Firm Heterogeneity, Capital Misallocation and Optimal Monetary Policy*

Beatriz González	Galo Nuño	Dominik Thaler	Silvia Albrizio
BdE	BdE	ECB	IMF

First version: October 2020. This version: May 2023

Abstract

Does a decline in interest rates increase capital misallocation? To answer this question, we introduce a New Keynesian model with heterogeneous firms and financial frictions. We show how the response of misallocation depends on the joint dynamics of the real and the natural rate. We also analyze optimal monetary policy. The central bank has a time-inconsistent incentive to engineer an unexpected monetary expansion to increase productivity in the medium run, otherwise the "divine coincidence" holds approximately. Finally, we present empirical evidence supporting the theoretical findings, and highlighting the importance of firm heterogeneity in monetary policy transmission.

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

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

---

*Beatriz González (Banco de España): beatrizgonzalez@bde.es. Galo Nuño (Banco de España and CEMFI): galo.nuno@bde.es. Dominik Thaler (European Central Bank): dominik.thaler@eui.eu. Silvia Albrizio (International Monetary Fund): salbrizio@imf.org.

Two version of this paper were previously circulated as "Monetary Policy and Capital Misallocation" and "Optimal Monetary Policy with Heterogeneous Firms". The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the Bank of Spain or the Eurosystem. We are specially grateful to Andy Atkenson, Anmol Bhandari, Saki Bigio, Jesús Fernández-Villaverde, and Dmitry Khametshin for their suggestions. We thank Klaus Adam, Adrien Auclert, Bence Bardóczy, Paco Buera, Matias Covarrubias, Jim Costain, Wei Cui, Ester Faia, Luis Franjo, Veronica Guerrieri, Masashige Hamano, Agnieszka Markiewicz, Ben Moll, Morten Ravn, Ricardo Reis, Kjetil Storesletten, Gianluca Violante, Thomas Winberry, Chris Wolf, 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 International Monetary Fund, the ECB, Banco de España or the Eurosystem. All errors are our own.

1 Introduction

Recent research (Reis, 2013; Gopinath et al., 2017; Asriyan et al., 2021) shows that a decline in real interest rates may worsen the allocation of capital across firms and thus reduce Total Factor Productivity (TFP).[1] This mechanism has potentially relevant implications for the conduct of monetary policy. First, central banks implement monetary policy by changing nominal interest rates. In the presence of nominal rigidities, changes in nominal rates also affect real interest rates. Does it imply that, when central banks lower interest rates, they are increasing capital misallocation? This seems to be at odds with a large body of literature analyzing responses to monetary policy shocks, which has documented how TFP *increases* after a monetary policy expansion.[2] Second, if central banks can influence capital misallocation through their actions, how does this affect the optimal design of monetary policy?

To answer these two questions, we introduce a framework that combines the workhorse model of monetary policy – the New Keynesian model – with a tractable 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 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. 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. This economy allows for an aggregate representation akin to that in the standard New Keynesian model with capital, except that in this case TFP is endogenous. The dynamics of aggregate TFP are determined by the evolution of

[1] In the presence of financial frictions, the distribution of capital across firms affects aggregate productivity, as documented by the literature on capital misallocation (e.g. Hsieh and Klenow, 2009; Midrigan and Xu, 2014; or Restuccia and Rogerson, 2017).

[2] See Evans (1992), Christiano et al. (2005), Garga and Singh (2021), Jordà et al. (2020), Moran and Queralto (2018), Meier and Reinelt (2020) or Baqaee et al. (2021), among others.

both the distribution of net-worth among entrepreneurs and the productivity threshold above which entrepreneurs are constrained and produce. We calibrate the model to replicate key firm-level moments of Spanish firms.

Turning to our first question, we find that our model resolves the apparent conflict between the findings that low real rates increase misallocation and reduce TFP, while expansionary monetary policy increases TFP. In particular, we show that a shock to the households' time preference that reduces the real interest rate leads to a worsening of misallocation and a drop in TFP, while a monetary policy shock that also reduces the real rate ameliorate misallocation and improves TFP. For the time-preference shock, the drop in TFP is brought about by a increase in the share of less productive entrepreneurs in total investment, similar to Reis (2013), Gopinath et al. (2017), or Asriyan et al. (2021). The opposite happens for the monetary policy shock. This difference in the behavior of misallocation is due to differences in the natural rate dynamics:[3] though in both cases the *real* rate drops, the *natural* rate falls only for the discount factor shock, remaining constant for the monetary policy shock. Just observing the dynamics of real interest rates is not sufficient to infer whether misallocation improves or worsens. It is the joint dynamics of real and natural rates that matter.

We turn then to the second, normative, question. We analyze the Ramsey problem of a benevolent central bank. The steady-state of the Ramsey plan features zero inflation, as in the complete-markets case.[4] We study first optimal policy in the absence of shocks when the initial state coincides with the zero-inflation steady state. Whereas the optimal policy in the case with complete markets is time consistent, financial frictions introduce a new source of *time inconsistency*: the central bank engineers a temporary monetary expansion in the short run while committing to price stability in the long run. This strategy allows the central bank to temporarily increase TFP. That is, the central bank expands aggregate supply by expanding aggregate demand, even if it means tolerating positive inflation during a certain period of time. For our calibration, we find this source of time inconsistency to be quantitatively relevant when measured by the inflation level under the optimal policy: it averages 3% in the first year.

We analyze next the optimal response of monetary policy to shocks from a 'timeless

[3]The natural rate is the rate that would pertain in the absence of nominal rigidities.

[4]Our model nests the complete-market New Keynesian model as a particular case. If the borrowing limit is removed, the most productive firm carries out all the production. In this case the allocation of capital is efficient, TFP is exogenous and the distribution of net worth across firms becomes irrelevant.

perspective' (Woodford, 2003), in which the central bank respects the commitments that it has optimally made at a date far in the past. We consider again a shock that reduces households' discount rate. In this case, the optimal response is price stability (zero inflation). This is the same policy as with complete markets. The "divine coincidence" (Blanchard and Gali, 2007) approximately extends to our model, as the output gap remains roughly zero.

This last result implies that firm heterogeneity and financial frictions do not play a major role in optimal monetary policy design from a timeless perspective. The implementation of this policy, however, differs with respect to the complete-market case. Under incomplete markets, the shock leads to an endogenous fall in TFP through the increase of capital-misallocation. The fall in TFP, 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 nominal rate mimics the natural rate, rates decline more, and more persistently, than in the standard New Keynesian model. This difference in implementation matters when nominal rates are constrained by the zero lower bound (ZLB). The optimal policy in this case, originally proposed by Eggertsson and Woodford (2004), is "low for longer": nominal rates should remain at the ZLB longer than what would be prescribed if the ZLB were not present.[5] In the case with incomplete markets, the larger and more persistent decline in natural rates due to the endogenous fall in TFP leads to what we call a "low for even longer" optimal policy: nominal rates should remain at the ZLB for significantly longer than they should under complete markets.

Finally, we examine empirically the impact of monetary policy on the allocation of capital. Combining micro-level panel data on the quasi-universe of Spanish firms with monetary policy shocks identified using the high-frequency approach of Jarociński and Karadi (2020), we find that firms with a high marginal revenue product of capital (MRPK) increase their investment relatively more than low-MRPK firms in response to an expansionary monetary policy shock, in line with our theory. This implies a shift in the distribution of capital towards high-MPRK firms, improving the capital allocation.[6] This evidence supports the model prediction that capital misallocation

[5]See also Eggertsson et al. (2003); Adam and Billi (2006), and Nakov et al. (2008).

[6]This finding shares some similarities with the results in Bau and Matray (2023), who find that, after a foreign capital liberalization, high-MRPK firms also increase their investment more than low-MRPK firms.

indeed contributes to the previously documented increase of TFP.

This paper contributes to two strands of the literature.

First, an emerging literature analyzes how financial frictions and firm heterogeneity affect the transmission of monetary policy to investment. Researchers have found that firms are 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, 2020), when they are young and do not pay dividends (Cloyne et al., 2018), when their excess bond premia is low (Ferreira et al., 2022) , or when a higher fraction of their debt matures (Jungherr et al., 2022).[7] We contribute to this literature by showing that firms with high-MRPK are more responsive to monetary policy shocks, and we build a model consistent with this finding which demonstrates that this mechanism is relevant for the optimal conduct of monetary policy. [8]

Second, this paper contributes to the literature analyzing optimal monetary policy problems in heterogeneous agent economies. This is a nontrivial endeavor, as the state variables of the central bank include a distribution. A number of recent papers, such as Bhandari et al. (2021), Acharya et al. (2020), Bilbiie and Ragot (2020), Le Grand et al. (2020), Mckay and Wolf (2022), or Auclert et al. (2022) propose different approaches to solve optimal policy problems with heterogeneous agents. In this paper, we follow Nuño and Thomas (2022), and set up the problem as one of infinite-dimensional optimal control. Bigio and Sannikov (2021), Smirnov (2022), and Dávila and Schaab (2022) adopt this approach as well. A technical contribution of our paper is to propose a new, simple and broadly applicable algorithm to solve these kinds of problems, which leverages the computational advantages of continuous time. The key novelty of our algorithm is to discretize the Ramsey planners' continuous-time, continuous-state problem using finite differences (as in Achdou et al., 2021), and then to use symbolic differentiation to obtain the planner's first-order conditions. This produces a high-dimensional nonlinear dynamic system, which is efficiently solved in the sequence space using a Newton

[7] Other papers have analyzed the links between monetary policy and firm heterogeneity through heterogeneity in markups and entry-exit (e.g. Meier and Reinelt, 2020, Bilbiie et al., 2014, Zanetti and Hamano, 2020, Andrés et al., 2021, Nakov and Webber, 2021 or Baqaee et al., 2021), in cyclicality (David and Zeke, 2021), in firm-level productivity trends (Adam and Weber, 2019), or the importance of the price elasticity of investment (Koby and Wolf, 2020).

[8] In complementary empirical work, Albrizio et al. (2023) 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.

algorithm.[9]

The structure of the paper is as follows. Section 2 presents the model and Section 3 analyzes the drivers of misallocation. Section 5 addresses the first question, analyzing the response of TFP to different shocks, whereas Section 6 addresses the second one, analyzing optimal monetary policy. Section 7 provides some supporting empirical evidence. Finally, section 8 concludes.

2 Model

We propose a New Keynesian closed economy model with financial frictions and heterogeneous firms. Time is continuous and there is no aggregate uncertainty. The economy is populated by five types of agents: households, the central bank, entrepreneurs that operate input-good firms, retail, and final goods producers. Entrepreneurs are heterogeneous in their net worth and productivity. They combine capital and labor to produce the input good. The input good is differentiated by imperfectly competitive retail goods producers facing sticky prices, whose output is aggregated by the final good producer. The latter two types of firms are standard in New Keynesian models.

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

[9]The algorithm can be implemented using several available software packages. To make our method accessible to a large audience, we employ Dynare.

[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 can be seen as equivalent from a positive perspective.

[11]For notational simplicity, we use x_t instead of $x(t)$ for the variables depending on time. Further-

(CRS) technology which uses capital k_t and labor l_t to produce a homogeneous input good y_t:

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

The capital share $\alpha \in (0, 1)$ is the same across entrepreneurs. 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 wort (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. Firms sell the input good at the real price $m_t = p_t^y/P_t$, which is the inverse of the gross markup associated to 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 use the return on their activities 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

$$\dot{a}_t = \frac{1}{q_t}\left[\underbrace{m_t f_t(z_t, k_t, l_t) - w_t l_t - R_t k_t}_{\text{Firm's profits}} + \underbrace{(R_t/q_t - \delta)}_{\text{Return on net worth}}\ q_t a_t - \underbrace{d_t}_{\text{Dividends}}\right]. \tag{3}$$

Note that we have rearranged the budget constraint to yield the law of motion of net worth *in units of capital.*

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

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

more, we suppress the input goods firm's index.

[12]The process is bounded, with some reflective barriers.

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

$$V_0(z,a) = \max_{k_t,l_t,d_t} \mathbb{E}_0 \int_0^{\infty} e^{-\eta t}\Lambda_{0,t}\left(\underbrace{d_t}_{\text{Dividends}} + \eta \underbrace{q_t a_t}_{\text{Terminal value}}\right) dt, \tag{5}$$

subject to the budget constraint (3), the collateral constraint (4), and the process followed by productivity (2). 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 a 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 its capital. From now on, we refer to the two groups of entrepreneurs as 'constrained' and 'unconstrained'. That is, firm's demand for capital and labor is:

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

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

Firm's profits are then given by

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

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

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

where this expression just tells us that the MRPK of the marginal firm equals the marginal cost of capital. Also note that factor demands and profits are linear in net worth. A consequence of the CRS technology, this result makes the model significantly more tractable.

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 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* $(R_t - \delta q_t)$, while for the household the return of this unit of capital is *exactly* $(R_t - \delta q_t)$. It is thus always worthwhile for entrepreneurs to keep their funds. 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), the law of motion of an entrepreneur's net worth (3) can thus be rewritten as

$$\dot{a}_t = \frac{1}{q_t}\left[(\gamma \max\{z_t\varphi_t - R_t, 0\} + R_t - \delta q_t)a_t\right]. \tag{11}$$

2.2 Households

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

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

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

where S_t^N is the investment into nominal bonds and T_t are the profits received by the household, which is the sum of the profits of the capital and retail-goods producers (discussed below) and net dividends received from entrepreneurs ($(1-\psi)\eta q_t A_t$).

We assume separable utility of CRRA form, 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 made use of 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), we can define the stochastic discount factor $\Lambda_{0,t}$ as

$$\Lambda_{0,t} \equiv e^{-\int_0^t \rho_t^h ds} \frac{u_c'(C_t)}{u_c'(C_0)} = e^{-\int_0^t r_s ds}.$$

2.3 Final good producers

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

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

where $\varepsilon > 0$ 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. We assume that monopolistic competition occurs at the retail level. Retailers purchase input goods from the input-good firms, differentiate them and sell them to final good producers. Each retailer j chooses the sales price $p_{j,t}$ to maximize profits subject to price adjustment costs as in Rotemberg (1982), taking as given the demand curve $y_{j,t}\left(p_{j,t}\right)$ and the price of input goods, p_t^y. We assume the government pays a proportional constant subsidy $\tau = \frac{\varepsilon}{\varepsilon-1}$ on the input good, so that the net real price for the retailer is $\tilde{m}_t = m_t(1-\tau)$. This subsidy is financed by a lump-sum tax on the retailers Ψ_t. This fiscal scheme is introduced to eliminate the distortions caused by imperfect competition in steady state, as common in the optimal policy literature. 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 $\theta > 0$. 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 producers

A representative capital producer owned by the representative household produces capital and sells it to the household and the firms at price q_t, which she takes as given. Her cost function is given by $\left(\iota_t + \Xi\left(\iota_t\right)\right) K_t$ where ι_t is the investment rate and $\Xi\left(\iota_t\right)$ is a capital adjustment cost function. She maximizes the expected profit stream, discounted at the stochastic discount factor of the household. Profits are paid in a lump-sum fashion to the household.

$$W_t = \max_{\iota_t, K_t} \int_0^\infty \Lambda_{0,t}\left(q_t \iota_t - \iota_t - \Xi\left(\iota_t\right)\right) K_t dt. \tag{22}$$

$$\text{s.t.} \quad \dot{K}_t \quad = \quad \left(\iota_t - \delta\right) K_t. \tag{23}$$

The optimality conditions imply (see Appendix B.4)

$$r_t = \left(\iota_t - \delta\right) + \frac{\dot{q}_t - \Xi''\left(\iota_t\right) \dot{\iota}_t}{q_t - 1 - \Xi'\left(\iota_t\right)} - \frac{q_t \iota_t - \iota_t - \Xi\left(\iota_t\right)}{q_t - 1 - \Xi'\left(\iota_t\right)}. \tag{24}$$

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

$$\frac{\partial g_t(z,a)}{\partial t} = \underbrace{-\frac{\partial}{\partial a}[g_t(z,a)s_t(z)a]}_{\text{Retained earnings}} \underbrace{-\frac{\partial}{\partial z}[g_t(z,a)\mu(z)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[g_t(z,a)\sigma^2(z)]}_{\text{Productivity changing randomly}} \underbrace{-\eta g_t(z,a)}_{\text{Entrepreneurs retiring}} \underbrace{+\frac{\eta}{\psi}g_t(z,\frac{a}{\psi}))}_{\text{Entrepreneurs entering}} \quad , \tag{25}$$

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

$$s_t(z) \equiv \frac{\dot{a}_t}{a_t} = \frac{1}{q_t}(\underbrace{\gamma \max\{z\varphi_t - R_t, 0\}}_{\text{Profit rate from operating the firm}} + R_t - \delta q_t), \tag{26}$$

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 $\omega_t(z) \equiv \frac{1}{A_t}\int_0^\infty a g_t(z,a)da$. 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, which is why we shall work with it. 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 shares is given by (see Appendix B.5)

$$\frac{\partial \omega_t(z)}{\partial t} = \left[s_t(z) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta\right]\omega_t(z) - \frac{\partial}{\partial z}\mu(z)\omega_t(z) + \frac{1}{2}\frac{\partial^2}{\partial z^2}\sigma^2(z)\omega_t(z). \tag{27}$$

2.7 Market Clearing and Aggregation

Market clearing. Define aggregate capital used in production as $K_t = \int k_t(z,a)dG_t(z,a)$, aggregate firm net worth as $A_t = \int a_t dG_t(z,a)$, and 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

$$K_t = A_t + B_t, \tag{28}$$

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

$$B_t = D_t. \tag{29}$$

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 (28), (29), aggregating capital used by firms (7), and solving for A_t, we obtain

$$A_t = \frac{D_t}{\gamma(1-\Omega_t(z_t^*)) - 1}. \tag{30}$$

Labor market clearing implies

$$L_t = \int_0^\infty l_t(z,a)dG_t(z,a). \tag{31}$$

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

$$Y_t = Z_t K_t^\alpha {L_t}^{1-\alpha}, \tag{32}$$

where aggregate TFP Z_t is an endogenous variable given by

$$Z_t = \left(\mathbb{E}_{\omega_t(\cdot)}\left[z \mid z > z_t^*\right]\right)^\alpha = \left(\frac{\int_{z_t^*}^\infty x\omega_t(x)\,dx}{1-\Omega_t(z_t^*)}\right)^\alpha. \tag{33}$$

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

$$w_t = (1-\alpha) m_t Z_t K_t^\alpha {L_t}^{-\alpha}, \tag{34}$$

$$R_t = \alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} \frac{z_t^*}{\mathbb{E}_{\omega_t(\cdot)}\left[z \mid z > z_t^*\right]}. \tag{35}$$

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

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}\left[\gamma(1-\Omega_t(z_t^*))\left(\alpha m_t Z_t K_t^{\alpha-1} {L_t}^{1-\alpha} - R_t\right) + R_t - \delta q_t - q_t(1-\psi)\eta)\right]. \tag{36}$$

Appendix B.6 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 5 we assume that the central bank sets the nominal rate according to a standard Taylor rule of the form

$$di = -\upsilon\left(i_t - \left(\rho^h + \phi\left(\pi_t - \bar{\pi}\right) + \bar{\pi}\right)\right)dt, \tag{37}$$

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 6 we assume that the central bank implements the Ramsey-optimal policy.

2.9 Discussion of some of the key assumptions

As in every model, we are constrained to make some particular assumptions. Here we discuss some of them and the potential implications they may have for the results presented in the rest of the paper.

Constant returns to scale. As discussed by Moll (2014), the assumption of constant return to scale (CRS) in firms' production function (1) can be seen as the limiting case of decreasing returns to scale (DRS), $y_t = \left[(z_t k_t)^{\alpha}(l_t)^{1-\alpha}\right]^{\nu}$, $\nu < 1$, 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 its 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.[13]

[13] Ferreira et al., 2023 find that financially constrained firms are found across the entire firm-size

Free mobility of capital. In the model, there are no firm-level capital adjustment costs. Furthermore, due to our CRS assumption, changes in z^* imply that firms at the margin disinvest or reinvest fully instantaneously. In line with the previous point, in a DRS model the changes in the capital stock of a firm that switches from being constrained to being unconstrained and thus crossing the threshold $z^*(a)$ would be smaller. This is so since the optimal capital stock of unconstrained firms would be positive (and not zero, as in the limiting CRS case). These smaller changes in capital can be archived by reductions in the gross investment rate, without requiring the reselling of capital, provided the depreciation rate is high enough.[14]

No balance-sheet effects. In our continuous time model, debt contracts are instantaneous and entrepreneurs borrow capital directly. Firms balance sheets are hence not exposed to Fisherian debt deflation or financial accelerator effects (Bernanke et al., 1999). Asriyan et al. (2021) include the latter. This assumption keeps the model tractable. Allowing for such effects would reinforce the impact of the shocks discussed next.

Separation between heterogeneous input-good firms and retailers. Finally, we distinguish between heterogeneous input-good firms and retailers. This is standard practice in previous New Keynesian models featuring firm heterogeneity and nominal rigidities, such as Ottonello and Winberry (2020) or Jeenas (2020). Besides greater tractability, it avoids the possibly implausible countercyclical behaviour of New Keynesian markups interfering with our mechanism, which we see as an important advantage. Baqaee et al. (2021) consider a model without financial frictions in which monopolistically competitive firms are heterogeneous in productivity and nominal rigidities. They find that expansionary monetary policy shocks increase TFP, a result similar to the one we find in the Section 5 below.

3 Equilibrium prices and misallocation

As section 2 highlighted, misallocation is endogenous and evolves over time. Misallocation is driven by the investment dynamics within the heterogeneous input-goods

distribution.

[14] Cui et al. (2022) study economies where firms acquire capital in primary markets then retrade it in secondary markets that incorporate bilateral trade with search, bargaining and liquidity frictions.

firms block of the model, which in turn depend on the factor prices determined in general equilibrium. In this section, we explore the theoretical mechanism through which changes in factor prices may affect misallocation.

As discussed above, by equation (33) which we reproduce here, TFP depends on the allocation of capital across entrepreneurs, which is proportional to the net worth of constrained entrepreneurs:

$$Z_t = \left(\mathbb{E}_{\omega_t(\cdot)}\left[z \mid z > z_t^*\right]\right)^{\alpha}, \tag{38}$$

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 3). Entrepreneurs below the cut-off z^* are unconstrained, operate at their optimal size $k(z) = 0$, and lend their net worth to constrained entrepreneurs above the cut-off. Equation (38) 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 analyzing the case in which the dynamics of TFP are driven purely by changes in the net worth distribution, which happens when the cut-off z_t^* is constant or responds very little to monetary policy. In this case, the 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

$$\tilde{\Phi}_t(z) \equiv \frac{\Phi_t}{q_t a_t} \quad = \quad \max\left\{\frac{\gamma\varphi_t}{q_t}\left(z - z_t^*\right), 0\right\}, \tag{39}$$

where $\tilde{\Phi}_t(z)$ is the return on equity 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 $z \leq z_t^*$.[15]

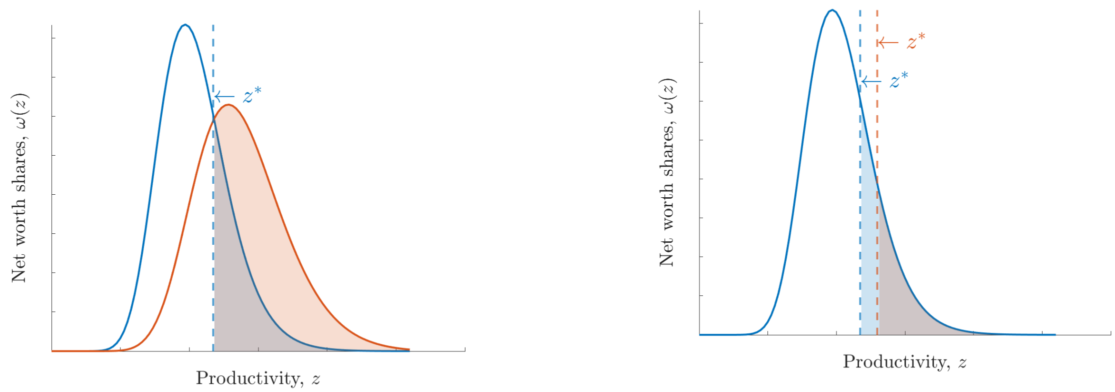
Proposition 1. ***(TFP response to the slope)***. *Conditional on a constant cutoff* z^*,

[15]It is also a measure of how constrained a firm is, since this expression is the Lagrange multiplier of the borrowing constraint in the firms' maximization problem. From the first order conditions of the firm, we get that $MRPK_t = R_t + q_t\lambda^{BC}$, where λ^{BC} is the multiplier on the borrowing constraint. Hence, $\lambda^{BC} = \tilde{\Phi}_t(z)$.

Figure 1: Changes in aggregate TFP in response to a changes in equilibrium prices.

(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.

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:

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

All the proofs in the section can be found in Appendix B.8. This proposition states that the slope $\frac{\gamma\varphi_t}{q_t}$ determines how, conditional on a constant cut-off z^*, the net-worth share distribution moves, and hence in which direction TFP evolves. $\frac{\gamma\varphi_t}{q_t}$ is a function of prices. If the slope increases, then high-productivity firms grow faster than low-productivity firms, the net worth distribution shifts rightwards, the allocation of capital improves, and TFP increases, as represented in Panel a of Figure 3. 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 other polar case, namely when the distribution remains constant and the cut-off changes in response to price changes. This happens in the limit of iid shocks, that is, the limit as $\varsigma_z\to\infty$, as discussed in Itskhoki and Moll (2019). In this case the net worth distribution $\omega(\cdot)$ is constant, and the response of TFP growth depends exclusively on the changes in 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:*

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

This proposition implies that a change in prices that raises the threshold $z_t^*=R_t/\varphi_t$ increases TFP. Panel (b) in Figure 3 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 firms 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 mechanism: 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.

Summing up, if changes in equilibrium prices are such that the slope of the excess investment function and/or the threshold increase, then high-MRPK firms increase their share of the total capital stock, while low-MRPK firms reduce it, such that TFP increases. To determine those general equilibrium price movements, we next turn to numerical analysis.

4 Numerical solution and calibration

Numerical algorithm. We solve the model numerically using a new method, described in Appendix D. 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. This can be easily implemented using Dynare's perfect foresight solver.[16]

Our solution approach is different from the one in Winberry (2018) or Ahn et al. (2018). These papers analyze heterogeneous-agent models with aggregate shocks building on the seminal contribution by Reiter (2009). To this end, they linearize the model around the deterministic 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. Here, instead, we compute the *nonlinear* transitional dynamics in the the sequence space, as Boppart et al. (2018) or Auclert et al. (2021). Boppart et al. (2018) show how the perfect-foresight transitional dynamics to a (small) MIT shock, such as the ones we compute here, coincide with the impulse responses obtained by a first-order perturbation approach in the model with aggregate uncertainty.

Our method solves for the same approximate solution as the nonlinear version of the sequence space Jacobian approach by Auclert et al. (2021). An important technical difference is that we solve simultaneously for all variables (prices, aggregates and distributions) in a single loop without decomposing the model in blocks.

Calibration. Table 1 summarizes our calibration. We calibrate the parameters of the heterogeneous firms block to match data on Spanish firms, described in Appendix A.1. 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).[17] The fraction of assets of exiting entrepreneurs reinvested (ψ) is 0.1, so that new entrant's account for 1 percent of the total capital stock, in line with the Spanish firm level dataset described in section 7. 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-Uhlenbeck process in logs,[18] with a

[16]Notice that the variables of the model include the distribution $\omega(z)$, which is an infinite-dimensional object. The finite-difference discretization turns this continuous variable into a finite dimensional vector.

[17]Specifically, the data comes from the *Directorio Central de Empresas*, which is a dataset maintained by INE, and it contains aggregate data on all firms operating in Spain, and its status (incumbent, entrant or exiter). The dataset can be accessed here.

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

reflective lower (upper) barriers at some value close to 0 (very high value).[19]

$$d\log(z) = -\varsigma_z \log(z)dt + \sigma_z dW_t. \tag{40}$$

We estimate this process using our firm panel data set as explained in Appendix A.4. The estimate for ς_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
ς_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 factor	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	Intertemporal elasticity of substitution HH	1	Log utility in consumption
ϑ	Inverse Frisch Elasticity	1	Kaplan et al. (2018)
Υ	Constant in disutility of labor	0.71	Normalization $L = 1$
ϕ^k	Capital adjustment costs	8	VAR evidence Christiano et al. (2016)
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 such that aggregate labor supply in

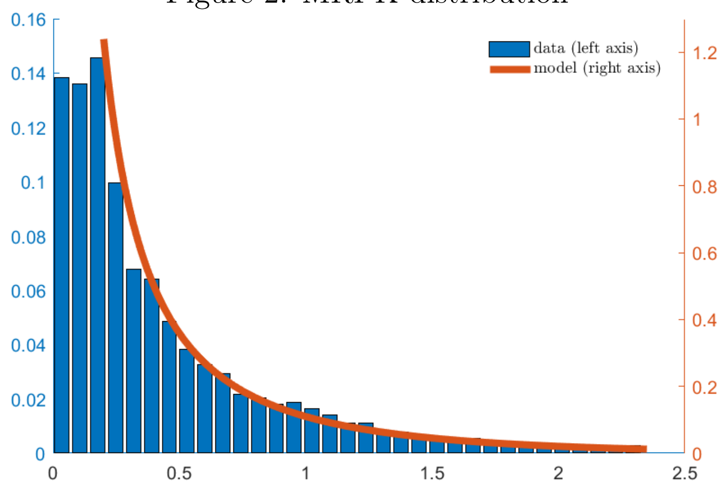
[19] We truncate the process for z at 48. This corresponds to truncating the MRPK distribution at the same level as in the data.

steady state is normalized to one. We assume capital adjustment costs are quadratic, i.e. $\Xi(\iota_t) = \frac{\phi^k}{2}(\iota_t - \delta)^2$, and set ϕ^k to 8, such that the peak response of investment to output after a monetary policy shock is around 2, in line with the VAR evidence of Christiano et al. (2016).

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 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 $\upsilon = 0.2$.

The model generates the steady state distribution shown in Figure 4, where blue bars show the capital-weighted MRPK distribution in the data, and the orange line shows the model-generated MRPK distribution. The model does a good job at replicating the MRPK distribution.

Figure 2: MRPK distribution



Notes: The figure shows the steady state distribution of firms MRPK in the model (orange solid line) and compares it to the data (histogram with blue bars). See section 7 for more details on the data. Note that in both model and data the MRPK is proxied by y_t/k_t. In the model, this proxy is equal to the MRPK divided by α. We drop observations above an MRPK of 2.3, 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 ($R = r + \delta$ in steady state).

5 On the relationship of interest rates and capital-misallocation

In this Section, we address the first of the two questions raised at the beginning of this paper: is a fall in interest rates necessarily associated to an increase in misallocation? To this end, we compute the impulse responses to different shocks that trigger a reduction in real rates, and analyze the dynamics of misallocation in general equilibrium.

Households' time preference shock without nominal rigidities. We first analyze the response to a time preference shock, that is a temporary fall in the household's discount factor ρ^h of 1 b.p.., abstracting for the moment from nominal rigidities (dotted yellow line in Figure 5). This case leads to a reduction in real interest rates (panel d) and a persistent decline in TFP (panel g).

The drop in TFP results from a deterioration of the capital allocation due to two effects. First, the distribution of entrepreneurial net worth shifts to the left, because the slope of the excess investment rate $\frac{\gamma \varphi_t}{q_t}$ decreases (panel i). This decrease is largely brought about by the increase in capital prices (panel e), since the MRPK of a firm with productivity one, $\varphi_t = \alpha \left(\frac{(1-\alpha)}{w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}}$, is largely unaffected by the shock, which barely moves wages and input good prices (panels b and c). Second, the cut-off $z_t^* = \frac{R_t}{\varphi_t}$ (eq. 10) also falls, as the rental rate, $R_t = r_t q_t + \delta q_t - \dot{q}_t$ declines with the real interest rate and the price of capital. In sum, the decline in real rates and the increase in the price of capital increase low-MRPK firms' share of the total capital stock, leading to a decrease in TFP. [20]

This result echoes the findings of Reis (2013), Gopinath et al. (2017), or Asriyan et al. (2021).

Monetary policy shock. Now we turn to the baseline model with nominal rigidities and analyze a monetary policy shock, that is, a 1 b.p. reduction in the nominal rate for the baseline economy (dashed red line). The decline in the nominal rate (not shown) leads to a fall in the real rate, and an increase in output and inflation (panels a, d, f) through the standard new Keynesian mechanism. 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 as the MRPK of a firm with productivity one, φ_t (not shown), now

[20]The responses to a permanent time preference shock are qualitatively similar. See Appendix B.10.

rises by more than the price of capital, q_t (panel e).[21] Second, and maybe somewhat surprisingly, the cut-off z_t^* also increases (panel h), because the increase in the rental rate R_t (not shown) overcompensates the increase in the MRPK of a firm with productivity one. The increase in the rental rate is a consequence of the dynamics of the price of capital. In sum, given the dynamics of the slope and the cutoff, the share of high-MRPK firms increases and so does TFP (panel g).

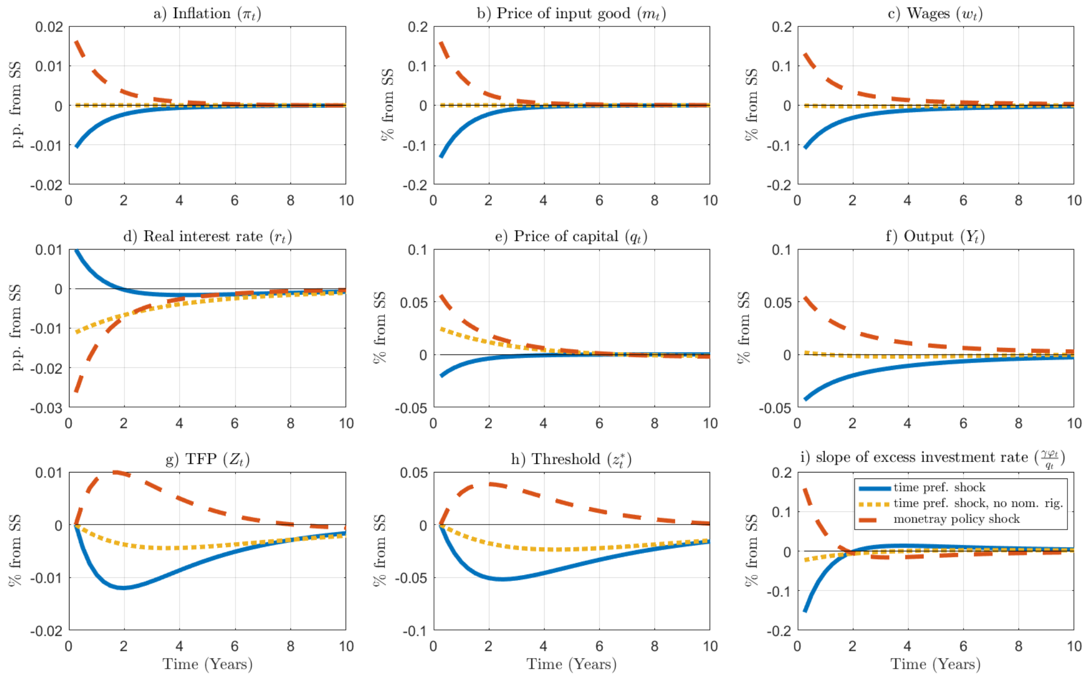
The fact that expansionary monetary policy shocks raise TFP has been widely documented (see Evans, 1992; Christiano et al., 2005; Garga and Singh, 2021; Jordà et al., 2020; Moran and Queralto, 2018; Meier and Reinelt, 2020 or Baqaee et al., 2021). Our results suggest that the improvement in capital allocation may explain a part of this effect, complementing the alternative mechanisms proposed by these authors such as R&D, hysteresis effects, or markup heterogeneity. The peak effect on TFP predicted by the model (0.87 p.p.) 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. We discuss further the empirical plausibility of these results in Section 7.

Households' time preference shock with nominal rigidities. With nominal rigidities, the response to the time preference shock is basically a combination of the response to the time preference shock absent price rigidities and a negative monetary policy shock (solid blue line). This is so because, under a Taylor rule, the central bank does not perfectly track the natural rate, which is defined as the real interest rate in the counterfactual economy without nominal rigidities (dotted yellow line in panel d). In particular, the central bank does not reduce nominal rates fast enough such that the real interest rate exceeds the natural rate (panel b), which has contractionary effects. The drop in TFP hence is larger than in the absence of price rigidities and it is brought about by a combination of the effects just described for the two shocks (with an inverted sign for the monetary shock).

Explaining the difference: Having understood how prices determine TFP, one may wonder why prices move so differently depending on the source of the shock, such that the TFP response to the two real-rate-reducing shocks is of different sign.

[21] Input-good prices and wages affect φ_t in opposite directions, as the former increases excess profits whereas the latter reduces 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 the MRPK φ_t increases.

Figure 3: 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 a time preference shock of 1 basis point. The dotted yellow line is the response of the economy to the same time preference shock in the absence of nominal rigidities.The orange dotted line is the response to a monetary policy shock .

Since prices are determined in general equilibrium, an analytical characterization is not possible, but we can offer some intuition.

First note that, the movements in prices are similar to those in the complete-market version of our model, that is in the standard representative agent New Keynesian model (RANK).[22] As a result, the slope of the excess investment rate $\frac{\gamma\varphi_t}{q_t}$ behaves very much as it would if prices were determined in the RANK. That is, the general equilibrium responses present in the RANK model explain the dependence of the dynamics of $\frac{\gamma\varphi_t}{q_t}$ on the source of the shock and drive the dynamics of TFP to a significant extent.

However, the same is not true for the cut-off $z_t^* = \frac{R_t}{\varphi_t}$. In RANK, the rental rate R_t would equal the unit-productivity MRPK φ_t, so that z_t^* would be constant. To

[22] The complete-market economy is the standard representative agent New Keynesian model with capital. It represents a special case of the baseline economy where the borrowing constraint is infinitely loose, so that the productivity/net-worth distribution becomes irrelevant and only the most productive entrepreneur operates. In this case, capital allocation is efficient (no misallocation) and TFP is exogenous. Appendix B.9 compares the baseline and complete-market models.

understand the dynamics of the cut-off z_t^*, it is instructive to consider some of the general equilibrium relationships that determine the prices that pin down the cut-off. Market clearing in the goods market by (30) requires

$$A_t \gamma \int_{z_t^*}^{\infty} \omega_t(x)\, dx = D_t + A_t. \tag{41}$$

That is, the threshold z_t^* has to adjust to equate capital supply and demand. Since capital holdings are a state variable, it is predetermined, and hence only adjusts gradually.

Consider the time preference shock in the absence of price rigidities. In this case the shock directly increases the savings of the household D_t. To absorb this additional capital supply, and given that high-productivity firms are at their borrowing constraint, previously unconstrained entrepreneurs need to employ more capital, i.e. z_t^* has to drop. This requires a drop in the real rate, which coincides with the natural rate in this case.

Now consider the monetary policy shock. The central bank lowers nominal rates below the natural rate, which remains constant. As a result, inflation picks up, which not only lowers the real rate further, but also erodes retailers prices. The latter increases the input-firms' profit opportunities such that the demand of capital by constrained high-productivity firms, $\gamma A_t \int_{z_t^*}^{\infty} \omega_t(x)\, dx$, grows faster than the combined capital holdings of the households, D_t, and the unconstrained low-productivity entrepreneurs. Since the capacity of constrained entrepreneurs to absorb capital increases faster that the combined supply of capital from unconstrained entrepreneurs and households, the threshold, z_t^*, must increase to clear the capital market.

Our results solve the apparent tension between the misallocation literature and the estimates of the effect of monetary policy shocks on TFP. Just looking at the real interest rate is not enough to derive any meaningful prescription about the future dynamics of misallocation: what matters is the joint dynamics of the real interest rate and the natural rate. In models without nominal rigidities, such as Gopinath et al. (2017), or Asriyan et al. (2021), the real rate always coincide with the natural one. These models are suitable to analyze medium-term dynamics, in which nominal rigidities become less relevant.

Having answered the first question, we turn next to the second one on how endogenous capital misallocation affects optimal monetary policy design.

6 Optimal monetary policy

6.1 Central bank objective and numerical approach

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:

$$\max_{\{\omega(z),w,r,q,\varphi,R,K,A,L,C,D,Z,\mathbb{E}[z|z>z_t^*],\Omega,z^*,\iota,\pi,m,\tilde{m},i,Y,T\}_{t\geq 0}} \mathbb{E}_0 \int_0^\infty e^{-\rho^h t} u(C_t, L_t) dt \tag{42}$$

subject to the all the private equilibrium conditions derived above and listed in Appendix B.7 and the initial conditions $\{\omega_0(z), K_0, D_0, A_0\}$. Relative to the standard New Keynesian model – i.e. the complete market version of our model – the problem of the central bank is richer by one dimension: the central bank understands that its policy affects TFP through the misallocation channel, and has to account for that.

Algorithm. This additional richness also makes the problem harder to solve computationally. The central bank's controls include the net-worth distribution $\omega_t(z)$, as the central bank internalizes the impact of its decisions on it. Notice that the density $\omega_t(z)$ not only depends on time, but also on individual productivity. This poses a challenge when solving optimal monetary policy, as we need to compute the first order conditions (FOCs) with respect to this 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. (2020) 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), Bigio and Sannikov (2021), Smirnov (2022) and Dávila and Schaab (2022) 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).

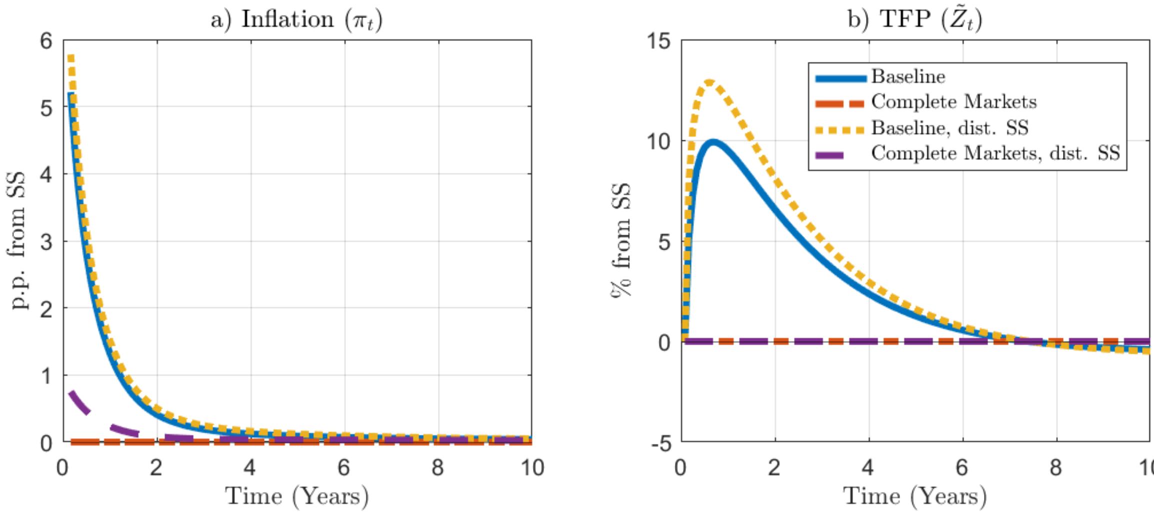
Here we propose a new algorithm, detailed in Appendix E. Instead of determining the FOCs for the planner's continuous space problem, we first discretize the planner's objective and constraints (the private equilibrium conditions) using finite differences. This transforms the original infinite-dimensional problem into a high-dimensional 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 (200 in our application) used to approximate the individual state space. Second, we find the planner's FOCs *by symbolic differentiation*. This delivers a large-dimensional system of difference equations. Third, we find the Ramsey steady state by solving this system at steady state. To do so, we compute the steady state of the model conditional on the steady-state level of the policy instrument with a conventional iterative method, and then use this function to find the Ramsey steady state using the Newton method. Fourth, we solve the system of difference equations non-linearly in the sequence space using the Newton method, as already described in Section 4 and Appendix D. The symbolic differentiation and the two applications of the Newton algorithm can conveniently be automated using several available software packages. In our case, we employ Dynare. This algorithm can be employed to compute optimal policies in a large class of heterogeneous agent models. Compared to other techniques, it stands out for being extremely easy to implement. In appendix E we present a proposition showing that our algorithm delivers the same results as computing the FOCs by hand using calculus of variations and then discretizing the model, as the time step gets smaller. We also apply the algorithm to solve the model in Nuño and Thomas (2022) in order to illustrate its generality.

6.2 Optimal Ramsey policy

Steady state. Let us focus first on the steady state of the Ramsey problem. It is well known that in the standard (complete-market) New Keynesian economy without steady state distortions inflation is zero in the Ramsey steady state. Due to capital misallocation, our baseline (incomplete-market) economy does not feature steady state efficiency. Yet, inflation is still zero in the steady state of the Ramsey problem.[23] This

[23] This is a numerical result. We have consider a wide range of parameters and optimal steady-state zero inflation holds.

Figure 4: Time 0 optimal monetary policy.



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 market 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.

result mirrors a similar result from the textbook New Keynesian model with a distorted steady state (Woodford, 2003; Gali, 2008). Though the long-run Phillips curve allows monetary policy to affect misallocation in the long run through positive trend inflation, the benefits of this policy are compensated for by the cost of the anticipation of this policy.

Time-0 optimal policy. We turn next to the deterministic dynamics under the Ramsey optimal plan. We solve for the Ramsey plan when the initial state of the economy coincides with the steady state under the optimal policy, i.e., that with zero inflation. The Ramsey planner faces no pre-commitments. This is commonly referred to as the "time-0 optimal policy" (Woodford, 2003).

We compare our baseline incomplete-market economy with a complete-market economy. The Ramsey plan in the model with complete markets is time-consistent. Hence, inflation and the rest of variables remain constant at their steady state values. This is displayed by the dashed red lines in Figure 4. Market incompleteness, however, introduces a new motive for time inconsistency, inducing the central bank to temporally deviate from the zero-inflation policy. The solid blue lines in Figure 4 show how the central bank engineers a sizable surprise monetary expansion, increasing inflation (panel a). The resulting dynamics are similar to those caused by an expansionary monetary policy shock, which were described in detail in Section 5: the change in factor prices increases TFP (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.

It is well known that the Ramsey policy in the complete market economy with a steady-state mark-up distortion also features inflationary time inconsistency. Comparing the optimal policy above with the optimal policy when there is no subsidy to correct for the mark-up distortion reveals that the time inconsistency problem caused by the incomplete market distortion is much larger: the optimal inflation level due to market incompleteness is more than six times higher than that due to the mark-up (dashed purple line). We hence conclude that the time inconsistency problem is not only large in absolute terms (with an average inflation of 3% during the first year), but also dwarfs the one resulting form markups in the standard New Keynesian model.

The desire of the central bank to redistribute resources towards 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. In our case, and given the lack of fiscal instruments, it is the central bank who engineers this redistribution through an expansion in aggregate demand.

6.3 Timeless optimal policy response

Next, we analyze the optimal policy response when an unexpected shock hits the economy that was previously in its zero-inflation steady state. In this case, we adopt a "timeless perspective" (Woodford, 2003, Gali, 2008). Timelessly optimal Ramsey policy implies that the central bank sticks to pre-commitments, implementing the policy that it would have chosen to implement if it had been optimizing from a time period far in the past.[24] As discussed in Section 4, building on the argument by Boppart et al. (2018) one can reinterpret the timeless response to MIT shocks as a first-order approximation to the response in a model with aggregate uncertainty under the, ex-ante optimal, time-invariant state-contingent policy rule.

Households' discount factor shock. We analyze the optimal response to a

[24]The Lagrange multipliers associated to forward-looking equations in the planner's FOCs in this case are initially set to their steady state values.

households' discount factor shock. Figure 5 shows that the optimal response in the baseline economy (blue solid line) mimics that under complete markets (orange dashed line). The response both with complete and incomplete markets is characterized by what Blanchard and Gali (2007) described as "divine coincidence": the optimal response by the central bank is to stabilize inflation at its steady state value of zero (panel a), which also keeps the output gap at its optimal steady state value of zero (panel b).[25]

Notice how, compared to the dynamics under the Taylor rule, the optimal policy is able to deliver price stability and a smaller decline in TFP. In fact, dynamics under the optimal policy reproduce the case without nominal rigidities discussed above. This result again reinforces the message that we need to jointly account for real and natural rates to understand misallocation dynamics.

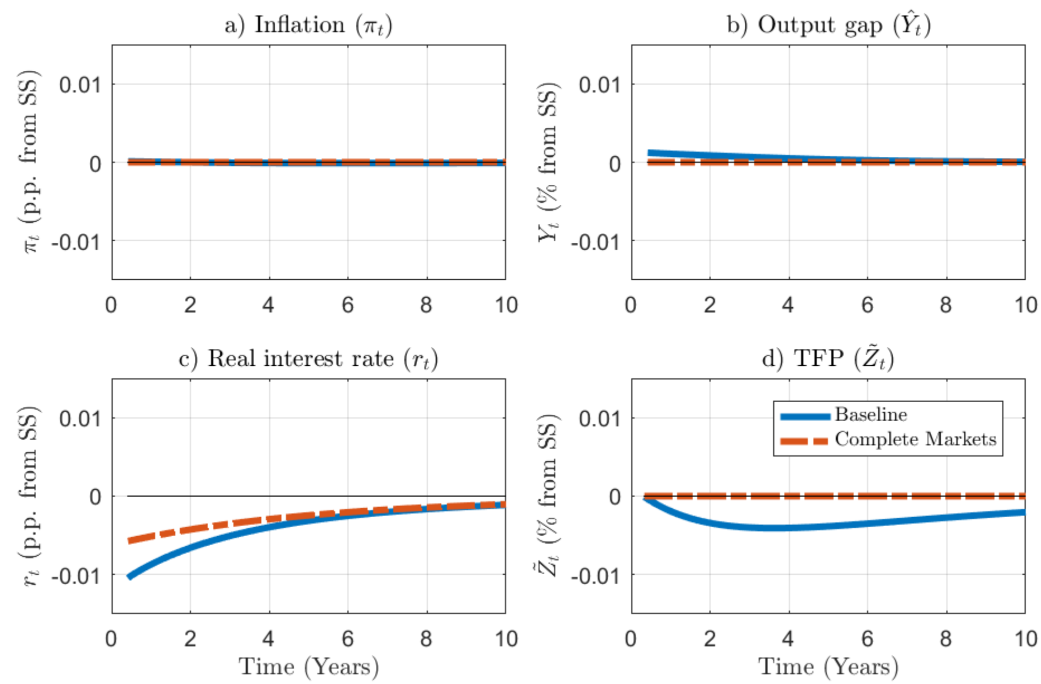
In order to implement the optimal policy, the central bank should lower the real and nominal rates. However, in the baseline with incomplete markets, this requires that the central bank acts more forcefully than under complete markets (panel c). The reason is that the original demand shock leads to a negative 'supply shock' through its impact on aggregate TFP (panel d), which depresses output and natural rates relative to the complete market case. As the central bank has to replicate the path of natural rates, it is forced to reduce real rates more persistently than with complete markets.

Zero lower bound. The fact that the central bank responds more persistently to the demand shock under the optimal policy has important implications when the zero lower bound constrains its room for maneuver. Figure 6 displays the optimal response from a timeless perspective to a large negative demand shock that drives the natural rate below the zero lower bound (ZLB). The optimal policy under complete markets, as shown by Eggertsson et al. (2003), is to adopt a "low for longer" strategy: The nominal rate (dotted yellow line) should remain at the ZLB for a longer period than it would in the absence of the ZLB (dashed orange line).

In the baseline economy with incomplete markets, optimal policy is also characterized by a low for longer strategy (dotted light blue line). However, the lift-off date is now delayed more than four times as much as with complete markets. We call this a "low for *even* longer" policy. The reason is simple. As discussed above, natural rates fall more persistently in the case with incomplete markets, and so do nominal rates under the optimal policy without the ZLB (solid blue line). To compensate for the inability

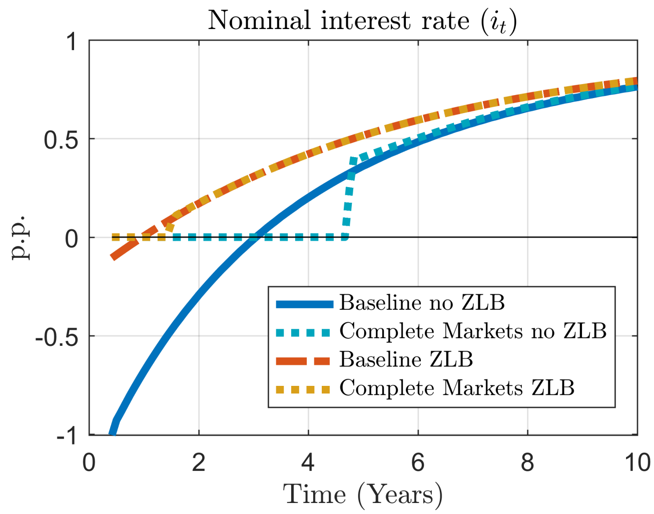
[25]The divine coincidence holds only in a approximate sense, but the deviation is negligibly small.

Figure 5: Optimal monetary policy response to a households' discount factor shock.



Notes: The figure shows the optimal response from a timeless perspective (in deviations from steady state) to a 1 b.p. decrease in the rate of time preference of the household ρ^h that is mean reverting with a yearly persistence of 0.8. The baseline economy is the solid blue line, and the complete market economy the dashed orange line. The output gap is defined as the difference between observed output and the counterfactual output in an economy without nominal rigidities.

Figure 6: Optimal monetary policy response to a demand shock when the zero lower bound is binding.



Notes: The figure shows the optimal response from a timeless perspective (in deviations from steady state) to a 4 p.p. decrease in the rate of time preference of the household ρ^h that is mean reverting with a yearly persistence of 0.8. The baseline economy without the zero lower bound is the solid blue line, and the complete market economy without the zero lower bound is the dashed orange line. The dotted light blue line is the optimal response in the baseline economy with the zero lower bound. The yellow dotted line is the optimal response in the complete market economy with the zero lower bound.

to move rates into negative territory, the central bank commits to stay low for even longer.

The answer to our second question thus crucially depends on the existence of pre-commitments. If there are none, surprising agents by engineering a monetary expansion is optimal, as it increases TFP at the cost of a temporary rise in inflation. If instead the central bank is bound by previous commitments, price stability remains as the best option, as in the standard case with complete markets.

7 Supporting evidence

A key insight of our theory is that a expansionary monetary policy shock increases TFP by reducing misallocation. Albeit we view the main contribution of the paper as being conceptual, it is nonetheless important to highlight that the theory is consistent with several pieces of empirical evidence. One is the estimated response of TFP to monetary policy shocks, already discussed in Section 5. This evidence, however, is silent about the particular mechanism: in our model the allocation of capital improves as firms with a high MRPK expand their share of the aggregate capital stock. In this Section we complement existing evidence first by testing this mechanism on Spanish micro data, and second by linking our model directly to the data.

Empirical strategy. For our empirical analysis we combine granular Spanish firm-level panel data with Jarociński and Karadi (2020) exogenous monetary policy shocks. We use yearly balance-sheet and cash-flow data from the quasi-universe of Spanish firms from 2000 to 2016 from the *Central de Balances Integrada*. The main advantage of this dataset is that it covers the quasi-universe of Spanish firms, including 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). These are generally large firms with access to the equity market, which can potentially behave very differently from the rest of firms in the economy, as documented for example by Caglio et al. (2021). Our key variable of interest is firm's MRPK, which we proxy by value added over capital.[26] Appendix A.1 details the data definition and the cleaning process. The monetary policy shock ε_t is taken from Jarociński and Karadi (2020), who use sign restrictions to decompose unexpected high frequency movements of interest rates around policy announcements into an information surprise and a monetary policy surprise component. We use the latter component, and we aggregate these shocks to yearly frequency following the methodology employed by Ottonello and Winberry (2020). Appendix A.2 provides more details on the identification and aggregation of the monetary policy shock.

To analyze the impact of an unexpected monetary policy expansion on a firm's capital growth as a function of its MRPK, we estimate the following specification:

$$\log k_{j,t} - \log k_{j,t-1} = \beta_0 + \beta_1 \log\left(MRPK_{j,t-1}\right) + \beta_2 \log\left(MRPK_{j,t-1}\right)\varepsilon_t + \gamma_{s,t} + u_{j,t}, \quad (43)$$

where $k_{j,t}$ is the tangible capital of an individual firm j at time t, $MRPK_{j,t-1}$ its lagged MRPK, and $\log\left(MRPK_{j,t-1}\right)\varepsilon_t$ is the interaction between firm's lagged MRPK and the monetary policy shock. $\gamma_{s,t}$ is a vector of sector-year fixed effects, which control for

[26]In a Cobb-Douglas production framework, such as the one presented in Section 2, under the assumption that firms have the same time-invariant input share in production, the MRPK is proportional to the the average revenue product of capital (ARPK). This is why we use ARPK as an empirical measure for MRPK, following the literature (see for instance Bau and Matray, 2023).

potential confounders that are year- and sector-specific and absorb the average effect of the monetary policy shock. Standard errors are clustered at the sector and year level. The main coefficient of interest is β_2: a positive value indicates that high-MRPK invest relative more than low-MRPK firms after an expansionary monetary policy surprise.

Table 2 reports the estimates of our coefficient of interest β_2. Column 1 shows that after a 1 p.p. expansionary monetary policy shock, a firm with an MRPK that is 1% higher than that of another firm increases its capital stock by 0.028 p.p. more. Since our panel is highly unbalanced, in Column 2 we perform the same regression, but restricting the sample to firms that we observe for more than five consecutive years, which is the effective sub-sample that we use in Section 5. In this latter case, the coefficient is equal to 0.044.

Table 2: Response of firm-level investment to an expansionary monetary policy shock

	(1)	(2)
β_2	0.028^{***}	0.044^{***}
	(0.01)	(0.02)
Obs	$3,692,188$	$1,253,505$
R^2	0.02	0.03
$\gamma_{i,t}$	Yes	Yes
Panel	Full	$N_i > 5$

Notes: The first column displays the OLS estimate of β_2 for the full unbalanced panel. The second column reports the same specification, but restricting the sample to firms that we observe for more than 5 consecutive years as in the analysis in section 5. Both regressions include sector-year fixed effects and standard errors clustered at the sector-year level.

While this evidence supports *qualitatively* the mechanism, we also evaluate it *quantitatively*: we simulate a monetary policy shock in the model and calculate the response of firms' capital stock as a function of their pre-shock MRPK. The model predicts that this response is *near-linear* in the logarithm of pre-shock MRPK, as Figure 10 in Appendix C illustrates. The slope of this function is the exact model counterpart to the coefficient β_2 which we estimated above. Its value is 0.045, very close to the empirical value of 0.044.

We perform a battery of robustness tests, which we report in Appendix A.3. In particular, we add firm-level fixed effects, firm-level controls, and introduce aggregate controls interacted with firm-level controls. In all cases the coefficient β_2 remains positive and statistically significant, and of similar magnitude. We also demean the MRPK at the firm level, to ensure that the results are not driven by permanent heterogeneity in responsiveness across firms – which means we test the same specification as in Ottonello and Winberry (2020) – and we still find a positive and significant coefficient. Furthermore, these results are not driven by the smaller firms in our sample: if we restrict the analysis to large firms, the coefficient is still positive and significant, and even of greater magnitude.

A follow-up empirical paper by Albrizio et al. (2023) analyzes in detail this result. They show that the higher investment by high-MRPK firms is possible thanks to an increase in their debt holdings, which points at a relaxation of financial frictions. They also document that common proxies for tighter financial frictions, such as firm age, leverage, or liquidity, matter for investment sensitivity to monetary policy only as long as firms have a high MRPK.

Linking the model to the data. In the model, the individual investment decisions aggregate up to changes in misallocation, such that aggregate TFP increases after an expansionary monetary policy shock. To test this prediction quantitatively, we need an empirical measure of TFP that abstracts from any changes in TFP that are brought about by anything but changes in the allocation of capital. For this purpose we define *dynamic weighted average MRPK*, $WAM_{t,\tau}$, as

$$WAM_{t,\tau} \equiv \sum_{j=0}^{J} MRPK_t^j \frac{k_{j,t+\tau}}{K_{t+\tau}},$$

where j indexes the firm, J is the number of firms, and $K_{t+\tau}$ is the aggregate capital. We approximate the growth rate of $WAM_{t,\tau}$ from time t to $t+\tau$ as

$$\Delta \log WAM_{t,\tau} \equiv \log WAM_{t,\tau} - \log WAM_{t,0},$$

where $\Delta WAM_{t,\tau}$ tells us how much the economy-wide average MRPK has changed from period t to $t+\tau$ *only* due to changes in the distribution of capital across firms, holding constant the MRPK of the firm at the initial level. As we show in Appendix A.5, in our model $\Delta \log WAM_{t,\tau}$ is approximately proportional to the growth rate of

TFP Z_t:[27]

$$\Delta \log WAM_{t,\tau} \approx 1/\alpha \Delta \log Z_{t,\tau}$$

Through the lens of the model, our empirical measure $\Delta \log WAM_{t,\tau}$ can hence be interpreted as a measure of changes in TFP that are brought about purely through changes in the allocation of capital, muting any other channels through which a monetary policy shock may simultaneously affect standard measures of TFP.

We use this variable as dependent variable in the following simple local projection, using the data set introduced in section 7:[28]

$$\Delta \log WAM_{t-1,\tau,s} = \alpha_{s,\tau} + \beta_\tau \varepsilon_t + u_{j,t,\tau} \text{ for } \tau = 1, 2, 3, 4.$$

We estimate this regression at the sector level s to account for potential sectoral differences. The variable $\alpha_{s,\tau}$ denotes horizon specific sector fixed effects. ε_t is the monetary policy shock. The regression coefficient β_τ thus tells us the cumulative change in our measure of capital misallocation at different horizons τ after a 1 p.p. monetary policy easing surprise.

Figure 7 reports our estimates for β_τ at different horizons τ (black line), with confidence intervals shaded in gray. Standard error are clustered at sector level. A 1 p.p. expansionary monetary policy shock causes an *increase* of the dynamic weighted average MRPK of 3 p.p. at impact, and of 7 p.p. at peak after 3 years. The effect is significant throughout 4 years at the 95% level.[29]

The dashed-dotted orange line in Figure 7 shows that the model produces a similar path for the dynamic weighted average MRPK, albeit of smaller magnitude.[30] The model explains about half of the observed increase in $\Delta \log WAM_{t,\tau}$ in the data. The model can hence be interpreted as being conservative with regards to the strength of

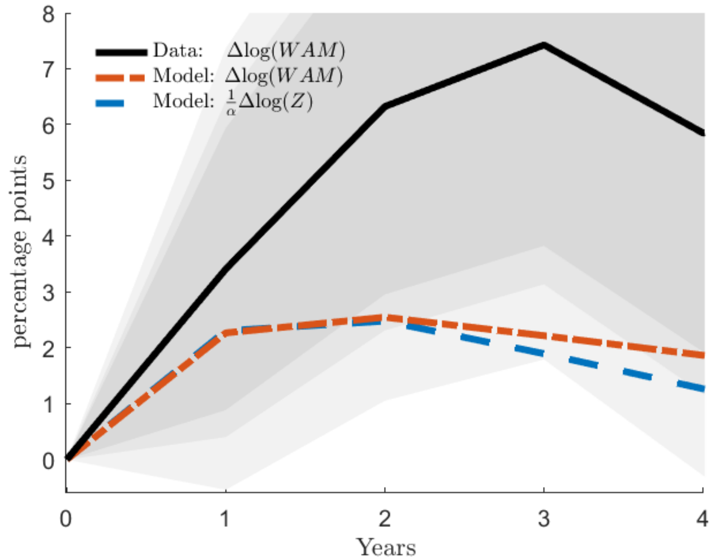
[27] The approximate relation between $\Delta WAM_{t,\tau}$ and TFP requires firm-level productivity shocks to be very persistent.

[28] Note that to construct the empirical measure of $\Delta \log WAM_{t,\tau}$, we need to restrict our sample to firms that we observe for several consecutive years. This is the same subsample of firms as in Column 2 of Table 2.

[29] Albrizio et al. (2023), using the sector-level variance of MRPK (see Hsieh and Klenow 2009) as a measure of misallocation for Spanish data, also find that expansionary monetary policy shocks decrease misallocation.

[30] The model counterpart is constructed by feeding a 1 p.p. monetary policy easing surprise into the model as a temporary deviation from the Taylor Rule, and then constructing $\Delta WAM_{t,\tau}$ from the model's outcome.

Figure 7: Response of average MRPK to an expansionary monetary policy shock.



Notes: The Figure shows the estimated impulse response function after a 1 p.p. expansionary monetary policy shock of $\Delta \log WAM_{t,\tau}$ on the data (black line), and the shaded area marks the 90%, 95% and 99% confidence intervals of the data estimates. It also shows response after a 1 p.p. expansionary monetary policy shock in the model of $\Delta \log WAM_{t,\tau}$ (orange broken line), and the log changes of model TFP (scaled by $1/\alpha$) (blue dashed line).

the capital-misallocation channel. The peak increase in $WAM_{t,\tau}$ of almost 2.5 p.p. predicted by the model (dashed-dotted orange line of Figure 7) corresponds to an increase of TFP of 0.87 p.p. (dashed blue line of Figure 7).

Taking together, these findings are consistent with expansionary monetary policy increasing TFP by increasing the share of high-MRPK firms in the capital stock.

8 Conclusions

This paper introduces a tractable model with heterogeneous firms, financial frictions, and nominal rigidities in order to answer two interrelated questions. The first one is whether falling interest rate foster capital misallocation. We show how a decline in real interest rates per se can affect positively or negatively misallocation. It is the joint dynamics of real and natural rates what determines the outcome.

The second question is how misallocation affects the optimal design of monetary policy. We analyze optimal monetary policy for a benevolent central bank with commitment. We show how a central bank without pre-commitments engineers an unexpected monetary expansion to increase TFP in the medium run. We also illustrate how, when faced with a temporary demand shock, price stability is the optimal policy, just as under complete markets. If the ZLB constraints the path of nominal rates, then the optimal policy is a "low for even longer", delaying the lift-off in nominal rates much longer than under complete markets. The paper also makes a methodological contribution. It introduces a new algorithm to compute optimal policies in heterogeneous-agent models. The algorithm leverages the numerical advantages of continuous time and will allow researchers to solve optimal policy in heterogeneous-agent models in an efficient and simple way using Dynare.

Finally, we provide some supporting empirical evidence. In particular, we present firm-level evidence that expansionary monetary policy reduces misallocation by inducing high-MRPK firms to increase their investment relatively more than low-MRPK firms.

The model presented in this paper abstracts from several relevant mechanisms driving firm dynamics, such as endogenous default, size-varying capital constraints, or decreasing returns to scale, among many others. This helps us to provide a clear understanding of the different 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 (2020). Optimal Monetary Policy According to HANK. Staff Reports 916, Federal Reserve Bank of New York. 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. 1, 4, 6.1, D.1, 38, E.1

Adam, K. and R. M. Billi (2006). Optimal monetary policy under commitment with a zero bound on nominal interest rates. *Journal of Money, credit and Banking*, 1877–1905. 5

Adam, K. and H. Weber (2019). Optimal trend inflation. *American Economic Review 109*(2), 702–37. 7

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. 4, E.1

Albrizio, S., B. González, and D. Khametshin (2023). A tale of two margins: Monetary policy and capital misallocation. *Documentos de trabajo N. 2302, Banco de España*. 8, 7, 29, A.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

Andrés, J., O. Arce, and P. Burriel (2021). Market polarization and the phillips curve. *Documentos de trabajo N.2106, Banco de España*. 7

Armenter, R. and V. Hnatkovska (2017). Taxes and capital structure: Understanding firms' savings. *Journal of Monetary Economics 87*, 13–33. A.1

Asriyan, V., L. Laeven, A. Martin, A. Van der Ghote, and V. Vanasco (2021). Falling interest rates and credit misallocation: Lessons from general equilibrium. Technical report. 1, 2.9, 5, 5

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. 4, D, E.1

Auclert, A., M. Cai, M. Rognlie, and L. Straub (2022). Optimal Policy with Heterogeneous Agents: A Sequence-Space Approach. Mimeo. 1

Auclert, A., M. Rognlie, and L. Straub (2020). Micro jumps, macro humps: Monetary policy and business cycles in an estimated hank model. Technical report, National Bureau of Economic Research. 36, 42

Baqaee, D., E. Farhi, and K. Sangani (2021). The supply-side effects of monetary policy. Technical report, National Bureau of Economic Research. 2, 7, 2.9, 5

Bau, N. and A. Matray (2023). Misallocation and capital market integration: Evidence from india. *Econometrica 91*(1), 67–106. 6, 26

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. 2.9

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2021). Inequality, business cycles, and monetary-fiscal policy. *Econometrica, forthcoming*. 1, 6.1

Bigio, S. and Y. Sannikov (2021). A model of credit, money, interest, and prices. Technical report, National Bureau of Economic Research. 1, 6.1

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. 7

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, 6.3

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. 4, 6.3, E.1

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. 7

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, 5

Christiano, L. J., M. S. Eichenbaum, and M. Trabandt (2016). Unemployment and business cycles. *Econometrica 84*(4), 1523–1569. 4

Cloyne, J., C. Ferreira, M. Froemel, and P. Surico (2018, December). Monetary Policy, Corporate Finance and Investment. NBER Working Papers 25366, National Bureau of Economic Research, Inc. 1

Cui, W., R. Wright, and Y. Zhu (2022). Endogenous Liquidity and Capital Reallocation. Staff Working Papers 22-27, Bank of Canada. 14

David, J. M. and D. Zeke (2021). Risk-taking, capital allocation and optimal monetary policy. Technical report. 7

Dávila, E. and A. Schaab (2022). Optimal monetary policy with heterogeneous agents: A timeless ramsey approach. *Working Paper*. 1, 6.1

Eggertsson, G. B. et al. (2003). Zero bound on interest rates and optimal monetary policy. *Brookings papers on economic activity 2003*(1), 139–233. 5, 6.3

Eggertsson, G. B. and M. Woodford (2004). Policy options in a liquidity trap. *American Economic Review 94*(2), 76–79. 1

Evans, C. L. (1992, April). Productivity shocks and real business cycles. *Journal of Monetary Economics 29*(2), 191–208. 2, 5

Ferreira, M., T. Haber, and C. Rorig (2023, May). Financial Constraints and Firm Size: Micro-Evidence and Aggregate Implications. Working Papers 777, DNB. 13

Ferreira, T. R., D. A. Ostry, and J. Rogers (2022). Firm Financial Conditions and the Transmission of Monetary Policy. Working Papers 2316, Cambridge. 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*. 6.2, 6.3

Garga, V. and S. R. Singh (2021). Output hysteresis and optimal monetary policy. *Journal of Monetary Economics 117*, 871–886. 2, 5

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

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, 4, 5, 5

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, 29

Itskhoki, O. and B. Moll (2019). Optimal development policies with financial frictions. *Econometrica 87*(1), 139–173. 3, 6.2, B.8

Jarociński, M. and P. Karadi (2020). Deconstructing monetary policy surprises: the role of information shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43. 1, 7, A.2, 8

Jeenas, P. (2020). Firm balance sheet liquidity, monetary policy shocks, and investment dynamics. Technical report, Universidad Pompeu Fabra. 1, 2.9

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, 5

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. D, E.1

Jungherr, J., M. Meier, T. Reinelt, and I. Schott (2022). Corporate Debt Maturity Matters For Monetary Policy. Discussion paper series, University of Bonn and University of Mannheim, Germany. 1

Kaplan, G., B. Moll, and G. L. Violante (2018). Monetary policy according to hank. *American Economic Review 108*(3), 697–743. 4, B.3

Koby, Y. and C. Wolf (2020). Aggregation in heterogeneous-firm models: Theory and measurement. Technical report. 7

Le Grand, F., A. Martin-Baillon, and X. Ragot (2020). What is monetary policy about? optimal monetary policy with heterogeneous agents. Technical report, Paris School of Economics. 1, 6.1

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

Meier, M. and T. Reinelt (2020). Monetary policy, markup dispersion, and aggregate tfp. Technical report, University of Bonn and University of Mannheim, Germany. 2, 7, 5

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, 2.1, 10, 2.9

Moran, P. and A. Queralto (2018). Innovation, productivity, and monetary policy. *Journal of Monetary Economics 93*, 24–41. 2, 5

Nakov, A. et al. (2008). Optimal and simple monetary policy rules with zero floor on the nominal interest rate. *International Journal of Central Banking 4*(2), 73–127. 5

Nakov, A. and H. Webber (2021). Micro heterogeneity, misallocation and monetary policy. Technical report. 7

Nuño, G. and C. Thomas (2022). Optimal Redistributive Inflation. *Annals of Economics and Statistics* (146), 3–63. 1, 6.1, E.1, E.2, E.3, 11

Ottonello, P. and T. Winberry (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica 88*(6), 2473–2502. 1, 2.9, 7, 7, A.2, A.3

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. 6.1

Reis, R. (2013). The portuguese slump and crash and the euro crisis. *Brookings Papers on Economic Activity*, 143–193. 1, 5

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control 33*(3), 649–665. 4, E.1

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, 6.1

Trimborn, T., K.-J. Koch, and T. Steger (2008). Multidimensional transitional dynamics: A simple numerical procedure. *Macroeconomic Dynamics 12*(3), 301â319. D, E.1

Winberry, T. (2018). A method for solving and estimating heterogeneous agent macro models. *Quantitative Economics 9*(3), 1123–1151. 4, E.1, E.1

Woodford, M. (2003). Interest and prices: Foundations of a theory of monetary policy. *Princeton University Press*. 1, 6.2, 6.3

Zanetti, F. and M. Hamano (2020). Monetary policy, firm heterogeneity, and product variety. Technical report, University of Oxford, Department of Economics. 7

Online appendix

A Empirical Appendix

A.1 Firm level data

The empirical exercise relies on annual firm balance-sheet data from the *Central de Balances Integrada* database (Integrated Central Balance Sheet Data Office Survey). We use an unbalanced panel of firms from 1999 to 2016, since these are the years for which the monetary policy shocks are available. Being a detailed administrative dataset, the main advantage is that it covers the quasi-universe of Spanish firms (see Almunia et al., 2018 for further details on the representativeness of this dataset). We use for our analysis only high quality observations, as defined by the Integrated Central Balance Sheet Data Office.

Our main variable of interest, firm's marginal revenue product of capital (MRPK), is proxied by the log of the ratio of value added over tangible capital.[31] We drop firms in the 5% upper tail of the capital-weighted MRPK distribution, so as to focus on firms holding a non-negligible capital share. Variables are deflated using industry price levels to preserve the firms' price-level changes and consider a revenue-based measure of MRPK (Foster et al., 2008). The capital-weighted MRPK distribution in the data is shown in blue bars in Figure 4.

Our dependent variable, the investment rate (or capital growth), is defined as the difference of firm's tangible capital, in logarithm, between periods t and $t-1$. We also use other firm-level information as controls in the robustness section below. Leverage is computed as total debt (short-term plus long-term debt) divided by total assets, and it is trimmed below 0 and over 10. Net financial assets are constructed as the log difference between financial assets and financial liabilities, where financial assets include short-term financial investment, trade receivables, inventories and cash holdings; and financial liabilities include short-term debt, trade payables and long-term debt. We trim this variable below at -10 and above at 10. This variable controls for firms' savings, following Armenter and Hnatkovska (2017). We proxy for size using the logarithm of total assets. Real sales growth is defined as the log-difference of sales in two consecutive

[31] Implicitly, this is restricts our sample to observations with positive value added.

years, the previous and the current one.[32] We use the value-added price deflator for value added and sales, and the investment price deflator for capital and total assets. Descriptive statistics are reported in Table 3.

Table 3: **Descriptive statistics**

	mean	sd	p5	p95
Capital growth (1 period)	0.00	0.29	-0.30	0.49
MRPK (logs)	-0.87	1.40	-3.57	0.73
MRPK (levels)	0.77	0.66	0.03	2.07
Total Assets	6.02	1.58	3.56	8.64
Leverage	0.31	0.33	0.00	0.96
Net financial Assets	0.07	0.51	-0.71	0.68
Sales growth	1.73	792.52	-0.51	1.25
Observations	5184233			

Notes: The table shows the mean (column 1), standard deviation (column 2), 5th and 95th percentile value (column 3 and 4 respectively) of the main variables used in the calibration and empirical analysis. MRPK is shown in logs and in levels. The table also displays total assets, leverage, net financial assets; and the log difference of the capital stock (capital growth) and output (sales growth). The number of observations are those for which the variable MRPK is available.

A.2 Monetary policy shocks

The key idea behind the identification strategy by Jarociński and Karadi (2020) is that movements of interest rates and stock markets within a narrow window around monetary policy announcements can help disentangle monetary policy shocks from information surprises. While an unexpected policy tightening raises interest rates and reduces stock prices, a positive central bank information shock (i.e. unexpected positive assessment of the economic outlook) raises both. Since our firm-level panel is at annual frequency, we aggregate the monthly monetary policy shocks following the scheme of Ottonello and Winberry (2020). However, instead of aggregating daily shocks into quarterly series, we apply a monthly-to-yearly transformation. This scheme accounts for the fact that firms have less time to react to shocks happening at the end of the year then to shocks happening earlier on. In particular, a monthly shock enters both the current year and the following year's annual shock, with the split between the current

[32]Both sales growth and capital growth are winsorized at 0.5%.

and the next year depending on the timing of the monthly shock within the current year.[33] Concretely, we construct the monetary policy shock as

$$\varepsilon_t = \sum_m \omega_{\text{past}}(m)\varepsilon_{m,t-1} + \sum_m \omega_{\text{current}}(m)\varepsilon_{m,t} \quad \omega_{\text{past}}(m) = \frac{m-1}{12}, \quad \omega_{\text{current}}(m) = \frac{12-(m-1)}{12}$$

where ε_t is the aggregated annual monetary policy shock in year t, and $\varepsilon_{m,t}$ is the high-frequency shock in month $m = 1, \dots 12$ of year t (see (Albrizio et al., 2023) for a formal derivation of this weighting scheme). Note that we multiply the original shocks by (-1), so that positive monetary policy shocks corresponds to expansionary monetary policy. Figure 8 shows the time series of the shock.

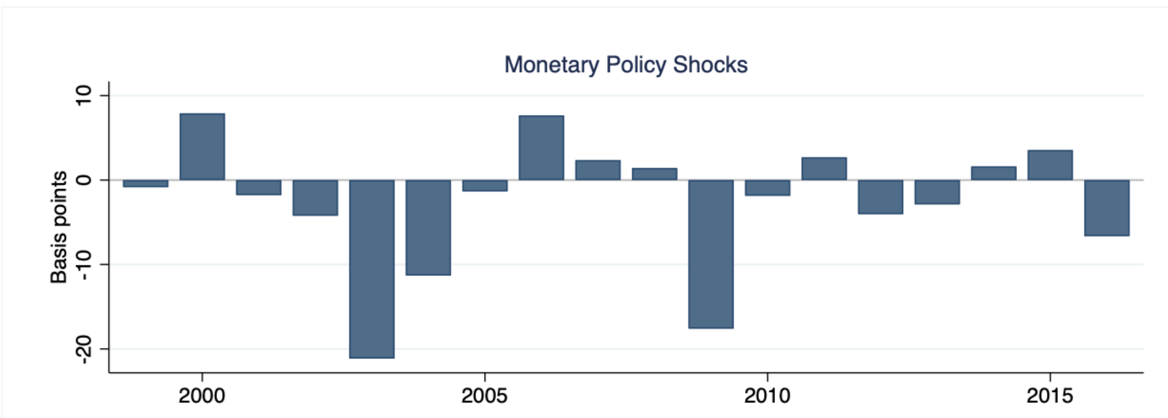


Figure 8: Monetary policy shocks at annual frequency.
Source: Jarociński and Karadi (2020) and own calculations.

A.3 Robustness of the firm level regression

The high degree of granularity of the firm-level data allows us to enhance our empirical specification with a large number of controls. In this section we exploit this feature to check the robustness of our finding that high-MRPK firms' investment responds more to monetary shocks. We consider variations of the main empirical specification explained in the main text, equation (43), which we repeat here expanded to include

[33]For instance, a high-frequency surprise happening in January is entirely attributed to the current year, while the one occurring in December mainly contributes to the following year's annual shock.

the robustness specifications we perform below

$$\log k_{j,t} - \log k_{j,t-1} = \beta_0 + \beta_1 \log \left(MRPK_{j,t-1}\right) + \beta_2 \log \left(MRPK_{j,t-1}\right) \varepsilon_t + \lambda' Z_{j,t-1} + \gamma_{i,t} + \kappa_j + u_{j,t}, \tag{44}$$

where $Z_{j,t-1}$ includes a vector of firm-level controls (total assets, sales growth, leverage and net financial assets), and in some specification it also includes their interaction with aggregate GDP growth; $\gamma_{i,t}$ are sector-year fixed effects; and κ_j are firm-level fixed effects.

Column (1) in Table 4 reproduces the results of our main specification, reproducing column 1 of Table 2 of the main text. It does not include firm-level nor aggregate controls, and only includes sector-time fixed effects. Column (2) includes firm fixed-effects, and Column (3) adds firm-level controls. The results remain positive, significant and of similar magnitude. Column (4) reports results for the same specifications as Column (3), just replacing the main variable of interest $\log \left(MRPK_{j,t-1}\right)$ with its demeaned value at the firm level, to make sure that results are not driven by permanent heterogeneity in MRPK levels (following Ottonello and Winberry, 2020). Results remain positive and significant.

One of the advantages of our dataset is that it also includes small and privately held firms. But precisely because of this, it could be the case that small firms are the ones driving these results and one may wonder if the same empirical pattern holds for large firms. To address this concern, we replicate the analysis of our baseline specification, Column (1), but keeping only firms below the 90th percentile of employment (Column 5), and keeping only firms above the 90th percentile (Column 6). The coefficient on the slope of MRPK is positive, significant, and even quantitatively larger for larger firms. The 90th percentile of employment in the Spanish distribution of firms is relatively low (15 employees), so that we repeat the same regression but keeping only firms with at least 100 employees in Column (7), and reach the same conclusion. Column (8) adds to the main specification shown in Column (1) an extra control: the interaction of aggregate GDP growth with $\log \left(MRPK_{j,t-1}\right)$, to control for variation in investment coming from different responses of high-MRPK firms along the business cycle. The coefficient of interest remains significant and of similar magnitude to that of the main specification. Column (9) augments this specification by adding firm fixed-effects and firm-level controls as in (3), and the interactions of these controls with aggregate GDP growth. The coefficient remains positive and significant, and just slightly lower quantitatively.

Finally, since our panel is highly unbalanced, we run our baseline specification, but only for firms that we observe for at least for 6 consecutive years (from $t-1$ to $t+4$) (Column 10), hence restricting the sample as in the aggregate analysis performed in Section 5. The coefficient is nearly twice as large, and it is much closer to the model counterpart. Summing up, all these exercises point at the robustness of the empirical result of a higher heterogeneous response of investment for high-MRPK firms to a monetary policy shock.

Table 4: **Robustness**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\varepsilon * mrpk$	0.0284***	0.0214**	0.0221**		0.0273**	0.0608***	0.0680***	0.0283***	0.0234**	0.0443**
	(0.01)	(0.01)	(0.01)		(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
$\varepsilon * \overline{mrpk}$				0.566***						
				(0.06)						
Observations	3692188	3538432	2964354	2964354	3290824	401359	40932	3692188	2964354	1253505
R^2	0.020	0.255	0.297	0.305	0.020	0.035	0.063	0.022	0.298	0.033
Time-sector FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm FE	NO	YES	YES	YES	NO	NO	NO	NO	YES	NO
Firm Controls	NO	NO	YES	YES	NO	NO	NO	NO	YES	NO
MRPK dem.	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
Agg. Control	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO
Panel	FULL	FULL	FULL	FULL	EMP<p90	EMP>p90	LARGE	FULL	FULL	$N > 5$

Notes: Results of estimating equation (44), departing from some of the specifications of the estimation in the main text of equation (43). Column (1) runs the baseline regression, that is, that of equation (43). Column (2) includes firm fixed-effects, and Column (3) also includes firm-level controls (total assets, sales growth, leverage and net financial assets). Column (4) runs the same specification as Column (3), just replacing the main variable of interest $\log(MRPK_{j,t-1})$ with its demeaned value at the firm level. Columns (5), (6) and (7) run the baseline specification, but including only firms with less employees than the 90th percentile (15 employees), firms with more employees than 90th percentile, and firms with more than 100 employees, respectively. Column (8) adds to the main specification an extra control, the interaction of aggregate GDP growth with $\log(MRPK_{j,t-1})$, and Column (9) further includes firm fixed-effects and firm-level controls as in (3), and the interactions of these controls with aggregate GDP growth. Column (10) shows the baseline specification, but only for firms that we observe at least for 6 consecutive years (from t-1 to t+4).

A.4 Estimating the process for idiosyncratic productivity z

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

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

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

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 $\sigma = 0.73$. As the data frequency is annual, $\Delta t = 1$, we back out the implied to continuous time parameter $\varsigma_z = -\log(\rho_z) = 0.189$.

A.5 Derivation of the approximate correspondence of $\triangle Z_t$ and $\triangle WAM_{t,s}$

We define $k_t(z) = \int_0^\infty k(z,a)\, g_t(z,a)da$, and $a_t(z) = \int_0^\infty a g_t(z,a)da$. Manipulating the definition of TFP (33) by subsequently using the definitions of $\Omega(z_t^*)$, $\omega_t(z)$ and the linearity of $k_t(z)$ in $a_t(z)$ when $z > z_t^*$, we get

$$\begin{aligned} Z_t^{1/\alpha} &= \frac{\int_{z_t^*}^{\infty} z\omega_t(z)dz}{(1-\Omega(z_t^*))}dz \\ &= \int_{z_t^*}^{\infty} z\frac{\omega_t(z)}{\int_{z_t^*}^{\infty} \omega_t(z)\, dz} \\ &= \int_{z_t^*}^{\infty} z\frac{a_t(z)/A_t}{\int_{z_t^*}^{\infty} a_t(z)/A_t dz} \\ &= \int_0^{\infty} z\frac{k_t(z)}{\int_0^{\infty} k_t(z)\, dz}dz \\ &= \int_0^{\infty} z\frac{k_t(z)}{K_t}dz \end{aligned}$$

Now consider two points in time t and $t+\tau$ where $t < t+\tau$. Since z follows a persistent process we can approximate a firm's j productivity level at $t+\tau$ by its productivity at t, $z_{j,t} \approx z_{j,t+\tau}$. This approximation holds exactly in the limit as the process slows down ($\varsigma_z \to 1$ and $\sigma_z \to 0$) or as the time difference shrinks ($\tau \to 0$). We can thus write

$$Z_{t+\tau}^{1/\alpha} \approx \int_0^1 z_{j,t}\frac{k_{j,t+\tau}}{K_{t+\tau}}dj$$

where $k_{j,t+\tau}$ denotes the period $t+\tau$ capital of an active firm with initial productivity level $z_{j,t}$ in period t. Using the definition of the MRPK,

$$MRPK_t(z_{j,t}) \equiv \varphi_t z_{j,t}$$

we arrive at

$$Z_{t+\tau}^{1/\alpha} \approx \tfrac{1}{\varphi_t} \int_0^1 MRPK_t(z_{j,t})\frac{k_{j,t+\tau}}{K_{t+\tau}}dj.$$

To understand how a monetary policy shock affects $Z_t^{1/\alpha}$, we are interested in the evolution of $\left(\log Z_{t+\tau}^{1/\alpha} - \log Z_t^{1/\alpha}\right) = \frac{1}{\alpha}\left(\log Z_{t+\tau} - \log Z_t\right)$ where t now denotes the period of the shock arrival. Using the above relationship and defining $WAM_{t,\tau} \equiv \int_0^1 MRPK_t(z_{j,t})\frac{k_{j,t+\tau}}{K_{t+\tau}}dj$

we can write:

$$\begin{aligned}\frac{1}{\alpha}\left(\log Z_{t+\tau}-\log Z_t\right) &\approx \log\int_0^1 MRPK_t(z_{j,t})\frac{k_{j,t+\tau}}{K_{t+\tau}}dj-\log\int_0^1 MRPK_t(z_{j,t})\frac{k_{j,t}}{K_t}dj\\ &= \log WAM_{t,\tau}-\log WAM_{t,0}\end{aligned}$$

The empirical counterpart of $\int_0^1 MRPK_t(z_{j,t})\frac{k_{j,t+\tau}}{K_{t+\tau}}dj$ is the expression in the main text $\sum_0^J MRPK_t^j\frac{k_{t+\tau}^j}{K_{t+\tau}}$. In the main text we report both the $\frac{1}{\alpha}\left(\log Z_{t+\tau}-\log Z_t\right)$ and $\log WAM_{t,\tau}-\log WAM_{t,0}$.[34] The approximation is indeed good, especially for the first years, as we show in figure 7.

[34] $WAM_{t,\tau}$ can be computed analytically given the transitional dynamics of equilibrium prices and z_t^*.

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 $V_t(z,a) = \kappa_t\left(z\right) q_t a$. The first order condition is

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

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 $\kappa_t\left(z\right)$ can be obtained from

$$\begin{aligned}&\left(r_t + \eta\right)\kappa_t\left(z\right) q_t = \\ &\eta q_t + \left(\gamma \max\left\{z_t\varphi_t - R_t, 0\right\} + R_t - \delta q_t\right)\kappa_t\left(z\right) + \mu(z) q_t \frac{\partial \kappa_t}{\partial z} + \frac{\sigma^2(z)}{2} q_t \frac{\partial^2 \kappa_t}{\partial z^2} + \frac{\partial\left(q_t \kappa_t\right)}{\partial t}.\end{aligned} \tag{45}$$

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

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

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

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

$$\mathbb{E}_0 a_t = \mathbb{E}_0 a_0 e^{\int_0^t s_u^a du} > a_0 e^{\int_0^t \frac{R_s - \delta q_s}{q_s} ds}. \tag{46}$$

The value function is then

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

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 (46). Hence $\kappa_{t_0}(z) > 1$ for any t_0.

B.2 Household's problem

We can rewrite the household's problem as

$$W_t = \max_{C_t, L_t, D_t, B_t^N, S_t^N} \mathbb{E}_0 \int_0^\infty e^{-\rho_t^h t} \left(\frac{C_t^{1-\zeta}}{1-\zeta} - \Upsilon \frac{L_t^{1+\vartheta}}{1+\vartheta} \right) dt. \tag{47}$$

$$\begin{aligned}
s.t.\quad \dot{D}_t &= \left[(R_t - \delta q_t) D_t + w_t L_t - C_t - S_t^N + \Pi_t \right] / q_t, &(48)\\
\dot{B_t^N} &= S_t^N + (i_t - \pi_t) B_t^N, &(49)
\end{aligned}$$

where S_t^N is the investment into nominal bonds.

The Hamiltonian is

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

The first order conditions are

$$C_t^{-\zeta} - \varrho_t/q_t = 0 \tag{50}$$

$$-\Upsilon L_t^{\vartheta} + \varrho_t w_t/q_t = 0 \tag{51}$$

$$-\varrho_t/q_t + \eta_t = 0 \tag{52}$$

$$\dot{\varrho}_t = \rho_t^h \varrho_t - \varrho_t \left(R_t - \delta q_t\right)/q_t \tag{53}$$

$$\dot{\eta}_t = \rho_t^h \eta_t - \eta_t \left[(i_t - \pi_t)\right] \tag{54}$$

(50) and (51) combine to the optimality condition for labor

$$w_t = \frac{L_t^{\vartheta}}{C_t^{-\eta}},$$

(50) can be rewritten as

$$\varrho_t = C_t^{-\eta} q_t$$

Now take derivative with respect to time

$$\dot{\varrho}_t = -\eta C_t^{-\eta-1}\,\dot{C}_t q_t + C_t^{-\eta}\,\dot{q}_t$$

and plug this into (53) and rearrange to get the first Euler equation

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

(52) can be rewritten as

$$\eta_t = \varrho_t/q_t$$

Now take derivative with respect to time

$$\dot{\eta}_t = \frac{\dot{\varrho}_t q_t - \varrho_t \dot{q}}{q_t^2}$$

Use these two expressions and the definition of $\dot{\varrho}_t$ in (54) to get the second Euler equation

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

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 Philips curve

The proof is similar to that of Lemma 1 in Kaplan et al. (2018). The Hamilton-Jacobi-Bellman (HJB) equation of the retailer's problem is

$$r_t V_t^r(p) = \max_{\pi} \left(\frac{p - p_t^y (1-\tau)}{P_t} \right) \left(\frac{p}{P_t} \right)^{-\varepsilon} Y_t - \frac{\theta}{2} \pi^2 Y_t + \pi p \frac{\partial V^r}{\partial p} + \frac{\partial V^r}{\partial t},$$

where where $V_t^r(p)$ is the real value of a retailer with price p. The first order and envelope conditions for the retailer are

$$\begin{aligned} \theta \pi Y_t &= p \frac{\partial V^r}{\partial p}, \\ (r - \pi) \frac{\partial V^r}{\partial p} &= \left(\frac{p}{P_t} \right)^{-\varepsilon} \frac{Y_t}{P_t} - \varepsilon \left(\frac{p - p_t^y (1-\tau)}{P_t} \right) \left(\frac{p}{P_t} \right)^{-\varepsilon - 1} \frac{Y_t}{P_t} + \pi p \frac{\partial^2 V^r}{\partial p^2} + \frac{\partial^2 V^r}{\partial t \partial p}. \end{aligned}$$

In a symmetric equilibrium we will have $p = P$, and hence

$$\begin{aligned} \frac{\partial V^r}{\partial p} &= \frac{\theta \pi Y_t}{p}, \\ (r - \pi) \frac{\partial V^r}{\partial p} &= \frac{Y_t}{p} - \varepsilon \left(\frac{p - p_t^y (1-\tau)}{p} \right) \frac{Y_t}{p} + \pi p \frac{\partial^2 V^r}{\partial p^2} + \frac{\partial^2 V^r}{\partial t \partial p}. \end{aligned} \tag{55}$$

Deriving (55) with respect to time gives

$$\pi p \frac{\partial^2 V^r}{\partial p^2} + \frac{\partial^2 V^r}{\partial t \partial p} = \frac{\theta \pi \dot{Y}}{p} + \frac{\theta \dot{\pi} Y}{p} - \frac{\theta \pi^2 Y}{p},$$

and substituting into the envelope condition and dividing by $\frac{\theta Y}{p}$ we obtain

$$\left(r-\frac{\dot{Y}}{Y}\right)\pi=\frac{1}{\theta}\left(1-\varepsilon\left(1-\frac{p_t^y(1-\tau)}{p}\right)\right)+\dot{\pi}.$$

Finally, rearranging we obtain the New Keynesian Phillips curve

$$\left(r-\frac{\dot{Y}}{Y}\right)\pi=\frac{\varepsilon}{\theta}\left(\frac{1-\varepsilon}{\varepsilon}+\tilde{m}\right)+\dot{\pi}.$$

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

$$\Pi_t=\left(1-m_t\right)Y_t-\frac{\theta}{2}\pi_t^2 Y_t. \tag{56}$$

B.4 Capital producers' problem

The problem of the capital producer is

$$W_t=\max_{\iota_t,K_t}\mathbb{E}_0\int_0^{\infty}e^{-\int_0^t r_s ds}\left(q_t\iota_t-\iota_t-\Xi\left(\iota_t\right)\right)K_t dt. \tag{57}$$

$$\dot{K}_t \;\;=\;\; \left(\iota_t-\delta\right)K_t, \tag{58}$$

We construct the Hamiltonian

$$H=\left(q_t\iota_t-\iota_t-\Xi\left(\iota_t\right)\right)K_t+\lambda_t\left(\iota_t-\delta\right)K_t$$

with first-order conditions

$$\left(q_t-1-\Xi'\left(\iota_t\right)\right)+\lambda_t=0 \tag{59}$$

$$\left(q_t\iota_t-\iota_t-\Xi\left(\iota_t\right)\right)+\lambda_t\left(\iota_t-\delta\right)=r_t\lambda_t-\dot{\lambda}_t \tag{60}$$

Taking the time derivative of equation (59)

$$\dot{\lambda}_t = -\left(\dot{q}_t - \Xi''\left(\iota_t\right) i_t\right)$$

which, combined with (60), yields

$$\left(q_t \iota_t - \iota_t - \Xi\left(\iota_t\right)\right) - \left(q_t - 1 - \Xi'\left(\iota_t\right)\right)\left(\iota_t - \delta - r_t\right) = \left(\dot{q}_t - \Xi''\left(\iota_t\right) i_t\right)$$

Rearranging we get

$$r_t = \left(\iota_t - \delta\right) + \frac{\dot{q}_t - \Xi''\left(\iota_t\right) i_t}{q_t - 1 - \Xi'\left(\iota_t\right)} - \frac{q_t \iota_t - \iota_t - \Xi\left(\iota_t\right)}{q_t - 1 - \Xi'\left(\iota_t\right)}.$$

B.5 Distribution

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

$$\frac{\partial g_t(z,a)}{\partial t} = -\frac{\partial}{\partial a}[g_t(z,a)s_t(z)a] - \frac{\partial}{\partial z}[g_t(z,a)\mu(z)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[g_t(z,a)\sigma^2(z)] - \eta g_t(z,a) + \eta/\psi g_t(z,a/\psi), \tag{61}$$

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

$$\frac{\partial \omega_t(z)}{\partial t} = -\frac{\dot{A}_t}{A_t^2}\int_0^\infty a g_t(z,a)da + \frac{1}{A_t}\int_0^\infty a \frac{\partial g_t(z,a)}{\partial t} da. \tag{62}$$

Next, we plug in the derivative of $g_t(z,a)$ wrt time from equation(61) into equation (62),

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

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,

$$\frac{\partial \omega_t(z)}{\partial t} = \left[s_t(z) - \frac{\dot{A}_t}{A_t} - (1-\psi)\eta \right] \omega_t(z) - \frac{\partial}{\partial z}\mu(z)\omega_t(z) + \frac{1}{2}\frac{\partial^2}{\partial z^2}\sigma^2(z)\omega_t(z). \quad (63)$$

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

$$0 = (s(z) - (1-\psi)\eta)\,\omega(z) - \frac{\partial}{\partial z}\mu(z)\omega(z) + \frac{1}{2}\frac{\partial^2}{\partial z^2}\sigma^2(z)\omega(z). \quad (64)$$

B.6 Market clearing and aggregation

Define the cumulative function of net-worth shares as

$$\Omega_t(z) = \int_0^z \omega_t(z)dz. \quad (65)$$

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

$$K_t = \int k_t(z,a)dG_t(z,a) = \int_{z_t^*}^{\infty} \int \gamma a \frac{1}{A_t} g_t(z,a) da dz A_t = \gamma(1 - \Omega(z_t^*))A_t. \quad (66)$$

By combining equations (28), (29) and (66), and solving for A_t,we obtain

$$A_t = \frac{D_t}{\gamma(1-\Omega(z_t^*)) - 1}, \quad (67)$$

Labor market clearing implies

$$L_t = \int_0^{\infty} l_t(z,a)dG_t(z,a). \quad (68)$$

Define the following auxiliary variable,

$$X_t \equiv \int_{z_t^*}^{\infty} z\omega_t(z)dz = \mathbb{E}\left[z \mid z > z_t^*\right](1 - \Omega(z_t^*)). \quad (69)$$

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

$$L_t = \int_0^{\infty} \left(\frac{\varphi_t}{\alpha m_t}\right)^{\frac{1}{1-\alpha}} z_t \gamma a_t dG_t(z,a) = \left(\frac{\varphi_t}{\alpha m_t}\right)^{\frac{1}{1-\alpha}} \gamma A_t X_t. \tag{70}$$

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

$$Y_t = \int \underbrace{\frac{z_t \varphi_t}{\alpha m_t} \gamma a}_{y_t(z,a)} \, dG_t(z,a) = \frac{\varphi_t}{\alpha m_t} X_t \gamma A_t = Z_t A_t^{\alpha} {L_t}^{1-\alpha}, \tag{71}$$

where in the last equality we have used equation (70), and we have defined

$$Z_t = (\gamma X_t)^{\alpha}. \tag{72}$$

Aggregate profits of retailers are given by

$$\Phi_t^{Agg} = \int \gamma \max\{z_t \varphi_t - R_t, 0\} \, a_t dG_t(z,a) = [\varphi_t X_t - R_t (1 - \Omega(z^*))] \, \gamma A_t. \tag{73}$$

We can also write the aggregate production in terms of physical capital,

$$Y_t = Z_t K_t^{\alpha} {L_t}^{1-\alpha}, \tag{74}$$

where the TFP term Z_t is defined as

$$Z_t = \left(\frac{X_t}{(1 - \Omega(z_t^*))}\right)^{\alpha} = \left(\mathbb{E}\left[z \mid z > z_t^*\right]\right)^{\alpha}. \tag{75}$$

$$Z_t^{1/a}(1 - \Omega(z_t^*)) = (X_t) = \left(\mathbb{E}\left[z \mid z > z_t^*\right]\right)^{\alpha}. \tag{76}$$

Aggregating the budget constraint of all input good firms, using the linearity of savings policy (11) and using (67), we obtain

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

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

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}(\gamma\varphi_t X_t - R_t\gamma(1-\Omega(z_t^*)) + R_t - \delta q_t - q_t(1-\psi)\eta). \tag{77}$$

Using the definition of X_t, and substituting φ_t using equation (70), we can simplify equation (77) as

$$\frac{\dot{A}_t}{A_t} = \frac{1}{q_t}(1-\Omega(z_t^*))\gamma(\alpha m_t Z_t {L_t}^{1-\alpha}\left((1-\Omega(z_t^*))\gamma A_t\right)^{\alpha-1} - R_t) + R_t - \delta q_t - q_t(1-\psi)\eta). \tag{78}$$

Using (66) and (67) we can replace $(1-\Omega(z_t^*))\gamma A_t$ by K_t, which delivers equation (36).

Finally, we can obtain factor prices

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

$$R_t = \alpha m_t Z_t A_t^{\alpha-1} {L_t}^{1-\alpha}\frac{z_t^*}{\gamma X_t} \tag{80}$$

where wages come from substituting the definition of φ_t into equation (70); and interest rates come from plugging in the wage expression (79) into the cut-off rule (10) and using equation (67). We could equivalently write equation (80) in terms of real rate of return r_t :

$$r_t = \frac{1}{q_t}\left(\alpha m_t Z_t A_t^{\alpha-1} {L_t}^{1-\alpha}\frac{z_t^*}{\gamma X_t}\right) - \delta + \frac{\dot{q}}{q_t} \tag{81}$$

We can easily get these equations in terms of capital instead of net worth by simply using equation (66), i.e. $A_t = \frac{K_t}{\gamma(1-\Omega(z_t^*))}$, and using that $\mathbb{E}\left[z \mid z > z_t^*\right] = \frac{X_t}{(1-\Omega(z_t^*))} = \frac{\int_{z_t^*}^{\infty} z\omega_t(z)dz}{(1-\Omega(z_t^*))}$ (see equation (72) and (75)).

B.7 Full set of equations

The competitive equilibrium economy is described by the following 22 equations, for the 22 variables $\{\omega(z), w, r, q, \varphi, R, K, A, L, C, D, Z, \mathbb{E}\left[z \mid z > z_t^*\right], \Omega, z^*, \iota, \pi, m, \tilde{m}, i, Y, T\}$. Remember that $\mu(z) = z\left(-\varsigma_z \log z + \frac{\sigma^2}{2}\right)$ and $\sigma(z) = \sigma_z z$, and that government bonds are in zero net supply ($B_t^N = 0$, hence $S_t^N = 0$). Except from the last equation (Taylor rule), the other 21 equations are the constraints of the Ramsey problem described in Section 2.8.

$$
\begin{aligned}
\frac{\partial \omega_t(z)}{\partial t} &= \left(s_t(z) - (1-\psi)\eta - \frac{\dot{A}_t}{A_t}\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 \text{where } s_t(z) \equiv \frac{1}{q_t}(\gamma \max\{z_t\varphi_t - R_t, 0\} + R_t - \delta q_t) \\
\Omega_t(z^*) &= \int_0^{z^*} \omega_t(z)\, dz \\
\varphi_t &= \alpha\left(\frac{(1-\alpha)}{w_t w_t}\right)^{(1-\alpha)/\alpha} m_t^{\frac{1}{\alpha}} \\
\tilde{m}_t &= m_t(1-\tau) \\
w_t &= (1-\alpha) m_t Z_t K_t{}^{\alpha} L_t{}^{-\alpha} \\
R_t &= \alpha m_t Z_t K_t^{\alpha-1} L_t{}^{1-\alpha} \frac{z_t^*}{\mathbb{E}\left[z \mid z > z_t^*\right]} \\
\frac{\dot{A}_t}{A_t} &= \frac{1}{q_t}\left[\gamma(1-\Omega(z_t^*))\left(\alpha m_t Z_t K_t^{\alpha-1} L_t{}^{1-\alpha} - R_t\right) + R_t - \delta q_t - q_t(1-\psi)\eta)\right] \\
K_t &= A_t + D_t \\
\dot{K}_t &= (\iota_t - \delta) K_t \\
A_t &= \frac{D_t}{\gamma(1-\Omega(z_t^*)) - 1} \\
Z_t &= (\mathbb{E}\left[z \mid z > z_t^*\right])^{\alpha} \\
\mathbb{E}\left[z \mid z > z_t^*\right] &= \frac{\int_{z_t^*}^{\infty} z\omega_t(z) dz}{(1-\Omega(z_t^*))} \\
\frac{\dot{C}_t}{C_t} &= \frac{r_t - \rho_t^h}{\eta} \\
w_t &= \frac{\Upsilon L_t^{\vartheta}}{C_t^{-\eta}} \\
\dot{D}_t &= \left[(R_t - \delta q_t) 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} \\
(q_t - 1 - \Phi'(\iota_t))(r_t - (\iota_t - \delta)) &= \dot{q}_t - \Phi''(\iota_t)\, \dot{\iota}_t - (q_t\iota_t - \iota_t - \Phi(\iota_t)) \\
\left(r_t - \frac{\dot{Y}_t}{Y_t}\right)\pi_t &= \frac{\varepsilon}{\theta}(\tilde{m}_t - m^*) + \dot{\pi}_t, \qquad m^* = \frac{\varepsilon - 1}{\varepsilon} \\
Y_t &= Z_t K_t^{\alpha} L_t{}^{1-\alpha} \\
T_t &= (1-m_t) Y_t - \frac{\theta}{2}\pi_t^2 Y_t + (1-\psi)\eta A_t + \left[\iota_t q_t - \iota_t - \frac{\phi^k}{2}(\iota_t - \delta)^2\right] K_t \\
di &= -\upsilon\left(i_t - \left(\rho_t^h + \phi(\pi_t - \bar{\pi}) + \bar{\pi}\right)\right) dt.
\end{aligned}
$$

B.8 Proofs of subsection 5

TFP is given by equation (33)

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

We compute the growth rate of TFP

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

where the dynamics of the density are

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

From there we can analyze two limit cases.

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

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

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

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

where

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

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

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

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 the cost-of-capital-adjusted unitary firms MRPK $\frac{\gamma\varphi_t}{q_t}$. 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

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

To uncover the sign, we analyze the term

$$\frac{\int_{z^*}^{\infty}\left(z-z^*\right)z\omega_t\left(z\right)}{\int_{z^*}^{\infty} z\omega_t\left(z\right)dz} - \frac{\int_{z^*}^{\infty}\left(z-z^*\right)\omega_t\left(z\right)}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz} = \frac{\int_{z^*}^{\infty} z^2\omega_t\left(z\right)}{\int_{z^*}^{\infty} z\omega_t\left(z\right)dz} - \frac{\int_{z^*}^{\infty} z\omega_t\left(z\right)}{\int_{z^*}^{\infty}\omega_t\left(z\right)dz}. \tag{82}$$

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

$$\frac{\int_{z^*}^{\infty} z\omega_t(z)}{\int_{z_t^*}^{\infty}\omega_t(z)\,dz} = \mathbb{E}_{\bar{\omega}_t(z)}[z] = \int_{z^*}^{\infty} z\bar{\omega}_t(z)zdz \quad < \quad \int_{z_t^*}^{\infty} z\tilde{\omega}_t(z)\,dz = \mathbb{E}_{\tilde{\omega}_t(z)}[z] = \frac{\int_{z^*}^{\infty} z^2\omega_t(z)}{\int_{z_t^*}^{\infty} z\omega_t(z)\,dz}.$$

Therefore, equation (82) 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):

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

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 $\omega(z)$ is constant and the growth rate of TFP simplifies to

$$\frac{1}{Z_t}\frac{dZ_t}{dt} = \alpha\omega(z_t^*)\frac{\int_{z_t^*}^{\infty}(z - z_t^*)\,\omega(z)\,dz}{\int_{z_t^*}^{\infty}\omega(z)\,dz\int_{z_t^*}^{\infty} z\omega(z)\,dz}\frac{dz_t^*}{dt}.$$

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 cut-off. In this case, the growth rate of TFP depends linearly with the growth rate of the cut-off: if the later increases, so does the former.

B.9 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 capital (complete markets). Note first that the baseline economy collapses to the standard complete market 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, $Z = (\mathbb{E}[z \mid z > z_t^*])^{\alpha}$. The rest of the agents (retailers, final good producers, capital producers) are identical in both economies.

Below we report the equilibrium conditions in the complete markets economy. Comparing them with those of the baseline economy reveals that they are identical up to the fact that in the baseline Z_t is endogenous (and determined by a bunch of extra equations) and up to a term in the condition equating the cost of capital R_t with the marginal return on capital.

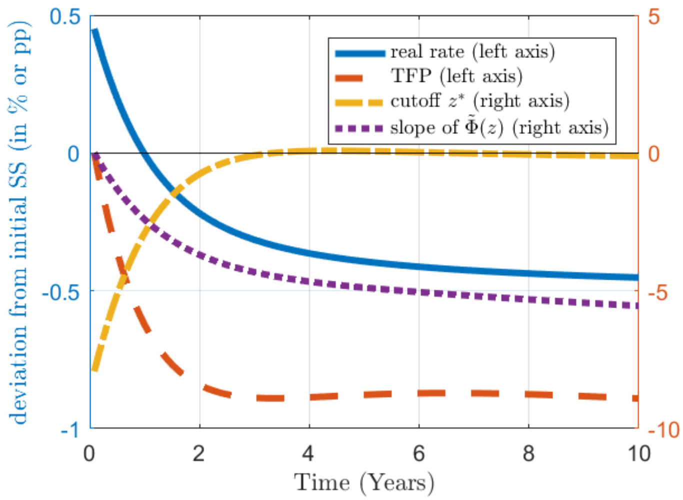
The competitive equilibrium of the complete market model with capital consists of the following 15 equations for the 15 variables $\{w, r, q, \varphi, K, L, C, D, \iota, \pi, m, \tilde{m}, i, Y, T\}$:

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

B.10 Dynamics after permanent real interest rate declines

Figure 9 displays the impulse responses to a permanent decline in the household's discount factor ρ^h, from to 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 (dashed-dotted yellow line) and the lower slope of the excess investment rate (dotted purple 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.[35]

Figure 9: Transition to a low-real-rate steady state.

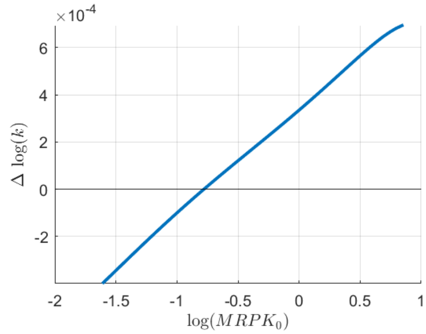


Notes: The figure shows the paths after an unexpected and permanent decline in the household's discount factor from 1% to 0.5% expressed in deviations from the initial steady state. The lines depict real rates r (solid blue), TFP Z (dashed orange), the threshold z^* (dashed-dotted yellow) and the slope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

[35]In an economy without nominal rigidities, real rates would always be below the initial value, and TFP would also fall (not shown) through the same channels.

C Additional figures

Figure 10: Firm-level capital growth as a function of initial MRPK one year after the shock.



Notes: The figure displays the average effect of an 1 b.p. expansionary monetary policy shock on the growth rate of the capital stock in the first year in p.p.. $-\ 100 * (\log k_{j,1} - \log k_{j,0})$ – as a function of the firms log MRPK before the shock $\log(MRPK_{j,0})$. We start the simulation at the steady state. To isolate the effect of monetary policy we subtract the evolution of this variable in the steady state., that is we plot $100 * \big((\log k_{j,1;MP} - \log k_{j,0}) - (\log k_{j,1;SS} - \log k_{j,0})\big)$

D Numerical Appendix

We discretize the model using a finite difference approach and compute non-linearly the responses to temporary change in parameters (an "MIT shock") using a Newton algorithm. Instead of time iterations over guesses for aggregate sequences, as is common in the literature, we use a global relaxation algorithm. This approach has been made popular in discrete-time models by Juillard et al. (1998) thanks to Dynare, but it is somewhat less common in continuous-time models (e.g. Trimborn et al., 2008). This approach 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. Recently Auclert et al. (2021) have exploited a particularly efficient variant of this approach in the context of heterogeneous-agent models.[36] We build on these contributions when we compute the optimal transition path. Again we make use of Dynare. We use its nonlinear Newton solver to compute both the steady state of the Ramsey problem and the optimal transition path under perfect foresight. To find the steady state, we provide Dynare with the steady state of the private equilibrium conditions as a function of the policy instrument.

D.1 Finite difference approximation of the Kolmogorov Forward equation

The KF equation is solved by a finite difference scheme following Achdou et al. (2021). It approximates the density $\omega_t(z)$ on a finite grid $z \in \{z_1, ..., z_J\}$, $t \in \{t_1, ..., t_N\}$ with steps Δz and time steps Δt. We use the notation $\omega_j^n := \omega_{n\Delta t}(z_j)$, $j = 1, ..., J$, $n = 0, .., N$. The KF equation is then approximated as

$$\frac{\omega_j^n - \omega_j^{n-1}}{\Delta t} = \left(s_n(z_j) - \frac{\dot{A}_n}{A_n} - (1-\psi)\eta\right)\omega_n(z_j) - \frac{\omega_j^n \mu(z_j) - \omega_{j-1}^n \mu(z_{j-1})}{\Delta z} + \frac{\omega_{j+1}^n \widetilde{\sigma}^2(z_{j+1}) + \omega_{j-1}^n \widetilde{\sigma}^2(z_{j-1}) - 2\omega_j^n \widetilde{\sigma}^2(z_j)}{2\left(\Delta z\right)^2},$$

[36] Compared to Auclert et al. (2020), who break the solution procedure into two steps, first solving for the idiosyncratic variables given the aggregate variables, we solve for the path of all aggregate and idiosyncratic variables at once. Note that, besides the nonlinear perfect foresight method we refer to here (see their Section 6), they also propose a linear method.

which, grouping, results in

$$\frac{\omega_j^n - \omega_j^{n-1}}{\Delta t} = \underbrace{\left[\left(s_n(z_j) - \frac{\dot{A}_n}{A_n} - (1-\psi)\eta\right) - \frac{\mu(z_j)}{\Delta z} - \frac{\widetilde{\sigma}^2(z_j)}{(\Delta z)^2}\right]}_{\beta_j^n}\omega_j^n$$
$$+ \underbrace{\left[\frac{\mu(z_{j-1})}{\Delta z} + \frac{\widetilde{\sigma}^2(z_{j-1})}{2\,(\Delta z)^2}\right]}_{\varrho_{j-1}^n}\omega_{j-1}^n + \underbrace{\left[\frac{\widetilde{\sigma}^2(z_{j+1})}{2\,(\Delta z)^2}\right]}_{\chi_{j+1}^n}\omega_{j+1}^n.$$

The boundary conditions are the ones associated with a reflected process z at the boundaries:[37]

$$\frac{\omega_1^n - \omega_1^{n-1}}{\Delta t} = (\beta_1^n + \chi_1^n)\,\omega_1^n + \chi_2^n \omega_{j+1}^n,$$
$$\frac{\omega_J^n - \omega_J^{n-1}}{\Delta t} = (\beta_J^n + \varrho_J^n)\,\omega_J^n + \varrho_{J-1}^n \omega_{J-1}^n.$$

If we define matrix

$$\mathbf{B}^n = \begin{bmatrix} \beta_1^n + \chi_1^n & \chi_2^n & 0 & 0 & \cdots & 0 & 0 & 0 \\ \varrho_1^n & \beta_2^n & \chi_3^n & 0 & \cdots & 0 & 0 & 0 \\ 0 & \varrho_2^n & \beta_3^n & \chi_4^n & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \varrho_{J-2}^n & \beta_{J-1}^n & \chi_J^n \\ 0 & 0 & 0 & 0 & \cdots & 0 & \varrho_{J-1}^n & \beta_J^n + \varrho_J^n \end{bmatrix},$$

then we can express the KF equation as

$$\frac{\boldsymbol{\omega}^n - \boldsymbol{\omega}^{n-1}}{\Delta t} = \mathbf{B}^{n-1}\boldsymbol{\omega}^n,$$

or

$$\boldsymbol{\omega}^n = \left(\mathbf{I} - \Delta t \mathbf{B}^{n-1}\right)^{-1} \boldsymbol{\omega}^{n-1}, \tag{83}$$

where $\boldsymbol{\omega}^n = \begin{bmatrix} \omega_1^n & \omega_2^n & \ldots & \omega_{J-1}^n & \omega_J^n \end{bmatrix}^T$, and $\mathbf{I}$ is the identity matrix of dimension J.

[37]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.

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.[38]

Be $z = \begin{bmatrix} z_1, & z_2, & \dots & z_{J-1} & z_J \end{bmatrix}$ the grid. Define $\Delta z_{a,b} = z_b - z_a$ and let $\Delta z = \frac{1}{2}\begin{bmatrix} \Delta z_{1,2}, & \Delta z_{1,3}, & \Delta z_{2,4}, & \dots, & \Delta z_{J-2,J} & \Delta z_{J-1,J} \end{bmatrix}$. We approximate the KFE (27) using central difference for both the first derivative and the second derivative.

$$
\begin{aligned}
\frac{\omega_j^n - \omega_j^{n-1}}{\Delta t} &= \left(s_n(z) - (1-\psi)\eta - \frac{\dot{A}_n}{A_n}\right)\omega^n(z_j) - \left[\frac{\mu(z_{j+1})\omega^n(z_{j+1}) - \mu(z_{j-1})\omega^n(z_{j-1})}{\Delta z_{j-1,j+1}}\right] \\
&\quad + \frac{1}{2}\frac{\Delta z_{j-1,j}\sigma^2(z_{j+1})\omega^n(z_{j+1}) + \Delta z_{j,j+1}\sigma^2(z_{j-1})\omega^n(z_{j-1}) - \Delta z_{j-1,j+1}\sigma^2(z_j)\omega^n(z_j)}{\frac{1}{2}(\Delta z_{j-1,j+1})\Delta z_{j,j+1}\Delta z_{j-1,j}}
\end{aligned}
$$

which, grouping, results in

$$
\begin{aligned}
\frac{\omega_j^n - \omega_j^{n-1}}{\Delta t} &= \underbrace{\left[\left(s_n(z) - (1-\psi)\eta - \frac{\dot{A}_n}{A_n}\right) - \frac{\sigma^2(z_j)}{\Delta z_{j,j+1}\Delta z_{j-1,j}}\right]}_{\beta_j^n}\omega^n(z_j) \\
&\quad + \underbrace{\left[\frac{\mu(z_{j-1})}{\Delta z_{j-1,j+1}} + \frac{\sigma^2(z_{j-1})}{(\Delta z_{j-1,j+1})\Delta z_{j,j+1}}\right]}_{\varrho_{j-1}^n}\omega^n(z_{j-1}) \\
&\quad + \underbrace{\left[-\frac{\mu(z_{j+1})}{\Delta z_{j-1,j+1}} + \frac{\sigma^2(z_{j+1})}{(\Delta z_{j-1,j+1})\Delta z_{j,j+1}}\right]}_{\chi_{j+1}^n}\omega^n(z_{j+1}).
\end{aligned}
$$

The boundary conditions are the ones associated with a reflected process z at the

[38] Our approach builds on the one in the appendix to 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.

boundaries:

$$
\begin{aligned}
\frac{\omega_1^n - \omega_1^{n-1}}{\Delta t} &= \left(\beta_1^n + \chi_1^n\right)\omega_n(z_1) + \chi_2^n \omega_{j+1}^n, \\
\frac{\omega_J^n - \omega_J^{n-1}}{\Delta t} &= \left(\beta_J^n + \varrho_J^n\right)\omega_n(z_J) + \varrho_{J-1}^n \omega_{j-1}^n.
\end{aligned}
$$

where we define $\Delta z_{0,1} \equiv \Delta z_{1,2}$ 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

$$\frac{\boldsymbol{\omega}^n - \boldsymbol{\omega}^{n-1}}{\Delta t} = \mathbf{B}^{n-1}\boldsymbol{\omega}^n$$

where

$$
\mathbf{B}^n = \begin{bmatrix}
\beta_1^n + \chi_1^n & \chi_2^n & 0 & 0 & \cdots & 0 & 0 & 0 \\
\varrho_1^n & \beta_2^n & \chi_3^n & 0 & \cdots & 0 & 0 & 0 \\
0 & \varrho_2^n & \beta_3^n & \chi_4^n & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & \varrho_{J-2}^n & \beta_{J-1}^n & \chi_J^n \\
0 & 0 & 0 & 0 & \cdots & 0 & \varrho_{J-1}^n & \beta_J^n + \varrho_J^n
\end{bmatrix},
$$

Abstracting for brevity from the term $\left(s_n(z) - (1-\psi)\eta - \frac{\dot{A}_n}{A_n}\right)$, which is independent of the grid, and spelling out $\mathbf{B}^n$ we have

$$
\frac{\boldsymbol{\omega}^n - \boldsymbol{\omega}^{n-1}}{\Delta t} = \begin{bmatrix}
\frac{\sigma^2(z_1)}{\Delta z_{1,2}\Delta z_{1,2}} - \frac{\mu(z_1)}{\Delta z_{1,2}} - \frac{2\sigma^2(z_1)}{\Delta z_{1,2}\Delta z_{1,2}} & -\frac{\mu(z_2)}{\Delta z_{1,2}} + \frac{\sigma^2(z_2)}{\Delta z_{1,2}\Delta z_{1,2}} & 0 & \cdots \\
\frac{\mu(z_1)}{\Delta z_{1,3}} + \frac{\sigma^2(z_1)}{\Delta z_{1,3}\Delta z_{1,2}} & -\frac{\sigma^2(z_2)}{\Delta z_{1,2}\Delta z_{2,3}} & -\frac{\mu(z_3)}{\Delta z_{1,3}} + \frac{\sigma^2(z_3)}{\Delta z_{1,3}\Delta z_{2,3}} & \cdots \\
0 & \frac{\mu(z_2)}{\Delta z_{2,4}} + \frac{\sigma^2(z_2)}{\Delta z_{2,4}\Delta z_{2,3}} & -\frac{\sigma^2(z_3)}{\Delta z_{2,3}\Delta z_{3,4}} & \cdots \\
0 & 0 & \frac{\mu(z_3)}{\Delta z_{3,5}} + \frac{\sigma^2(z_3)}{\Delta z_{3,4}\Delta z_{3,5}} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix} \boldsymbol{\omega}^n.
$$

We can rewrite this as follows

$$\frac{\boldsymbol{\omega}^n-\boldsymbol{\omega}^{n-1}}{\Delta t}=\begin{bmatrix} -\frac{\mu(z_1)}{\boldsymbol{\Delta z_{1,2}}}-\frac{\sigma^2(z_1)}{\mathbf{\Delta z_{1,2}}\Delta z_{1,2}} & -\frac{\mu(z_2)}{\boldsymbol{\Delta z_{1,2}}}+\frac{\Delta z_{2,3}\sigma^2(z_2)}{\Delta z_{2,3}(\mathbf{\Delta z_{1,2}}\Delta z_{1,2})} & 0 & \cdots \\ \frac{\mu(z_1)}{\boldsymbol{\Delta z_{1,3}}}+\frac{\sigma^2(z_1)}{\mathbf{\Delta z_{1,3}}\Delta z_{1,2}} & -\frac{(\Delta z_{1,2}+\Delta z_{2,3})\sigma^2(z_2)}{\mathbf{\Delta z_{1,3}}(\Delta z_{1,2}\Delta z_{2,3})} & -\frac{\mu(z_3)}{\boldsymbol{\Delta z_{1,3}}}+\frac{\Delta z_{3,4}\sigma^2(z_3)}{\Delta z_{3,4}(\mathbf{\Delta z_{1,3}}\Delta z_{2,3})} & \cdots \\ 0 & \frac{\mu(z_2)}{\boldsymbol{\Delta z_{2,4}}}+\frac{\Delta z_{1,2}\sigma^2(z_2)}{\Delta z_{1,2}(\mathbf{\Delta z_{2,4}}\Delta z_{2,3})} & -\frac{(\Delta z_{2,3}+\Delta z_{3,4})\sigma^2(z_3)}{\mathbf{\Delta z_{2,4}}(\Delta z_{2,3}\Delta z_{3,4})} & \cdots \\ 0 & 0 & \frac{\mu(z_3)}{\boldsymbol{\Delta z_{3,5}}}+\frac{\Delta z_{2,3}\sigma^2(z_3)}{\Delta z_{2,3}(\Delta z_{3,4}\mathbf{\Delta z_{3,5}})} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}\boldsymbol{\omega}^n.$$

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 $\Delta z\mathbf{B}^n$ yields a vector of ones and the operation is mass preserving, in the sense that the above relationship guarantees that

$$\sum_j \omega_j^n \Delta z_j=\sum_j \omega_j^{n-1}\Delta z_j=1$$

where $\sum_j \omega_j^n \Delta z_j$ is a trapezoid approximation of the integral $\int \omega^n(z)dz$.

D.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 $z_j \le \bar{z} \le z_{j+1}$ then the integral from the closest lower gridpoint is given by

$$\int_{z_j}^{\bar{z}} f(z)\,dz=\left[f(z_j)+\frac{1}{2}\left[f(z_{j+1})-f(z_j)\right]\frac{\bar{z}-z_j}{z_{j+1}-z_j}\right](\bar{z}-z_j)$$

We use this formula to construct the integrals over a larger range piecewise. For example:

$$\int_{z_1}^{z_N} f(z)\,dz=\begin{bmatrix}\frac{1}{2} & 1 & 1 & \cdots & 1 & \frac{1}{2}\end{bmatrix}\begin{bmatrix} f(z_1)\\ f(z_2)\\ \vdots \\ f(z_N)\end{bmatrix}$$

and

$$\int_{z_1}^{z^*} f(z)\,dz = \begin{bmatrix} \frac{1}{2} & 1 & 1 & \cdots & 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} f(z_1) \\ f(z_2) \\ \vdots \\ f(z_{j^*-1}) \end{bmatrix}$$
$$+ \left[f(z_{j^*-1}) + \frac{1}{2}\left[f(z_{j^*}) - f(z_{j^*-1}) \right] \frac{z^* - z_{j^*-1}}{z_{j^*} - z_{j^*-1}} \right] (z^* - z_{j^*-1})$$
$$\text{where } j^* = \arg\min_j \{ j \leq J | z_{j^*} > z^* \}$$

D.3 Algorithm to solve for the SS

Here we present how to solve for the SS of the private equilibrium, that is for the SS when the central bank sets a certain level of the nominal interest rate in SS i^{ss}.

We know that in SS consumption does not grow, hence from (14)

$$r^{ss} = \rho^h. \tag{84}$$

We also know that in SS, the investment rate is equal to the depreciation,

$$\iota^{ss} = \delta. \tag{85}$$

This means that, from equation (17) and the functional form we assumed for the capital adjustment costs

$$(q_t - 1 - \Phi'(\iota_t))(r_t - (\iota_t - \delta)) = \dot{q}_t - \Phi''(\iota_t)\, \dot{\iota}_t - (q_t \iota_t - \iota_t - \Phi(\iota_t)) \tag{86}$$

$$\left(q^{ss} - 1 - \phi^k(\iota^{ss} - \delta)\right)\left(\rho^{hh} - (\iota^{ss} - \delta)\right) = 0 - \phi^k * 0 - \left(q^{ss}\iota^{ss} - \iota^{ss} - \phi^k(\iota^{ss} - \delta)\right)$$

$$\rho^{hh}(q^{ss} - 1) = \delta(1 - q^{ss})$$

.From here we can solve for the steady state value of q^{ss}, which is given by

$$q^{ss} = 1. \tag{87}$$

Furthermore, combining (84) with the fisher equation and the fact that the planner sets a certain nominal rate i^{ss} we get that

$$\pi^{ss} = i^{ss} - \rho^h. \tag{88}$$

In SS, $\dot{\pi}_t = 0$ and $\dot{Y}_t = 0$. Hence, from equation (21) we obtain

$$m^{ss} = \left(m^* + \rho^h \pi^{ss} \frac{\theta}{\varepsilon} \right). \tag{89}$$

Using equation (35) and (84),

$$\rho^h = \frac{1}{q^{ss}} \left(\alpha m_t Z_t A_t^{\alpha-1} L^{1-\alpha} \frac{z_t^*}{\gamma X_t} \right) - \delta \tag{90}$$

From equation (36) and (84),

$$\frac{\dot{A}_t}{A_t} = 0 = \frac{1}{q_t}(\alpha m_t Z_t {A_t}^{\alpha-1} {L_t}^{1-\alpha} - R_t(1-\Omega(z_t^*)) + R_t - \delta q_t - q_t(1-\psi)\eta). \tag{91}$$

Plugging the latter equation into the former, using $q^{SS} = 1$ and using the definition of r_t we obtain:

$$\rho^h + \delta = \left[(\rho^h + \delta) \left(\gamma(1 - \Omega(z_t^*)) - 1 \right) + (1-\psi)\eta + \delta \right] \frac{z^*}{\gamma X^*}. \tag{92}$$

In the algorithm, we use a non-linear equation solver to obtain z^* from this equation. *The Algorithm.*

- Get $r^{ss} = \rho^h$, $\pi^{ss} = \bar{\pi}$ and $i^{ss} = \rho^h + \pi^{ss}$ and $R^{ss} = q^{ss}(\rho^h + \delta)$ and $m^{ss} = m^* + \rho^h \pi^{ss} \frac{\theta}{\epsilon}$.
- Given that our calibration target for $L^{ss} = 1$, we "guess" $L^{ss} = 1$
- 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 (92) to obtain z^* from equation (92).
    2. Obtain $Z_n = (\gamma_n X_n^*)^{\alpha}$.

3. Find A from equation (34),

$$A^n = \left[\frac{q^{ss}\rho^h + \delta q^{ss}}{\alpha m_n Z_n {L_m}^{1-\alpha} \frac{z_t^*}{\gamma X_t}}\right]^{\frac{1}{\alpha-1}}.$$

4. Find the stocks $K_n = \gamma(1 - \Omega^n(z^*))A^n$, $D_n = K_n - A_n$.
5. Compute $w_n = (1-\alpha)m^{ss} Z_n {A_n}^{\alpha} {L_n}^{-\alpha}$,$\varphi_n = \alpha\left(\frac{(1-\alpha)}{w_n}\right)^{(1-\alpha)/\alpha} {m^{ss}}^{\frac{1}{\alpha}}$.
6. Get aggregate output $Y = Z_n A_n^{\alpha} {L_n}^{1-\alpha}$, transfers $T_n = (1 - m^{ss}) Y_n - \frac{\theta}{2}(\pi^{ss})^2 Y_n + (1-\psi)\eta A_t$, and consumption $C_n = w_n L_m + r^{ss} D_n + T_n$.
7. Update $\hat{s}_j^n = \frac{1}{q^{ss}}(\gamma \max\{z\varphi_n - R_n, 0\} + R_n - \delta q^{ss})$ and employ it to construct matrix $\mathbf{B}^{n-1}$.
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. If the net worth distribution do not coincide with the guess, set $n = n + 1$ and return to point 1

- Set $\Upsilon = \left(w_{L=1} C_{L=1}^{-\eta}\right)$ to ensure our "guess" for L^{ss} is correct.

E Computing optimal policies in heterogeneous-agent models

E.1 General algorithm

Solving for the optimal policy in models with heterogeneous agents poses a certain challenge since the state in such a model contains a distribution, which is an infinite-dimensional object. In this section, we explain how such models can be solved in a relatively straightforward manner. Our approach relies on three main conceptual ingredients: (i) finite difference approximation of continuous time and continuous idiosyncratic states, (ii) symbolic derivation of the planner's first-order conditions, and (iii) use of a Newton algorithm to solve the optimal policy problem non-linearly in the sequence space. Here we present a general overview which goes beyond the particular model presented in the paper.

(i) Finite difference approximation A continuous-time, continuous-space heterogeneous-agent model discretized using an upwind finite-difference method becomes a discrete-time, discrete-space model. In this discretized model the dynamics of the (now finite-dimensional) distribution $\boldsymbol{\mu}_t$ at period t are given by

$$\boldsymbol{\mu}_t = \left(\mathbf{I} - \Delta t \boldsymbol{A}_t^T\right)^{-1} \boldsymbol{\mu}_{t-1}, \tag{93}$$

where Δt is the time step between periods and $\boldsymbol{A}_\mathbf{t}$ is a matrix whose entries depend *nonlinearly* and in *closed form* on the idiosyncratic and aggregate variables in period t.[39] Similarly, the HJB equation is approximated as[40]

$$\rho \mathbf{v}_{t+1} = \boldsymbol{u}_{t+1} + \mathbf{A}_{t+1}\mathbf{v}_{t+1,} - \left(\mathbf{v}_{t+1} - \mathbf{v}_t\right)/\Delta t. \tag{94}$$

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, they define the

[39]Technically, this matrix results from the discretization of the *infinitesimal generator* of the idiosyncratic states. In the example of Section 2, $\boldsymbol{\mu}_t = \boldsymbol{\omega}_\mathbf{t}$ and $\boldsymbol{A}_t = \mathbf{B}_\mathbf{t}$.

[40]In the model presented in this paper the HJB can be solved analytically and hence there is no need to solve it computationally.

discretized model.

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 (93) and Bellman equation (94) 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 Π_t sparse.[41] 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).

(ii) Symbolic derivation of planner's FOCs Once we have a finite-dimensional discrete-time discrete-space model, we can derive the planner's FOCs by *symbolic differentiation* using standard software packages. For convenience, we rely on Dynare's toolbox for Ramsey optimal policy to do this task for us. To this end, we simply provide the discretized version of our model's private equilibrium conditions to Dynare (the discretized counterpart to the equations in Appendix B.7), making use of loops for the heterogeneous-agent block, as in Winberry (2018). We furthermore provide the discretized objective function, and Dynare then takes symbolic derivatives to construct the set of optimality conditions of the planner for us.

A natural question at this stage is under which conditions the optimal policies of the discrete-time, discrete-space problem coincide with those of the original problem. The following proposition shows that, if the time interval is small enough (the standard condition when approximating continuous-time models), then the two solutions coincide.

Proposition E.1 : *Provided that all the Lagrange multipliers associated to the equilibrium conditions are continuous for* $t > 0$*, the solution of the "discretize-optimize" and the "optimize-discretize" algorithms converge to each other as the time step* Δt *goes towards 0.*

Proof: See Appendix E.2.

The proposition guarantees that both strategies coincide when Δt goes towards zero. This proposition is quite general, as most continuous-time, perfect-foresight, general

[41] The introduction of Poisson shocks would not change the sparsity of matrix Π_t.

equilibrium models do not feature discontinuities for $t > 0$.

The model presented in Section 2 is arguably simpler than the general heterogenous-agent model covered by Proposition 1, as it features an analytic solution for the HJB equation. To get an idea of the performance of our method in a case in which the HJB is also a constraint in the planner's problem, as well as to showcase its generality in dealing with different problems, we compute the optimal monetary policy in the HANK model of Nuño and Thomas (2022) using our method in Dynare (see Appendix E.3). We compare our results with those using their "optimize-discretize" algorithm at monthly frequency $\triangle t = 1/12$. We conclude that both approaches essentially coincide.

(iii) 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 change in parameters (an "MIT shock") using a Newton algorithm. Instead of time iterations over guesses for aggregate sequences, as is common in the literature, we use a global relaxation algorithm. This approach has been made popular in discrete-time models by Juillard et al. (1998) thanks to Dynare, but it is somewhat less common in continuous-time models (e.g. Trimborn et al., 2008). This approach 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. Recently Auclert et al. (2021) have exploited a particularly efficient variant of this approach in the context of heterogeneous-agent models.[42] We build on these contributions when we compute the optimal transition path. Again we make use of Dynare. We use its nonlinear Newton solver to compute both the steady state of the Ramsey problem and the optimal transition path under perfect foresight.[43] Our hope is that the convenience of using Dynare will make optimal policy problems in heterogeneous-agent models easily accessible to a large audience of researchers.

The solution to the perfect foresight problem can be easily adapted to the case

[42]Compared to Auclert et al. (2020), who break the solution procedure into two steps, first solving for the idiosyncratic variables given the aggregate variables, we solve for the path of all aggregate and idiosyncratic variables at once. Note that, besides the nonlinear perfect foresight method we refer to here (see their Section 6), they also propose a linear method.

[43]To find the steady state, we provide Dynare with the steady state of the private equilibrium conditions as a function of the policy instrument.

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. We follow this approach to analyze the optimal response to a cost-push shock below.

As discussed in Section 4, our solution approach 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. These methods, however, were not created to deal with the problem of finding the optimal policies, the focus of our algorithm, as the first stage requires the computation of the steady state, which in our case is the steady state of the problem under optimal policies. Our algorithm finds the steady state of the planner's problem, including the Lagrange multipliers. Naturally, this steady does not need to coincide with the steady state that can be found by looking for the value of the planner's policy that maximizes steady-state welfare.

E.2 Proof of proposition E.1

Proof: 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. Second, we derive the continuous time optimality conditions of the planner's problem and discretize them. Third, we discretize the planners problem and the derive the optimality conditions. Fourth, we compare the two sets of discretized optimality conditions.

1. The generic problem The planner's problem in an economy with heterogeneity among one agent type (e.g. households or firms) can be written as

$$
\begin{aligned}
\max_{Z_t,u_t(x),\mu_t(x),v_t(x)} & & \int_0^\infty \exp(-\varrho t) f_0(Z_t) dt & \quad (95)\\
& s.t.\ \forall t & & \\
\dot{X}_t & = & f_1(Z_t) & \quad (96)\\
\dot{U}_t & = & f_2(Z_t) & \quad (97)\\
0 & = & f_3(Z_t) & \quad (98)\\
\tilde{U}_t & = & \int f_4\left(x, u_t(x), Z_t\right) \mu_t\left(x\right)\ dx & \quad (99)\\
\rho v_t(x) & = & \dot{v}_t(x) + f_5(x, u_t(x), Z_t) & \quad (100)\\
& & + \sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right) \frac{\partial v_t(x)}{\partial x_i} + \sum_{i=1}^{I}\sum_{k=1}^{I} \frac{\left(\sigma(x)\sigma(x)^\top\right)_{i,k}}{2} \frac{\partial^2 v_t(x)}{\partial x_i \partial x_k},\ \forall x & \\
0 & = & \frac{\partial f_5}{\partial u_{j,t}} + \sum_{i=1}^{I} \frac{\partial b_i}{\partial u_{j,t}} \frac{\partial v_t(x)}{\partial x_i}, \quad j = 1, ..., J, \forall x. & \quad (101)\\
\dot{\mu}_t\left(x\right) & = & -\sum_{i=1}^{I} \frac{\partial}{\partial x_i}\left[b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\right] & \quad (102)\\
& & +\frac{1}{2}\sum_{i=1}^{I}\sum_{k=1}^{I} \frac{\partial^2}{\partial x_i \partial x_k}\left[\left(\sigma(x)\sigma(x)^\top\right)_{i,k}\mu_t\left(x\right)\right],\ \forall x & \\
X_0 & = & \bar{X}_0 & \quad (103)\\
\mu_0\left(x\right) & = & \bar{\mu}_0\left(x\right) & \quad (104)\\
\lim_{t\to\infty} U & = & \bar{U}_\infty & \quad (105)\\
\lim_{t\to\infty} v(x) & = & \bar{v}(x)_\infty & \quad (106)
\end{aligned}
$$

where we have adopted the following notation:

- Variables (capitals are reserved for aggregate variables):
  - x individual state vector with I elements
  - 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)
- $\hat{U}$ aggregate control vector of purely contemporaneous variables
- U aggregate control vector of intertemporal variables
- $\tilde{U}$ 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 (95) is the planner's objective function.[44] Equations (96)-(98) are the aggregate equilibrium conditions for aggregate states, jump variables and contemporaneous variables. In our model, examples for each of these three types of equations are the law of motion of aggregate capital, the household's Euler equation and the household's labor supply condition, respectively. Equation (99) links aggregate and individual variables, such as the definition of aggregate TFP in our model. Equations (100) and (101) 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 (25) determines the evolution of the distribution of agents. Finally (103)-(106) 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.

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.

[44] Notice that the planner's discount factor, ϱ, can be different to that of individual agents, ρ.

2.a The Lagrangian The Lagrangian for this problem is given by:[45]

$$
\begin{aligned}
\mathcal{L} &= \int_0^\infty \Big\{ e^{-\varrho t}\, f_0(Z_t) \\
&+ \lambda_{1,t}\left(\dot{X}_t - f_1(Z_t)\right) \\
&+ \lambda_{2,t}\left(\dot{U}_t - f_2(Z_t)\right) \\
&+ \lambda_{3,t}\left(f_3(Z_t)\right) \\
&+ \lambda_{4,t}\left(\tilde{U}_t - \int f_4\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\; dx\right) \\
&+ \int \left[\lambda_{5,t}(x)\left(-\rho v_t(x) + \dot{v}_t(x) + f_5(x, u_t(x), Z_t) + \sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right)\frac{\partial v_t(x)}{\partial x_i} + \sum_{i=1}^{I}\frac{\sigma_i^2(x)}{2}\frac{\partial^2 v_t(x)}{\partial^2 x_i}\right)\right] dx \\
&+ \sum_{j=1}^{J}\int\left[\lambda_{6,j,t}(x)\left(\frac{\partial f_5}{\partial u_{j,t}} + \sum_{i=1}^{I}\frac{\partial b_i}{\partial u_{j,t}}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx \\
&+ \int\left[\lambda_{7,t}(x)\left(-\dot{\mu}_t\left(x\right) + \left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\right] + \frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\mu_t\left(x\right)\right]\right)\right)\right] dx \Big\}\, dt
\end{aligned}
$$

where λ_1 to λ_7 denote the multipliers on the respective constraints. For convenience,

[45] 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.

we write the time derivatives in a separate line at the end. The Lagrangian becomes:

$$
\begin{aligned}
\mathcal{L} &= \int_0^\infty \{e^{-\varrho t}\, f_0(Z_t) \\
&+ \lambda_{1,t}\left(-f_1(Z_t)\right) \\
&+ \lambda_{2,t}\left(-f_2(Z_t)\right) \\
&+ \lambda_{3,t}\left(-f_3(Z_t)\right) \\
&+ \lambda_{4,t}\left(\tilde{U}_t - \int f_4\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\; dx\right) \\
&+ \int\left[\lambda_{5,t}(x)\left(-\rho v_t(x) + f_5(x, u_t(x), Z_t) + \sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right)\frac{\partial v_t(x)}{\partial x_i} + \sum_{i=1}^{I}\frac{\sigma_i^2(x)}{2}\frac{\partial^2 v_t(x)}{\partial^2 x_i}\right)\right] dx \\
&+ \sum_{j=1}^{J}\int\left[\lambda_{6,j,t}(x)\left(\frac{\partial f_{5,t}}{\partial u_{j,t}} + \sum_{i=1}^{I}\frac{\partial b_i}{\partial u_{j,t}}\frac{\partial v_t(x)}{\partial x_i}\right)\right] dx \\
&+ \int\left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\right] + \frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\mu_t\left(x\right)\right]\right)\right] dx\Bigg\}\, dt \\
&+ \int_0^\infty\left\{e^{-\varrho t}\;\lambda_{1,t}\dot{X}_t + \lambda_{2,t}\dot{U}_t + \int\left[\lambda_{5,t}\dot{v}_t(x)\right] dx - \int\left[\lambda_{7,t}\dot{\mu}_t\left(x\right)\right] dx\right\} dt.
\end{aligned}
$$

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

$$
\begin{aligned}
& \lim_{t\to\infty} e^{-\varrho t}\lambda_{1,t}X_t - \lambda_{1,0}X_0 - \int_0^\infty e^{-\varrho t}(\dot\lambda_{1,t} - \varrho\lambda_{1,t})X_t dt + \lim_{t\to\infty} e^{-\varrho t}\lambda_{2,t}U_t - \lambda_{2,0}U_0 \\
& - \int_0^\infty e^{-\varrho t}(\dot\lambda_{2,t} - \varrho\lambda_{2,t})U_t dtx \\
+ \quad & \int \left(\lim_{t\to\infty} e^{-\varrho t}\lambda_{5,t}(x)v_t(x) - \lambda_{5,0}(x)v_0(x)\right) dx - \int\int_0^\infty e^{-\varrho t}(\dot\lambda_{5,t}(x) - \varrho\lambda_{5,t}(x))v_t(x)dtdx \\
- \quad & \int \lim_{t\to\infty} e^{-\varrho t}\lambda_{7,t}(x)\mu_t(x) - \lambda_{7,0}(x)\mu_0(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot\lambda_{7,t}(x) - \varrho\lambda_{7,t}(x))\mu_t\left(x\right)dtdx
\end{aligned}
$$

Now we use the initial and terminal conditions to drop some $\lim_{t\to\infty}$ and $t = 0$ terms,

$$
\begin{aligned}
+ \quad & \lim_{t\to\infty} e^{-\varrho t}\lambda_{1,t}X_t - \lambda_{2,0}U_0 - \int_0^\infty e^{-\varrho t}(\dot\lambda_{1,t} - \varrho\lambda_{1,t})X_t dt - \int_0^\infty e^{-\varrho t}(\dot\lambda_{2,t} - \varrho\lambda_{2,t})U_t dt \\
- \quad & \int \lambda_{5,0}(x)v_0(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot\lambda_{5,t}(x) - \varrho\lambda_{5,t}(x))v_t(x)dtdx \\
- \quad & \int \lim_{t\to\infty} e^{-\varrho t}\lambda_{7,t}(x)\mu_t(x)dx + \int\int_0^\infty e^{-\varrho t}(\dot\lambda_{7,t}(x) - \varrho\lambda_{7,t}(x))\mu_t\left(x\right)dtdx
\end{aligned}
$$

Next we integrate lines 6 to 8 by parts with respect to x. This yields:

$$
\begin{aligned}
+ \quad & \int\Bigg\{\left[\left(-\rho\lambda_{5,t}(x)v_t(x) + f_5(x,u_t(x),Z_t) - \sum_{i=1}^{I}\frac{\partial b_i\left(x,u_t(x),Z_t\right)\lambda_{5,t}(x)}{\partial x_i}v_t(x)\right)\right]dx \\
& + \int\left[\left(+\frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\lambda_{5,t}(x)\right]v_t(x)\right)\right]dx \\
+ \quad & \sum_{j=1}^{J}\int\left[\lambda_{6,j,t}(x)\frac{\partial f_{5,t}}{\partial u_{j,t}} - \sum_{i=1}^{I}\frac{\partial\left[\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}\right]}{\partial x_i}v_t(x)\right]dx \\
+ \quad & \int\left[\left(\sum_{i=1}^{I}\frac{\partial\lambda_{7,t}(x)}{\partial x_i}\left[b_i\left(x,u_t(x),Z_t\right)\mu_t\left(x\right)\right] + \sum_{i=1}^{I}\frac{\partial^2\lambda_{7,t}(x)}{\partial^2 x_i}\frac{\sigma_i^2(x)}{2}\mu_t\left(x\right)\right)\right]dx\Bigg\}dt
\end{aligned}
$$

Putting this all together the Lagrangian has become:

$$
\begin{aligned}
\mathcal{L} &= \int_0^{\infty} \left\{ e^{-\varrho t}\, f_0(Z_t) \right. \\
&+ \lambda_{1,t}\left(-f_1(Z_t)\right) \\
&+ \lambda_{2,t}\left(-f_2(Z_t)\right) \\
&+ \lambda_{3,t}\left(-f_3(Z_t)\right) \\
&+ \lambda_{4,t}\left(\tilde{U}_t - \int f_4\left(x, u_t(x), Z_t\right) \mu_t\left(x\right)\ dx\right) \\
&+ \int \left(-\rho\lambda_{5,t}(x)v_t(x) + \lambda_{5,t}(x) f_5(x, u_t(x), Z_t) - \sum_{i=1}^{I} \frac{\partial\left[b_i\left(x, u_t(x), Z_t\right)\lambda_{5,t}(x)\right]}{\partial x_i} v_t(x)\right) dx \\
&+ \int \left(\frac{1}{2}\sum_{i=1}^{I} \frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\lambda_{5,t}(x)\right] v_t(x)\right) dx \\
&+ \sum_{j=1}^{J} \int \left[\lambda_{6,j,t}(x)\frac{\partial f_{5,t}}{\partial u_{j,t}} - \sum_{i=1}^{I} \frac{\partial\left[\lambda_{6,j,t}(x)\frac{\partial b_i}{\partial u_{j,t}}\right]}{\partial x_i} v_t(x)\right] dx \\
&+ \left. \int_0^{\infty} \left[\left(\sum_{i=1}^{I} \frac{\partial\lambda_{7,t}(x)}{\partial x_i}\left[b_i\left(x, u_t(x), Z_t\right)\mu_t\left(x\right)\right] + \sum_{i=1}^{I} \frac{\partial^2\lambda_{7,t}(x)}{\partial^2 x_i}\frac{\sigma_i^2(x)}{2}\mu_t\left(x\right)\right)\right] dx \right\} dt \\
&+ \lim_{t\to\infty} e^{-\varrho t}\lambda_{1,t}X_t - \lambda_{2,0}U_0 - \int_0^{\infty} e^{-\varrho t}(\dot{\lambda}_{1,t} - \varrho\lambda_{1,t})X_t dt - \int_0^{\infty} e^{-\varrho t}(\dot{\lambda}_{2,t} - \varrho\lambda_{2,t})U_t dt \\
&+ \int -\lambda_{5,0}(x)v_0(x)dx + \int\int_0^{\infty} e^{-\varrho t}(\dot{\lambda}_{5,t}(x) - \varrho\lambda_{5,t}(x))v_t(x)dtdx \\
&- \int \lim_{t\to\infty} e^{-\varrho t}\lambda_{7,t}(x)\mu_t(x)dx + \int\int_0^{\infty} e^{-\varrho t}(\dot{\lambda}_{7,t}(x) - \varrho\lambda_{7,t}(x))\mu_t\left(x\right)dtdx.
\end{aligned}
$$

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:

$$
\begin{aligned}
U_t: \quad 0 = \quad & -(\dot{\lambda}_{2,t} - \varrho\lambda_{2,t}) && (107)\\
& +\frac{\partial f_{0,t}}{\partial U_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial U_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial U_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial U_t} - \lambda_{4,t}\int \frac{\partial f_{4,t}}{\partial U_t}\mu_t\left(x\right)\ dx && (108)\\
& +\int\left[\lambda_{5,t}(x)\left(\frac{\partial f_{5,t}}{\partial U_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial U_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right]dx && (109)\\
& +\sum_{j=1}^{J}\int\left[\lambda_{6,j,t}(x)\left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t}\partial U_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial u_{j,t}\partial U_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right]dx && (110)\\
& +\int\left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[\frac{\partial b_{i,t}}{\partial U_t}\mu_t\left(x\right)\right]\right)\right]dx, && (111)\\
& \forall t > 0, && (112)
\end{aligned}
$$

$$
0 = \quad \lambda_{2,0}. \tag{113}
$$

$$
\begin{aligned}
X_t: \quad 0 = \quad & -(\dot{\lambda}_{1,t} - \varrho\lambda_{1,t})\\
& +\frac{\partial f_{0,t}}{\partial X_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial X_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial X_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial X_t} - \lambda_{4,t}\int \frac{\partial f_{4,t}}{\partial X_t}\mu_t\left(x\right)\ dx\\
& +\int\left[\lambda_{5,t}(x)\left(\frac{\partial f_{5,t}}{\partial X_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial X_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right]dx\\
& +\sum_{j=1}^{J}\int\left[\lambda_{6,j,t}(x)\left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t}\partial X_t} + \sum_{i=1}^{I}\frac{\partial b_{i,t}}{\partial u_{j,t}\partial X_t}\frac{\partial v_t(x)}{\partial x_i}\right)\right]dx\\
& +\int\left[\lambda_{7,t}(x)\left(-\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left[\frac{\partial b_{i,t}}{\partial X_t}\mu_t\left(x\right)\right]\right)\right]dx,\\
& \forall t \geq 0,
\end{aligned}
$$

$$
0 = \quad \lim_{t\to\infty} e^{-\varrho t}\lambda_{1,t}(x).
$$

$$
\begin{aligned}
\hat{U}_t: \quad 0 = \quad & 0 \\
& + \frac{\partial f_{0,t}}{\partial \hat{U}_t} - \lambda_{1,t} \frac{\partial f_{1,t}}{\partial \hat{U}_t} - \lambda_{2,t} \frac{\partial f_{2,t}}{\partial \hat{U}_t} - \lambda_{3,t} \frac{\partial f_{3,t}}{\partial \hat{U}_t} - \lambda_{4,t} \int \frac{\partial f_{4,t}}{\partial \hat{U}_t} \mu_t\left(x\right) \; dx \\
& + \int \left[\lambda_{5,t}(x) \left(\frac{\partial f_{5,t}}{\partial \hat{U}_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}}{\partial \hat{U}_t} \frac{\partial v_t(x)}{\partial x_i} \right) \right] dx \\
& + \sum_{j=1}^{J} \int \left[\lambda_{6,j,t}(x) \left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t} \partial \hat{U}_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}}{\partial u_{j,t} \partial \hat{U}_t} \frac{\partial v_t(x)}{\partial x_i} \right) \right] dx \\
& + \int \left[\lambda_{7,t}(x) \left(- \sum_{i=1}^{I} \frac{\partial}{\partial x_i} \left[\frac{\partial b_{i,t}}{\partial \hat{U}_t} \mu_t\left(x\right) \right] \right) \right] dx, \\
& \forall t \geq 0.
\end{aligned}
$$

$$
\begin{aligned}
\tilde{U}_t: \quad 0 = \quad & \lambda_{4,t} \\
& + \frac{\partial f_{0,t}}{\partial \tilde{U}_t} - \lambda_{1,t} \frac{\partial f_{1,t}}{\partial \tilde{U}_t} - \lambda_{2,t} \frac{\partial f_{2,t}}{\partial \tilde{U}_t} - \lambda_{3,t} \frac{\partial f_{3,t}}{\partial \tilde{U}_t} - \lambda_{4,t} \int \frac{\partial f_{4,t}}{\partial \tilde{U}_t} \mu_t\left(x\right) \; dx \\
& + \int \left[\lambda_{5,t}(x) \left(\frac{\partial f_{5,t}}{\partial \tilde{U}_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}}{\partial \tilde{U}_t} \frac{\partial v_t(x)}{\partial x_i} \right) \right] dx \\
& + \sum_{j=1}^{J} \int \left[\lambda_{6,j,t}(x) \left(\frac{\partial^2 f_{5,t}}{\partial u_{j,t} \partial \tilde{U}_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}}{\partial u_{j,t} \partial \tilde{U}_t} \frac{\partial v_t(x)}{\partial x_i} \right) \right] dx \\
& + \int \left[\lambda_{7,t}(x) \left(- \sum_{i=1}^{I} \frac{\partial}{\partial x_i} \left[\frac{\partial b_{i,t}}{\partial \tilde{U}_t} \mu_t\left(x\right) \right] \right) \right] dx, \\
& \forall t \geq 0.
\end{aligned}
$$

Value function, distribution and policy functions

$$
\begin{aligned}
v_t\left(x\right): \quad 0 = \quad & \left(-\lambda_{5,t}(x)\rho - \sum_{i=1}^{I} \frac{\partial\left[\lambda_{5,t}(x)b_i\left(x, u_t(x), Z_t\right)\right]}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{I}\frac{\partial^2}{\partial^2 x_i}\left[\sigma_i^2(x)\lambda_{5,t}(x)\right]\right) \\
& - \sum_{j=1}^{J}\sum_{i=1}^{I}\frac{\partial}{\partial x_i}\left(\lambda_{6,j,t}(x)\frac{\partial b_i\left(x, u_t(x), Z_t\right)}{\partial u_{j,t}}\right) \\
& -(\dot{\lambda}_{5,t}(x) - \varrho\lambda_{5,t}(x)), \\
& \forall t > 0,
\end{aligned}
$$

$$
0 = \quad \lambda_{5,0}(x).
$$

$$
\begin{aligned}
\mu_t\left(x\right): \quad 0 = \quad & -\lambda_{4,t} f_4\left(x, u_t(x), Z_t\right) \\
& + \lambda_{7,t}(x)\left(\sum_{i=1}^{I}\frac{\partial\lambda_{7,t}(x)}{\partial x_i}b_i\left(x, u_t(x), Z_t\right) + \sum_{i=1}^{I}\frac{\partial^2\lambda_{7,t}(x)}{\partial^2 x_i}\frac{\sigma_i^2(x)}{2}\right) \\
& +(\dot{\lambda}_{7t}(x) - \varrho\lambda_{7,t}(x)), \\
& \forall t \geq 0,
\end{aligned}
$$

$$
0 = \quad \lim_{t\to\infty} e^{-\varrho t}\lambda_{7,t}(x).
$$

$$
\begin{aligned}
u_{l,t}\left(x\right): \quad 0 = \quad & -\lambda_{4,t}\frac{\partial f_4}{\partial u_{l,t}}\mu_t\left(x\right) \\
+ \quad & \overbrace{\left[\lambda_{5,t}(x)\left(\frac{\partial f_5}{\partial u_{l,t}} + \sum_{i=1}^{I}\frac{\partial b_i}{\partial u_{l,t}}\frac{\partial v_t(x)}{\partial x_i}\right)\right]}^{=0} \\
+ \quad & \sum_{j=1}^{J}\lambda_{6,k,t}(x)\left(\frac{\partial^2 f_5}{\partial u_{l,t}\partial u_{j,t}} + \sum_{i=1}^{I}\frac{\partial^2 b_i}{\partial u_{l,t}\partial u_{j,t}}\frac{\partial v_t(x)}{\partial x_i}\right) \\
- \quad & \left(\sum_{i=1}^{I}\frac{\partial\lambda_{7,t}(x)}{\partial x_i}\frac{\partial b_{i,t}}{\partial u_{l,t}}\mu_t\left(x\right)\right).
\end{aligned}
$$

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 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 $x^{n_i+1} \equiv (x_{1,n_1}, ..., x_{i,n_i+1}, ..., x_{I,n_I})$, $x^{n_i-1} \equiv (x_{1,n_1}, ..., x_{i,n_i-1}, ..., x_{I,n_I})$ $f_t^n \equiv f(x^n, u_t^n, Z_t)$, $f_t^{n_i-1} \equiv f(x^{n_i-1}, u_t^n, Z_t)$ and $f_t^{n_i+1} \equiv f(x^{n_i+1}, u_t^n, Z_t)$. 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 replace (i) time derivatives of multipliers by backward derivatives, (ii) integrals by sums (iii) derivatives with respect to x by the upwind derivatives ∇ or $\hat{\nabla}$:

$$\begin{aligned}
\nabla_i [v_t^n] &\equiv \left[\mathbb{I}_{b_{i,t}^n>0} \frac{v_t^{n_i+1} - v_t^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n<0} \frac{v_t^n - v_t^{n_i-1}}{\Delta x_i}\right], \\
\hat{\nabla}_i [\mu_t^n] &\equiv \left[\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\mu_t^{n_i+1} - \mathbb{I}_{b_{i,t}^n<0}\mu_t^n}{\Delta x_i} + \frac{\mathbb{I}_{b_{i,t}^n>0}\mu_t^n - \mathbb{I}_{b_{i,t}^{n_i-1}>0}\mu_t^{n_i-1}}{\Delta x_i}\right],
\end{aligned}$$

for any discretized functions v_t^n, μ_t^n. We simplify the notation for sums $\sum_n \equiv \sum_{n_1 \in \{1,..,N_1\},..,n_I \in \{1,..,N_I\}}$.

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 $X_{t+\triangle t}$. The second-order derivative is approximated as

$$\triangle_i [v_t^n] \equiv \left[\frac{\left(v_t^{n_i+1}\right) + \left(v_t^{n_i-1}\right) - 2\left(v_t^n\right)}{\left(\Delta x_i\right)^2}\right].$$

We start with the optimality condition for U_t

$$
\begin{aligned}
U_t: \quad 0 = \quad & -\left(\frac{\lambda_{2,t} - \lambda_{2,t-1}}{\Delta t} - \varrho \lambda_{2,t}\right) && (114)\\
& + \frac{\partial f_0}{\partial U_t} - \lambda_{1,t} \frac{\partial f_1}{\partial U_t} - \lambda_{2,t} \frac{\partial f_2}{\partial U_t} - \lambda_{3,t} \frac{\partial f_3}{\partial U_t} - \lambda_{4,t} \sum_{n=1}^{N} \frac{\partial f_4^n}{\partial U_t} \mu_t^n && (115)\\
& + \sum_n \left[\lambda_{5,t}^n \left(\frac{\partial f_5^n}{\partial U_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial U_t} \nabla_i \left[v_t^n\right]\right)\right]\\
& + \sum_{j=1}^{J} \sum_n \left[\lambda_{6,j,t}^n \left(\frac{\partial^2 f_5^n}{\partial u_j \partial U_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial u_j \partial U_t} \nabla_i \left[v_t^n\right]\right)\right]\\
& + \sum_n \left[-\lambda_{7,t}^n \sum_{i=1}^{I} \hat{\nabla}_i \left[\frac{\partial b_{i,t}^n}{\partial U_t} \mu_t^n\right]\right] && (116)\\
& \forall t \geq 0.
\end{aligned}
$$

The optimality conditions for the other aggregate variables look very much alike:

$$
\begin{aligned}
X_t: \quad 0 = \quad & -\left(\frac{\lambda_{1,t} - \lambda_{1,t-1}}{\Delta} - \varrho \lambda_{1,t}\right)\\
& + \frac{\partial f_0}{\partial X_t} - \lambda_{1,t} \frac{\partial f_1}{\partial X_t} - \lambda_{2,t} \frac{\partial f_2}{\partial X_t} - \lambda_{3,t} \frac{\partial f_3}{\partial X_t} - \lambda_{4,t} \sum_n \frac{\partial f_4^n}{\partial X_t} \mu_t^n\\
& + \sum_n \left[\lambda_{5,t}^n \left(\frac{\partial f_5^n}{\partial X_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial X_t} \nabla_i \left[v_t^n\right]\right)\right]\\
& + \sum_{j=1}^{J} \sum_n \left[\lambda_{6,j,t}^n \left(\frac{\partial^2 f_5^n}{\partial u_j \partial X_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial u_j \partial X_t} \nabla_i \left[v_t^n\right]\right)\right]\\
& + \sum_n \left[-\lambda_{7,t}^n \sum_{i=1}^{I} \hat{\nabla}_i \left[\frac{\partial b_{i,t}^n}{\partial X_t} \mu_t^n\right]\right]\\
& \forall t > 0.
\end{aligned}
$$

$$
\begin{aligned}
\hat{U}_t: \quad 0 = \quad & 0 \\
& + \frac{\partial f_0}{\partial \hat{U}_t} - \lambda_{1,t} \frac{\partial f_1}{\partial \hat{U}_t} - \lambda_{2,t} \frac{\partial f_2}{\partial \hat{U}_t} - \lambda_{3,t} \frac{\partial f_3}{\partial \hat{U}_t} - \lambda_{4,t} \sum_n \frac{\partial f_4^n}{\partial \hat{U}_t} \mu_t^n \\
& + \sum_n \left[\lambda_{5,t}^n \left(\frac{\partial f_5^n}{\partial \hat{U}_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial \hat{U}_t} \nabla_i \left[v_t^n \right] \right) \right] \\
& + \sum_{j=1}^{J} \sum_n \left[\lambda_{6,j,t}^n \left(\frac{\partial^2 f_5^n}{\partial u_j \partial \hat{U}_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial u_j \partial \hat{U}_t} \nabla_i \left[v_t^n \right] \right) \right] \\
& + \sum_n \left[-\lambda_{7,t}^n \sum_{i=1}^{I} \hat{\nabla}_i \left[\frac{\partial b_{i,t}^n}{\partial \hat{U}_t} \mu_t^n \right] \right] \\
& \forall t \geq 0.
\end{aligned}
$$

$$
\begin{aligned}
\tilde{U}_t: \quad 0 = \quad & \lambda_{4,t} \\
& + \frac{\partial f_0}{\partial \tilde{U}_t} - \lambda_{1,t} \frac{\partial f_1}{\partial \tilde{U}_t} - \lambda_{2,t} \frac{\partial f_2}{\partial \tilde{U}_t} - \lambda_{3,t} \frac{\partial f_3}{\partial \tilde{U}_t} - \lambda_{4,t} \sum_{n=1}^{N} \frac{\partial f_4^n}{\partial \tilde{U}_t} \mu_t^n \\
& + \sum_n \left[\lambda_{5,t}^n \left(\frac{\partial f_5^n}{\partial \tilde{U}_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial \tilde{U}_t} \nabla_i \left[v_t^n \right] \right) \right] \\
& + \sum_{j=1}^{J} \sum_n \left[\lambda_{6,j,t}^n \left(\frac{\partial^2 f_5^n}{\partial u_j \partial \tilde{U}_t} + \sum_{i=1}^{I} \frac{\partial b_i^n}{\partial u_j \partial \tilde{U}_t} \nabla_i \left[v_t^n \right] \right) \right] \\
& + \sum_n \left[-\lambda_{7,t}^n \sum_{i=1}^{I} \hat{\nabla}_i \left[\frac{\partial b_{i,t}^n}{\partial \tilde{U}_t} \mu_t^n \right] \right] \\
& \forall t \geq 0.
\end{aligned}
$$

The discretized optimality condition with respect to the value function $v_t(x)$, the distribution $\mu_t(x)$ and the individual jump variable $u_{j,t}(x)$ are.

$$
\begin{aligned}
v_t(x): \quad 0 = \quad & -\lambda_{5,t}^n \rho - \sum_{i=1}^{I} \hat{\nabla}_i \left[\lambda_{5,t}^n b_{i,t}^n\right] \\
& + \frac{1}{2} \sum_{i=1}^{I} \sum_{k=1}^{I} \nabla_i \left[\sigma_{i,k}^n \lambda_{5,t}^n\right] \\
& - \sum_{j=1}^{J} \sum_{i=1}^{I} \left(\hat{\nabla}_i \left[\lambda_{6,j,t}^n \frac{\partial b_{i,t}^n}{\partial u_{j,t}^n} \right] \right) \\
& - \left(\frac{\lambda_{5,t}^n - \lambda_{5,t-1}^n}{\Delta t} - \varrho \lambda_{5,t}^n \right).
\end{aligned} \tag{117}
$$

$$
\begin{aligned}
\mu_t(x): \quad 0 = \quad & -\lambda_{4,t} f_{4,t}^n \\
& + \lambda_{7,t}(x) \left(\sum_{i=1}^{I} b_i\left(x, u_t(x), Z_t\right) \nabla_i \left[\lambda_{7,t}^n\right] + \frac{1}{2} \sum_{i=1}^{I} \left(\sigma_i^2\right)^n \triangle_i^2 \left[\lambda_{7,t}^n\right] \right) \\
& + \frac{\lambda_{7,t}^n - \lambda_{7,t-1}^n}{\Delta t} - \varrho \lambda_{7,t}^n
\end{aligned} \tag{118}
$$

$$
\begin{aligned}
u_{l,t}(x): \quad 0 = \quad & -\lambda_{4,t} \frac{\partial f_4}{\partial u_{l,t}} \mu_t^n \\
& + \sum_{j=1}^{J} \lambda_{6,k,t}^n \left(\frac{\partial^2 f_{5,t}^n}{\partial u_{l,t}^n \partial u_{j,t}^n} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{l,t}^n \partial u_{j,t}^n} \nabla_i \left[v_t^n\right] \right) \\
& - \sum_{i=1}^{I} \nabla_i \left[\lambda_{7,t}^n\right] \frac{\partial b_{i,t}^n}{\partial u_{l,t}} \mu_t^n
\end{aligned} \tag{119}
$$

3. Discretize, then optimize We follow here the reverse approach, discretizing first and optimizing next.3.a The discretized planner's problem

Now first discretize the optimization problem with respect to time (time step Δt) and the idiosyncratic state (N grid points, grid step Δx_i). We define the discount factor $\beta \equiv (1+\varrho\Delta t)^{-1}$.

$$
\begin{aligned}
\max_{Z_t, u_t^n, \mu_t^n, v_t^n} & \qquad \sum_t \beta^t f_0(Z_t) \\
& s.t.\ \forall t \\
\frac{X_{t+1}-X_t}{\Delta t} &= f_1(Z_t) && (120) \\
\frac{U_{t+1}-U_t}{\Delta t} &= f_2(Z_t) && (121) \\
0 &= f_3(Z_t) && (122) \\
\tilde{U}_t &= \sum_{n=1}^{N} f_4\left(x^n, u_t^n, Z_t\right)\mu_t^n && (123) \\
\rho v_t^n &= \frac{v_{t+1}^n - v_t^n}{\Delta t} + f_5(x^n, u_t^n, Z_t) + \sum_{i=1}^{I} b_i\left(x^n, u_t^n, Z_t\right)\nabla_i\left[v_t^n\right] && (124) \\
& + \frac{1}{2}\sum_{i=1}^{I}\left(\sigma_i^2\right)^n \triangle_i^2\left[v_t^n\right],\ \forall n \\
0 &= \frac{\partial f_{5,t}^n}{\partial u_{j,t}^n} + \sum_{i=1}^{I}\frac{\partial b_{i,t}^n}{\partial u_{j,t}^n}\nabla_i\left[v_t^n\right], \quad \forall j, n. && (125) \\
\frac{\mu_{t+1}^n - \mu_t^n}{\Delta t} &= -\sum_{i=1}^{I}\hat{\nabla}_i\left[b_{i,t}^n \mu_t^n\right] && (126) \\
& + \frac{1}{2}\sum_{i=1}^{I}\triangle_i\left[\sigma_i^2 \mu_t^n\right] && (127) \\
X_0 &= \bar{X}_0 && (128) \\
\mu_0^n &= \bar{\mu}_0^n && (129)
\end{aligned}
$$

3.b The Lagrangian The Lagrangian is

$$
\begin{aligned}
L \;=\; & \sum_t \beta^t f_0(Z_t) \\
& + \sum_t \beta^t \lambda_{1,t} \left\{ \frac{X_{t+1} - X_t}{\Delta t} - f_1(Z_t) \right\} \\
& + \sum_t \beta^t \lambda_{2,t} \left\{ \frac{U_{t+1} - U_t}{\Delta t} - f_2(Z_t) \right\} \\
& + \sum_t \beta^t \lambda_{3,t} \left\{ -f_3(Z_t) \right\} \\
& + \sum_t \beta^t \lambda_{4,t} \left\{ \tilde{U}_t - \sum_n f_4\left(x^n, u_t^n, Z_t\right) \mu_t^n \right\} \\
& + \sum_t \sum_n \beta^t \lambda_{5,t}^n \left\{ \begin{array}{c} -\rho v_t^n + \frac{v_{t+1}^n - v_t^n}{\Delta t} + f_5(x^n, u_t^n, Z_t) + \sum_{i=1}^I b_i\left(x^n, u_t^n, Z_t\right) \nabla_i\left[v_t^n\right] \\ + \sum_{i=1}^I \triangle_i^2 \left[v_t^n\right] \end{array} \right\} \\
& + \sum_t \sum_n \sum_{j=1}^J \beta^t \lambda_{6,j,t}^n \left\{ \frac{\partial f_{5,t}^n}{\partial u_{j,t}^n} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial u_{j,t}^n} \nabla_i\left[v_t^n\right] \right\} \\
& + \sum_t \sum_n \beta^t \lambda_{7,t}^n \left\{ \begin{array}{c} -\frac{\mu_{t+1}^n - \mu_t^n}{\Delta t} - \sum_{i=1}^I \hat{\nabla}_i\left[b_{i,t}^n \mu_t^n\right] \\ + \frac{1}{2} \sum_{i=1}^I \triangle_i\left[\sigma_i^2 \mu_t^n\right] \end{array} \right\}
\end{aligned}
$$

3.c The optimality conditions The FOCs are

$$
\begin{aligned}
\frac{\partial L}{\partial U_t} : 0 \;=\; & \frac{\partial f_{0,t}}{\partial U_t} - \lambda_{1,t} \frac{\partial f_{1,t}}{\partial U_t} + \lambda_{2,t} \left\{ -\frac{1}{\Delta t} - \frac{\partial f_{2,t}}{\partial U_t} \right\} + \beta^{-1} \lambda_{2,t-1} \frac{1}{\Delta t} - \lambda_{3,t} \frac{\partial f_{3,t}}{\partial U_t} - \lambda_{4,t} \sum_n \frac{\partial f_{4,t}^n}{\partial U_t} \mu_t^n \\
& + \sum_n \lambda_{5,t}^n \left\{ + \frac{\partial f_{5,t}^n}{\partial U_t} + \sum_{i=1}^I \frac{\partial b_{i,t}^n}{\partial U_t} \nabla_i\left[v_t^n\right] \right\} \\
& + \sum_n \sum_{j=1}^J \lambda_{6,j,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial U_t} + \sum_{i=1}^I \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial U_t} \nabla_i\left[v_t^n\right] \right\} \\
& + \sum_n \left\{ \sum_{i=1}^I \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right) \left[\mathbb{I}_{b_{i,t}^n < 0} \frac{\partial b_{i,t}^n}{\partial U_t} \frac{\mu_t^n}{\Delta x_i} \right] + \sum_{i=1}^I \left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right) \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{\partial b_{i,t}^n}{\partial U_t} \frac{\mu_t^n}{\Delta x_i} \right] \right\} \\
& \forall t \geq 0
\end{aligned} \tag{130}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial X_t}:0 \;=\;& \frac{\partial f_{0,t}}{\partial X_t} - \lambda_{1,t}\left\{\frac{1}{\Delta t} + \frac{\partial f_{1,t}}{\partial X_t}\right\} + \beta^{-1}\lambda_{1,t-1}\frac{1}{\Delta t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial X_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial X_t} - \lambda_{4,t}\sum_n \frac{\partial f_{4,t}^n}{\partial X_t}\mu_t^n \\
&+ \sum_n \lambda_{5,t}^n \left\{\frac{\partial f_{5,t}^n}{\partial X_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}^n}{\partial X_t}\nabla_i\left[v_t^n\right]\right\} \\
&+ \sum_n \sum_j \lambda_{6,j,t}^n \left\{\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial X_t} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial X_t}\nabla_i\left[v_t^n\right]\right\} \\
&+ \sum_n \left\{\sum_{i=1}^{I}\left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial X_t}\frac{\mu_t^n}{\Delta x_i}\right] + \sum_{i=1}^{I}\left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right)\left[\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial X_t}\frac{\mu_t^n}{\Delta x_i}\right]\right\} \\
&\forall t > 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial \tilde{U}_t}:0 \;=\;& \frac{\partial f_{0,t}}{\partial \tilde{U}_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial \tilde{U}_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial \tilde{U}_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial \tilde{U}_t} - \lambda_{4,t}\sum_n \frac{\partial f_{4,t}^n}{\partial \tilde{U}_t}\mu_t^n \\
&+ \sum_n \lambda_{5,t}^n \left\{+\frac{\partial f_{5,t}^n}{\partial \tilde{U}_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}^n}{\partial \tilde{U}_t}\nabla_i\left[v_t^n\right]\right\} \\
&+ \sum_n \sum_j \lambda_{6,j,t}^n \left\{\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial \tilde{U}_t} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial \tilde{U}_t}\nabla_i\left[v_t^n\right]\right\} \\
&+ \sum_n \left\{\sum_{i=1}^{I}\left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial \tilde{U}_t}\frac{\mu_t^n}{\Delta x_i}\right] + \sum_{i=1}^{I}\left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right)\left[\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial \tilde{U}_t}\frac{\mu_t^n}{\Delta x_i}\right]\right\} \\
&\forall t \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial \hat{U}_t}:0 \;=\;& \frac{\partial f_{0,t}}{\partial \hat{U}_t} - \lambda_{1,t}\frac{\partial f_{1,t}}{\partial \hat{U}_t} - \lambda_{2,t}\frac{\partial f_{2,t}}{\partial \hat{U}_t} - \lambda_{3,t}\frac{\partial f_{3,t}}{\partial \hat{U}_t} - \lambda_{4,t}\sum_n \frac{\partial f_{4,t}^n}{\partial \hat{U}_t}\mu_t^n \\
&+ \sum_n \lambda_{5,t}^n \left\{+\frac{\partial f_{5,t}^n}{\partial \hat{U}_t} + \sum_{i=1}^{I} \frac{\partial b_{i,t}^n}{\partial \hat{U}_t}\nabla_i\left[v_t^n\right]\right\} \\
&+ \sum_n \sum_j \lambda_{6,j,t}^n \left\{\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial \hat{U}_t} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial \hat{U}_t}\nabla_i\left[v_t^n\right]\right\} \\
&+ \sum_n \left\{\sum_{i=1}^{I}\left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial \hat{U}_t}\frac{\mu_t^n}{\Delta x_i}\right] + \sum_{i=1}^{I}\left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right)\left[\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial \hat{U}_t}\frac{\mu_t^n}{\Delta x_i}\right]\right\} \\
&\forall t \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial v_t^n}:0 \quad = \quad & \lambda_{5,t}^n \left\{ -\rho - \frac{1}{\Delta t} + \sum_{i=1}^{I} b_{i,t}^n \frac{\mathbb{I}_{b_t^n<0} - \mathbb{I}_{b_t^n>0}}{\Delta x_i} - \sum_{i=1}^{I} \frac{2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \qquad (131) \\
& + \lambda_{5,t-1}^n \beta^{-1} \frac{1}{\Delta t} \\
& + \sum_{i=1}^{I} \lambda_{5,t}^{n_i-1} b_{i,t}^{n_i-1} \frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i} + \sum_{i=1}^{I} \lambda_{5,t}^{n_i-1} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \\
& - \sum_{i=1}^{I} \lambda_{5,t}^{n_i+1} b_{i,t}^{n_i+1} \frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i} + \sum_{i=1}^{I} \lambda_{5,t}^{n_i+1} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \\
& + \sum_{j=1}^{J} \sum_{i=1}^{I} \left\{ \lambda_{6,j,t}^n \left\{ \frac{\partial b_{i,t}^n}{\partial u_{j,t}^n} \frac{\mathbb{I}_{b_{i,t}^n<0} - \mathbb{I}_{b_{i,t}^n>0}}{\Delta x_i} \right\} + \lambda_{6,j,t}^{n_i-1} \left\{ \frac{\partial b_{i.t}^{n_i-1}}{\partial u_{j,t}^{n_i-1}} \frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i} \right\} \right\} \\
& - \sum_{j=1}^{J} \sum_{i=1}^{I} \lambda_{6,j,t}^{n_i+1} \left\{ \frac{\partial b_{i,t}^{n_i+1}}{\partial u_{j,t}^{n_i+1}} \frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i} \right\} \qquad (132) \\
& \forall t \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial \mu_t^n}:0 \quad = \quad & -\lambda_{4,t} f_{4,t}^n \qquad (133) \\
& + \lambda_{7,t}^n \left\{ \frac{1}{\Delta t} - \sum_{i=1}^{I} \left[\left(\mathbb{I}_{b_{i,t}^n>0} - \mathbb{I}_{b_{i,t}^n<0} \right) \frac{b_{i,t}^n}{\Delta x_i} \right] - \sum_{i=1}^{I} \frac{-2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \\
& + \left\{ -\sum_{i=1}^{I} \lambda_{7,t}^{n_i-1} \left[\frac{\mathbb{I}_{b_{i,t}^n<0} b_{i,t}^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \\
& + \left\{ -\sum_{i=1}^{I} \lambda_{7,t}^{n_i+1} \left[\frac{-\mathbb{I}_{b_{i,t}^n>0} b_{i,t}^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \\
& + \beta^{-1} \lambda_{7,t-1}^n \left\{ -\frac{1}{\Delta t} \right\} \\
& \forall t > 0
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial u_{l,t}^n} : 0 \quad = \quad & -\lambda_{4,t} \frac{\partial f_{4,t}^n}{\partial u_{l,t}^n} \mu_t^n \qquad (134) \\
& + \beta^t \lambda_{5,t}^n \left\{ \frac{\partial f_{5,t}^n}{\partial u_{l,t}^n} + \sum_{i=1}^{I} \frac{\partial b_{i,t}^n}{\partial u_{l,t}^n} \nabla_i \left[v_t^n \right] \right\} \\
& + \sum_j \lambda_{6,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} \nabla_i \left[v_t^n \right] \right\} \\
& + \sum_{i=1}^{I} \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i - 1} \right) \left[\mathbb{I}_{b_{i,t}^n < 0} \frac{\partial b_{i,t}^n}{\partial u_{l,t}^n} \frac{\mu_t^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \left(\lambda_{7,t}^{n_i + 1} - \lambda_{7,t}^n \right) \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{\partial b_{i,t}^n}{\partial u_{l,t}^n} \frac{\mu_t^n}{\Delta x_i} \right] \\
& \forall t \geq 0
\end{aligned}
$$

By the individual agents' optimality condition, line 2 of this expression is equal to 0.

4. Compare Finally, by comparing the respective discretized optimality conditions, we show that the two procedures yield the same equilibrium conditions in the limit. Consider first the condition for U_t. The optimize-discretize condition is given by (114), 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 discretize-optimize condition (130), rearranges to

$$
\begin{aligned}
\frac{\partial L}{\partial U_t}:0 \quad = \quad & -\left(\frac{\lambda_{2,t}-\lambda_{2,t-1}}{\Delta t}-\frac{\beta^{-1}-1}{\Delta t}\lambda_{2,t-1}\right) \\
& +\frac{\partial f_{0,t}}{\partial U_t}-\lambda_{1,t}\frac{\partial f_{1,t}}{\partial U_t}-\lambda_{2,t}\frac{\partial f_{2,t}}{\partial U_t}-\lambda_{3,t}\frac{\partial f_{3,t}}{\partial U_t}-\lambda_{4,t}\sum_{n=1}^{N}\frac{\partial f_{4,t}^{n}}{\partial U_t}\mu_t^n \\
& +\sum_{n=1}^{N}\lambda_{5,t}^{n}\left\{\frac{\partial f_{5,t}^{n}}{\partial U_t}+\sum_{i=1}^{I}\frac{\partial b_i^n}{\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& +\sum_{n=1}^{N}\sum_{j=1}^{J}\lambda_{6,j,t}^{n}\left\{\frac{\partial^2 f_{5,t}^{n}}{\partial u_{j,t}^{n}\partial U_t}+\frac{\partial^2 b_t^n}{\partial u_{j,t}^{n}\partial U_t}\nabla_i\left[v_t^n\right]\right\} \\
& +\sum_{n}\left\{\sum_{i=1}^{I}\left(\lambda_{7,t}^{n}-\lambda_{7,t}^{n_i-1}\right)\left[\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial U_t}\frac{\mu_t^n}{\Delta x_i}\right]+\sum_{i=1}^{I}\left(\lambda_{7,t}^{n_i+1}-\lambda_{7,t}^{n}\right)\left[\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial U_t}\frac{\mu_t^n}{\Delta x_i}\right]\right\} \\
& \forall t\geq 0
\end{aligned}
$$

The second to fourth lines are evidently identical. The last lines also coincide once we take into account the definition of $\hat{\nabla}_i\left[\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n\right]=\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\frac{\partial b_{i,t}^{n_i+1}}{\partial U_t}\mu_t^{n_i+1}-\mathbb{I}_{b_{i,t}^n<0}\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n}{\Delta x_i}+\frac{\mathbb{I}_{b_{i,t}^n>0}\frac{\partial b_{i,t}^n}{\partial U_t}\mu_t^n-\mathbb{I}_{b_{i,t}^{n_i-1}>0}\frac{\partial b_{i,t}^{n_i-1}}{\partial U_t}\mu_t^{n_i-1}}{\Delta x_i}$.

Finally compare the first lines. Since $\beta\equiv(1+\varrho\Delta t)^{-1}$ we have that $\frac{\beta^{-1}-1}{\Delta t}=\varrho$. The difference between these two equations hence is $\|\varrho\left(\lambda_{2,t}-\lambda_{2,t-1}\right)\|$. In the limit as $\Delta t\to 0$, and provided that $\lambda_{2,t}$ features no jumps for $t>0$,this difference converges to zero.The same argument applies to the optimality conditions with respect to X_t with the difference now proportional to $\|\varrho\left(\lambda_{1,t}-\lambda_{1,t-1}\right)\|$. 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 (117) and (131). After some rearranging they are given by

$$
\begin{aligned}
v_t(x): \quad 0 = \quad & -\sum_{i=1}^{I}\left(\frac{\mathbb{I}_{b_{i,t}^{n}>0}\lambda_{5,j,t}^{n}b_{i,t}^{n}-\mathbb{I}_{b_{i,t}^{n_i-1}>0}\lambda_{5,j,t}^{n_i-1}b_{i,t}^{n_i-1}}{\Delta x_i}+\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\lambda_{5,j,t}^{n_i+1}b_{i,t}^{n_i+1}-\mathbb{I}_{b_{i,t}^{n}<0}\lambda_{5,j,t}^{n}b_{i,t}^{n}}{\Delta x_i}\right) \\
& +\frac{1}{2}\sum_{i=1}^{I}\frac{\left(\sigma_i^2\right)^{n_i+1}\lambda_{5,t}^{n_i+1}+\left(\sigma_i^2\right)^{n_i-1}\lambda_{5,t}^{n_i-1}-2\left(\sigma_i^2\right)^{n}\lambda_{5,t}^{n}}{\left(\Delta x_i\right)^2} \\
& -\sum_{j=1}^{J}\sum_{i=1}^{I}\left(\frac{\mathbb{I}_{b_{i,t}^{n}>0}\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}-\mathbb{I}_{b_{i,t}^{n_i-1}>0}\lambda_{6,j,t}^{n_i-1}\frac{\partial b_{i,t}^{n_i-1}}{\partial u_{j,t}^{n_i-1}}}{\Delta x_i}+\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}\lambda_{6,j,t}^{n_i+1}\frac{\partial b_{i,t}^{n_i+1}}{\partial u_{j,t}^{n_i+1}}-\mathbb{I}_{b_{i,t}^{n}<0}\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}}{\Delta x_i}\right) \\
& -\lambda_{5,t}^{n}\rho-\left(\frac{\lambda_{5,t}^{n}-\lambda_{5,t-1}^{n}}{\Delta t}-\varrho\lambda_{5,t}^{n}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial L}{\partial v_t^n}: 0 \quad = \quad & \lambda_{5,t}^{n}\left\{\sum_{i=1}^{I}b_{i,t}^{n}\frac{\mathbb{I}_{b_t^n<0}-\mathbb{I}_{b_t^n>0}}{\Delta x_i}-\sum_{i=1}^{I}\frac{\left(\sigma_i^2\right)^n}{\left(\Delta x_i\right)^2}\right\} \\
& +\left\{\sum_{i=1}^{I}\lambda_{5,t}^{n_i-1}b_{i,t}^{n_i-1}\frac{\mathbb{I}_{b_{i,t}^{n_i-1}>0}}{\Delta x_i}+\sum_{i=1}^{I}\lambda_{5,t}^{n_i-1}\frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\right\} \\
& +\left\{-\sum_{i=1}^{I}\lambda_{5,t}^{n_i+1}b_{i,t}^{n_i+1}\frac{\mathbb{I}_{b_{i,t}^{n_i+1}<0}}{\Delta x_i}+\sum_{i=1}^{I}\lambda_{5,t}^{n_i+1}\frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2}\right\} \\
& +\sum_{j=1}^{J}\sum_{i=1}^{I}\left(\lambda_{6,j,t}^{n}\frac{\partial b_{i,t}^{n}}{\partial u_{j,t}^{n}}\frac{\mathbb{I}_{b_t^n<0}-\mathbb{I}_{b_t^n>0}}{\Delta x_i}+\lambda_{6,j,t}^{n_i-1}\frac{\partial b_{i,t}^{n_i-1}}{\partial u_{j,t}^{n_i-1}}\frac{\mathbb{I}_{b_t^{n_i-1}>0}}{\Delta x_i}-\lambda_{6,j,t}^{n_i+1}\frac{\partial b_{i,t}^{n_i+1}}{\partial u_{j,t}^{n_i+1}}\frac{\mathbb{I}_{b_t^{n_i+1}<0}}{\Delta x_i}\right) \\
& -\rho\lambda_{5,t}^{n}-\left(\frac{\lambda_{5,t}^{n}-\lambda_{5,t-1}^{n}}{\Delta t}-\frac{\beta^{-1}-1}{\Delta t}\lambda_{5,t-1}^{n}\right) \qquad (135)
\end{aligned}
$$

Again these, two expressions are identical up to the last time index in the last line (λ_5^n), and thus the difference is $\|\varrho\left(\lambda_{5,t}-\lambda_{5,t-1}\right)\|$.

Next, consider the two discretized optimality conditions with respect to μ_t^n (118)

and (133). After some rearranging they are given by

$$\begin{aligned} \mu_t(x): \quad 0 = \quad & -\lambda_{4,t} f_{4,t}^n \\ & + \sum_{i=1}^{I} b_{i,t}^n \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}}{\Delta x_i} \right] + \frac{1}{2} \sum_{i=1}^{I} \left(\sigma_i^2\right)^n \frac{\lambda_{7,t}^{n_i+1} + \lambda_{7,t}^{n_i-1} - 2\lambda_{7,t}^n}{2\left(\Delta x_i\right)^2} \\ & + \frac{\lambda_{7,t}^n - \lambda_{7,t-1}^n}{\Delta t} - \varrho \lambda_{7,t}^n \end{aligned} \tag{136}$$

$$\begin{aligned} \frac{\partial L}{\partial \mu_t^n}: 0 \quad = \quad & -\lambda_{4,t} f_{4,t}^n \\ & + \lambda_{7,t}^n \left\{ -\sum_{i=1}^{I} \left[\left(\mathbb{I}_{b_{i,t}^n > 0} - \mathbb{I}_{b_{i,t}^n < 0} \right) \frac{b_{i,t}^n}{\Delta x_i} \right] - \sum_{i=1}^{I} \frac{-2\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \right\} \\ & - \sum_{i=1}^{I} \left[\lambda_{7,t}^{n_i-1} \frac{\mathbb{I}_{b_{i,t}^n < 0} b_{i,t}^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \lambda_{7,t}^{n_i-1} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \\ & + \sum_{i=1}^{I} \left[\lambda_{7,t}^{n_i+1} \frac{\mathbb{I}_{b_{i,t}^n > 0} b_{i,t}^n}{\Delta x_i} \right] + \sum_{i=1}^{I} \lambda_{7,t}^{n_i+1} \frac{\left(\sigma_i^2\right)^n}{2\left(\Delta x_i\right)^2} \\ & + \frac{\lambda_{7,t}^n - \lambda_{7,t-1}^n}{\Delta t} - \frac{\beta^{-1} - 1}{\Delta t} \lambda_{7,t-1}^n, \end{aligned} \tag{137}$$

which again differ in $\left\| \varrho \left(\lambda_{7,t} - \lambda_{7,t-1} \right) \right\|$.

Finally, consider the two discretized optimality conditions with respect to $u_{l,t}^n(x)$, (119) and (134). After some rearranging they are given by

$$\begin{aligned} u_{l,t}(x): \quad 0 = \quad & -\lambda_{4,t} \frac{\partial f_4}{\partial u_{l,t}} \mu_t^n \\ + \quad & \sum_{j=1}^{J} \lambda_{6,l,t}^n \left(\frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} + \sum_{i=1}^{I} \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{v_t^{n_i+1} - v_t^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{v_t^n - v_t^{n1}}{\Delta x_i} \right] \right) \\ - \quad & \sum_{i=1}^{I} \left[\mathbb{I}_{b_{i,t}^n > 0} \frac{\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n < 0} \frac{\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}}{\Delta x_i} \right] \frac{\partial b_{i,t}^n}{\partial u_{l,t}} \mu_t^n \end{aligned} \tag{138}$$

$$
\begin{aligned}
\frac{\partial L}{\partial u_{l,t}^n}:0 \;\; = \;\; & -\lambda_{4,t}\frac{\partial f_{4,t}^n}{\partial u_{l,t}^n}\mu_t^n \\
& + \sum_j \lambda_{6,t}^n \left\{ \frac{\partial^2 f_{5,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} + \sum_{i=1}^I \frac{\partial^2 b_{i,t}^n}{\partial u_{j,t}^n \partial u_{l,t}^n} \left[\mathbb{I}_{b_{i,t}^n>0} \frac{v_t^{n_i+1}-v_t^n}{\Delta x_i} + \mathbb{I}_{b_{i,t}^n<0} \frac{v_t^n - v_t^{n_i-1}}{\Delta x_i} \right] \right\} \\
& + \left[\sum_{i=1}^I \left(\lambda_{7,t}^n - \lambda_{7,t}^{n_i-1}\right) \left[\mathbb{I}_{b_{i,t}^n<0} \frac{1}{\Delta x_i} \right] + \sum_{i=1}^I \left(\lambda_{7,t}^{n_i+1} - \lambda_{7,t}^n\right) \left[\mathbb{I}_{b_{i,t}^n>0} \frac{1}{\Delta x_i} \right] \right] \frac{\partial b_{i,t}^n}{\partial u_{l,t}} \mu_t^n,
\end{aligned}
$$

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 and then derives the optimality conditions, one arrives to a set of optimality conditions that coincide in everything but the timing of the multiplier in the term $\varrho\lambda_t$. Provided that multipliers experience no jumps, the difference between the two approaches goes to 0 as $\Delta t \to 0$. Note that this issue has nothing to do with heterogeneity.

E.3 Solving the Nuño and Thomas model using Dynare

Here we apply the "discretize-optimize" methodology outlined in Section E to the heterogeneous-agent model introduced in Nuño and Thomas (2022). This is a model *à la* Aiyagari-Bewley-Huggett with non-state-contingent long-term nominal debt contracts. Finding the optimal policy in this problem requires that the central bank takes into account not only the dynamics of the state distribution (given by the KF equation) but also the HJB equation. Figure 11 displays the time-0 optimal policy (inflation) in this case, compared to the one obtained through the "optimize-discretize" methodology employed in Nuño and Thomas (2022). Optimal inflation coincides in both cases, up to a numerical error that is reduced as we increase the number of grid points and we reduce the time step.

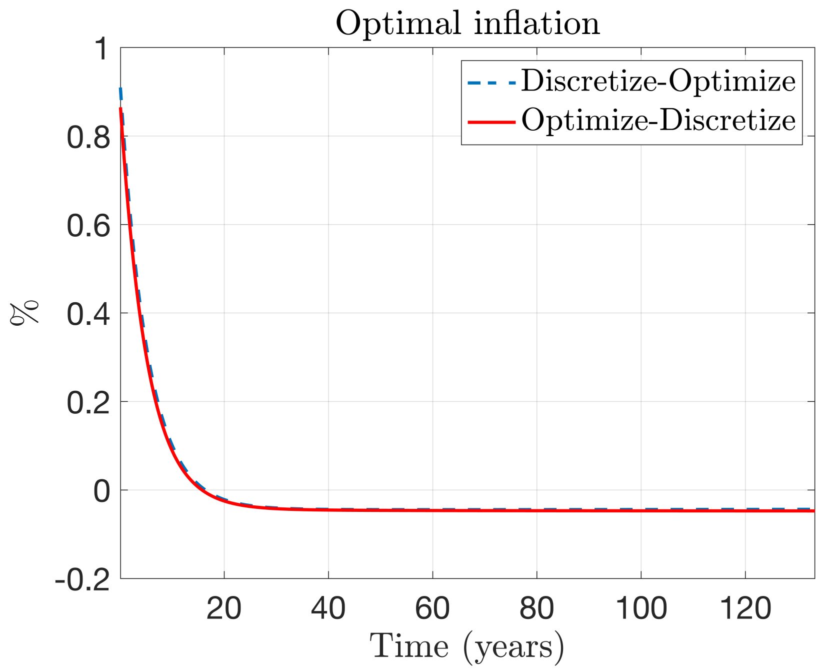


Figure 11: Time-0 optimal monetary policy using the two approaches.
Notes: The figure shows the optimal path of inflation in the Nuño and Thomas (2022) model using the "discretize-optimize" and "optimize-discretize" methods.