\lambda}_{CA,t} &= \frac{\Psi_3 \tau^0 e^{-\tau^1 t}}{r_2 - r_1} \left(\frac{1}{r_1 + \tau^1} - \frac{1}{r_2 + \tau^1} \right) + c_1 e^{r_1 t} && (140) \\
&\quad + \frac{1}{(r_2 - r_1)} \left(\int_t^\infty \left[e^{-r_1(s-t)} - e^{-r_2(s-t)} \right] \Psi_4 \hat{Y}_s ds \right), \\
\dot{\hat{\lambda}}_{CA,t} &= -\frac{\Psi_3 \tau^0 \tau^1 e^{-\tau^1 t}}{r_2 - r_1} \left(\frac{1}{r_1 + \tau^1} - \frac{1}{r_2 + \tau^1} \right) + r_1 c_1 e^{r_1 t} && (141) \\
&\quad + \frac{1}{(r_2 - r_1)} \left(\int_t^\infty \left[r_1 e^{-r_1(s-t)} - r_2 e^{-r_2(s-t)} \right] \Psi_4 \hat{Y}_s ds \right), \\
\ddot{\hat{\lambda}}_{CA,t} &= \frac{\Psi_3 \tau^0 \left(\tau^1\right)^2 e^{-\tau^1 t}}{r_2 - r_1} \left(\frac{1}{r_1 + \tau^1} - \frac{1}{r_2 + \tau^1} \right) + r_1^2 c_1 e^{r_1 t} && (142)
\end{aligned}$$

$$+\frac{1}{(r_2-r_1)}\left(\int_t^{\infty}\left[r_1^2 e^{-r_1(s-t)}-r_2^2 e^{-r_2(s-t)}\right]\Psi_4\hat{Y}_s ds\right).$$

Using the initial condition $\hat{\lambda}_{CA,0}=0$, we compute the constant c_1,

$$c_1=-\frac{\Psi_3\tau^0}{(r_2-r_1)}\left(\frac{1}{r_1+\tau^1}-\frac{1}{r_2+\tau^1}\right)-\frac{1}{(r_2-r_1)}\left(\int_0^{\infty}\left[e^{-r_1 s}-e^{-r_2 s}\right]\Psi_4\hat{Y}_s ds\right). \tag{143}$$

Price level targeting In the limit $t\to\infty$, $\hat{\tau}_t=0$, because the shock is temporary. In the limit $t\to\infty$, since the 2nd order ODE (138) has no unit root ($r\neq 0$),[58] by (140) we also have that $\hat{\lambda}_{CA,t}=\frac{d\hat{\lambda}_{CA,t}}{dt}=\frac{d^2\hat{\lambda}_{CA,t}}{dt^2}=0$. Thus, by equation (134), in the limit $t\to\infty$ it must be that $\hat{P}_t=0$.

Weaker "*Leaning against the wind*" Consider equation (133). It must also hold at $t=0$. Since these variables are predetermined that $\hat{\lambda}_{CA,0}=\hat{P}_0=0$. Solving it for Y_0 it (133) reads

$$\hat{Y_0}=\frac{\Omega_3}{\Omega_{14}}\tau^0-\frac{\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}}{\Omega_{14}}\frac{d\hat{\lambda}_{CA,0}}{dt}.$$

Plugging in (141) and (143), we get

$$\begin{aligned}
\hat{Y_0} &= \frac{\Omega_3}{\Omega_{14}}\hat{\tau_0}-\frac{\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}}{\Omega_{14}}\times \\
&\quad \frac{1}{(r_2-r_1)}\left(\begin{array}{c} -\Psi_3\tau^0\tau^1\left(\frac{1}{r_1+\tau^1}-\frac{1}{r_2+\tau^1}\right)+\left(\int_0^{\infty}\left[r_1e^{-r_1(s)}-r_2e^{-r_2(s)}\right]\Psi_4\hat{Y}_s ds\right) \\ +r_1\left\{-\Psi_3\tau^0\left(\frac{1}{r_1+\tau^1}-\frac{1}{r_2+\tau^1}\right)-\left(\int_0^{\infty}\left[e^{-r_1 s}-e^{-r_2 s}\right]\Psi_4\hat{Y}_s ds\right)\right\}\end{array}\right) \\
\hat{Y_0} &= \tau^0\left\{\frac{\Omega_3}{\Omega_{14}}+\frac{\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}}{\Omega_{14}}\frac{\Psi_3}{(r_2-r_1)}\left(1-\frac{r_1+\tau^1}{r_2+\tau^1}\right)\right\} \\
&\quad +\frac{\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}}{\Omega_{14}}\left(\Psi_4\int_0^{\infty}\left[e^{-r_2 s}\right]\hat{Y}_s ds\right).
\end{aligned}$$

Now using (135) and (136):

[58]By (137) $\Psi_2=\Omega_5\frac{\Omega_{13}}{\Omega_{12}}\frac{\Omega_9}{\Omega_{10}}\left(1-\frac{\Omega_8}{\Omega_6}\right)=\underbrace{\Omega_5\frac{\Omega_{13}}{\Omega_{12}}\frac{\Omega_9}{\Omega_{10}}}_{>0}\left(\frac{1-\zeta\frac{Y}{C}}{1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}}\right)$, a unit root exists only if $1-\zeta\frac{Y}{C}=0$. However, in that case no Ramsey SS exists by 127.

$$\hat{Y}_0 = \underbrace{\frac{1}{\Omega_{14}}\left(\Omega_3 - \frac{\left(\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\frac{\Omega_{13}}{\Omega_{12}}\Omega_3\left(1-\frac{\Omega_8}{\Omega_6}\right)}{r_2+\tau^1}\right)}_{\Theta_1}\tau^0$$
$$-\underbrace{\left(\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\frac{\Omega_8}{\Omega_6}}_{\Theta_1}\int_0^\infty \left[e^{-r_2 s}\right]\hat{Y}_s ds.$$

Using the definitions of the Ωs, and assuming that $1 > \zeta\frac{Y}{C}$, we can evaluate the terms as follows

$$\begin{aligned}\Theta_1 &\equiv \frac{1}{\Omega_{14}}\left(\Omega_3 - \frac{\left(\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\frac{\Omega_{13}}{\Omega_{12}}\Omega_3\left(1-\frac{\Omega_8}{\Omega_6}\right)}{r_2+\tau^1}\right)\\ &= \frac{1}{\frac{Y}{C}\zeta\left(\frac{Y}{C}-1\right)}\left(\frac{\vartheta+1}{1-\alpha} - \frac{\frac{K}{A}\rho^h\left(\frac{1}{\xi-1}+\frac{A}{K}\right)\frac{\alpha}{\xi}\left(\frac{\vartheta+1}{1-\alpha}\right)^2 C^{-\zeta}Y\frac{1-\zeta\frac{Y}{C}}{\left(1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}\right)}}{r_2+\tau^1}\right) > 0,\end{aligned}$$

and

$$\begin{aligned}\Theta_2 &\equiv -\left(\Omega_4+\Omega_1-\Omega_2\frac{\Omega_9}{\Omega_{10}}\right)\frac{\Omega_8}{\Omega_6}\\ &= Y^\zeta K C^{-\zeta}\rho^h\left(\frac{1}{\xi-1}+\frac{A}{K}\right)\frac{\left(\frac{\vartheta+1}{1-\alpha}\right)^2}{\left(1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}\right)} < 0.\end{aligned}$$

Therefore,

$$\hat{Y}_0 = \underbrace{\Theta_1}_{+}\tau^0 + \underbrace{\Theta_2}_{-}\int_0^\infty \left[e^{-r_2 s}\right]\hat{Y}_s ds$$

where

$$r_2 = \frac{-\Psi_1+\sqrt{\Psi_1^2-4\Psi_2}}{2} > 0$$

where, in turn,

$$\Psi_2 = \left(\frac{\vartheta+1}{1-\alpha}\right)\frac{\alpha}{\xi}\rho^h\left(\frac{Y}{C}\right)^\zeta(1-\psi)\eta\left(\frac{1-\zeta\frac{Y}{C}}{1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}}\right) < 0,$$

$$\Psi_1 \;=\; \left[\left(\frac{Y}{C}\right)^{\zeta}\frac{\alpha}{\xi}\frac{-1+\zeta\frac{Y}{C}}{\left(1-\zeta\frac{Y}{C}-\frac{\vartheta+1}{1-\alpha}\right)}\frac{\vartheta+1}{1-\alpha}-1\right](1-\psi)\eta+\left(\left(\frac{Y}{C}\right)^{\zeta}-1\right)\rho^h<0,$$

This means that either $\hat{Y}_0 > 0$ or $\int_0^\infty \left[e^{-r_2 s}\right]\hat{Y}_s ds > 0$ or both. That is $\hat{Y}_t$ either jumps up on impact or is positive when integrated across time with discount rate $r_2 > 0$, which requires it to be positive at least for some t.

E.5.3 Comparison to the case with complete-markets economy (RANK)

For brevity we focus on the case with and undistorted steady state, that is when the appropriate subsidy is in place. The planner chooses $\hat{\pi}_t, \tilde{Y}_t, \hat{\chi}_{F,t}$ to maximize welfare subject to the Phillipscurve. The planner's Lagrangian is

$$\begin{aligned}
&\int_0^\infty e^{-\rho^h t}Y^{1-\zeta}\left[-\pi_t^2\frac{\theta}{2}-\frac{1}{2}\left\{\frac{\vartheta+1}{1-\alpha}-(1-\zeta)\right\}\tilde{Y}_t^2\right]\\
&+e^{-\rho^h t}\hat{\lambda}_{PC,t}\left[\rho^h\hat{\chi}_{F,t}-\frac{d\hat{\chi}_{F,t}}{dt}-\frac{1-\varepsilon}{\theta}Y^{1-\zeta}\left(\left[(1-\zeta)-\frac{\vartheta+1}{1-\alpha}\right]\tilde{Y}_t-\hat{\tau}_t\right)\right]\\
&+e^{-\rho^h t}\hat{\lambda}_{F,t}\left[Y^{1-\zeta}\hat{\pi}_t-\hat{\chi}_{F,t}\right]dt,
\end{aligned}$$

where the two constraints together are the Phillips curve. If we integrate by parts, we get

$$\begin{aligned}
&\int_0^\infty e^{-\rho^h t}Y^{1-\zeta}\left[-\pi_t^2\frac{\theta}{2}-\frac{1}{2}\left\{\frac{\vartheta+1}{1-\alpha}-(1-\zeta)\right\}\tilde{Y}_t^2\right]\\
&+e^{-\rho^h t}\hat{\lambda}_{PC,t}\left[-\frac{1-\varepsilon}{\theta}Y^{1-\zeta}\left(\left[(1-\zeta)-\frac{\vartheta+1}{1-\alpha}\right]\tilde{Y}_t-\hat{\tau}_t\right)\right]\\
&+e^{-\rho^h t}\hat{\chi}_{F,t}\frac{d\hat{\lambda}_{PC,t}}{dt}dt\\
&+e^{-\rho^h t}\hat{\lambda}_{F,t}\left[Y^{1-\zeta}\hat{\pi}_t-\hat{\chi}_{F,t}\right]dt,
\end{aligned}$$

The FOCs are

$$\begin{aligned}
-\theta\pi_t+\hat{\lambda}_{F,t} &= 0,\\
-\tilde{Y}_t+\hat{\lambda}_{PC,t}\left[\frac{1-\varepsilon}{\theta}\right] &= 0,\\
\frac{d\hat{\lambda}_{PC,t}}{dt}-\hat{\lambda}_{F,t} &= 0.
\end{aligned}$$

Simplifying, we get:

$$\theta\pi_t = \hat{\lambda}_{F,t} = \frac{d\hat{\lambda}_{PC,t}}{dt} = \frac{\theta}{1-\varepsilon}\frac{d\tilde{Y}_t}{dt},$$

or equivalently:

$$\pi_t = \frac{1}{1-\varepsilon}\frac{d\tilde{Y}_t}{dt}. \tag{144}$$

This target rule, by which the central bank accepts a reduction in output gap in exchange to contain inflation is known as leaning against the wind .

Now, using the Phillips curve (equation 131), we solve for $\tilde{Y}_t$

$$\tilde{Y}_t = \frac{\frac{\theta}{1-\varepsilon}\left(\rho^h\pi_t - \dot{\pi}_t\right) + \hat{\tau}_t}{(1-\zeta) - \frac{\vartheta+1}{1-\alpha}}.$$

Next take the time derivative of $\tilde{Y}_t$ and plug it into the target rule (144) to get:

$$\pi_t = \frac{1}{\varepsilon-1}\frac{\frac{\theta}{1-\varepsilon}\left(\ddot{\pi}_t - \rho^h\dot{\pi}_t\right) - \frac{d\hat{\tau}_t}{dt}}{(1-\zeta) - \frac{\vartheta+1}{1-\alpha}}.$$

Assume again that $\hat{\tau}_t = \tau^0\exp\left(-\tau^1 t\right)$, we can write this condition as

$$\ddot{\pi}_t + \Xi_1\dot{\pi}_t + \Xi_2\pi_t = \Xi_3\exp\left(-\tau^1 t\right),$$

where

$$\begin{aligned}
\Xi_1 &= -\rho^h < 0,\\
\Xi_2 &= -\frac{(1-\varepsilon)^2}{\theta}\left[(\zeta-1) + \frac{\vartheta+1}{1-\alpha}\right] < 0,\\
\Xi_3 &= -\frac{(1-\varepsilon)}{\theta}\tau^0\tau^1 > 0.
\end{aligned}$$

The characteristic polynomial

$$x = \frac{-\Xi_1 \pm \sqrt{\Xi_1^2 - 4\Xi_2}}{2},$$

which has two real roots $x_1 < 0 < x_2$. The homogeneous solution has a solution

$$\pi_t^h = p_1 e^{x_1 t} + p_2 e^{x_2 t},$$

where p_1 and p_2 are constants. The limit condition $\lim_{t\to\infty}\pi_t = 0$ implies that $p_2 = 0$. The solution is of the form

$$\begin{aligned}\hat{\pi}_t &= p_1 e^{x_1 t} + \frac{\Xi_3}{(x_2 - x_1)}\left[e^{x_1 t}\int_t^\infty e^{-\left(x_1+\tau^1\right)s}ds - e^{x_2 t}\int_t^\infty e^{-\left(x_s+\tau^1\right)s}ds\right] \\ &= p_1 e^{x_1 t} + \frac{\Xi_3}{(x_2 - x_1)}\left[\frac{1}{x_1+\tau^1} - \frac{1}{x_2+\tau^1}\right]e^{-\tau^1 t},\end{aligned}$$

or

$$\hat{\pi}_t = \tau^0 \Upsilon_3 e^{-\tau^1 t} + p_1 e^{x_1 t}, \tag{145}$$

where

$$\Upsilon_3 = \frac{\frac{(\varepsilon-1)}{\theta}\tau^1}{\left(x_2+\tau^1\right)\left(x_1+\tau^1\right)}.$$

To find p_1 we use the initial condition on π_t implied by the initial condition on the PC multiplier $\hat{\lambda}_{PC,t}$:

$$\begin{aligned}0 &= \hat{\lambda}_{PC,0}, \\ &\Rightarrow \\ 0 &= \left[\frac{1-\varepsilon}{\theta}\right]\hat{\lambda}_{PC,0} = \tilde{Y}_0 \\ &= -\frac{\frac{\theta}{1-\varepsilon}\left(\dot{\hat{\pi}}_0 - \rho^h \hat{\pi}_0\right) - \tau^0}{(1-\zeta) - \frac{\vartheta+1}{1-\alpha}}, \\ &= -\frac{\frac{\theta}{1-\varepsilon}\left(-\tau^1\tau^0\Upsilon_3 + x_1 p_1 - \rho^h\tau^0\Upsilon_3 - \rho^h p_1\right) - \tau^0}{(1-\zeta) - \frac{\vartheta+1}{1-\alpha}} \\ &\Rightarrow \\ p_1 &= \tau^0\frac{\frac{\theta}{1-\varepsilon}\left(\tau^1+\rho^h\right)\Upsilon_3 + 1}{\frac{\theta}{1-\varepsilon}\left(x_1-\rho^h\right)}\end{aligned}$$

where the first line is the initial condition for the PC multiplier $\hat{\lambda}_{PC,0}$, the second line uses the planners FOC for $\tilde{Y}_t$, the third line uses the PC to substitute for $\tilde{Y}_0$ and the fourth line plugs in the solution for pi 145, which we can then solve for p_1. Notice that the second line implies that $\tilde{Y}_t$ (and thus for cost-push shocks $\hat{Y}_t$) can't jump on impact.

So

$$\hat{\pi}_t = \tau^0 \left(\Upsilon_3 e^{-\tau^1 t} + \Upsilon_4 e^{x_1 t} \right), \tag{146}$$

where $\Upsilon_4 = p_1/\tau^0 = \frac{1-\frac{\left(\tau^1+\rho^h\right)\tau^1}{(x_2+\tau^1)(x_1+\tau^1)}}{\frac{\theta}{1-\varepsilon}\left(x_1-\rho^h\right)}$.

Price level targeting Integrating the target rule 144 we get

$$\int_0^t \pi_t dt = \frac{1}{1-\varepsilon}\int_0^t \frac{d\tilde{Y}_t}{dt} = \frac{1}{1-\varepsilon}\tilde{Y}_t,$$

where the second equality is because $\tilde{Y}_t$ can't jump on impact under OMP. Note that $\int_0^t \pi_t = \hat{P}_t$ is the price level. Since $\tilde{Y}_t$ (and thus $\hat{Y}_t$) has to return to its steady state of 0 for $t \to \infty$, and the price level must also return to 0. This is the well known price level targeting feature of the RANK economy under Ramsey policy.

The sign of the IRFs We can now determine the impact value of inflation as

$$\begin{aligned} \hat{\pi_0} &= (\Upsilon_3 + \Upsilon_4)\,\tau^0, \\ &= \underbrace{\frac{\varepsilon-1}{\theta(x_2+\tau^1)}\left[\frac{x_2}{-x_1+\rho^h}\right]}_{>0}\tau^0, \end{aligned}$$

That is if τ^0 is positive (a reduction in the subsidy) then inflation jumps up on impact. Since inflation is the sum of two exponentials, and since the price level has to return to its original value, inflation then gradually falls below zero before it converges to zero from below.

Since $\int_0^t \pi_t dt = \frac{1}{1-\varepsilon}\tilde{Y}_t$ as just discussed, this means that output gap, which is constant on impact, gradually drops below 0 before it returns to zero from below. It is thus always negative along the impulse response.

F Numerical Appendix

In this Appendix we present the new computational method used to (i) solve continuous-time heterogeneous-agent models in sequence space, and (ii) compute optimal policies within those models.

F.1 Overview

The central bank's Ramsey problem in the heterogeneous-firm economy is computationally more demanding than in a standard representative-agent framework, because the distribution of net worth, $\omega_t(z)$, is now both a state variable in the private equilibrium and a control variable for the planner.[59] Notice that the density $\omega_t(z)$ varies over time and across idiosyncratic productivity z. This poses a challenge when solving optimal monetary policy, since the first-order conditions (FOCs) must be taken with respect to an infinite-dimensional object.

There are a number of proposals in the literature to deal with this problem. Bhandari et al. (2021) make the continuous cross-sectional distribution finite-dimensional by assuming that there are N agents instead of a continuum. They then derive standard FOCs for the planner. In order to cope with the large dimensionality of their problem, they employ a perturbation technique. Le Grand et al. (2025) employ the finite-memory algorithm proposed by Ragot (2019). It requires changing the original problem such that, after K periods, the state of each agent is reset. This way the cross-sectional distribution becomes finite-dimensional. Nuño and Thomas (2022), Smirnov (2022) and Dávila and Schaab (2023) deal with the full infinite-dimensional planner's problem. This implies that the continuous Kolmogorov forward (KF) and the Hamilton-Jacobi-Bellman (HJB) equations are constraints faced by the central bank. They derive the planner's FOCs using calculus of variations, thus expanding the original problem to also include the Lagrange multipliers, which in this case are also infinite-dimensional. These papers solve the resulting differential equation system using the upwind finite-difference method of Achdou et al. (2021).

While we follow the latter approach in spirit – and we indeed use the analytical planner's FOCs to derive some analytical results – we propose a new algorithm that

[59] In our model, we do not need to keep track of the firm's value function since the firm's problem has a simple closed form solution. In a more complex heterogeneous agent models, one would additionally have to keep track also of the value function. Our method would accommodate that.

automates this procedure. Below we summarize the main steps of the algorithm, with the following subsections providing detailed explanations of each step.

1. First, we *discretize* the planner's problem with regards to the idiosyncratic state space using finite differences using a finite difference approach. This transforms the original infinite-dimensional-state continuous-time problem into a high-dimensional-state continuous-time problem, in which the value function and the state density are replaced by large vectors with a dimensionality equal to the number of grid points used to approximate the individual state space (200 in our application).

2. Next we use *symbolic algebra* to determine the continuous-time, discrete-state first-order conditions. Then we discretize time. This delivers a large-dimensional system of difference equations.

3. Now we find the Ramsey steady state by solving this system at steady state. To do so, we compute the *steady state* of the private economy conditional on the steady-state level of the policy instrument with a conventional iterative method, and then back out the associated multipliers.[60]

4. Finally, we solve for the nonlinear transitional dynamics under perfect foresight after unexpected "MIT" shocks. That is, we solve the non-linear system of difference equations resulting from step 2 in the *sequence space* using the Newton method, as already described in Section 3.

This algorithm can be employed to compute optimal policies in a large class of continuous-time models with either heterogeneous or representative agents. We provide a toolbox that allows to apply this algorithm in a largely automated manner.

F.2 Detailed general algorithm

1. Finite difference approximation of the idiosyncratic state Consider an arbitrary heterogeneous-agent model in continuous time continuous-time with a continuous idiosyncratic state space. Its private equilibrium conditions will typically contain

[60] In our application we know the value of the instrument in the Ramsey steady state ($\pi = 0$). In other applications this may not be the case, then finding the steady state boils down to solving a nonlinear equation (system) in x variables, where x is the number of instruments (1 in our case).

a Hamilton-Jacobi-Bellman (HJB) and a Kolmogorov forward (KF) equation. That is the equilibrium is defined by a set of partial differential equations.

We first **discretize with respect to the continuous state** the KF equation. The dynamics of the (now finite-dimensional) distribution $\boldsymbol{\mu}_t$ at period t are given by

$$\frac{\partial \boldsymbol{\mu}_t}{\partial t} = \boldsymbol{B}_t \boldsymbol{\mu}_t, \tag{147}$$

where $\boldsymbol{B}_\mathbf{t}$ is a matrix whose entries depend *nonlinearly* and in *closed form* on the idiosyncratic and aggregate variables in period t.[61] Similarly, the HJB equation is approximated as[62]

$$\rho \mathbf{v}_t = \boldsymbol{u}_t + \mathbf{A}_t \mathbf{v}_{t,} - \frac{\partial \mathbf{v}_t}{\partial t}. \tag{148}$$

Together with additional static equations, such as market clearing conditions or budget constraints, and aggregate dynamic equations, including the Euler equations of representative agents (if any) and the dynamics of aggregate states, these conditions characterize the private equilibrium in continuous time with a discretized state space.. We have transformed the equilibrium from a set of partial differential equations (PDEs) to a set of ordinary differential equations (ODEs).

2. Symbolic derivation of planner's FOCs Once we have a continuous-time discrete-space model, we can derive the planner's FOCs by *symbolic differentiation.* To do so we take the discrete-state-continuous-time equilibrium conditions from the previous step, define the planner's objective and then use symbolical math (in MATLAB) to derive the planner's FOCs. For this purpose we provide a toolbox, which basically follows the same steps as those followed in D.1, but in an automated fashion and without having to deal with the continuous state dimension. The result is a large set of FOCs in continuous time, which together with the private equilibrium conditions, form a large system of ODEs.

We then discretize time. We use the following convention: Forward-looking variables are approximated by forward differences and vice versa: $\frac{\partial \mathbf{v}_t}{\partial t} \simeq \frac{\mathbf{v}_{t+1} - \mathbf{v}_t}{\Delta t}$ and $\frac{\partial \boldsymbol{\mu}_t}{\partial t} \simeq \frac{\boldsymbol{\mu}_t - \boldsymbol{\mu}_{t-1}}{\Delta t}$. In other words, the KF equation is approximated implicitly, and the HJB

[61] Technically, this matrix results from the discretization of the *infinitesimal generator* of the idiosyncratic states. In the example of Section 2, $\boldsymbol{\mu}_t = \boldsymbol{\omega}_\mathbf{t}$.

[62] In the model presented in this paper the HJB can be solved analytically and hence there is no need to solve it computationally.

equation explicitly. We end up with a large set of difference equations.

A natural question at this stage is whether the result of our algorithm differs from that of the approach of Nuño and Thomas (2022), Bigio and Sannikov (2021), Smirnov (2022) and Dávila and Schaab (2023), who derive the planners FOCs from the original continuous-time-continuous-state Ramsey problem and the discretize time and state. The following Proposition shows that two solutions coincide in models where exogenous states feature reflective boundaries and where the endogenous state is unconstrained. While many models in the heterogeneous agent literature do feature constraints also on endogenous variables (e.g. borrowing constraints), those can often be approximated by a smooth cost function, in which case the below proof applies again

Proposition E.1 : *The system of difference equations resulting from the our algorithm and the "first optimize then discretize" algorithm used in* Nuño and Thomas (2022), Smirnov (2022) and Dávila and Schaab (2023) *coincide in models where exogenous states feature reflective boundaries and where the endogenous state is unconstrained.*

Proof: See Appendix F.7.

The Proposition shows that our approach adds no additional approximation error relative to the manual method. But it allows to automate the finding of first-order conditions with standard symbolic math.

Though we have ended up with a discrete-time approximation, casting the original model in continuous time is central to our method. The discretized dynamics of the distribution and Bellman equation present two advantages compared to their counterparts in the discrete-time continuous-state formulation typically employed in the literature. First, the analytical tractability of the original continuous-time model implies that the agents' optimal choices in the discretized version are always "on the grid", avoiding the need for interpolation, and are "one step at a time" making the matrix $\boldsymbol{B}_{\mathbf{t}}$ sparse.[63] Second, the private agent's FOCs hold with equality even at the exogenous boundaries (see Achdou et al. (2021) for a detailed discussion of these advantages).

3. Solving for the Ramsey steady state Before we can solve for the dynamics we need to know the **Ramsey steady state**, both to determine initial and terminal conditions and to provide an initial guess for the transition paths of the dynamic problem. In our setup, we determine the value of the policy instrument (inflation) in the

[63]The introduction of Poisson shocks would not change the sparsity of matrix $\boldsymbol{B}_{\mathbf{t}}$.

Ramsey steady state analytically. Given this inflation rate, we use a standard iterative procedure (explained in Section F.6) to solve the (steady state) private equilibrium conditions for the steady state values of the prices and quantities. Given those values, we solve the (steady state) planner's first-order conditions for the steady state Lagrange multipliers. This step is trivial, since this system is linear in the multipliers. Our toolbox automatically produces a function that delivers those values. If the steady state value of the policy instrument were unknown, it would be easy to find it by solving a uni-variate nonlinear equation, using the automatically generated function.

4. Newton algorithm to solve the optimal policy problem non-linearly in the sequence space Finally, we use the discretized optimality conditions of the planner to compute non-linearly the *optimal responses* to a temporary, unexpected change in parameters (an "MIT shock") under perfect foresight. Finding the optimal policy in the **sequence space** (and not in the state space) helps to overcome the curse of dimensionality since in the sequence space the complexity of the problem grows only linearly in the number of aggregate variables, whereas the complexity of the state-space solution grows exponentially in the number of state variables. This approach has been made popular in discrete-time models by Juillard et al. (1998) (Dynare with perfect foresight) and more recently been extended to heterogeneous agent economies by Auclert et al. (2021) (Sequence Space Jacobian toolbox in its nonlinear version). We build on these contributions when we compute the optimal transition path. The solution to the perfect foresight problem can be easily reinterpreted for the case with aggregate shocks. As Boppart et al. (2018) show, the perfect-foresight transitional dynamics to an "MIT shock" coincides with the solution of the model with aggregate uncertainty using a first-order perturbation approach.

To find the paths for all our equilibrium variables, including our discretized distributions and value functions, we use a **global relaxation algorithm**. This approach has traditionally been applied to discrete-time representative agent models, especially since it became easily available through Dynare's perfect foresight solver (Juillard et al. (1998)). It is somewhat less common in continuous-time models (e.g. Trimborn et al., 2008). We apply this method to our continuous time heterogeneous agent economy under Ramsey policy.

For convenience, and in order to reach a wider audience, we use Dynare to perform this step. Specifically, our tool automatically translates the system of difference equations obtained in step 2 into Dynare syntax. We then employ Dynare's nonlinear Newton solver to compute the optimal transition path under perfect foresight. Although this is the most computationally intensive part of the algorithm, it typically takes only 1–2 minutes to run for our model on a standard laptop.

Our hope is that by relying on the popular and time-tested Dynare toolbox where possible, and by providing a new, user-friendly extension that broadens its scope to optimal policy problems in continuous-time settings with heterogeneous agents, we make the analysis of optimal policy in such models more accessible to a wider community of researchers.

F.3 The continuous time Ramsey tool for Dynare

Our tool allows solving Ramsey problems in a large class of continuous-time heterogeneous agent models.[64] The toolbox requires Dynare and Matlab. It is available from the authors' websites, and in the GitHub repository https://github.com/beagzl/OMPH.git. As inputs, the user has to provide the following:

- The dynamic private equilibrium conditions and the planner's objective, discretized with respect to the state space according to a valid discretization scheme. Equations are entered very much as they would typically appear in a paper, with `dot_X` denoting the time-derivative of a variable `X`, sorted by their type (backward-looking, forward-looking, static). Parameters and variables, which can include vectors for discretized idiosyncratic variables / parameters, need to be declared as such. Variables need to be declared by type (backward-looking, forward-looking, static). Dynamic auxiliary variables, which are defined purely for notational convenience (such as a real interest rate in our case), must not be used, as this would create incorrect initial commitments under the timeless Ramsey plan.

- A file that computes the steady state of the private equilibrium, conditional on (a guess for) the steady state value for the policy instrument and parameter values.

- A set of parameter values, including the models parameters, the parameters of the discretization scheme and the aggregate shock sequence.

[64] The toolbox could also be useful for continuous-time models with exogenous policy or without heterogeneity.

The user has to manipulate two files: the file “buildfile.m”, where he has to declare the model and the parameters and variables; and the file “HANK.mod” and “RANK.mod”for the baseline model and the complete-markets model, respectively, where the steady state of the private equilibrium and the particular simulation need to be defined.

F.4 Comparison to other solution methods

We are aware of four methods to solve Ramsey problems in general heterogeneous agents models. We have already mentioned them above, and we compare them in more detail in this table.[65] The purpose is not to argue that one approach dominates the others, but to show similarities and differences. An advantage of our method is that it allows for an arbitrary path of the distribution while allowing for occasionally binding constraints and being highly automatized.

The approach most similar to ours is that of Nuño and Thomas (2022), Smirnov (2022) and Dávila and Schaab (2023). Relative to them, we deviate in two important ways. On the one hand, we *first* discretize the planner's problem with respect to the continuous state and *then* find the FOCs, as opposed to *first* deriving the FOCs by hand using calculus of variations and *then* discretizing them. Above we presented a Proposition showing that the two approaches are equivalent. On the other hand, we automate much of the process. All the user has to do is (i) to provide the dynamic continuous-state continuous-time private equilibrium conditions and to discretize them with respect to the continuous state variable (step 1), and (ii) to find the steady state of the private equilibrium given (a guess for) the steady state instrument (part of step 3). Deriving the planners FOCs, computing the Ramsey steady state, and computing transitional dynamics is all automated. We provide a toolbox that automates the second step and translates the resulting equation system into dynare, which we then use to solve for the Ramsey steady state and the transitional dynamics.

When it comes to solving the nonlinear perfect foresight dynamics (step 4), we use a global relaxation algorithm. This approach has been made popular in discrete-time representative agent models by Juillard et al. (1998) (dynare), but it is somewhat less common in continuous-time models (e.g. Trimborn et al., 2008). We build on these contributions when we compute the optimal transition path. In doing so we deviate from the two loop approach traditionally used in the heterogeneous agent literature,

[65] Note that grouping 4 papers as one conceptual method glosses over some differences.

which solves the heterogeneous agents block conditional on a guess for aggregates and then updates the guess for the aggregates. Recently Auclert et al. (2021) have proposed a particularly efficient variant of this approach, the Sequence Space Jacobian toolbox. In principle, our approach would also be compatible with the non-linear sequence space Jacobian toolbox. Indeed using the SSJ toolbox (in its non-linear mode) in step 4 instead of dynare would deliver the exact same solution. We leave linking our toolbox with the SSJ toolbox for future work.

Our approach to solving for the equilibrium dynamics is different from the one in Winberry (2018) or Ahn et al. (2018). These papers expand the seminal contribution by Reiter (2009), based on a two-stage algorithm that (i) first finds the nonlinear solution of the steady state of the model and (ii) then applies perturbation techniques to produce a linear system of equations describing the dynamics around the steady state. Winberry (2018) illustrates how this can be also implemented using Dynare, and Ahn et al. (2018) extend the methodology to continuous-time problems. The key difference is that these methods are linear in aggregates, while ours is non-linear. That said, in principle one can easily modify our algorithm to employ a linear approach in step 4.

None of the methods discussed in the last two paragraphs, however, are meant for finding Ramsey optimal policies, which is the focus and main contribution of our algorithm.

	(1)	(2)	(3)	(4)
	Nuño and Thomas (2022), Smirnov (2022) and Dávila and Schaab (2023)	Le Grand et al. (2025)	Bhandari et al. (2021)	This paper
time	continuous	discrete	discrete	continuous
how to cope with infinite dimensional distribution in planners problem	infinite-dimensional calculus	discrete number of truncated histories N	discretize distribution by considering large number N of individuals	discretize HJB, KFE and other equations containing the distribution with N finite elements
planner optimizes	continuous-state continuous time problem	discrete-state discrete-time problem	discrete-state discrete-time problem	discrete-state continuous-time problem
how to find planner's FOCs	by hand	by hand	custom-built code	automated
sequence or state space method	sequence space	sequence space	state space	sequence space
representation of distribution as	finite difference approximation of distribution function with N elements	population of N histories	population of N agents	finite difference approximation of distribution function with N elements
how to tame curse of dimensionality	perfect foresight	perfect foresight	local approximation	perfect foresight
solving dynamic equilibrium conditions under optimal policy	custom build iterative procedure where the heterogeneous agents block is solved conditional on prices and the prices are solved for conditional on the heterogeneous agents block until convergence	Newton method, automated in Dynare	custom-built code	Newton method, automated in Dynare
occasionally binding constraints and non-linearities	possible	possible	not possible	possible

F.5 Finite difference approximation of the Kolmogorov Forward equation

The KF equation is discretized with respect to the idiosyncratic state space by a finite difference scheme following Achdou et al. (2021). It approximates the density $\omega_t(z)$ on a finite grid $z \in \{z^1, ..., z^J\}$with steps Δz. We use the notation $\omega_t^j := \omega_t(z^j)$, $j = 1, ..., J$. The KF equation is then approximated as

$$\begin{aligned}\frac{\partial \omega_t^j}{\partial t} = &\left(s_t(z^j) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta\right)\omega_t^j \\ &- \frac{\omega_t^j \mu(z^j) - \omega_t^{j-1}\mu(z^{j-1})}{\Delta z} + \frac{\omega_t^{j+1}\widetilde{\sigma}^2(z^{j+1}) + \omega_t^{j-1}\widetilde{\sigma}^2(z^{j-1}) - 2\omega_t^j\widetilde{\sigma}^2(z^j)}{2\left(\Delta z\right)^2},\end{aligned}$$

which, grouping, results in

$$\begin{aligned}\frac{\partial \omega_t^j}{\partial t} = &\underbrace{\left[\left(s_t(z^j) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta\right) - \frac{\mu(z^j)}{\Delta z} - \frac{\widetilde{\sigma}^2(z^j)}{\left(\Delta z\right)^2}\right]}_{\beta_j^n}\omega_t^j \\ &+ \underbrace{\left[\frac{\mu(z^{j-1})}{\Delta z} + \frac{\widetilde{\sigma}^2(z^{j-1})}{2\left(\Delta z\right)^2}\right]}_{\varrho_t^{j-1}}\omega_t^{j-1} + \underbrace{\left[\frac{\widetilde{\sigma}^2(z^{j+1})}{2\left(\Delta z\right)^2}\right]}_{\chi_t^{j+1}}\omega_t^{j+1}.\end{aligned}$$

The boundary conditions are the ones associated with a reflected process z at the boundaries:[66]

$$\begin{aligned}\frac{\partial \omega_t^1}{\partial t} &= \left(\beta_t^1 + \chi_t^1\right)\omega_t^1 + \chi_t^2\omega_t^{j+1}, \\ \frac{\partial \omega_t^J}{\partial t} &= \left(\beta_t^J + \varrho_t^J\right)\omega_t^J + \varrho_t^{J-1}\omega_t^{J-1}.\end{aligned}$$

[66]It is easy to check that this formulation preserves the fact that matrix $\mathbf{B}^n$ below is the transpose of the matrix associated with the infinitesimal generator of the process.

If we define matrix

$$\mathbf{B}_t = \begin{bmatrix} \beta_t^1 + \chi_t^1 & \chi_t^2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \varrho_t^1 & \beta_t^2 & \chi_t^3 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \varrho_t^2 & \beta_t^3 & \chi_t^4 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \varrho_t^{J-2} & \beta_t^{J-1} & \chi_t^J \\ 0 & 0 & 0 & 0 & \cdots & 0 & \varrho_t^{J-1} & \beta_t^J + \varrho_t^J \end{bmatrix},$$

then we can express the KF equation as

$$\frac{\partial \boldsymbol{\omega}_t}{\partial t} = \mathbf{B}_t \boldsymbol{\omega}_t, \tag{149}$$

where $\boldsymbol{\omega}^t = \begin{bmatrix} \omega_t^2 & \omega_t^2 & ... & \omega_t^{J-1} & \omega_t^J \end{bmatrix}^T$.

F.5.1 Extension to non-homogeneous grids

Our model can be solved using a homogeneous grid. However, we use a non-homogeneous grid for the state z to economize on grid points. This is useful for two reasons: First, it allows us to concentrate grid points around z_t^*, which is convenient since z_t^* does not live on the grid, which introduces additional approximation error. Second, numerical error may pile up at the lower end of the grid. We could not find a universally applicable way to implement non-homogeneous grids in the economics literature, so we propose the following discretization scheme.[67]

Be $z = \left[z^1, ..., z^J\right]$ the grid. Define $\Delta z^{a,b} = z^b - z^a$ and let

$$\Delta z = \frac{1}{2} \begin{bmatrix} \Delta z^{1,2}, & \Delta z^{1,3}, & \Delta z^{2,4}, & ..., & \Delta z^{J-2,J} & \Delta z^{J-1,J} \end{bmatrix}.$$

We approximate the KFE (24) using central difference for both the first derivative and the second derivative.

$$\frac{\partial \omega_t^j}{\partial t} = \left(s_t(z^j) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta \right) \omega_t^j - \left[\frac{\mu(z^{j+1})\omega_t^{j+1} - \mu(z^{j-1})\omega_t^{j-1}}{\Delta z^{j-1,j+1}} \right]$$

[67] Our approach builds on the one in the Appendix of Achdou et al., 2021. It differs from theirs in two ways. First, it can be derived as a finite difference scheme over the KFE. Second, it relies on central differences for the first order derivative, and hence it is not an upwind scheme.

$$+ \quad \frac{1}{2}\frac{\Delta z^{j-1,j}\sigma^2(z^{j+1})\omega_t^{j+1} + \Delta z^{j,j+1}\sigma^2(z^{j-1})\omega_t^{j-1} - \Delta z^{j-1,j+1}\sigma^2(z_j)\omega^n(z_j)}{\frac{1}{2}\left(\Delta z^{j-1,j+1}\right)\Delta z^{j,j+1}\Delta z^{j-1,j}},$$

which, grouping, results in

$$\begin{aligned}
\frac{\partial \omega_t^j}{\partial t} \quad = \quad & \underbrace{\left[\left(s_t(z^j) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta\right) - \frac{\sigma^2(z^j)}{\Delta z^{j,j+1}\Delta z^{j-1,j}}\right]}_{\beta_{jt}^j}\omega_t^j \\
& + \underbrace{\left[\frac{\mu(z^{j-1})}{\Delta z^{j-1,j+1}} + \frac{\sigma^2(z^{j-1})}{\left(\Delta z^{j-1,j+1}\right)\Delta z^{j,j-1}}\right]}_{\varrho_t^{j-1}}\omega_t^{j-1} \\
& + \underbrace{\left[-\frac{\mu(z^{j+1})}{\Delta z^{j-1,j+1}} + \frac{\sigma^2(z^{j+1})}{\left(\Delta z^{j-1,j+1}\right)\Delta z^{j,j+1}}\right]}_{\chi_t^{j+1}}\omega_t^{j+1}.
\end{aligned}$$

The boundary conditions are the ones associated with a reflected process z at the boundaries:

$$\begin{aligned}
\frac{\partial \omega_t^1}{\partial t} &= \left(\beta_t^1 + \chi_t^1\right)\omega_t^1 + \chi_t^2\omega_t^2, \\
\frac{\partial \omega_t^J}{\partial t} &= \left(\beta_t^J + \varrho_t^J\right)\omega_n^J + \varrho_t^{J-1}\omega_t^{J-1}.
\end{aligned}$$

where we define $\Delta z^{0,1} \equiv \Delta z^{11,2}$ and $\Delta z^{J,J+1} \equiv \Delta z^{J-1,J}$.

The law of motion of ω can equivalently be written in matrix form

$$\frac{\partial \boldsymbol{\omega}_t}{\partial t} = \mathbf{B}_t\boldsymbol{\omega}_t,$$

where

$$\mathbf{B}_t = \begin{bmatrix}
\beta_t^1 + \chi_t^1 & \chi_t^2 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\varrho_t^1 & \beta_t^2 & \chi_t^3 & 0 & \cdots & 0 & 0 & 0 \\
0 & \varrho_t^2 & \beta_t^3 & \chi_t^4 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & \varrho_t^{J-2} & \beta_t^{J-1} & \chi_t^J \\
0 & 0 & 0 & 0 & \cdots & 0 & \varrho_t^{J-1} & \beta_t^J + \varrho_t^J
\end{bmatrix},$$

Abstracting for brevity from the term $\left(s_t(z^j) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta\right)$, which is independent of the grid, and spelling out $\mathbf{B}_t$ we have

$$\frac{\partial \boldsymbol{\omega}_t}{\partial t} = \begin{bmatrix} \frac{\sigma^2(z_1)}{\Delta z_{1,2}\Delta z_{1,2}} - \frac{\mu(z_1)}{\Delta z_{1,2}} - \frac{2\sigma^2(z_1)}{\Delta z_{1,2}\Delta z_{1,2}} & -\frac{\mu(z_2)}{\Delta z_{1,2}} + \frac{\sigma^2(z_2)}{\Delta z_{1,2}\Delta z_{1,2}} & 0 & \cdots \\ \frac{\mu(z_1)}{\Delta z_{1,3}} + \frac{\sigma^2(z_1)}{\Delta z_{1,3}\Delta z_{1,2}} & -\frac{\sigma^2(z_2)}{\Delta z_{1,2}\Delta z_{2,3}} & -\frac{\mu(z_3)}{\Delta z_{1,3}} + \frac{\sigma^2(z_3)}{\Delta z_{1,3}\Delta z_{2,3}} & \cdots \\ 0 & \frac{\mu(z_2)}{\Delta z_{2,4}} + \frac{\sigma^2(z_2)}{\Delta z_{2,4}\Delta z_{2,3}} & -\frac{\sigma^2(z_3)}{\Delta z_{2,3}\Delta z_{3,4}} & \cdots \\ 0 & 0 & \frac{\mu(z_3)}{\Delta z_{3,5}} + \frac{\sigma^2(z_3)}{\Delta z_{3,4}\Delta z_{3,5}} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \boldsymbol{\omega}_t.$$

We can rewrite this as follows

$$\frac{\partial \boldsymbol{\omega}_t}{\partial t} = \begin{bmatrix} -\frac{\mu(z_1)}{\boldsymbol{\Delta z_{1,2}}} - \frac{\sigma^2(z_1)}{\boldsymbol{\Delta z_{1,2}}\Delta z_{1,2}} & -\frac{\mu(z_2)}{\boldsymbol{\Delta z_{1,2}}} + \frac{\Delta z_{2,3}\sigma^2(z_2)}{\Delta z_{2,3}(\boldsymbol{\Delta z_{1,2}}\Delta z_{1,2})} & 0 & \cdots \\ \frac{\mu(z_1)}{\boldsymbol{\Delta z_{1,3}}} + \frac{\sigma^2(z_1)}{\boldsymbol{\Delta z_{1,3}}\Delta z_{1,2}} & -\frac{(\Delta z_{1,2}+\Delta z_{2,3})\sigma^2(z_2)}{\boldsymbol{\Delta z_{1,3}}(\Delta z_{1,2}\Delta z_{2,3})} & -\frac{\mu(z_3)}{\boldsymbol{\Delta z_{1,3}}} + \frac{\Delta z_{3,4}\sigma^2(z_3)}{\Delta z_{3,4}(\boldsymbol{\Delta z_{1,3}}\Delta z_{2,3})} & \cdots \\ 0 & \frac{\mu(z_2)}{\boldsymbol{\Delta z_{2,4}}} + \frac{\Delta z_{1,2}\sigma^2(z_2)}{\Delta z_{1,2}(\boldsymbol{\Delta z_{2,4}}\Delta z_{2,3})} & -\frac{(\Delta z_{2,3}+\Delta z_{3,4})\sigma^2(z_3)}{\boldsymbol{\Delta z_{2,4}}(\Delta z_{2,3}\Delta z_{3,4})} & \cdots \\ 0 & 0 & \frac{\mu(z_3)}{\boldsymbol{\Delta z_{3,5}}} + \frac{\Delta z_{2,3}\sigma^2(z_3)}{\Delta z_{2,3}(\Delta z_{3,4}\boldsymbol{\Delta z_{3,5}})} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \boldsymbol{\omega}_t.$$

Note that the bold terms in row i are equal to $1/\Delta z_i$, where Δz_i is the i-th element of Δz. Furthermore note that, up to the bold terms, the columns sum up to 0. Thus $\Delta z \mathbf{B}^n = 0$. Therefore

$$\Delta z \frac{\partial \boldsymbol{\omega}_t}{\partial t} = \frac{\partial \Delta z \boldsymbol{\omega}_t}{\partial t} = \Delta z \mathbf{B}_t \boldsymbol{\omega}_t = 0.$$

Thus our central difference approximation scheme on an non-homogenous grid is mass preserving. That is, given that we start with a properly defined distribution of mass one $\Delta z \boldsymbol{\omega}_0 = 1$ we have $\Delta z \boldsymbol{\omega}_t = 1 \forall t > 0$, where $\Delta z \boldsymbol{\omega}_t = \sum_j \omega_t^j \Delta z^j$ is the trapezoid-rule approximation of the integral $\int \omega_t(z) dz$.

F.5.2 Finite difference approximation of the integrals

To approximate the integrals in $\int_0^z \omega_t(z)\, dz$ and $\int_{z_t^*}^\infty z\omega_t(z) dz$ we use the trapezoid rule. I.e. if $f(z)$ is either $\omega_t(z)$ or $z\omega_t(z)$ and $z^j \leq \bar{z} \leq z^{j+1}$ then the integral from the closest gridpoint below $\bar{z}$ to $\bar{z}$ is given by

$$\int_{z^j}^{\bar{z}} f(z)\, dz = \left[f\left(z^j\right) + \frac{1}{2}\left[f\left(z^{j+1}\right) - f\left(z^j\right) \right] \frac{\bar{z} - z^j}{z^{j+1} - z^j} \right] \left(\bar{z} - z^j\right)$$

We use this formula to construct the integrals over a larger range piecewise. For example:

$$\int_{z^1}^{z^N} f(z)\,dz = \Delta z(z^1, z^N) \begin{bmatrix} f(z^1) \\ f(z^2) \\ \vdots \\ f(z^N) \end{bmatrix}$$

and

$$\begin{aligned} \int_{z^1}^{z^*} f(z)\,dz &= \Delta z(z^1, z^{j^*-1}) \begin{bmatrix} f(z^1) \\ f(z^2) \\ \vdots \\ f(z^{j^*-1}) \end{bmatrix} \\ &+ \left[f\left(z^{j^*-1}\right) + \frac{1}{2}\left[f\left(z^{j^*}\right) - f\left(z^{j^*-1}\right)\right] \frac{z^* - z^{j^*-1}}{z^{j^*} - z^{j^*-1}} \right] (z^* - z^{j^*-1}) \\ & \text{where } j^* = \arg\min_j \left\{ j \leq J | z^{j^*} > z^* \right\}, \end{aligned}$$

where $\Delta z(z^1, z^J) = \frac{1}{2} \begin{bmatrix} \Delta z^{1,2}, & \Delta z^{1,3}, & \Delta z^{2,4}, & ..., & \Delta z^{J-2,J} & \Delta z^{J-1,J} \end{bmatrix}$ is defined as in the previous Section.

F.6 Algorithm to solve for the steady state of the private equilibrium

Here we present how to solve for the steady state of the private equilibrium, that is for the steady state when the central bank sets a certain level of the nominal interest rate in SS i. We normalize the price of capital $q = 1$ and a labor to $\bar{L} = 1$. [68]In SS consumption does not grow, hence from (14)

$$r = \rho^h. \tag{150}$$

Combining (150) with the Fisher equation and the fact that the planner sets a certain nominal rate i we get that

$$\pi = i - \rho^h. \tag{151}$$

[68]We calibrate the capital stock K and the constant Υ to ensure that the conditions $q = 1$ and $L = \bar{L} = 1$ hold respectively.

In steady state, $\dot{\pi}_t = 0$ and $\dot{Y}_t = 0$. Hence, from equation (21) we obtain

$$m = \left(m^* + \rho^h \pi \frac{\theta}{\varepsilon}\right). \tag{152}$$

Using equation (32) and (150),

$$\rho^h = \frac{1}{q}\left(\alpha m Z A^{\alpha-1} L^{1-\alpha} \frac{z^*}{\gamma X}\right) - \delta \tag{153}$$

From equation (33) and (150),

$$\frac{\dot{A}_t}{A_t} = 0 = \left[\gamma(1-\Omega(z^*))\left(\alpha m Z K^{\alpha-1} L^{1-\alpha} - R\right) + R - \delta q - q(1-\psi)\eta)\right], \tag{154}$$

Plugging the latter equation into the former, we obtain:

$$\rho^h + \delta = \left[(\rho^h + \delta)\left(\gamma(1-\Omega(z^*)) - 1\right) + (1-\psi)\eta + \delta\right] \frac{z^*}{\gamma X}. \tag{155}$$

In the algorithm, we use a non-linear equation solver to obtain z^* from this equation. *The Algorithm.*

- Get $r = \rho^h$, $\pi = \bar{\pi}$ and $i = \rho^h + \pi$ and $R = q(\rho^h + \delta)$ and $m = m^* + \rho^h \pi \frac{\theta}{\epsilon}$.
- Set $L = \bar{L}$ and $q = 1$.
- Let n now denote the iteration counter. Make an initial guess for the net worth distribution $\boldsymbol{\omega_0}$.
  1. Use a non-linear equation solver on equation (155) to obtain z_n^* from equation (155) and obtain $X_n \equiv \int_{z_n^*}^{\infty} z\omega_n(z)dz$.
  2. Get $Z_n = (\gamma X_n)^{\alpha}$.
  3. Find A_n from equation (31),

$$A_n = \left[\frac{q\rho^h + \delta q}{\alpha m Z_n L^{1-\alpha} \frac{z_n^*}{\gamma X_n}}\right]^{\frac{1}{\alpha-1}}.$$

  4. Find the stocks $K_n = \gamma(1-\Omega_n(z_n^*))A_n$, $D_n = K_n - A_n$.

5. Compute $w_n = (1-\alpha)mZ_nA_n{}^{\alpha}L^{-\alpha}$,$\varphi_n = \alpha\left(\frac{(1-\alpha)}{w_n}\right)^{(1-\alpha)/\alpha} m^{\frac{1}{\alpha}}$.
6. Get aggregate output $Y = Z_nA_n^{\alpha}L^{1-\alpha}$, transfers $T_n = (1-m)\,Y_n - \frac{\theta}{2}\,(\pi)^2\,Y_n + q(1-\psi)\eta A_n - K_n\delta(1-q)$, and consumption $C_n = w_nL + (R_n - \delta q)D_n + T_n$.
7. Update $\hat{s}_j^n = \frac{1}{q}(\gamma \max\{z\varphi_n - R, 0\} + R - \delta q)$ and employ it to construct matrix $\mathbf{B}^n$.
8. Update $\boldsymbol{\omega}^{n+1}$ using equation $\frac{\boldsymbol{\omega}^{n+1}-\boldsymbol{\omega}^n}{\Delta t} = \mathbf{B}^n\boldsymbol{\omega}^{n+1}$.
9. Check for convergence. If the net worth distribution does not coincide with the guess, set $n = n+1$ and return to point 1.

- Set $\Upsilon = \frac{wC^{-\varsigma}}{L^{\vartheta}}$ to ensure our guess for L is correct.

F.7 Proof of Proposition F.2

The proof has the following structure. First, we set up a generic planner's problem in a continuous-time heterogeneous-agent economy without aggregate uncertainty. This problem features reflective boundaries on exogenous states. The endogenous states are unconstrained. Second, we derive the continuous-continuous-time-continuous-state optimality conditions of the planner's problem and discretize them. Third, we discretize the planners problem with respect to the idiosyncratic state space and then derive the optimality conditions and then discretize them with respect to time. Fourth, we compare the two sets of discretized optimality conditions.

1. The generic planner's problem The planner's problem in an economy with heterogeneity among one agent type (e.g. households or firms) can be written as[69]

$$\max_{Z_t,u_t(x),\mu_t(x),v_t(x)} \int_0^{\infty} \exp(-\varrho t) f_0(Z_t)dt \tag{156}$$

$$s.t.\ \forall t$$

$$\dot{X}_t = f_1(Z_t) \tag{157}$$

$$\dot{U}_t = f_2(Z_t) \tag{158}$$

$$0 = \lambda_{5,t}(x)v_t(x)f_3(Z_t) \tag{159}$$

[69]We abuse notation and employ the dot both as a partial derivative with respect to time,$\dot{v}_t(x) \equiv \frac{\partial v_t(x)}{\partial t}$,and as derivative with respect to time $\dot{X}_t \equiv \frac{dX_t}{dt}$.

$$\tilde{U}_t = \int f_4\left(x, u_t(x), Z_t\right) \mu_t\left(x\right) \; dx \tag{160}$$

$$\rho v_t(x) = \dot{v}_t(x) + f_5(x, u_t(x), Z_t) \tag{161}$$
$$+\sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right) \frac{\partial v_t(x)}{\partial x_i} + \sum_{i=1}^{I}\sum_{k=1}^{I} \frac{\left(\sigma(x)\sigma(x)^\top\right)_{i,k}}{2} \frac{\partial^2 v_t(x)}{\partial x_i \partial x_k}, \; \forall x$$

$$0 = \frac{\partial f_5}{\partial u_{j,t}} + \sum_{i=1}^{I} \frac{\partial b_i}{\partial u_{j,t}} \frac{\partial v_t(x)}{\partial x_i}, \quad j = 1, ..., J, \forall x. \tag{162}$$

$$\dot{\mu}_t\left(x\right) = -\sum_{i=1}^{I} \frac{\partial}{\partial x_i}\left[b_i\left(x, u_t(x), Z_t\right) \mu_t\left(x\right)\right] \tag{163}$$
$$+\frac{1}{2}\sum_{i=1}^{I}\sum_{k=1}^{I} \frac{\partial^2}{\partial x_i \partial x_k}\left[\left(\sigma(x)\sigma(x)^\top\right)_{i,k} \mu_t\left(x\right)\right], \; \forall x$$

$$X_0 = \bar{X}_0, \tag{164}$$

$$\mu_0\left(x\right) = \bar{\mu}_0\left(x\right), \tag{165}$$

$$0 = \lim_{t\to\infty} \exp(-\varrho t) U_t, \tag{166}$$

$$0 = \lim_{t\to\infty} \exp(-\varrho t) v_t(x), \tag{167}$$

$$0 = \left(b\left(x, u_t(x), Z_t\right) \mu_t\left(x\right) - \frac{1}{2}\frac{\partial}{\partial x}\left[\sigma^2(x)\mu_t\left(x\right)\right]\right)_{|x_{x,i}=\bar{x}_{x,i}}, \tag{168}$$

$$0 = \frac{\partial v\left(x\right)}{\partial x}_{|x_{x,i}=\bar{x}_{x,i}}, \tag{169}$$

where we have adopted the following notation:

- Variables (capitals are reserved for aggregate variables):
  - x individual state vector with I elements. x contains I_n endogenous states and I_x exogenous states $x = [x_n, x_x]$, $I = I_n + I_x$.
  - u individual control vector with J elements
  - v individual value function vector with 1 element
  - $u(x)$ control vector as function of individual state
  - $\mu(x)$ distribution of agents across states
  - $v(x)$ value function as function of individual state
  - X aggregate state vector (other than μ)
  - $\bar{U}$ aggregate control vector of purely contemporaneous variables

- U aggregate control vector of forward-looking variables. These are typically the multipliers of representative agents in the economy.
- $\tilde{U}$ control vector of aggregator variables
- $Z_t = \left\{\tilde{U}_t, U_t, \bar{U}_t, X_t\right\}$ vector of all aggregate variables

- Functions
  - b function that determines the drift of x
  - f_0 welfare function
  - f_1, f_2, f_3 aggregate equilibrium conditions
  - f_4 aggregator function
  - f_5 individual utility function

Line (156) is the planner's objective function.[70] Equations (157)-(159) are the aggregate equilibrium conditions for aggregate states, jump variables and contemporaneous variables. Notice that the variable U should be the multiplier associated to a backward-looking constraint in the optimization problem of a representative agent. In our model, examples for each of these three types of equations are the budget constraint of the household, the household's Euler equation and the household's labor supply condition, respectively; and $U = [\chi_H, \chi_F]$ are the multipliers of the household (on the budget constraint) and the retail firm (on the LOM of its price). Equation (160) links aggregate and individual variables, such as the definition of aggregate TFP in our model. Equations (161) and (162) are the individual agent's value function and first-order conditions, which must hold across the whole individual state vector x. In our model we do not have these two types of equations since we can analytically solve the individual optimal choice. The Kolmogorov Forward equation (22) determines the evolution of the distribution of agents. Finally (164)-(167) are the initial and terminal conditions for the aggregate and individual state and dynamic control variables. In our model these are the initial capital stock and firm distribution and the terminal conditions for variables such as consumption. The last two lines (168,169) are the boundary conditions of the KFE and the HJB associated to reflective boundaries on the exogenous states.

[70] Notice that the planner's discount rate, ϱ, can be different to that of individual agents, ρ.

2. Optimize, then discretize First we consider the approach introduced in Nuño and Thomas (2022), namely to compute the first-order conditions using calculus of variations and then to discretize the problem using an upwind finite difference scheme.

2.a The Lagrangian The Lagrangian for this problem is given by:[71]

$$
\begin{aligned}
\mathcal{L} &= \int_0^\infty e^{-\varrho t} \left\{ f_0(Z_t) \right. \\
&+ \lambda_{1,t} \left(\dot{X}_t - f_1(Z_t) \right) \\
&+ \lambda_{2,t} \left(\dot{U}_t - f_2(Z_t) \right) \\
&+ \lambda_{3,t} \left(f_3(Z_t) \right) \\
&+ \lambda_{4,t} \left(\tilde{U}_t - \int f_4\left(x, u_t(x), Z_t\right) \mu_t\left(x\right) \ dx \right) \\
&+ \int \left[\lambda_{5,t}(x) \left(-\rho v_t(x) + \dot{v}_t(x) + f_5(x, u_t(x), Z_t) \right) \right] dx \\
&+ \int \left[\lambda_{5,t}(x) \left(\sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right) \frac{\partial v_t(x)}{\partial x_i} + \sum_{i=1}^{I} \frac{\sigma_i^2(x)}{2} \frac{\partial^2 v_t(x)}{\partial^2 x_i} \right) \right] dx \\
&+ \sum_{j=1}^{J} \int \left[\lambda_{6,j,t}(x) \left(\frac{\partial f_5}{\partial u_{j,t}} + \sum_{i=1}^{I} \frac{\partial b_i}{\partial u_{j,t}} \frac{\partial v_t(x)}{\partial x_i} \right) \right] dx \\
&+ \left. \int \left[\lambda_{7,t}(x) \left(-\dot{\mu}_t\left(x\right) + \left(-\sum_{i=1}^{I} \frac{\partial}{\partial x_i} \left[b_i\left(x, u_t(x), Z_t\right) \mu_t\left(x\right) \right] + \frac{1}{2} \sum_{i=1}^{I} \frac{\partial^2}{\partial^2 x_i} \left[\sigma_i^2(x) \mu_t\left(x\right) \right] \right) \right) \right] dx \right\} dt
\end{aligned}
$$

where λ_1 to λ_7 denote the multipliers on the respective constraints. For convenience, we write the time derivatives in a separate line at the end. The Lagrangian becomes:

$$
\begin{aligned}
\mathcal{L} &= \int_0^\infty e^{-\varrho t} \left\{ f_0(Z_t) \right. \\
&+ \lambda_{1,t} \left(-f_1(Z_t) \right) \\
&+ \lambda_{2,t} \left(-f_2(Z_t) \right) \\
&+ \lambda_{3,t} \left(-f_3(Z_t) \right) \\
&+ \lambda_{4,t} \left(\tilde{U}_t - \int f_4\left(x, u_t(x), Z_t\right) \mu_t\left(x\right) \ dx \right) \\
&+ \int \left[\lambda_{5,t}(x) \left(-\rho v_t(x) + f_5(x, u_t(x), Z_t) + \sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right) \frac{\partial v_t(x)}{\partial x_i} + \sum_{i=1}^{I} \frac{\sigma_i^2(x)}{2} \frac{\partial^2 v_t(x)}{\partial^2 x_i} \right) \right] dx
\end{aligned}
$$

[71] For simplicity, we assume that the Wiener processes driving the dynamics of the state x are independent, though the proof can be trivially extended to that case, at the cost of a more cumbersome notation.

$$
\begin{aligned}
&+ \quad \sum_{j=1}^{J} \int \left[\lambda_{6,j,t}(x) \left(\frac{\partial f_{5,t}}{\partial u_{j,t}} + \sum_{i=1}^{I} \frac{\partial b_i}{\partial u_{j,t}} \frac{\partial v_t(x)}{\partial x_i} \right) \right] dx \\
&+ \quad \int \left[\lambda_{7,t}(x) \left(- \sum_{i=1}^{I} \frac{\partial}{\partial x_i} \left[b_i \left(x, u_t(x), Z_t \right) \mu_t\left(x \right) \right] + \frac{1}{2} \sum_{i=1}^{I} \frac{\partial^2}{\partial^2 x_i} \left[\sigma_i^2(x) \mu_t\left(x \right) \right] \right) \right] dx \Bigg\} dt \\
&+ \quad \int_0^\infty \left\{ e^{-\varrho t} \ \lambda_{1,t} \dot{X}_t + \lambda_{2,t} \dot{U}_t + \int \left[\lambda_{5,t} \dot{v}_t(x) \right] dx - \int \left[\lambda_{7,t} \dot{\mu}_t\left(x \right) \right] dx \right\} dt.
\end{aligned}
$$

We have ignored the terminal and initial conditions but we will account for them later on. Now we manipulate the Lagrangian using integration by parts in order to bring it into a more convenient form. We start with the last line. Switching the order of integration, the last line becomes

$$
\begin{aligned}
\int_0^\infty e^{-\varrho t} \ \lambda_{1,t} \dot{X}_t \ dt + \int_0^\infty e^{-\varrho t} \ \lambda_{2,t} \dot{U}_t \ dt + \int \int_0^\infty \left[e^{-\varrho t} \lambda_{5,t}(x) \dot{v}_t(x) \right] dt \ dx \\
- \int \int_0^\infty \left[e^{-\varrho t} \lambda_{7,t}(x) \dot{\mu}_t\left(x \right) \right] dt \ dx
\end{aligned}
$$

Now we integrate this expression by parts with respect to time t, using

$$
\begin{aligned}
\int_0^\infty e^{-\varrho t} \ a_t \dot{b}_t \ dt &= \left[e^{-\varrho t} a_t b_t \right]_0^\infty - \int_0^\infty e^{-\varrho t} (\dot{a}_{1,t} - \varrho a_{1,t}) b_t \ dt \\
&= \lim_{t \to \infty} e^{-\varrho t} a_t b_t - a_0 b_0 - \int_0^\infty e^{-\varrho t} (\dot{a}_t - \varrho a_t) b_t \ dt
\end{aligned}
$$

to get

$$
\begin{aligned}
& \lim_{t\to\infty} e^{-\varrho t} \lambda_{1,t} X_t - \lambda_{1,0} X_0 - \int_0^\infty e^{-\varrho t} (\dot{\lambda}_{1,t} - \varrho \lambda_{1,t}) X_t dt + \lim_{t\to\infty} e^{-\varrho t} \lambda_{2,t} U_t - \lambda_{2,0} U_0 \\
& - \int_0^\infty e^{-\varrho t} (\dot{\lambda}_{2,t} - \varrho \lambda_{2,t}) U_t dtx \\
+ \quad & \int \left(\lim_{t\to\infty} e^{-\varrho t} \lambda_{5,t}(x) v_t(x) - \lambda_{5,0}(x) v_0(x) \right) dx - \int \int_0^\infty e^{-\varrho t} (\dot{\lambda}_{5,t}(x) - \varrho \lambda_{5,t}(x)) v_t(x) dt dx \\
- \quad & \int \left(\lim_{t\to\infty} e^{-\varrho t} \lambda_{7,t}(x) \mu_t(x) - \lambda_{7,0}(x) \mu_0(x) \right) dx + \int \int_0^\infty e^{-\varrho t} (\dot{\lambda}_{7,t}(x) - \varrho \lambda_{7,t}(x)) \mu_t\left(x \right) dt dx
\end{aligned}
$$

Now we use the initial and terminal conditions to drop some $\lim_{t\to\infty}$ and $t = 0$ terms,[72]

$$
\lim_{t\to\infty} e^{-\varrho t} \lambda_{1,t} X_t - \lambda_{2,0} U_0 - \int_0^\infty e^{-\varrho t} (\dot{\lambda}_{1,t} - \varrho \lambda_{1,t}) X_t dt - \int_0^\infty e^{-\varrho t} (\dot{\lambda}_{2,t} - \varrho \lambda_{2,t}) U_t dt
$$

[72]We focus on the case that multipliers do not explode.

$$
\begin{aligned}
- & \int \lambda_{5,0}(x)v_0(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot{\lambda}_{5,t}(x) - \varrho\lambda_{5,t}(x))v_t(x)dtdx \\
- & \int \lim_{t\to\infty} e^{-\varrho t}\lambda_{7,t}(x)\mu_t(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot{\lambda}_{7,t}(x) - \varrho\lambda_{7,t}(x))\mu_t\left(x\right)dtdx
\end{aligned}
$$

Next we integrate lines 6 to 8 by parts with respect to x. Using the notational convention that $x_{\bar{i}} = x_{|x_i=\bar{x}_i}$, this yields:

$$
\begin{aligned}
& \int_0^\infty e^{-\rho t}\int\left(-\rho\lambda_{5,t}(x)v_t(x) + \lambda_{5,t}(x)f_5(x,u_t(x),Z_t)\right)dx \qquad (170)\\
& \int_0^\infty e^{-\rho t}\int\left(-\sum_{i=1}^{I}\frac{\partial b_i\left(x,u_t(x),Z_t\right)\lambda_{5,t}(x)}{\partial x_i}v_t(x) + \frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\lambda_{5,t}(x)\right]v_t(x)\right)dx \qquad (171)\\
& +\sum_{i=1}^{I_n}\int\left[\lim_{x_{n,i}\to\infty} b_i\left(x,u_t(x),Z_t\right)\lambda_{5,t}(x)v_t(x) - \lim_{x_{n,i}\to-\infty} b_i\left(x,u_t(x),Z_t\right)\lambda_{5,t}(x)v_t(x)\right]dx_{-i} \qquad (172)\\
& +\sum_{i=I_n+1}^{I}\int\left(b_i\left(x_{\bar{i}}\right)\lambda_{5,t}(x_{\bar{i}})v_t(x_{\bar{i}}) - b_2\left(x_{\underline{i}}\right)\lambda_{5,t}(x_{\underline{i}})v_t(x_{\underline{i}})\right)dx_{-i} \qquad (173)\\
& +\sum_{i=I_n+1}^{I}\int\left(\frac{\lambda_{5,t}(x_{\bar{i}})\sigma^2(x_{\bar{i}})}{2}\frac{\partial v_t\left(x\right)}{\partial x_i}_{|x=x_{\bar{i}}} - \frac{\lambda_{5,t}(x_{\underline{i}})\sigma^2(x_{\underline{i}})}{2}\frac{\partial v_t\left(x\right)}{\partial x_i}_{|x=x_{\underline{i}}}\right)dx_{-i} \qquad (174)\\
& +\sum_{i=I_n+1}^{I}\int\left(-\frac{\partial\lambda_{5,t}(x)\sigma_i^2(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{v_t\left(x_{\bar{i}}\right)}{2} + \frac{\partial\lambda_{5,t}(x)\sigma_i^2(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{v_t\left(x_{\bar{i}}\right)}{2}\right)dx_{-i} \qquad (175)\\
+ \; & \sum_{j=1}^{J}\int\left(\lambda_{6,j,t}(x)\frac{\partial f_{5,t}}{\partial u_{j,t}} - \sum_{i=1}^{I_n}\frac{\partial\left[\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}\right]}{\partial x_i}v_t(x)\right)dx\\
& +\sum_{j=1}^{J}\sum_{i=1}^{I_n}\int\left(\lim_{x_i\to\infty}\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}v_t(x) - \lim_{x_i\to\infty}\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}v_t(x)\right)dx_{-i} \qquad (176)\\
+ \; & \int\sum_{i=1}^{I}\left(\frac{\partial\lambda_{7,t}(x)}{\partial x_i}\left[b_i\left(x,u_t(x),Z_t\right)\mu_t\left(x\right)\right] + \frac{\partial^2\lambda_{7,t}(x)}{\partial^2 x_i}\frac{\sigma_i^2(x)}{2}\mu_t\left(x\right)\right)dx \qquad (177)\\
& +\sum_{i=1}^{I_n}\int\left(-\lim_{x_i\to\infty}\lambda_{7,t}(x)b_i\left(x,u_t(x),Z_t\right)\mu_t\left(x\right) + \lim_{x_i\to-\infty}\lambda_{7,t}(x)b_i\left(x,u_t(x),Z_t\right)\mu_t\left(x\right)\right)dx_{-i}\\
& +\sum_{i=I_n+1}^{I}\int\left(-\lambda_{7,t}(x_{\bar{i}})b_i\left(x_{\bar{i}},u_t(x_{\bar{i}}),Z_t\right)\mu_t\left(x_{\bar{i}}\right) + \lambda_{7,t}(x)b_i\left(x_{\underline{i}},u_t(x_{\underline{i}}),Z_t\right)\mu_t\left(x_{\underline{i}}\right)\right)dx_{-i} \qquad (178)\\
& +\sum_{i=I_n+1}^{I}\int\left(\lambda_{7,t}(x_{\bar{i}})\frac{\partial}{\partial x_i}\left[\sigma_i^2(x)\mu_t\left(x\right)\right]_{|x=x_{\bar{i}}} - \lambda_{7,t}(x_{\bar{i}})\frac{\partial}{\partial x_i}\left[\sigma_i^2(x)\mu_t\left(x\right)\right]_{|x=x_{\bar{i}}}\right)dx_{-i} \qquad (179)\\
& +\sum_{i=I_n+1}^{I}\int\left(-\frac{\partial\lambda_{7,t}(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{\sigma_i^2(x_{\bar{i}})}{2}\mu_t\left(x_{\bar{i}}\right) + \frac{\partial\lambda_{7,t}(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{\sigma_i^2(x_{\bar{i}})}{2}\mu_t\left(x_{\bar{i}}\right)\right)dx_{-i}\Bigg\}dt. \qquad (180)
\end{aligned}
$$

Using the boundary conditions (168) and (169) we can drop the fourth, tenth and eleventh lines of this expression. Reordering equations, the Lagrangian becomes

$$
\begin{aligned}
\mathcal{L} &= \int_0^\infty e^{-\varrho t}\left\{ f_0(Z_t) \right. \\
&+ \lambda_{1,t}\left(-f_1(Z_t)\right) \\
&+ \lambda_{2,t}\left(-f_2(Z_t)\right) \\
&+ \lambda_{3,t}\left(-f_3(Z_t)\right) \\
&+ \lambda_{4,t}\left(\tilde{U}_t - \int f_4\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\ dx\right) \\
&+ \int\left(-\rho\lambda_{5,t}(x)v_t(x) + \lambda_{5,t}(x)f_5(x, u_t(x), Z_t) - \sum_{i=1}^{I}\frac{\partial\left[b_i\left(x, u_t(x), Z_t\right)\lambda_{5,t}(x)\right]}{\partial x_i}v_t(x)\right)dx \\
&+ \int\left(\frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\lambda_{5,t}(x)\right]v_t(x)\right)dx \\
&+ \sum_{j=1}^{J}\int\left[\lambda_{6,j,t}(x)\frac{\partial f_{5,t}}{\partial u_{j,t}} - \sum_{i=1}^{I}\frac{\partial\left[\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}\right]}{\partial x_i}v_t(x)\right]dx \\
&+ \int\left[\left(\sum_{i=1}^{I}\frac{\partial\lambda_{7,t}(x)}{\partial x_i}\left[b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\right] + \sum_{i=1}^{I}\frac{\partial^2\lambda_{7,t}(x)}{\partial^2 x_i}\frac{\sigma_i^2(x)}{2}\mu_t\left(x\right)\right)\right]dx \\
&+ \sum_{i=1}^{I_n}\int\left[\lim_{x_{n,i}\to\infty} b_i\left(x, u_t(x), Z_t\right)\lambda_{5,t}(x)v_t(x) - \lim_{x_{n,i}\to-\infty} b_i\left(x, u_t(x), Z_t\right)\lambda_{5,t}(x)v_t(x)\right]dx_{-i} \\
&+ \sum_{i=I_n+1}^{I}\int\left(b_i\left(x_{\bar{i}}\right)\lambda_{5,t}(x_{\bar{i}})v_t(x_{\bar{i}}) - b_2\left(x_{\underline{i}}\right)\lambda_{5,t}(x_{\underline{i}})v_t(x_{\underline{i}})\right)dx_{-i} \\
&+ \sum_{i=I_n+1}^{I_x}\int\left(-\frac{\partial\lambda_{5,t}(x)\sigma_i^2(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{v_t\left(x_{\bar{i}}\right)}{2} + \frac{\partial\lambda_{5,t}(x)\sigma_i^2(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{v_t\left(x_{\bar{i}}\right)}{2}\right)dx_{-i} \\
&+ \sum_{j=1}^{J}\sum_{i=1}^{I_n}\int\left(\lim_{x_i\to\infty}\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}v_t(x) - \lim_{x_i\to\infty}\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}v_t(x)\right)dx_{-i} \\
&+ \sum_{i=1}^{I_n}\int\left(-\lim_{x_i\to\infty}\lambda_{7,t}(x)b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right) + \lim_{x_i\to-\infty}\lambda_{7,t}(x)b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\right)dx_{-i} \\
&+ \left. \sum_{i=I_n+1}^{I}\int\left(-\frac{\partial\lambda_{7,t}(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{\sigma_i^2(x_{\bar{i}})}{2}\mu_t\left(x_{\bar{i}}\right) + \frac{\partial\lambda_{7,t}(x)}{\partial x_i}_{|x=x_{\bar{i}}}\frac{\sigma_i^2(x_{\bar{i}})}{2}\mu_t\left(x_{\bar{i}}\right)\right)dx_{-i}\right\}dt. \\
&+ \lim_{t\to\infty} e^{-\varrho t}\lambda_{1,t}X_t - \lambda_{2,0}U_0 - \int_0^\infty e^{-\varrho t}(\dot{\lambda}_{1,t} - \varrho\lambda_{1,t})X_t dt - \int_0^\infty e^{-\varrho t}(\dot{\lambda}_{2,t} - \varrho\lambda_{2,t})U_t dt \\
&+ \int -\lambda_{5,0}(x)v_0(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot{\lambda}_{5,t}(x) - \varrho\lambda_{5,t}(x))v_t(x)dtdx \\
&- \int\lim_{t\to\infty} e^{-\varrho t}\lambda_{7,t}(x)\mu_t(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot{\lambda}_{7,t}(x) - \varrho\lambda_{7,t}(x))\mu_t\left(x\right)dtdx
\end{aligned}
$$

2.b Optimality conditions in the continuous state space We take the Gateaux derivatives in direction $h_t(x)$ for each endogenous variable x. These derivatives have to be equal to zero for any $h_t(x)$ in the optimum. This implies the following optimality conditions:

Aggregate variables:

$$
\begin{aligned}
U_t: \quad 0 = \quad & -(\dot{\lambda}_{2,t} - \varrho\lambda_{2,t}) && (181)\\
& + \frac{\partial f_{0,t}}{\partial U_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial U_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial U_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial U_t} - \lambda_{4,t}\int \frac{\partial f_{4,t}}{\partial U_t}\mu_t(x)\ dx && (182)\\
& + \int \left[\lambda_{5,t}(x)\left(\frac{\partial f_{5,t}}{\partial U_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial U_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx && (183)\\
& + \sum_{j=1}^{J}\int \left[\lambda_{6,j,t}(x)\left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t}\partial U_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial u_{j,t}\partial U_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx && (184)\\
& + \int \left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[\frac{\partial b_{i,t}}{\partial U_t}\mu_t(x)\right]\right)\right] dx, && (185)\\
& \forall t > 0, && (186)
\end{aligned}
$$

$$
0 = \quad \lambda_{2,0}. \qquad (187)
$$

$$
\begin{aligned}
X_t: \quad 0 = \quad & -(\dot{\lambda}_{1,t} - \varrho\lambda_{1,t})\\
& + \frac{\partial f_{0,t}}{\partial X_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial X_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial X_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial X_t} - \lambda_{4,t}\int \frac{\partial f_{4,t}}{\partial X_t}\mu_t(x)\ dx\\
& + \int \left[\lambda_{5,t}(x)\left(\frac{\partial f_{5,t}}{\partial X_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial X_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx\\
& + \sum_{j=1}^{J}\int \left[\lambda_{6,j,t}(x)\left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t}\partial X_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial u_{j,t}\partial X_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx\\
& + \int \left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[\frac{\partial b_{i,t}}{\partial X_t}\mu_t(x)\right]\right)\right] dx,\\
& \forall t \geq 0,
\end{aligned}
$$

$$0 = \lim_{t\to\infty} e^{-\varrho t}\lambda_{1,t}(x).$$

$$\begin{aligned}
\hat{U}_t: \quad 0 = \; & 0 \\
& + \frac{\partial f_{0,t}}{\partial \hat{U}_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial \hat{U}_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial \hat{U}_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial \hat{U}_t} - \lambda_{4,t}\int \frac{\partial f_{4,t}}{\partial \hat{U}_t}\mu_t\left(x\right)\; dx \\
& + \int \left[\lambda_{5,t}(x)\left(\frac{\partial f_{5,t}}{\partial \hat{U}_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial \hat{U}_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx \\
& + \sum_{j=1}^{J}\int \left[\lambda_{6,j,t}(x)\left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t}\partial \hat{U}_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial u_{j,t}\partial \hat{U}_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx \\
& + \int \left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[\frac{\partial b_{i,t}}{\partial \hat{U}_t}\mu_t\left(x\right)\right]\right)\right] dx, \\
& \forall t \geq 0.
\end{aligned}$$

$$\begin{aligned}
\tilde{U}_t: \quad 0 = \; & \lambda_{4,t} \\
& + \frac{\partial f_{0,t}}{\partial \tilde{U}_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial \tilde{U}_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial \tilde{U}_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial \tilde{U}_t} - \lambda_{4,t}\int \frac{\partial f_{4,t}}{\partial \tilde{U}_t}\mu_t\left(x\right)\; dx \\
& + \int \left[\lambda_{5,t}(x)\left(\frac{\partial f_{5,t}}{\partial \tilde{U}_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial \tilde{U}_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx \\
& + \sum_{j=1}^{J}\int \left[\lambda_{6,j,t}(x)\left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t}\partial \tilde{U}_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial u_{j,t}\partial \tilde{U}_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx \\
& + \int \left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[\frac{\partial b_{i,t}}{\partial \tilde{U}_t}\mu_t\left(x\right)\right]\right)\right] dx, \\
& \forall t \geq 0.
\end{aligned}$$

Value function, distribution and policy functions

$$\begin{aligned}
v_t\left(x\right): \quad 0 = \; & \left(-\lambda_{5,t}(x)\rho - \sum_{i=1}^{I}\frac{\partial\left[\lambda_{5,t}(x)b_i\left(x, u_t(x), Z_t\right)\right]}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\lambda_{5,t}(x)\right]\right) \\
& - \sum_{j=1}^{J}\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left(\lambda_{6,j,t}(x)\frac{\partial b_i\left(x, u_t(x), Z_t\right)}{\partial u_{j,t}}\right) \\
& - (\dot{\lambda}_{5,t}(x) - \varrho\lambda_{5,t}(x))
\end{aligned}$$

$$\forall t > 0,\ \forall x_{|\underline{x}_i < x_i < \bar{x}_i \forall i < I_n},$$

$$
\begin{aligned}
0 =&\quad \lim_{x_i \to \infty} \lambda_{5,t}(x)\ \forall x_i | i \leq I_n, \\
0 =&\quad \lim_{x_i \to -\infty} \lambda_{5,t}(x)\ \forall x_i | i \leq I_n, \\
0 =&\quad b_2\left(x_{\bar{i}}\right)\lambda_{5,t}(x_{\bar{i}}) - \frac{1}{2}\frac{\partial}{\partial x_i}\left[\lambda_{5,t}(x)\sigma^2(x_2)\right]_{||x = x_{\bar{i}}}\ \ \forall x_i | I_n > i \geq I, \\
0 =&\quad b_2\left(x_{\underline{i}}\right)\lambda_{5,t}(x_{\underline{i}}) - \frac{1}{2}\frac{\partial}{\partial x_i}\left[\lambda_{5,t}(x)\sigma^2(x_2)\right]_{||x = x_{\underline{i}}}\ \ \forall x_i | I_n > i \geq I,
\end{aligned}
$$

$$0 = \quad \lambda_{5,0}(x).$$

$$
\begin{aligned}
\mu_t(x): \quad 0 =&\quad -\lambda_{4,t} f_4\left(x, u_t(x), Z_t\right) \\
&+ \left(\sum_{i=1}^{I} \frac{\partial \lambda_{7,t}(x)}{\partial x_i} b_i\left(x, u_t(x), Z_t\right) + \sum_{i=1}^{I} \frac{\partial^2 \lambda_{7,t}(x)}{\partial^2 x_i}\frac{\sigma_i^2(x)}{2}\right) \\
&+ (\dot{\lambda}_{7t}(x) - \varrho\lambda_{7,t}(x)), \\
&\forall t \geq 0,\ \forall x_{|\underline{x}_i < x_i < \bar{x}_i \forall i < I_n},
\end{aligned}
$$

$$
\begin{aligned}
0 =&\quad \lim_{x_1 \to \infty} \lambda_{7,t}(x)\ \forall x_i | i \leq I_n, \\
0 =&\quad \lim_{x_1 \to -\infty} \lambda_{7,t}(x)\ \forall x_i | i \leq I_n, \\
0 =&\quad \frac{\partial \lambda_{7,t}(x)}{\partial x}_{|x = x_{\bar{i}}}\ \ \forall x_i | I_n > i \geq I, \\
0 =&\quad \frac{\partial \lambda_{7,t}(x)}{\partial x}_{||x = x_{\underline{i}}}\ \ \forall x_i | I_n > i \geq I,
\end{aligned}
$$

$$0 = \quad \lim_{t \to \infty} e^{-\varrho t}\lambda_{7,t}(x).$$

$$
\begin{aligned}
u_{l,t}(x): \quad 0 =&\quad -\lambda_{4,t}\frac{\partial f_4}{\partial u_{l,t}}\mu_t(x) \\
&+ \quad \overbrace{\left[\lambda_{5,t}(x)\left(\frac{\partial f_5}{\partial u_{l,t}} + \sum_{i=1}^{I}\frac{\partial b_i}{\partial u_{l,t}}\frac{\partial v_t(x)}{\partial x_i}\right)\right]}^{=0}
\end{aligned}
$$

$$
\begin{aligned}
+ \quad & \sum_{j=1}^{J} \lambda_{6,j,t}(x) \left(\frac{\partial^2 f_5}{\partial u_{l,t} \partial u_{j,t}} + \sum_{i=1}^{I} \frac{\partial^2 b_i}{\partial u_{l,t} \partial u_{j,t}} \frac{\partial v_t(x)}{\partial x_i} \right) \\
- \quad & \left(\sum_{i=1}^{I} \frac{\partial \lambda_{7,t}(x)}{\partial x_i} \frac{\partial b_{i,t}}{\partial u_{l,t}} \mu_t\left(x\right) \right) \\
& \forall j,\ \forall t,\ \forall x,
\end{aligned}
$$

$$
\begin{aligned}
0 = \quad & \lim_{x_i \to \infty} \lambda_{6,j,t}(x)\ \forall x_i | i \leq I_n, \\
0 = \quad & \lim_{x_i \to -\infty} \lambda_{6,j,t}(x)\ \forall x_i | i \leq I_n.
\end{aligned}
$$

2.c Discretized optimality conditions Now we discretize these conditions with respect to time and idiosyncratic states.

The idiosyncratic state is discretized by a evenly-spaced grid of size $[N_1, ..., N_I]$ where $1, .., I$ are the dimensions of the state x. We assume that in each dimension there is no mass of agents outside the compact domain $[x_{i,1}, x_{i,N_i}]$. The state step size in dimension i is Δx_i. We define $x^n \equiv (x_{1,n_1}, ..., x_{i,n_i}, ..., x_{I,n_I})$, where $n_1 \in \{1, N_1\}, ..., n_I \in \{1, N_I\}$. We are assuming that, due to state constraints and/or reflecting boundaries, the dynamics of idiosyncratic states are constrained to the compact set $[x_{1,1}, x_{1,N_1}] \times [x_{2,1}, x_{2,N_2}] \times .... \times [x_{I,1}, x_{I,N_I}]$. We also define $x^{n_i+1} \equiv (x_{1,n_1}, ..., x_{i,n_i+1}, ..., x_{I,n_I})$, $x^{n_i-1} \equiv (x_{1,n_1}, ..., x_{i,n_i-1}, ..., x_{I,n_I})$ $f_t^n \equiv f\left(x^n, u_t^n, Z_t\right)$, $f_t^{n_i-1} \equiv f\left(x^{n_i-1}, u_t^n, Z_t\right)$ and $f_t^{n_i+1} \equiv f\left(x^{n_i+1}, u_t^n, Z_t\right)$. I.e. the superscript n indicates a particular grid point and the superscript n_i+1 and n_i-1 indicate neighboring grid points along dimension i.

To discretize the problem we now (i) approximate time derivatives of forward-looking (backward-looking) variables and multipliers by forward (backward) derivatives, (ii) replace integrals by sums (iii) approximate derivatives with respect to x by the upwind derivatives ∇ or $\hat{\nabla}$:

$$
\begin{aligned}
\nabla_i \left[v_t^n\right] \quad & \equiv \quad \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{v_t^{n_i+1} - v_t^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{v_t^n - v_t^{n_i-1}}{\Delta x_i} \right], \\
\hat{\nabla}_i \left[\mu_t^n\right] \quad & \equiv \quad \left[\frac{\mathbb{I}_{b_{i,t}^{n_i+1} < 0} \mu_t^{n_i+1} - \mathbb{I}_{b_{i,t}^n < 0} \mu_t^n}{\Delta x_i} + \frac{\mathbb{I}_{b_{i,t}^n > 0} \mu_t^n - \mathbb{I}_{b_{i,t}^{n_i-1} > 0} \mu_t^{n_i-1}}{\Delta x_i} \right],
\end{aligned}
$$

for any discretized functions v_t^n, μ_t^n. In particular, we use the first derivative ∇_i for the value function $v_t(x)$ and the multiplier on the KFE $\lambda_{7,t}(x)$, and the second derivative $\hat{\nabla}_i$ for the distribution $\mu_t(x)$ and the multiplier on the HJB $\lambda_{5,t}(x)$, since those are the appropriate upwind discretization schemes consistent with the respective boundary conditions in the x-dimension reported above. We simplify the notation for sums $\sum_n \equiv \sum_{n_1\in\{1,..,N_1\},..,n_I\in\{1,..,N_I\}}$.

The second-order derivative is approximated as

$$\triangle_i \left[v_t^n\right] \quad \equiv \quad \left[\frac{\left(v_t^{n_i+1}\right)+\left(v_t^{n_i-1}\right)-2\left(v_t^n\right)}{\left(\Delta x_i\right)^2}\right].$$

We maintain the subscript t even if it refers now to discrete time with a step Δt, that is, $X_{t+1}$is the shortcut for $X_{t+\triangle t}$. To discretize time, for the state variables $Z^s = \{X, \mu_t\left(x\right), \lambda_2, \lambda_5(x)\}$ we replace the time derivative $\dot{Z}_t^s$ by $\frac{Z_t^s - Z_{t-1}^s}{\Delta t}$. For the forward-looking control variables $Z^f = \{U, v_t\left(x\right), \lambda_1, \lambda_7(x)\}$ we replace the time derivatives $\dot{Z}_t^f$ by $\frac{Z_{t+1}^f - Z^f}{\Delta t}$.

We start with the optimality condition for U_t

$$\begin{aligned}
U_t: \quad 0 = \quad & -\left(\frac{\lambda_{2,t}-\lambda_{2,t-1}}{\Delta t}-\varrho\lambda_{2,t}\right) && (188)\\
& +\frac{\partial f_0}{\partial U_t}-\lambda_{1,t}\frac{\partial f_1}{\partial U_t}-\lambda_{2,t}\frac{\partial f_2}{\partial U_t}-\lambda_{3,t}\frac{\partial f_3}{\partial U_t}-\lambda_{4,t}\sum_{n=1}^{N}\frac{\partial f_4^n}{\partial U_t}\mu_t^n && (189)\\
& +\sum_n\left[\lambda_{5,t}^n\left(\frac{\partial f_5^n}{\partial U_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial U_t}\nabla_i\left[v_t^n\right]\right)\right]\\
& +\sum_{j=1}^{J}\sum_n\left[\lambda_{6,j,t}^n\left(\frac{\partial^2 f_5^n}{\partial u_j\partial U_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial u_j\partial U_t}\nabla_i\left[v_t^n\right]\right)\right]\\
& +\sum_n\left[-\lambda_{7,t}^n\sum_{i=1}^{I}\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n\right]\right] && (190)\\
& \forall t\geq 0.
\end{aligned}$$

The optimality conditions for the other aggregate variables look very much alike:

$$\begin{aligned}
X_t: \quad 0 = \quad & -\left(\frac{\lambda_{1,t+1}-\lambda_{1,t}}{\Delta}-\varrho\lambda_{1,t}\right)\\
& +\frac{\partial f_0}{\partial X_t}-\lambda_{1,t}\frac{\partial f_1}{\partial X_t}-\lambda_{2,t}\frac{\partial f_2}{\partial X_t}-\lambda_{3,t}\frac{\partial f_3}{\partial X_t}-\lambda_{4,t}\sum_n\frac{\partial f_4^n}{\partial X_t}\mu_t^n
\end{aligned}$$

$$+\sum_{n}\left[\lambda_{5,t}^{n}\left(\frac{\partial f_5^n}{\partial X_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial X_t}\nabla_i\left[v_t^n\right]\right)\right]$$
$$+\sum_{j=1}^{J}\sum_{n}\left[\lambda_{6,j,t}^{n}\left(\frac{\partial^2 f_5^n}{\partial u_j\partial X_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial u_j\partial X_t}\nabla_i\left[v_t^n\right]\right)\right]$$
$$+\sum_{n}\left[-\lambda_{7,t}^{n}\sum_{i=1}^{I}\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial X_t}\mu_t^n\right]\right]$$
$$\forall t>0.$$

$$\begin{aligned}\hat{U}_t:\quad 0 =\ & 0\\ &+\frac{\partial f_0}{\partial \hat{U}_t}-\lambda_{1,t}\frac{\partial f_1}{\partial \hat{U}_t}-\lambda_{2,t}\frac{\partial f_2}{\partial \hat{U}_t}-\lambda_{3,t}\frac{\partial f_3}{\partial \hat{U}_t}-\lambda_{4,t}\sum_{n}\frac{\partial f_4^n}{\partial \hat{U}_t}\mu_t^n\\ &+\sum_{n}\left[\lambda_{5,t}^{n}\left(\frac{\partial f_5^n}{\partial \hat{U}_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial \hat{U}_t}\nabla_i\left[v_t^n\right]\right)\right]\\ &+\sum_{j=1}^{J}\sum_{n}\left[\lambda_{6,j,t}^{n}\left(\frac{\partial^2 f_5^n}{\partial u_j\partial \hat{U}_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial u_j\partial \hat{U}_t}\nabla_i\left[v_t^n\right]\right)\right]\\ &+\sum_{n}\left[-\lambda_{7,t}^{n}\sum_{i=1}^{I}\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial \hat{U}_t}\mu_t^n\right]\right]\\ &\forall t\geq 0.\end{aligned}$$

$$\begin{aligned}\tilde{U}_t:\quad 0 =\ & \lambda_{4,t}\\ &+\frac{\partial f_0}{\partial \tilde{U}_t}-\lambda_{1,t}\frac{\partial f_1}{\partial \tilde{U}_t}-\lambda_{2,t}\frac{\partial f_2}{\partial \tilde{U}_t}-\lambda_{3,t}\frac{\partial f_3}{\partial \tilde{U}_t}-\lambda_{4,t}\sum_{n=1}^{N}\frac{\partial f_4^n}{\partial \tilde{U}_t}\mu_t^n\\ &+\sum_{n}\left[\lambda_{5,t}^{n}\left(\frac{\partial f_5^n}{\partial \tilde{U}_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial \tilde{U}_t}\nabla_i\left[v_t^n\right]\right)\right]\\ &+\sum_{j=1}^{J}\sum_{n}\left[\lambda_{6,j,t}^{n}\left(\frac{\partial^2 f_5^n}{\partial u_j\partial \tilde{U}_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial u_j\partial \tilde{U}_t}\nabla_i\left[v_t^n\right]\right)\right]\\ &+\sum_{n}\left[-\lambda_{7,t}^{n}\sum_{i=1}^{I}\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial \tilde{U}_t}\mu_t^n\right]\right]\\ &\forall t\geq 0.\end{aligned}$$

The discretized optimality condition with respect to the value function $v_t\left(x\right)$, the distribution $\mu_t\left(x\right)$ and the individual jump variable $u_{j,t}(x)$ are.

$$
\begin{aligned}
v_t\left(x\right): \quad 0 = \quad & -\lambda_{5,t}^{n}\rho - \sum_{i=1}^{I}\hat{\nabla}_i\left[\lambda_{5,t}^{n}b_{i,t}^{n}\right] \\
& +\frac{1}{2}\sum_{i=1}^{I}\triangle_i^2\left[\sigma_i^n\lambda_{5,t}^{n}\right] \\
& -\sum_{j=1}^{J}\sum_{i=1}^{I}\left(\hat{\nabla}_i\left[\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}\right]\right) \\
& -(\frac{\lambda_{5,t}^{n}-\lambda_{5,t-1}^{n}}{\Delta t}-\varrho\lambda_{5,t}^{n}). \\
& \forall t \geq 0.
\end{aligned} \tag{191}
$$

$$
\begin{aligned}
\mu_t\left(x\right): \quad 0 = \quad & -\lambda_{4,t}f_{4,t}^{n} \\
& +\left(\sum_{i=1}^{I}b_i\left(x,u_t(x),Z_t\right)\nabla_i\left[\lambda_{7,t}^{n}\right]+\frac{1}{2}\sum_{i=1}^{I}\left(\sigma_i^2\right)^n\triangle_i^2\left[\lambda_{7,t}^{n}\right]\right) \\
& +\frac{\lambda_{7,t+1}^{n}-\lambda_{7,t}^{n}}{\Delta t}-\varrho\lambda_{7,t}^{n} \\
& \forall t > 0.
\end{aligned} \tag{192}
$$

$$
\begin{aligned}
u_{l,t}\left(x\right): \quad 0 = \quad & -\lambda_{4,t}\frac{\partial f_4}{\partial u_{l,t}}\mu_t^n \\
& +\sum_{j=1}^{J}\lambda_{6,k,t}^{n}\left(\frac{\partial^2 f_{5,t}^{n}}{\partial u_{l,t}^{n}\partial u_{j,t}^{n}}+\sum_{i=1}^{I}\frac{\partial^2 b_{i,t}^{n}}{\partial u_{l,t}^{n}\partial u_{j,t}^{n}}\nabla_i\left[v_t^n\right]\right) \\
& -\sum_{i=1}^{I}\nabla_i\left[\lambda_{7,t}^{n}\right]\frac{\partial b_{i,t}^{n}}{\partial u_{l,t}}\mu_t^n \\
& \forall t \geq 0.
\end{aligned} \tag{193}
$$

3. Discretize the state first, then optimize, then discretize time Now we apply our alternative approach. We first discretize the state space, then we solve the planners problem, then we discretize time.

3.a The discretized planner's problem Now first discretize the optimization problem with respect to the idiosyncratic state (N grid points, grid step Δx_i).

$$\max_{Z_t,u_t^n,\mu_t^n,v_t^n} \int_0^\infty \exp(-\varrho t) f_0(Z_t) dt$$

$$s.t.\ \forall t$$

$$\dot{X}_t = f_1(Z_t) \tag{194}$$

$$\dot{U}_t = f_2(Z_t) \tag{195}$$

$$0 = f_3(Z_t) \tag{196}$$

$$\tilde{U}_t = \sum_{n=1}^{N} f_4\left(x^n, u_t^n, Z_t\right) \mu_t^n \tag{197}$$

$$\rho v_t^n = \dot{v}_t + f_5(x^n, u_t^n, Z_t) + \sum_{i=1}^{I} b_i\left(x^n, u_t^n, Z_t\right) \nabla_i \left[v_t^n\right] \tag{198}$$

$$+\frac{1}{2}\sum_{i=1}^{I}\left(\sigma_i^2\right)^n \triangle_i^2\left[v_t^n\right],\ \forall n$$

$$0 = \frac{\partial f_{5,t}^n}{\partial u_{j,t}^n} + \sum_{i=1}^{I} \frac{\partial b_{i,t}^n}{\partial u_{j,t}^n} \nabla_i \left[v_t^n\right], \quad \forall j, n \tag{199}$$

$$\dot{\mu}_t^n = -\sum_{i=1}^{I} \hat{\nabla}_i\left[b_{i,t}^n \mu_t^n\right] \tag{200}$$

$$+\frac{1}{2}\sum_{i=1}^{I} \triangle_i\left[\sigma_i^2 \mu_t^n\right] \tag{201}$$

$$X_0 = \bar{X}_0 \tag{202}$$

$$\mu_0^n = \bar{\mu}_0^n \tag{203}$$

$$0 = \lim_{t\to\infty} \exp(-\varrho t) U_t \tag{204}$$

$$0 = \lim_{t\to\infty} \exp(-\varrho t) v_t(x) \tag{205}$$

Note: The boundary conditions in the x-dimension are absorbed in the upwind derivatives.

3.b The Lagrangian The Lagrangian is

$$
\begin{aligned}
L \quad = \quad & \int_0^\infty e^{-\varrho t} f_0(Z_t)dt \\
& + \int_0^\infty e^{-\varrho t}\lambda_{1,t}\left\{\dot{X}_t - f_1(Z_t)\right\}dt \\
& + \int_0^\infty e^{-\varrho t}\lambda_{2,t}\left\{\dot{U}_t - f_2(Z_t)\right\}dt \\
& + \int_0^\infty e^{-\varrho t}\lambda_{3,t}\left\{-f_3(Z_t)\right\}dt \\
& + \int_0^\infty e^{-\varrho t}\lambda_{4,t}\left\{\tilde{U}_t - \sum_n f_4\left(x^n, u_t^n, Z_t\right)\mu_t^n\right\}dt \\
& + \int_0^\infty e^{-\varrho t}\sum_n \beta^t \lambda_{5,t}^n \left\{\begin{array}{c} -\rho v_t^n + \dot{v}_t^n + f_5(x^n, u_t^n, Z_t) + \sum_{i=1}^I b_i\left(x^n, u_t^n, Z_t\right)\nabla_i\left[v_t^n\right] \\ + \sum_{i=1}^I \left(\sigma_i^2\right)^n \triangle_i^2\left[v_t^n\right] \end{array}\right\}dt \\
& + \int_0^\infty e^{-\varrho t}\sum_n \sum_{j=1}^J \beta^t \lambda_{6,j,t}^n \left\{\frac{\partial f_{5,t}^n}{\partial u_{j,t}^n} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial u_{j,t}^n}\nabla_i\left[v_t^n\right]\right\}dt \\
& + \int_0^\infty e^{-\varrho t}\sum_n \beta^t \lambda_{7,t}^n \left\{\begin{array}{c} -\dot{\mu}_t^n - \sum_{i=1}^I \hat{\nabla}_i\left[b_{i,t}^n \mu_t^n\right] \\ +\frac{1}{2}\sum_{i=1}^I \triangle_i\left[\sigma_i^2 \mu_t^n\right] \end{array}\right\}dt
\end{aligned}
$$

3.c The optimality conditions The FOCs are

$$
\begin{aligned}
\frac{\partial L}{\partial U_t} : 0 \quad = \quad & \frac{\partial f_{0,t}}{\partial U_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial U_t} - \left(\dot{\lambda}_{2,t} - \varrho\lambda_{2,t}\right) - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial U_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial U_t} - \lambda_{4,t}\sum_n \frac{\partial f_{4,t}^n}{\partial U_t}\mu_t^n \\
& + \sum_n \lambda_{5,t}^n\left\{+\frac{\partial f_{5,t}^n}{\partial U_t} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& + \sum_n \sum_{j=1}^J \lambda_{6,j,t}^n\left\{\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial U_t} + \sum_{i=1}^I \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& + \sum_n\left\{\sum_{i=1}^I\left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial U_t}\frac{\mu_t^n}{\Delta x_i}\right] + \sum_{i=1}^I\left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right)\left[\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial U_t}\frac{\mu_t^n}{\Delta x_i}\right]\right\} \\
& \forall t \geq 0
\end{aligned} \tag{206}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial X_t} : 0 \quad = \quad & \frac{\partial f_{0,t}}{\partial X_t} - \left(\dot{\lambda}_{1,t} - \varrho\lambda_{1,t}\right) - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial X_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial X_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial X_t} - \lambda_{4,t}\sum_n \frac{\partial f_{4,t}^n}{\partial X_t}\mu_t^n \\
& + \sum_n \lambda_{5,t}^n\left\{\frac{\partial f_{5,t}^n}{\partial X_t} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial X_t}\nabla_i\left[v_t^n\right]\right\}
\end{aligned}
$$

$$+\sum_n \sum_j \lambda_{6,j,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial X_t} + \sum_{i=1}^I \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial X_t} \nabla_i \left[v_t^n \right] \right\}$$
$$+\sum_n \left\{ \sum_{i=1}^I \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1} \right) \left[\mathbb{I}_{b_{i,t}^n<0} \frac{\partial b_{i,t}^n}{\partial X_t} \frac{\mu_t^n}{\Delta x_i} \right] + \sum_{i=1}^I \left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n \right) \left[\mathbb{I}_{b_{i,t}^n>0} \frac{\partial b_{i,t}^n}{\partial X_t} \frac{\mu_t^n}{\Delta x_i} \right] \right\}$$
$$\forall t > 0$$

$$\frac{\partial L}{\partial \tilde{U}_t} : 0 = \frac{\partial f_{0,t}}{\partial \tilde{U}_t} - \lambda_{1,t} \frac{\partial f_{1,t}}{\partial \tilde{U}_t} - \lambda_{2,t} \frac{\partial f_{2,t}}{\partial \tilde{U}_t} - \lambda_{3,t} \frac{\partial f_{3,t}}{\partial \tilde{U}_t} + \lambda_{4,t} - \lambda_{4,t} \sum_n \frac{\partial f_{4,t}^n}{\partial \tilde{U}_t} \mu_t^n$$
$$+\sum_n \lambda_{5,t}^n \left\{ +\frac{\partial f_{5,t}^n}{\partial \tilde{U}_t} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial \tilde{U}_t} \nabla_i \left[v_t^n \right] \right\}$$
$$+\sum_n \sum_j \lambda_{6,j,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial \tilde{U}_t} + \sum_{i=1}^I \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial \tilde{U}_t} \nabla_i \left[v_t^n \right] \right\}$$
$$+\sum_n \left\{ \sum_{i=1}^I \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1} \right) \left[\mathbb{I}_{b_{i,t}^n<0} \frac{\partial b_{i,t}^n}{\partial \tilde{U}_t} \frac{\mu_t^n}{\Delta x_i} \right] + \sum_{i=1}^I \left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n \right) \left[\mathbb{I}_{b_{i,t}^n>0} \frac{\partial b_{i,t}^n}{\partial \tilde{U}_t} \frac{\mu_t^n}{\Delta x_i} \right] \right\}$$
$$\forall t \geq 0$$

$$\frac{\partial L}{\partial \hat{U}_t} : 0 = \frac{\partial f_{0,t}}{\partial \hat{U}_t} - \lambda_{1,t} \frac{\partial f_{1,t}}{\partial \hat{U}_t} - \lambda_{2,t} \frac{\partial f_{2,t}}{\partial \hat{U}_t} - \lambda_{3,t} \frac{\partial f_{3,t}}{\partial \hat{U}_t} - \lambda_{4,t} \sum_n \frac{\partial f_{4,t}^n}{\partial \hat{U}_t} \mu_t^n$$
$$+\sum_n \lambda_{5,t}^n \left\{ +\frac{\partial f_{5,t}^n}{\partial \hat{U}_t} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial \hat{U}_t} \nabla_i \left[v_t^n \right] \right\}$$
$$+\sum_n \sum_j \lambda_{6,j,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial \hat{U}_t} + \sum_{i=1}^I \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial \hat{U}_t} \nabla_i \left[v_t^n \right] \right\}$$
$$+\sum_n \left\{ \sum_{i=1}^I \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1} \right) \left[\mathbb{I}_{b_{i,t}^n<0} \frac{\partial b_{i,t}^n}{\partial \hat{U}_t} \frac{\mu_t^n}{\Delta x_i} \right] + \sum_{i=1}^I \left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n \right) \left[\mathbb{I}_{b_{i,t}^n>0} \frac{\partial b_{i,t}^n}{\partial \hat{U}_t} \frac{\mu_t^n}{\Delta x_i} \right] \right\}$$
$$\forall t \geq 0$$

$$\frac{\partial L}{\partial v_t^n} : 0 = \lambda_{5,t}^n \left\{ -\rho + \sum_{i=1}^I b_{i,t}^n \frac{\mathbb{I}_{b_t^n<0} - \mathbb{I}_{b_t^n>0}}{\Delta x_i} - \sum_{i=1}^I \frac{2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \tag{207}$$
$$-(\dot{\lambda}_{5,t}^n - \varrho \lambda_{5,t}^n)$$
$$+\sum_{i=1}^I \lambda_{5,t}^{n_i-1} b_{i,t}^{n_i-1} \frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i} + \sum_{i=1}^I \lambda_{5,t}^{n_i-1} \frac{\left(\sigma_i^2\right)^{n_i-1}}{2\left(\Delta x_i\right)^2}$$
$$-\sum_{i=1}^I \lambda_{5,t}^{n_i+1} b_{i,t}^{n_i+1} \frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i} + \sum_{i=1}^I \lambda_{5,t}^{n_i+1} \frac{\left(\sigma_i^2\right)^{n_i+1}}{2\left(\Delta x_i\right)^2}$$

$$
\begin{aligned}
&+\sum_{j=1}^{J}\sum_{i=1}^{I}\left\{\lambda_{6,j,t}^{n}\left\{\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}\frac{\mathbb{I}_{b_{i,t}^{n}<0}-\mathbb{I}_{b_{i,t}^{n}>0}}{\Delta x_i}\right\}+\lambda_{6,j,t}^{n_i-1}\left\{\frac{\partial b_{i.t}^{n_i-1}}{\partial u_{j,t}^{n_i-1}}\frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i}\right\}\right\}\\
&-\sum_{j=1}^{J}\sum_{i=1}^{I}\lambda_{6,j,t}^{n_i+1}\left\{\frac{\partial b_{i,t}^{n_i+1}}{\partial u_{j,t}^{n_i+1}}\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i}\right\} \qquad (208)\\
&\forall t\geq 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial \mu_t^n}:0 \quad = \quad & -\lambda_{4,t}f_{4,t}^{n} \qquad (209)\\
&+\lambda_{7,t}^{n}\left\{-\sum_{i=1}^{I}\left[\left(\mathbb{I}_{b_{i,t}^{n}>0}-\mathbb{I}_{b_{i,t}^{n}<0}\right)\frac{b_{i,t}^{n}}{\Delta x_i}\right]-\sum_{i=1}^{I}\frac{2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\right\}\\
&+\left\{-\sum_{i=1}^{I}\lambda_{7,t}^{n_i-1}\left[\frac{\mathbb{I}_{b_{i,t}^{n}<0}b_{i,t}^{n}}{\Delta x_i}\right]+\sum_{i=1}^{I}\frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\right\}\\
&+\left\{-\sum_{i=1}^{I}\lambda_{7,t}^{n_i+1}\left[\frac{-\mathbb{I}_{b_{i,t}^{n}>0}b_{i,t}^{n}}{\Delta x_i}\right]+\sum_{i=1}^{I}\frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\right\}\\
&+(\dot{\lambda}_{7t}^{n}-\varrho\lambda_{7,t}^{n}),\\
&\forall t>0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial u_{l,t}^n}:0 \quad = \quad & -\lambda_{4,t}\frac{\partial f_{4,t}^{n}}{\partial u_{l,t}^{n}}\mu_t^n \qquad (210)\\
&+\beta^t\lambda_{5,t}^{n}\left\{\frac{\partial f_{5,t}^{n}}{\partial u_{l,t}^{n}}+\sum_{i=1}^{I}\frac{\partial b_{i,t}^{n}}{\partial u_{l,t}^{n}}\nabla_i\left[v_t^n\right]\right\}\\
&+\sum_{j}\lambda_{6,t}^{n}\left\{\frac{\partial^2 f_{5,t}^{n}}{\partial u_{j,t}^{n}\partial u_{l,t}^{n}}+\sum_{i=1}^{I}\frac{\partial^2 b_{i,t}^{n}}{\partial u_{j,t}^{n}\partial u_{l,t}^{n}}\nabla_i\left[v_t^n\right]\right\}\\
&+\sum_{i=1}^{I}\left(\lambda_{7,t}^{n}-\lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^{n}<0}\frac{\partial b_{i,t}^{n}}{\partial u_{l,t}^{n}}\frac{\mu_t^n}{\Delta x_i}\right]+\sum_{i=1}^{I}\left(\lambda_{7,t}^{n_i+1}-\lambda_{7,t}^{n}\right)\left[\mathbb{I}_{b_{i,t}^{n}>0}\frac{\partial b_{i,t}^{n}}{\partial u_{l,t}^{n}}\frac{\mu_t^n}{\Delta x_i}\right]\\
&\forall t\geq 0
\end{aligned}
$$

By the individual agents' optimality condition, line 2 of this expression is equal to 0.

3.d Discretize time Finally we discretize time. As before, for the state variables $Z^s=\{X,\mu_t(x),\lambda_2,\lambda_5(x)\}$ we replace the time derivative $\dot{Z}_t^s$ by $\frac{Z_t^s-Z_{t-1}^s}{\Delta t}$. For the forward-looking control variables $Z^f=\{U,v_t(x),\lambda_1,\lambda_7(x)\}$ we replace the time derivatives $\dot{Z}_t^f$ by $\frac{Z_{t+1}^f-Z^f}{\Delta t}$. The resulting equations are skipped for brevity, since this transformation is trivial.

4. Compare Finally, by comparing the respective discretized optimality conditions, we show that the two procedures yield the same equilibrium conditions. Consider first the condition for U_t. The optimize-discretize condition is given by (188), which we reproduce here

$$
\begin{aligned}
U_t: \quad 0 = \quad & -\left(\frac{\lambda_{2,t}-\lambda_{2,t-1}}{\Delta} - \varrho\lambda_{2,t}\right) \\
& + \frac{\partial f_0}{\partial U_t} - \lambda_{1,t}\frac{\partial f_1}{\partial U_t} - \lambda_{2,t}\frac{\partial f_2}{\partial U_t} - \lambda_{3,t}\frac{\partial f_3}{\partial U_t} - \lambda_{4,t}\sum_{n=1}^{N}\frac{\partial f_4^n}{\partial U_t}\mu_t^n \\
& + \sum_n \lambda_{5,t}^n\left\{\frac{\partial f_5^n}{\partial U_t} + \sum_{i=1}^{I}\frac{\partial b_i^n}{\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& + \sum_n\sum_{j=1}^{J}\lambda_{6,j,t}^n\left\{\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n\partial U_t} + \sum_{i=1}^{I}\frac{\partial^2 b_t^n}{\partial u_{j,t}^n\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& + \sum_n\left[-\lambda_{7,t}^n\sum_{i=1}^{I}\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n\right]\right] \\
& \forall t \geq 0
\end{aligned}
$$

The respective conditions from the second approach, equation (206) (after discretizing time), is

$$
\begin{aligned}
\frac{\partial L}{\partial U_t}: 0 \quad = \quad & -\left(\frac{\lambda_{2,t}-\lambda_{2,t-1}}{\Delta} - \varrho\lambda_{2,t}\right) \\
& + \frac{\partial f_{0,t}}{\partial U_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial U_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial U_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial U_t} - \lambda_{4,t}\sum_{n=1}^{N}\frac{\partial f_{4,t}^n}{\partial U_t}\mu_t^n \\
& + \sum_{n=1}^{N}\lambda_{5,t}^n\left\{\frac{\partial f_{5,t}^n}{\partial U_t} + \sum_{i=1}^{I}\frac{\partial b_i^n}{\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& + \sum_{n=1}^{N}\sum_{j=1}^{J}\lambda_{6,j,t}^n\left\{\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n\partial U_t} + \frac{\partial^2 b_t^n}{\partial u_{j,t}^n\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& + \sum_n\left\{\sum_{i=1}^{I}\left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial U_t}\frac{\mu_t^n}{\Delta x_i}\right] + \sum_{i=1}^{I}\left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right)\left[\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial U_t}\frac{\mu_t^n}{\Delta x_i}\right]\right\} \\
& \forall t \geq 0
\end{aligned}
$$

The first to fourth lines are evidently identical. The last line also coincides once we take into account the definition of $\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n\right] = \frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\frac{\partial b_{i,t}^{n_i+1}}{\partial U_t}\mu_t^{n_i+1} - \mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n}{\Delta x_i} +$

$\frac{\mathbb{I}_{b_{i,t}^{n}>0}\frac{\partial b_{i,t}^{n}}{\partial U_t}\mu_t^n-\mathbb{I}_{b_{i,t}^{n_i-1}>0}\frac{\partial b_{i,t}^{n_i-1}}{\partial U_t}\mu_t^{n_i-1}}{\Delta x_i}$.

The same argument applies to the optimality conditions with respect to X_t . The optimality conditions with respect to $\hat{U}_t$ and $\tilde{U}_t$ are identical, that is, there is no difference.

Next consider the two discretized optimality conditions with respect to v_t^n (191) and (207). After some rearranging they are given by

$$
\begin{aligned}
v_t(x):\quad 0= \quad & -\sum_{i=1}^{I}\left(\frac{\mathbb{I}_{b_{i,t}^{n}>0}\lambda_{5,j,t}^{n}b_{i,t}^{n}-\mathbb{I}_{b_{i,t}^{n_i-1}>0}\lambda_{5,j,t}^{n_i-1}b_{i,t}^{n_i-1}}{\Delta x_i}+\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\lambda_{5,j,t}^{n_i+1}b_{i,t}^{n_i+1}-\mathbb{I}_{b_{i,t}^{n}<0}\lambda_{5,j,t}^{n}b_{i,t}^{n}}{\Delta x_i}\right)\\
&+\frac{1}{2}\sum_{i=1}^{I}\frac{\left(\sigma_i^2\right)^{n_i+1}\lambda_{5,t}^{n_i+1}+\left(\sigma_i^2\right)^{n_i-1}\lambda_{5,t}^{n_i-1}-2\left(\sigma_i^2\right)^{n}\lambda_{5,t}^{n}}{\left(\Delta x_i\right)^2}\\
&-\sum_{j=1}^{J}\sum_{i=1}^{I}\left(\frac{\mathbb{I}_{b_{i,t}^{n}>0}\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}-\mathbb{I}_{b_{i,t}^{n_i-1}>0}\lambda_{6,j,t}^{n_i-1}\frac{\partial b_{i,t}^{n_i-1}}{\partial u_{j,t}^{n_i-1}}}{\Delta x_i}\right)\\
&-\sum_{j=1}^{J}\sum_{i=1}^{I}\left(\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\lambda_{6,j,t}^{n_i+1}\frac{\partial b_{i,t}^{n_i+1}}{\partial u_{j,t}^{n_i+1}}-\mathbb{I}_{b_{i,t}^{n}<0}\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}}{\Delta x_i}\right)\\
&-\lambda_{5,t}^{n}\rho-\left(\frac{\lambda_{5,t}^{n}-\lambda_{5,t-1}^{n}}{\Delta t}-\varrho\lambda_{5,t}^{n}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial L}{\partial v_t^n}:0 \quad = \quad & \lambda_{5,t}^{n}\left\{\sum_{i=1}^{I}b_{i,t}^{n}\frac{\mathbb{I}_{b_t^n<0}-\mathbb{I}_{b_t^n>0}}{\Delta x_i}-\sum_{i=1}^{I}\frac{2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\right\}\\
&+\sum_{i=1}^{I}\lambda_{5,t}^{n_i-1}b_{i,t}^{n_i-1}\frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i}+\sum_{i=1}^{I}\lambda_{5,t}^{n_i-1}\frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\\
&-\sum_{i=1}^{I}\lambda_{5,t}^{n_i+1}b_{i,t}^{n_i+1}\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i}+\sum_{i=1}^{I}\lambda_{5,t}^{n_i+1}\frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\\
&+\sum_{j=1}^{J}\sum_{i=1}^{I}\left\{\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}\frac{\mathbb{I}_{b_{i,t}^{n}<0}-\mathbb{I}_{b_{i,t}^{n}>0}}{\Delta x_i}+\lambda_{6,j,t}^{n_i-1}\frac{\partial b_{i.t}^{n_i-1}}{\partial u_{j,t}^{n_i-1}}\frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i}-\lambda_{6,j,t}^{n_i+1}\frac{\partial b_{i,t}^{n_i+1}}{\partial u_{j,t}^{n_i+1}}\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i}\right\}\\
&-\lambda_{5,t}^{n}\rho-\left(\frac{\lambda_{5,t}^{n}-\lambda_{5,t-1}^{n}}{\Delta t}-\varrho\lambda_{5,t}^{n}\right)
\end{aligned}
\tag{211}
$$

Again these, two expressions are identical.

Next, consider the two discretized optimality conditions with respect to μ_t^n (192)

and (209). After some rearranging they are given by

$$\begin{aligned}\mu_t(x): \quad 0 = \quad & -\lambda_{4,t} f_{4,t}^n \\ & + \sum_{i=1}^{I} b_{i,t}^n \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}}{\Delta x_i} \right] + \frac{1}{2} \sum_{i=1}^{I} \left(\sigma_i^2\right)^n \frac{\lambda_{7,t}^{n_i+1} + \lambda_{7,t}^{n_i-1} - 2\lambda_{7,t}^n}{2\left(\Delta x_i\right)^2} \\ & + \frac{\lambda_{7,t}^n - \lambda_{7,t-1}^n}{\Delta t} - \varrho \lambda_{7,t}^n \end{aligned} \tag{212}$$

$$\begin{aligned}\frac{\partial L}{\partial \mu_t^n} : 0 \quad = \quad & -\lambda_{4,t} f_{4,t}^n \\ & + \lambda_{7,t}^n \left\{ -\sum_{i=1}^{I} \left[\left(\mathbb{I}_{b_{i,t}^n > 0} - \mathbb{I}_{b_{i,t}^n < 0} \right) \frac{b_{i,t}^n}{\Delta x_i} \right] - \sum_{i=1}^{I} \frac{2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \\ & - \sum_{i=1}^{I} \left[\lambda_{7,t}^{n_i-1} \frac{\mathbb{I}_{b_{i,t}^n < 0} b_{i,t}^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \lambda_{7,t}^{n_i-1} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \\ & + \sum_{i=1}^{I} \left[\lambda_{7,t}^{n_i+1} \frac{\mathbb{I}_{b_{i,t}^n > 0} b_{i,t}^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \lambda_{7,t}^{n_i+1} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \\ & + \frac{\lambda_{7,t}^n - \lambda_{7,t-1}^n}{\Delta t} - \varrho \lambda_{7,t-1}^n, \end{aligned} \tag{213}$$

which again are identical.

Finally, consider the two discretized optimality conditions with respect to $u_{l,t}^n(x)$, (193) and (210). After some rearranging they are given by

$$\begin{aligned} u_{l,t}(x): \quad 0 = \quad & -\lambda_{4,t} \frac{\partial f_4}{\partial u_{l,t}} \mu_t^n \\ + \quad & \sum_{j=1}^{J} \lambda_{6,l,t}^n \left(\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{v_t^{n_i+1} - v_t^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{v_t^n - v_t^{n1}}{\Delta x_i} \right] \right) \\ - \quad & \sum_{i=1}^{I} \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}}{\Delta x_i} \right] \frac{\partial b_{i,t}^n}{\partial u_{l,t}} \mu_t^n \end{aligned} \tag{214}$$

$$\begin{aligned} \frac{\partial L}{\partial u_{l,t}^n} : 0 \quad = \quad & -\lambda_{4,t} \frac{\partial f_{4,t}^n}{\partial u_{l,t}^n} \mu_t^n \\ & + \sum_{j} \lambda_{6,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{v_t^{n_i+1} - v_t^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{v_t^n - v_t^{n_i-1}}{\Delta x_i} \right] \right\} \\ & + \left[\sum_{i=1}^{I} \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1} \right) \left[\mathbb{I}_{b_{i,t}^n < 0} \frac{1}{\Delta x_i} \right] + \sum_{i=1}^{I} \left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n \right) \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{1}{\Delta x_i} \right] \right] \frac{\partial b_{i,t}^n}{\partial u_{l,t}} \mu_t^n, \end{aligned}$$

which are identical.

To summarize, whether one discretize the optimality conditions of the planner and then discretizes them, or one discretizes the planner's problem with respect to the idiosyncratic state and then derives the optimality conditions and finally discretizes time, one arrives to the same set of optimality conditions.