New Keynesian Economics with Household and Firm Heterogeneity*

Thomas Winberry
Wharton and NBER
twinb@wharton.upenn.edu

Adrien Auclert
Stanford and NBER
aauclert@stanford.edu

Matthew Rognlie
Northwestern and NBER
matthew.rognlie@northwestern.edu

Ludwig Straub
Harvard and NBER
ludwigstraub@g.harvard.edu

December 2025

Abstract

The Heterogeneous-Agent New Keynesian literature has revisited the transmission of monetary and fiscal policy to consumption using models where heterogeneous households face idiosyncratic income risk and borrowing constraints. We show that the key lessons from this literature also apply to investment using a model where heterogeneous firms face idiosyncratic productivity risk and financial frictions: constrained firms' investment depends on their free cash flow, generating indirect effects of monetary policy and implying that transfer payments stimulate investment demand. Quantitatively, the strength of these new mechanisms is governed by firms' marginal propensities to invest (MPIs), similar to the role of marginal propensities to consume (MPCs) for households. But unlike MPCs, we currently lack quasi-experimental evidence about MPIs that we can use to directly discipline the new mechanisms.

*In preparation for the *Journal of Economic Literature.* We thank the editor David Romer, five anonymous referees, as well as Adrien Bilal, Mark Gertler, Calvin He, Ben Hebert, Patrick Kehoe, Chen Lian, Hanno Lustig, Yueran Ma, Matteo Maggiori, Emi Nakamura, Pablo Ottonello, Jón Steinsson, David Thesmar, Christian Wolf, and participants at the Stanford Macro lunch, the Penn-Wharton Macro lunch, the Nemmers Prize Conference in Honor of Michael Woodford, and UT Austin for helpful comments. Tomáš Tapák provided exceptional research assistance.

1 Introduction

Over the past few decades, the New Keynesian framework has emerged as a workhorse to study the aggregate effects of monetary and fiscal policy on the economy. Simple versions of the model build on the "three-equation" specification in which consumption is the single endogenous component of aggregate demand (Clarida, Galí and Gertler, 1999; Woodford, 2003, chapter 4, Galí, 2008), while quantitative versions of the model used for policy analysis add investment and other features (Woodford, 2003, chapter 5, Christiano, Eichenbaum and Evans, 2005; Smets and Wouters, 2007). Underlying all these models are the twin assumptions of a representative household and a representative firm. Yet, micro data reveals extensive heterogeneity across individual households and firms, raising a natural question: how does micro-level heterogeneity matter for the transmission of monetary and fiscal policy?

Over the last ten years, the heterogeneous-agent New Keynesian (HANK) literature has developed to answer this question, focusing on the household side (e.g., McKay, Nakamura and Steinsson, 2016; Kaplan, Moll and Violante, 2018; Auclert, Rognlie and Straub, 2024; see Kaplan and Violante, 2018, McKay and Wolf, 2023 and Auclert, Rognlie and Straub, 2025b for recent reviews). Comparatively less work has been done to incorporate heterogeneous firms, with the existing literature primarily focusing on the heterogeneous effects of policy across firms (e.g. Ottonello and Winberry, 2020). From an aggregate perspective, this imbalance is surprising because investment accounts for a large share of changes in output over the business cycle and in response to monetary shocks.

Our goal in this paper is to bring heterogeneous firms centrally into HANK models, exploiting analogies with heterogeneous households. To do so, we develop a canonical heterogeneous-firm model in which firms face idiosyncratic productivity shocks and make investment decisions subject to capital adjustment costs and financial constraints, as in Khan and Thomas (2013) and Ottonello and Winberry (2020). In our model, constrained firms have a positive marginal propensity to invest (MPI), which measures how much investment increases in response to a one-time, unexpected increase in the firm's current cash flow, holding all other determinants of investment fixed.[1]

[1] This definition is parallel to the standard definition of MPCs on the household side. However, it differs from other definitions of MPIs in the literature, such as when current cash flows signal higher demand in

The presence of high-MPI firms activates new mechanisms governing the transmission of monetary and fiscal policy to investment, just as the presence of high-MPC households activated new mechanisms for consumption. In particular, monetary policy affects investment not just through traditional direct effects (lower interest rates decrease firms' discount rates) but also new indirect effects (equilibrium changes in cash flows affecting the investment of constrained firms). In addition, deficit-financed fiscal transfers to firms stimulate aggregate demand by increasing investment among constrained firms. Given the aggregate importance of investment, characterizing these new mechanisms on the firm side seems just as important as characterizing them on the household side.

Unlike MPCs, there are no quasi-experimental estimates of MPIs for a broad sample of firms that we can use to directly discipline these new mechanisms. Instead, we indirectly discipline the distribution of MPIs by matching other moments and relying on the structure of the model. Going forward, we believe that more empirical work directly estimating MPIs will spur progress in this space, similar to how direct estimates of MPCs have stimulated the development of heterogeneous-household models over the last decade.

We begin our analysis by developing a representative agent benchmark with capital investment. We extend the textbook three-equation New Keynesian model (Clarida et al., 1999, Woodford, 2003, Galí, 2008) to include capital investment subject to standard capital adjustment costs, and use sticky wages rather than sticky prices. The resulting model provides a launching point for incorporating micro-level heterogeneity throughout the paper.

Consumption in this representative agent benchmark follows the permanent income hypothesis (PIH). This property implies that monetary transmission is dominated by direct effects, i.e. changes in interest rates generating intertemporal substitution. In contrast, indirect effects resulting from changes in income are small because MPCs out of temporary income changes are low. Furthermore, Ricardian equivalence holds under the PIH, so transfer payments that are paid with future lump-sum taxes have no effect on consumption.

Investment in the representative agent benchmark follows a version of the Q theory, which has similarly stark predictions for policy transmission. Monetary transmission is dominated

the future (Alati, Fischer, Froemel and Ozturk, 2024), or households' marginal propensity to save in specific accounts or securities (e.g., Luetticke, 2021).

by direct effects, i.e. changes in interest rates affecting firms' discounting; indirect effects are small because the MPI out of changes in cash flows is zero. Furthermore, because the MPI is zero, deficit-financed transfer payments have no effect on investment and are instead simply paid out as dividends.

Both the PIH model of consumption and Q theory of investment have been routinely rejected by the data. Early work testing the PIH typically found "excess sensitivity," i.e. that consumption is positively correlated with temporary changes in income (e.g. Flavin, 1981, Campbell and Mankiw, 1989; see Jappelli and Pistaferri, 2010 for a review). More recently, well-identified empirical studies have confirmed this finding using exogenous variation in income (e.g. tax rebates in Johnson, Parker and Souleles, 2006, or lottery winners in Fagereng, Holm and Natvik, 2021). On the firm side, early work testing the Q theory typically found that average Q was not a sufficient statistic for investment, as predicted by the theory, with investment highly correlated with cash flows instead (e.g. Fazzari, Hubbard and Petersen, 1988). The promise of heterogeneous agent models is that, by moving beyond the PIH and Q theory, we can develop richer, more realistic models of policy transmission.

In response to the failure of the PIH, the heterogeneous-household literature has focused on models that can match the distribution of MPCs estimated in micro data. We summarize the main takeaways of this literature using a canonical HANK model consistent with that data. As in Kaplan et al. (2018), monetary transmission propagates through indirect effects because high-MPC households are not very sensitive to interest rates but are quite sensitive to changes in income. As a result, the power of monetary policy depends on the initial distribution of MPCs, as well as on the response of fiscal policy and the effects on the labor market. Furthermore, as in Auclert et al. (2024) and Angeletos, Lian and Wolf (2024), high-MPC households break Ricardian equivalence, so deficit-financed fiscal transfers increase consumption even when they must be paid back with future taxes. As a result, transfer payments are a potentially useful source of aggregate stimulus.

Our main contribution is to show that similar insights apply in a canonical heterogeneous-firm model of investment. In the spirit of the firm dynamics literature (Hopenhayn, 1992), heterogeneity in our model is driven by both life-cycle dynamics and idiosyncratic productivity shocks. In the spirit of the Q theory, investment is subject to quadratic capital adjustment costs. But given the empirical relevance of cash flows to investment, we also assume that firms face financial frictions in the form of collateral constraints (as in the earlier representative-agent work of Kiyotaki and Moore, 1997 and Bernanke, Gertler and Gilchrist, 1999). As a result, constrained firms have a positive MPI because additional cash flows provide additional resources which can be spent on investment; in fact, since firms can use the new capital as collateral for additional borrowing, their MPI can be larger than one.

The new transmission mechanisms in our heterogeneous-firm model are governed by two features of the MPI distribution. To understand these features, consider a deficit-financed transfer to firms. If the transfer is uniformly distributed, it directly increases aggregate investment according to the *unweighted MPI* aggregated across all firms.[2] In our calibration, the unweighted MPI is large (0.50), so the firm transfer multiplier is large—even larger than the household multiplier. Given our assumption of sticky wages, the resulting increase in aggregate demand lowers the real wage, raising firms' cash flows and therefore further stimulating investment. Since the increase in cash flows is in proportion to the firm's initial labor demand, this indirect effect is governed by the *size-weighted MPI*. In our calibration, the size-weighted MPI is small (0.05) because high-MPI firms tend to be small. By the same logic, the indirect effects of monetary policy are also relatively small in our calibration, accounting for approximately 15% of the aggregate response of investment.

We emphasize that the modest size-weighted MPI is the result of our particular calibration, not a general property of these models. In particular, our model features a strong negative correlation between a firm's size and its MPI, akin to how the first generation of household models features a strong negative correlation between a household's income and its MPC. However, direct estimates of the distribution of MPCs then emerged which indicated a much weaker correlation between income and MPCs, and this spurred a next generation of models (e.g. Kaplan and Violante, 2014). On the firm side, it may be that new direct evidence on the distribution of MPIs similarly spurs a next generation of models with larger size-weighted MPIs and therefore stronger indirect effects of monetary policy.

[2] While real-world fiscal transfers to firms generally do not take the form of a simple uniform lump-sum transfer, analyzing this policy is useful for us to highlight the parallel with household transfers. Studying more realistic policies, such as debt forgiveness programs or guaranteed loans, is an interesting avenue for future research.

These firm-level transfers also highlight new mechanisms generated by combining heterogeneous firms with heterogeneous households. In particular, because the firm transfers stimulate aggregate demand, they also raise households' income; constrained households with high MPCs respond by increasing their consumption demand, which further stimulates aggregate demand, and so on. In this sense, our model captures the intuition that providing fiscal support to constrained firms also supports household consumption. Quantitatively, we find that this additional stimulus through combining heterogeneous households and firms accounts for about 20% of the aggregate effect of the fiscal transfer.

These new mechanisms unlocked by incorporating heterogeneous firms into the New Keynesian model are currently as underexplored as their cousins were on the household side ten years ago. To reduce barriers to entry, we have developed code for solving models with both heterogeneous firms and heterogeneous households using the sequence-space approach of Auclert, Bardóczy, Rognlie and Straub (2021). The resulting model, with both sources of heterogeneity and three state variables for firms, can be solved in a matter of minutes.

Outline of paper. The rest of our paper is organized as follows. Section 2 develops the representative agent benchmark and shows that policy transmission is dominated by direct effects. Section 3 incorporates household heterogeneity and reviews the main takeaways of the HANK literature for policy transmission. Section 4 develops a canonical heterogeneous-firm model and shows how borrowing constraints generate positive MPIs for constrained firms. Section 5 describes how we use observed data to discipline MPIs and calls for more empirical research to directly estimate them. Section 6 shows how high-MPI firms generate indirect effects of monetary policy and imply that deficit-financed fiscal transfers stimulate aggregate investment. Section 7 studies the aggregate demand complementarities that arise from studying heterogeneous households and heterogeneous firms together. Finally, Section 8 summarizes the analysis by discussing various analogies between heterogeneous household and firm models and uses these analogies as a framework to suggest possible avenues for future research. Section 9 concludes.

2 Representative Agent Benchmark

We begin by studying policy transmission in a representative agent benchmark to serve as the foundation of our later models with heterogeneity. We show that this benchmark has stark predictions for policy transmission: monetary policy is driven by direct effects and fiscal transfers have no impact on aggregate spending. While these results are well-understood for consumption, we are not aware of a similar analysis for investment.

2.1 Model

Our benchmark extends the RBC model to incorporate sticky wages.[3] This benchmark is a useful middle ground between the simplest three-equation New Keynesian model, which lacks investment, and the quantitative versions, which add investment together with many other features, such as habit formation or adjustment costs to the flow of investment itself.

Time is discrete and infinite. Aggregate shocks are modeled as MIT shocks, i.e. small unexpected aggregate shocks at date 0 with perfect foresight after date 0. Due to aggregate certainty equivalence, the impulse responses to this model are the same as those in the equivalent model with aggregate uncertainty up to first order (see e.g. Auclert et al., 2021).[4]

Household. There is a representative household with preferences over sequences of consumption C_t and labor supply N_t described by the utility function $\sum_{t=0}^{\infty} \beta^t \left[u\left(C_t\right) - v\left(N_t\right)\right]$. Each period, the household receives the real wage $w_t = W_t/P_t$ per unit of labor N_t, may receive a lump-sum transfer T_t^h from the government, and can invest its wealth A_t in shares of a mutual fund. In turn, the mutual fund holds government bonds and shares of the firm.

The mutual fund provides a real realized return r_t^p per share at time t, so the household's budget constraint in period t is $C_t + A_t = \left(1 + r_t^p\right) A_{t-1} + w_t N_t + T_t^h$. Let $q_t \equiv \prod_{s=0}^{t-1} \frac{1}{1+r_{s+1}^p}$

[3] We use sticky wages, rather than sticky prices, for two reasons. First, wages change less than prices in the micro data: wages are reset on average every 12 months (Grigsby, Hurst and Yildirmaz, 2021), while prices are reset on average every 3 to 9 months (Bils and Klenow, 2004, Nakamura and Steinsson, 2008). Second, the opposite assumption of sticky prices but flexible wages generates highly (and counterfactually) countercyclical profits, and the distribution of these profits is well-understood to be key for aggregate outcomes once we add heterogeneity (Broer, Hansen, Krusell and Öberg, 2020, Bilbiie, 2025).

[4] The nonlinear impulse responses of this perfect-foresight model are also related to the generalized impulse responses of the model with aggregate uncertainty to higher order (see Auclert, Rigato, Rognlie and Straub, 2025c). How large these nonlinearities are is an open question, especially in the model with firm heterogeneity.

denote the discount factor associated with the mutual fund's expected returns. Iterating forward and using the transversality condition delivers the intertemporal budget constraint

$$\sum_{t=0}^{\infty} q_t C_t = \left(1 + r_0^p\right) A_{-1} + \sum_{t=0}^{\infty} q_t \left(w_t N_t + T_t^h\right). \tag{1}$$

The nominal wage W_t is sticky and labor is demand-determined at each date t. The household supplies that labor demand N_t, but allows a labor union to negotiate for a new level of the nominal wage. As in Erceg, Henderson and Levin (2000), the wage is reset infrequently according to the mechanism in Calvo (1983), leading to a Phillips curve for nominal wage inflation $\pi_t^w \equiv \frac{W_t}{W_{t-1}} - 1$, which is given to first order by[5]

$$\pi_t^w = \kappa \left(\frac{v'(N_t)}{u'(C_t)} - w_t\right) + \beta \pi_{t+1}^w. \tag{2}$$

Mutual fund. The mutual fund invests household wealth in government bonds B_t, which pay a promised real return r_t between time t and $t+1$, and firm shares s_t, which have (real) price p_t at time t and pay a (real) dividend D_t per share. The mutual fund's balance A_t evolves as: $p_t s_t + B_t + \left(1 + r_t^p\right) A_{t-1} = \left(p_t + D_t\right) s_{t-1} + \left(1 + r_{t-1}\right) B_{t-1} + A_t$. Since portfolios are unrestricted, equilibrium imposes the no-arbitrage condition

$$1 + r_t^p = 1 + r_{t-1} = \frac{p_t + D_t}{p_{t-1}}, \quad \text{all } t \geq 1. \tag{3}$$

The discount factors q_t can therefore also be written as $q_t \equiv \prod_{s=0}^{t-1} \frac{1}{1+r_s}$.

Normalizing the total number of shares to $s_t = 1$, the end-of-period mutual fund balance equals the value of the underlying assets $A_t = B_t + p_t$ in each period t. At date 0, when an unexpected shock may occur, the return paid to mutual fund holders must then be given by

$$1 + r_0^p = \left(\frac{p_0 + D_0}{p_{-1}}\right) \frac{p_{-1}}{A_{-1}} + \left(1 + r_{-1}\right) \frac{B_{-1}}{A_{-1}}, \tag{4}$$

where $\frac{p_{-1}}{A_{-1}} = 1 - \frac{B_{-1}}{A_{-1}}$ is the fraction of the mutual fund's portfolio allocated to stocks prior to

[5] We assume, as in Erceg et al. (2000), that employment is subsidized to eliminate the monopolistic distortion associated with the wage-setting power of unions.

date 0. The value of a stock at date 0 is given by the present discounted value of dividends

$$p_0 + D_0 = \sum_{t=0}^{\infty} q_t D_t. \tag{5}$$

Firm. There is a representative firm with production function $Y_t = K_{t-1}^{\alpha} N_t^{\nu}$, where $\alpha + \nu \leq 1$. The firm hires workers in a frictionless labor market at real wage w_t. Given that our model features nominal rigidities, some form of capital adjustment cost is necessary to limit the movement of investment in response to shocks that affect the real interest rate.[6] Specifically, we assume that the resources required to go from K_{t-1} to K_t are the sum of investment $I_t \equiv K_t - (1-\delta) K_{t-1}$ and adjustment costs $\varphi\left(\frac{K_t - K_{t-1}}{K_{t-1}}\right) K_{t-1}$, where $\varphi(\cdot)$ is a convex function satisfying $\varphi(0) = 0$ so that no adjustment costs are needed to keep the capital stock constant. The firm can also receive lump-sum transfers T_t^f from the government.

The objective of the firm is to maximize firm value (5) by choosing K_t and N_t subject to the flow of funds constraint

$$D_t = Y_t - w_t N_t - I_t - \varphi\left(\frac{K_t - K_{t-1}}{K_{t-1}}\right) K_{t-1} + T_t^f, \tag{6}$$

which determines dividend payments D_t.

Monetary and fiscal policy. The fiscal authority issues government debt B_t and provides lump sum transfers to households T_t^h and firms T_t^f (with negative transfers representing lump-sum taxes). The government budget constraint is $B_t = (1 + r_{t-1}) B_{t-1} + T_t^h + T_t^f$. We assume that the government has no initial debt, $B_{-1} = 0$, and that any change in the path of fiscal policy $\left\{T_t^h, T_t^f\right\}_{t=0}^{\infty}$ must respect the intertemporal budget constraint:

$$\sum_{t=0}^{\infty} q_t \left(T_t^h + T_t^f\right) = 0. \tag{7}$$

The monetary authority sets the nominal interest rate i_t to achieve a certain path for the

[6] Specifications of the New Keynesian model without adjustment costs suffer pathologies due to the extreme responsiveness of investment to changes in real interest rates; see, for instance, Dupor (2001) and Rupert and Šustek (2019).

ex-ante real interest rate r_t^* via

$$1 + i_t = \left(1 + r_t^*\right)\left(1 + \pi_{t+1}\right), \tag{8}$$

where π_t is consumer price inflation $1 + \pi_t = \left(1 + \pi_t^w\right)\frac{w_{t-1}}{w_t}$. This real rate rule, popularized by Woodford (2011), implies that the ex-ante real interest rate is $r_t = r_t^*$. As is well-known, a real rate rule generates equilibrium multiplicity; we select the unique equilibrium such that aggregates are back to steady state in the long-run.[7] We use this rule because it provides a simple description of monetary policy in our decompositions. It can be thought of as an intermediate case between very responsive monetary policy as under an active Taylor rule, and very passive monetary policy as at the zero lower bound (see e.g. Auclert et al., 2024).

Equilibrium. Given the monetary and fiscal policy $\left\{r_t^*, T_t^h, T_t^f\right\}$, equilibrium is a path of prices $\{w_t, \pi_t, \pi_t^w\}$ and quantities $\{C_t, I_t, K_t, Y_t, D_t, N_t\}$ such that all agents maximize and the goods market clears at all dates: $C_t + I_t + \varphi\left(\frac{K_t - K_{t-1}}{K_{t-1}}\right)K_{t-1} = Y_t$.

2.2 Equilibrium analysis in the sequence space

We now analyze the equilibrium using the sequence-space approach. This approach is particularly useful for us because it allows us to cleanly decompose the different transmission channels of macro policies both in this benchmark and in the heterogeneous agent models.

Consumption function. The representative household maximizes utility subject to its intertemporal budget constraint (1). This yields a set of Euler equations in terms of the realized return r_t^p, which combined with the no-arbitrage equations (3) imply

$$u'\left(C_t\right) = \beta\left(1 + r_t\right)u'\left(C_{t+1}\right). \tag{9}$$

[7]To implement this rule in practice, we can let the mutual funds trade nominal government bonds in addition to real bonds, and assume that the nominal bonds are in zero net supply. No arbitrage then implies a Fisher equation $1 + r_t = \frac{1+i_t}{1+\pi_{t+1}}$, which combined with (8) implies $r_t = r_t^*$. The equilibrium we obtain with our selection rule is the limit of equilibria such that $1 + i_t = \left(1 + r_t^*\right)\left(1 + \pi_{t+1}\right)^\phi$ as ϕ approaches 1 from above. This selection is natural because it implies that monetary policy has no long-run effects.

In addition, combining the household's intertemporal budget constraint (1), the mutual fund revaluation equation (4), the asset pricing equation (5), and the government's intertemporal budget constraint (7), we obtain $\sum_{t=0}^{\infty} q_t C_t = \sum_{t=0}^{\infty} q_t \left(w_t N_t + D_t\right)$, where recall that $q_t \equiv \prod_{s=0}^{t-1} \frac{1}{1+r_s}$. Note that government transfers T_t^h drop out of the intertemporal budget constraint because they are paid back with future taxes—an expression of *Ricardian equivalence*. Hence, the solution to the household problem only depends on the sequence of interest rates $\{r_s\}_{s=0}^{\infty}$ and real incomes $\{w_s N_s + D_s\}_{s=0}^{\infty}$. Let $\omega_t = w_t N_t + D_t$ be real income at t.

We denote by $\mathcal{C}_t\left(\{r_s, \omega_s\}_{s=0}^{\infty}\right)$ the implied path of aggregate consumption given these objects. This consumption function is useful to understand the forces shaping households' consumption: first, movements in r_t cause an intertemporal substitution effect through the Euler equation (9), and second, movements in total income $omega_t$ cause an income effect through the intertemporal budget constraint (1).

Since we know that consumption returns to steady state, we can also consider another consumption function $\widetilde{\mathcal{C}}_t\left(\{r_s\}_{s=0}^{\infty}\right)$ obtained by solving only the consumption Euler equation (9) given the condition at infinity $C_\infty = C_{-1}$. In equilibrium we must have at once that $C_t = \mathcal{C}_t\left(\{r_s, \omega_s\}_{s=0}^{\infty}\right) = \widetilde{\mathcal{C}}_t\left(\{r_s\}_{s=0}^{\infty}\right)$. The function $\widetilde{\mathcal{C}}_t\left(\cdot\right)$ is useful for solving for equilibrium, while $\mathcal{C}_t\left(\cdot\right)$ is useful for understanding the underlying drivers of consumption.

Investment function. The representative firm maximizes the present discounted value of dividends (5), with dividends given by (6). This delivers a version of the Q theory. Since labor is chosen statically, we have that

$$w_t = MPN_t = \nu K_{t-1}^{\alpha} N_t^{\nu-1}. \tag{10}$$

The first order condition for capital can be written as

$$1 + \varphi'\left(\frac{K_t - K_{t-1}}{K_{t-1}}\right) = \frac{1}{1+r_t}\left(MPK_{t+1} + 1 - \delta - \varphi\left(\frac{K_{t+1} - K_t}{K_t}\right) + \varphi'\left(\frac{K_{t+1} - K_t}{K_t}\right)\frac{K_{t+1}}{K_t}\right) \tag{11}$$

where $MPK_{t+1} = \alpha K_t^{\alpha-1} N_{t+1}^{\nu}$ is the expected marginal product of capital at date $t+1$, which depends on the wage at $t+1$ through the choice of labor N_{t+1}.[8] Note that government

[8]Specifically, we have $MPK_{t+1} = \alpha \nu^{\frac{\nu}{1-\nu}} K_t^{-\frac{1-(\alpha+\nu)}{1-\nu}} w_{t+1}^{-\frac{\nu}{1-\nu}}$.

transfers T_t^f do not enter either first order condition because they are additively separable from these choices. Hence, the solution to the firm problem only depends on the sequence of interest rates $\{r_s\}_{s=0}^{\infty}$ and real wages $\{w_s\}_{s=0}^{\infty}$.

We denote the implied path of investment by $\mathcal{I}_t(\{r_s, w_s\})$. Analogously to the consumption function, this investment function allows us to understand the forces shaping firms' investment. First, movements in r_t cause changes in discounting on the right-hand side of (11). Second, movements in the real wage w_t cause changes in the firm's expected marginal product of capital through the choice of labor N_{t+1}.

The firm's investment and hiring decisions determine the paths of aggregate capital and labor, and therefore aggregate output $Y_t = K_{t-1}^{\alpha} N_t^{\nu}$. We denote this path as a function of prices by $\mathcal{Y}_t(\{r_s, w_s\})$. Similarly, we denote the path of aggregate adjustment costs $\varphi\left(\frac{K_t - K_{t-1}}{K_{t-1}}\right) K_{t-1}$ by $\Phi_t(\{r_s, w_s\})$.

Equilibrium. Given a path of policies $\left\{r_t^*, T_t^h, T_t^f\right\}$ and recognizing that $r_t = r_t^*$, the key step in finding equilibrium is to obtain the sequence of real wages $\{w_s\}$ satisfying

$$\widetilde{\mathcal{C}}_t(\{r_s\}) + \mathcal{I}_t(\{r_s, w_s\}) + \Phi_t(\{r_s, w_s\}) = \mathcal{Y}_t(\{r_s, w_s\}). \tag{12}$$

Once the path of real wages and aggregates is obtained, one can find wage inflation through (2) and price inflation via $1 + \pi_t = (1 + \pi_t^w)\frac{w_{t-1}}{w_t}$. However, given our real rate rule, the paths of inflation are irrelevant for equilibrium dynamics.[9]

To first order, we can summarize the behavior of the model in terms of the derivatives of $\widetilde{\mathcal{C}}$, $\mathcal{I}$, Φ and $\mathcal{Y}$ functions with respect to the paths of prices, which are known as *sequence-space Jacobians* (Auclert et al., 2021). By the implicit function theorem, the path of real wages $d\mathbf{w} = (dw_0, dw_1, \ldots)$ that results from a change in the path of real interest rates $d\mathbf{r} = (dr_0, dr_1, \ldots)$ is given by

$$d\mathbf{w} = (\mathcal{Y}_w - \mathcal{I}_w - \Phi_w)^{-1}\left(\widetilde{\mathcal{C}}_r + \mathcal{I}_r + \Phi_r - \mathcal{Y}_r\right) d\mathbf{r}. \tag{13}$$

[9]The flexible price version of the model is determined by a sequence of real interest rates and wages $\{r_s, w_s\}$ such that the goods market clears and households' labor supply is $v'(N_t) = w_t u'(C_t)$ at all t.

The path of any other aggregate, such as output Y_t, follows from $d\mathbf{Y} = \mathcal{Y}_r d\mathbf{r} + \mathcal{Y}_w d\mathbf{w}$. The heterogeneous-agent models that we develop below also satisfy a version of (12), but differ in their aggregate consumption and/or investment functions. Viewing the models in this way makes the problem "modular" in the sense that we can easily swap out either form of heterogeneity, or bring them both together.

2.3 Calibration

We now parameterize the model in order to study the impulse responses to monetary and fiscal policy shocks. For this quantitative work, we choose standard functional forms for utility and adjustment costs:

$$u(C) = \frac{C^{1-\sigma}}{1-\sigma}, \qquad v(N) = \zeta \frac{N^{1+\psi}}{1+\psi}, \qquad \varphi(x) = \frac{\phi}{2} x^2. \tag{14}$$

The labor disutility parameters ζ and ψ are irrelevant for the aggregate dynamics of real variables given sticky wages and our real interest rate rule.

Table 1 summarizes the parameters and targeted moments in our calibration. We set the first four parameters to ensure that steady-state moments match long-run averages in the data. In particular, we set the discount factor $\beta = 0.99$ to match an annualized real interest rate of $r = 4\%$. We set the depreciation rate $\delta = 2.5\%$ quarterly to match the investment-capital ratio, and the labor coefficient $\nu = 0.6$ to match the labor share in the data. Finally, we calibrate the capital coefficient α so that there are aggregate diminishing returns to scale, $\alpha + \nu < 1$. This proxies for a return to managerial skill and makes the optimal scale of firms well-defined in the version of the model with heterogeneous productivity (e.g. Lucas, 1978). In the aggregate, decreasing returns generate rents which get capitalized into the value of the firm. We set $\alpha = 0.32$, which delivers an annual ratio of capitalized rents to GDP of 2, consistent with estimates of the markup and rent component of U.S. aggregate wealth (see, eg, Auclert, Malmberg, Rognlie and Straub, 2025a.)[10]

The remaining two parameters—the elasticity of intertemporal substitution (EIS) σ^{-1}

[10] In the steady state of our model, aggregate wealth is made of this rent-to-GDP ratio plus the capital-GDP ratio of $K/Y = 2.29$, so the total wealth-to-GDP ratio is equal to 4.29.

Table 1: Parameters and Targets for the Representative-Agent Model

Parameter	Description	Value	Target (all joint)	Value
β	Discount rate	0.99	Real interest rate	4%
ν	Labor exponent	0.6	Labor share of output	60%
α	Capital exponent	0.32	Rent/GDP ratio	200%
δ	Capital depreciation rate	0.025	Investment/Capital ratio	10%
σ^{-1}	EIS	0.8	C IRF to monetary shock	0.2%
ϕ	Capital adjustment cost	9	I IRF to monetary shock	1%

Notes: The calibration is at quarterly frequency. The parameters in the left panel are not annualized; the targets on the right are annualized.

and the adjustment cost parameter ϕ—control elasticities and therefore cannot be disciplined by steady-state moments alone.[11] Instead, we follow a long tradition in the literature of pinning down these parameters using the response of macro aggregates to identified monetary policy shocks; in particular, we match the peak impulse responses of consumption and investment from Christiano et al. (2005). While there exist many alternative ways of identifying monetary policy shocks in the data, these alternatives tend to deliver qualitatively similar aggregate responses.

In response to a typical monetary policy shock, Christiano et al. (2005) find that the real interest rate falls by 50 basis points, consumption rises by 0.2%, and investment rises by 1%.[12] We model their monetary shock as a decline in the real interest rate target $r_t^* = r - \frac{0.5}{400} \cdot (0.5)^t$ for $t \geq 0$, implying an impact effect at date 0 of -50 basis points (annualized) and an autocorrelation of 0.5. We then choose $\sigma^{-1} = 0.8$ and $\phi = 9$ to match the peak responses of consumption and investment.

The resulting paths for the real interest rate, output, consumption and investment are displayed in Figure 1. Early on, the decline in the real interest rate increases consumption and investment demand, resulting in excess demand in the goods market that must be resolved by a decline in the real wage. Later on, consumption and investment demand have

[11] In the heterogeneous agent models, individual consumption and investment decisions still vary in response to idiosyncratic shocks, even though aggregates are constant. In this sense, heterogeneous-agent models can also let one calibrate elasticities in steady state, though we do not pursue this route for σ^{-1} or ϕ.

[12] Since consumption represents 80% of private spending, this implies that investment makes up over 55% $(= \frac{0.2\times 1}{0.8\times 0.2+0.2\times 1})$ of the total response of private spending to monetary policy in the data. Hence, it is critical for quantitative models of monetary policy to feature investment as an endogenous component of aggregate demand.

returned to steady state but the capital stock is higher, resulting in excess supply in the goods market that must be resolved by an increase in the real wage.

In our model, as well as our later heterogeneous-agent models, the peak response to a monetary policy shock occurs on impact: the model does not feature ingredients to delay (and reduce) the consumption and investment responses, such as habit formation in consumption or adjustment costs on the change in investment. As a result, our calibrated adjustment costs are higher than typically estimated in the literature.[13]

Calibration approach across models. As we add micro-level heterogeneity in Sections 3 and 4, we will recalibrate the EIS σ^{-1} and adjustment costs ϕ to match the same aggregate responses to a monetary shock. We do so because aggregate data provide important discipline on the aggregate strength of various mechanisms operating at the micro level. In this sense, the exact numbers we target are less important than the fact that we maintain this target across all models we study.

Importantly, the fact that the aggregate response to monetary policy is the same across models does not mean that micro-level heterogeneity is irrelevant for those responses. First, the channels through which monetary policy affects aggregates are different across models, implying that the monetary authority (and researchers) need to consider a wider variety of issues when formulating policy. Second, given these different transmission mechanisms, the response to other shocks—in our context, fiscal transfers—will be different as well.

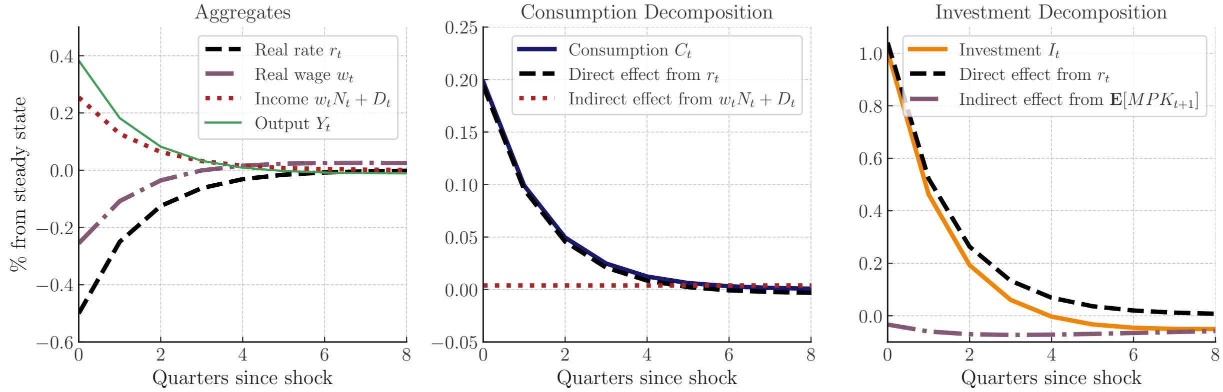
2.4 Policy Transmission in the Representative Agent Benchmark

We now use the consumption and investment functions to decompose the transmission channels of monetary and fiscal policy.

Consumption and the Permanent Income Hypothesis. We first consider the channels driving the response of aggregate consumption to a monetary shock. Using the consumption function, $C_t = \mathcal{C}_t(\{r_s, \omega_s\})$, we can write the first-order impulse response of consumption

[13] See Auclert, Rognlie and Straub (2020) for an example of how humps can be reconciled with household heterogeneity and MPCs.

FIGURE 1: Monetary Transmission in the Representative Agent Benchmark



Notes: impulse responses of aggregates (left), consumption (middle, blue line) and investment (right, orange line) to a monetary policy shock modeled as a 50bp annualized decline in the real interest rate with quarterly persistence of 0.5, similar to Christiano et al. (2005). All variables are in percent deviations from their steady state values, except the real interest rate which is in annualized percentage points. The middle panel also displays the decomposition of consumption into a direct and indirect effect following equation (15). The right panel also displays the decomposition of investment into a direct and indirect effect by considering the first order effects from equilibrium movements in r_t and MPK_{t+1} on I_t implied by equations (18)–(19).

$d\mathbf{C} = (dC_0, dC_1, \dots)$ as

$$d\mathbf{C} = \underbrace{\mathcal{C}_r d\mathbf{r}}_{\text{direct effect}} + \underbrace{\mathcal{C}_\omega d\boldsymbol{\omega}}_{\text{indirect effect}} . \tag{15}$$

This consumption decomposition writes the change in consumption as the sum of the effects from the change in the real interest rate, which the literature calls the "direct effect," and the effect from equilibrium changes in wages and dividend, called the "indirect effect."

In the representative-agent model, the sequence-space Jacobians $\mathcal{C}_r$ and $\mathcal{C}_\omega$ have simple analytical expressions, which imply:

$$(\mathcal{C}_r d\mathbf{r})_t = \frac{C}{\sigma}\left(-\sum_{s=t}^{\infty}\frac{dr_s^*}{1+r} + (1-\beta)\sum_{s=0}^{\infty}\beta^s\sum_{j\geq s}\frac{dr_j^*}{1+r}\right) \tag{16}$$

$$(\mathcal{C}_\omega d\boldsymbol{\omega})_t = (1-\beta)\sum_{s=0}^{\infty}\beta^s d\omega_s. \tag{17}$$

The direct effect is the substitution effect: consumption is increased at t by any decline in real interest rates at t or beyond, net of a constant term that ensures that the present discounted value of the direct effect is zero. The scale of this effect is governed by the elasticity of intertemporal substitution $\frac{1}{\sigma}$. The indirect effect is the income effect: the present value of labor and dividend income rises, which increases the household's permanent income in their intertemporal budget constraint—Friedman (1957)'s Permanent Income Hypothesis (PIH). Each period, the household consumes the annuity value of that income, such that their MPC is equal to $1-\beta$.

The middle panel of Figure 1 shows that nearly all of the response of consumption to the monetary shock is driven by the direct effect; indirect effects are small because the MPC is $1-\beta = 0.01$. To delve even deeper into this result, recall that our real interest rate rule implies that we have the alternative consumption function $\widetilde{\mathcal{C}}_t(\{r_s\})$, which implies $dC_t = -\frac{C}{\sigma}\sum_{s=t}^{\infty}\frac{dr_s^*}{1+r}$. It follows that the indirect effect in (17) is exactly offset by the second term in the direct effect in equation (16).

Turning to fiscal policy, since transfers $\{T_t^h, T_t^f\}$ do not enter the consumption function, it immediately follows that households do not respond at all to deficit-financed changes in transfers to either households and firms. Again, this is a reflection of Ricardian equivalence, whose logic is closely tied to the PIH: the present value of any change in transfers is zero, so transfers do not affect the household's intertemporal budget constraint; hence, the effective marginal propensity to consume out of current transfers is exactly zero.

We emphasize that the key mechanisms driving these results—the PIH and Ricardian equivalence—have been rejected by the empirical literature (see Jappelli and Pistaferri, 2010 for a survey). Instead, this literature has consistently found large marginal propensities to consume out of transitory windfall income shocks, as well as non-zero marginal propensities to consume out of deficit-financed government transfers. The heterogeneous-household model in Section 3 resolves these shortcomings.

Investment and the Q theory. We now study the transmission channels of monetary policy to investment. To motivate our decomposition, it is useful to go back to the equation determining investment in (11). Using our functional form for adjustment costs (14), we get

$$\frac{I_t}{K_{t-1}} - \delta = \frac{1}{\phi}\left(\frac{1}{1+r_t}V'_{t+1} - 1\right), \tag{18}$$

where V'_{t+1}, the marginal value of capital at $t+1$, is given by the recursion

$$V'_{t+1} = MPK_{t+1}\left(w_{t+1}, K_t\right) - \frac{I_{t+1}}{K_t} - \frac{\phi}{2}\left(\frac{I_{t+1}}{K_t} - \delta\right)^2 + \frac{1}{1+r_{t+1}}\frac{K_{t+1}}{K_t}V'_{t+2}. \tag{19}$$

These equations encode the standard forces from the Q theory of investment: equation (18) says that net investment is positive as long as the discounted marginal value of capital (also known as marginal Q_t) exceeds 1, while equation (19) governs the dynamic evolution of Q_t, capturing the present discounted value of the marginal product of capital in future periods.

Equations (18) and (19) clarify the direct and indirect effects of monetary policy in the representative-agent benchmark. The direct effect $\mathcal{I}_r d\mathbf{r}$ stems from the fact that lower real interest rates raise firm value, thereby increasing the incentive to invest.[14] This effect is highest on impact of the real interest rate decline, when the effect of the new path of interest rates on present discounted values is the highest. The indirect effect $\mathcal{I}_w d\mathbf{w}$ stems from the fact that higher real wages result in less hiring N_{t+1} and therefore a lower marginal product of capital in period $t+1$. Since this effect only runs through changes in MPK_{t+1}, we call this the indirect effect from the expected marginal product of capital. Note that changes in w_0 have no effect on investment at any date; they affect dividends but not investment.

The right panel of Figure 1 shows that the response of investment to monetary policy is primarily driven by the direct effects. In fact, the direct effects account for slightly more than 100% of the increase in investment, with the indirect effect from the change in the real wage being slightly negative over the first eight quarters of the shock. This occurs because firms incorporate future wages into (19) in a forward-looking way. While the initial period of low real wages would, on its own, raise investment, the following prolonged period of high real wages pushes investment down. However, this effect is quite small, since the path of wages does not significantly change the present discounted value of the marginal products.

Turning to fiscal policy, lump-sum transfers to firms at any date have no effect on investment because those transfers do not interact with the cost or benefit of investment—in other words, the MPI in this representative agent benchmark is 0. Since firms are already at

[14] For instance, consider a purely transitory change in the real interest rate dr. This essentially does not affect the present discounted value of capital V'. Therefore, we have that $\frac{dI}{K} = \frac{1}{\phi}dr$, so the semielasticity of investment to the real interest rate is approximately $\frac{1}{\phi}$.

an optimum of their investment plan, they instead distribute the proceeds from the transfer back to their shareholders, per equation (6).[15]

Just like the PIH, Q theory was subject to extensive empirical testing starting in the 1980s. A key challenge for that literature is that marginal Q, $V'/(1+r)$, is not directly observable in the data. In the special case with constant returns to scale $\alpha+\nu=1$, however, marginal Q is proportional to average Q, which we can compute as the firm's market value relative to its capital stock (Hayashi, 1982). In this case, equation (18) predicts that a regression of the investment rate on average Q should have an R^2 of 1 and a coefficient on average Q equal to $1/\phi$. By and large, the literature has found that these types of regressions produce very low R^2 because observables beyond Q, especially cash flows, tend to predict investment (see, e.g., Fazzari et al., 1988).

The investment literature largely proceeded along two branches following the empirical failures of the Q theory. The first branch interpreted the importance of cash flows as suggestive evidence that the MPI is positive and therefore studied various forms of financial frictions (e.g. Gomes, 2001, Hennessy and Whited, 2007). We will follow this branch in Section 4 given our focus on indirect effects due to changes in aggregate demand. The second branch of the literature questioned the assumption that the marginal cost of investment is smoothly increasing in the amount of investment, which is required to arrive at (18). This branch incorporated nonconvex costs of adjusting capital, motivated by the empirical lumpiness of investment at the micro level (e.g. Caballero and Engel, 1999, Cooper and Haltiwanger, 2006, and Khan and Thomas, 2008). While we will not review those models given the constraints of space, studying the role of fixed costs in monetary transmission is a potentially fruitful avenue for future research.

3 Heterogeneous Households: A Review of HANK

We now replace the representative household with heterogeneous households facing idiosyncratic income risk and borrowing constraints (Zeldes, 1989, Deaton, 1991, Carroll, 1997;

[15] If we gave firms access to a borrowing and savings vehicle at rate $1+r_t$, the timing of the distribution of dividends would also not be pinned down, a reflection of the Modigliani-Miller theorem.

Bewley, 1977, Huggett, 1996, Aiyagari, 1994.). Given that the material here is relatively standard (see, for instance, Auclert et al., 2025b for a recent review of this literature), our presentation will be fairly succinct and focus on aspects of the model that parallel the findings with heterogeneous firms.

3.1 Model

A mass 1 of ex-ante identical, but ex-post heterogeneous households are subject to idiosyncratic risk to their productivity e and their rate of time discounting β.[16] The joint evolution of (e, β) is governed by a Markov transition matrix Π, and we assume that the cross-sectional distribution of (e, β) starts at the stationary distribution of Π, normalizing $\mathbb{E}[e] = 1$. The environment is identical to that in Section 2, except that households' wage income in productivity state e is now ew_tN_t, and they face incomplete markets in the sense that the mutual fund payoffs cannot be made contingent on their realization of their idiosyncratic shocks (e, β). In addition, we assume that households face a borrowing constraint, which for simplicity we set to zero, so that the choice of assets a' must satisfy $a' \geq 0$.

The dynamic problem of a household with exogenous state (e, β) and incoming asset position a is described by the Bellman equation:

$$\begin{aligned} v_t(a, e, \beta) \quad &= \quad \max_{c,a'} u(c) + \beta \mathbb{E}[v_{t+1}(a', e', \beta')] \\ \text{s.t.} \quad & c + a' = (1 + r_t^p) a + ew_tN_t + T_t^h \\ & a' \geq 0. \end{aligned} \tag{20}$$

In the representative agent model, the solution was characterized by the Euler equation (9); here, the Euler equation only holds with the weak inequality

$$u'(c_t(a, e, \beta)) \geq \beta \left(1 + r_{t+1}^p\right) \mathbb{E}[u'(c_{t+1}(a', e', \beta'))].$$

This inequality is strict whenever the borrowing constraint $a' \geq 0$ binds; in this case, con-

[16] Shocks to the discount factor β help us jointly match the wealth distribution and average MPCs, as we describe below.

sumption is determined residually from the budget constraint so that $c = (1+r_t^p)a + ew_tN_t + T_t^h$. For these households, consumption responds one-for-one to changes in current income, generating an MPC of 1.

We represent the aggregate consumption sequence from this model with the new consumption function $\mathcal{C}_t\left(\left\{r_t^p, w_tN_t, T_t^h\right\}\right)$. Intuitively, this consumption function can be computed as follows. First, given the sequence of real interest rates $\{r_s^p\}$, labor income w_tN_t, and transfers T_t^h, compute households' decision rules by solving (20) as a function of individual states (a, e, β). Second, starting from the stationary distribution of households over (a, e, β), use the decision rules to iterate forward on the law of motion to compute the path of cross-sectional distributions over time. Finally, compute aggregate consumption by integrating individual consumption rules against the cross-sectional distribution for each $t \geq 0$.

3.2 Calibration and the MPC vs. wealth tradeoff

We now calibrate the model in order to assess its implications for monetary transmission. As discussed in Section 2, the MPC out of temporary changes in income is critical for determining the size of the indirect effects of monetary policy and the effects of fiscal transfers. The quasi-experimental empirical literature has settled on a consensus that the average MPC at a quarterly frequency is about 0.25 (see, e.g., the reviews in Kaplan and Violante, 2022 and Auclert, Bardóczy and Rognlie, 2023).

As discussed in Kaplan and Violante (2022), in a model with a single discount factor β, there is a fundamental tension between matching both the average level of the MPC and the aggregate amount of wealth in the economy. In our model, aggregate assets equal the value of firms because steady state government debt is 0; the value of firms to GDP is $\frac{A}{Y} = 4.29$, the sum of the capital-to-GDP and the rent-to-GDP ratio; given the labor share is $\frac{wN}{Y} = 0.6$, aggregate assets are 7.15 times average labor income. At our annualized real interest rate of $r = 4\%$, a single-β model calibrated to this amount of wealth features very patient agents that are away from the borrowing constraint and have on average low MPCs, while a model that hits the average MPC features very impatient agents that are close to the borrowing constraint and hold too little wealth on average.

Table 2: Parameters and Targets for the Heterogeneous Household Block

Parameter	Description	Value	Target (all joint)	Value
Markov Chain for (e, β) from Auclert et al. (2025b)				
r	Real interest rate	0.01	Real interest rate	4%
β_L	High discount rate	0.994	Assets/labor income	715%
β_H	Low discount rate	0.824	Average MPC	0.25
σ^{-1}	EIS	0.61	C IRF to monetary shock	0.2%

Notes: parameters in the left panel chosen to match moments in the right panel. The calibration is at quarterly frequency. The parameters in the left panel are not annualized; the targets on the right are annualized. Production side parameters are as in the representative agent benchmark from Section 2. The asset-to-labor income A/wN ratio and real interest rate are targeted to be the same as in that model, so that the aggregate steady-states are identical in both models.

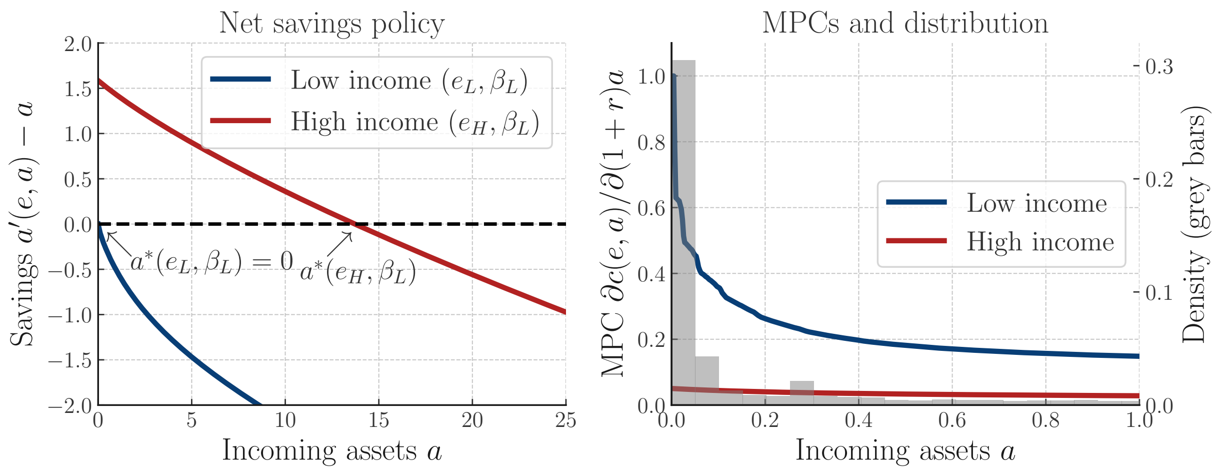
We resolve this tradeoff by assuming heterogeneity in β, as in Krusell and Smith (1998).[17] We calibrate the model using the Markov chain for income and impatience described in Auclert et al. (2025b): the income process is from Kaplan et al. (2018), featuring a high degree of income risk, and the impatience process is a two-point process for $\beta \in \{\beta_L, \beta_H\}$, describing the slow evolution of patience across generations as in Krusell and Smith (1998). We set (β_L, β_H) to simultaneously match an (unweighted) average MPC of $\mathbb{E}[MPC] = 0.25$ and an asset to wage income ratio $A/wN = 7.15$ at our calibrated annual $r = 4\%$.

Table 2 contains the parameter values for the heterogeneous household model. In addition to the parameters described above, we also re-calibrate the EIS σ^{-1} so that we continue to match the impact response of aggregate consumption to a monetary shock. With heterogeneous households, this choice requires a significantly lower EIS than in the representative agent benchmark (0.61 vs 0.8). The reason is that, as we will show below, the presence of high-MPC households generates an amplification mechanism: lower interest rates stimulate investment and therefore income, which in turn increases consumption (see Auclert et al., 2020 and Bilbiie, Känzig and Surico, 2022). Offsetting this effect requires consumption to be less responsive to interest rates in the first place, i.e., a lower EIS.[18]

[17] Other solutions in the literature certainly exist, with the leading approach being the introduction of a high-return illiquid account that households use for most of their savings, but that is not useful for smoothing period-to-period income fluctuations (e.g. Kaplan and Violante, 2014).

[18] If adjustment costs were infinite, so that investment had a completely inelastic response to monetary policy, then the Werning (2015) result of equivalence between the aggregate consumption behavior of the heterogeneous-agent model and representative-agent model would apply and we would not need to recalibrate the EIS. In other words, it is the elasticity of the investment response that activates this amplification mechanism. See Auclert et al. (2020), Appendix A, for more discussion of this point.

FIGURE 2: Savings Decision Rules, MPCs, and Distribution with Heterogeneous Households



Notes: households' decision rules in steady state of heterogeneous-household model calibrated as in table 2. The left panel plots net savings $a'(e,a) - a$ for impatient households ($\beta = \beta_L$) with two different levels of idiosyncratic productivity e as a function of a. The right panel plots the marginal propensities to consume $\frac{\partial c(a,e)}{\partial(1+r)a}$ for these households. The grey bars show the location of the stationary household distribution.

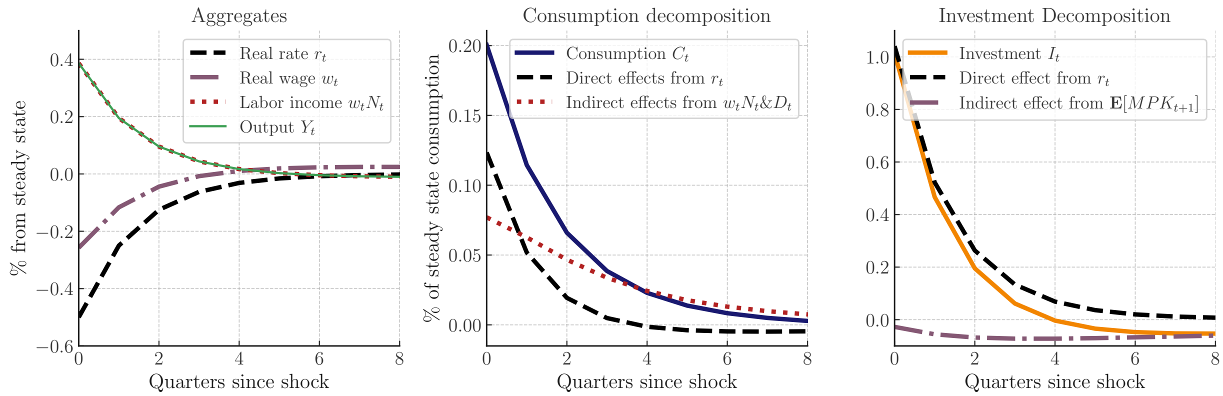
Savings decisions and MPCs. The left panel in Figure 2 visualizes the steady state net savings policy $a'(a,e,\beta) - a$ as a function of households' incoming asset position a, for two types of households with relatively low income e_L and relatively high income e_H (both households in this illustration are impatient, β_L).

The high-income household in red has positive net savings for low initial assets a in order to reduce the probability of becoming borrowing constrained, which generates costly consumption fluctuations. On the other hand, the household has negative net savings for high initial assets a because $\beta_L(1+r) < 1$. Therefore, there is a unique target level of assets $a^*(e,\beta)$ to which the household will converge if their idiosyncratic state (e,β) were to remain constant. This target, or *buffer stock*, of assets balances their impatience with their desire for precautionary savings against future declines in income (see, eg, Carroll, 1997).

The low-income household in blue dissaves for any positive level of initial assets a. In other words, when a household receives a negative shock and enters this low-income state, they use their previously accumulated buffer stock to limit the drop in their consumption induced by the decline in income.

The right panel of Figure 2 plots the implied MPCs of these agents. We compute the

Figure 3: Monetary Transmission in the Heterogeneous Household Model



Notes: impulse responses of aggregates (left), consumption (middle, blue line) and investment (right, orange line) to a monetary policy shock modeled as a 50bp annualized decline in the real interest rate with quarterly persistence of 0.5, similar to Christiano et al. (2005). All variables are in percent deviations from their steady state values, except the real interest rate which is in annualized percentage points. The middle panel also displays the decomposition of consumption into a direct and indirect effect in the spirit of equation (15). The right panel also displays the decomposition of investment into a direct and indirect effect by considering the first order effects from equilibrium movements in r_t and MPK_{t+1} on I_t implied by equations (18)–(19).

MPC as $\frac{1}{(1+r)}\frac{\partial c(a,e,\beta)}{\partial a}$, since a one-time transfer is equivalent to a one-time increase in current-period assets. As is well known, the consumption function $c(a,e,\beta)$ is concave in this model (Carroll and Kimball, 1996). This implies declining marginal propensities to consume, which equal 1 at the borrowing constraint and then fall for higher levels of a. The grey bars overlay the stationary distribution of households; because of the presence of impatient agents, our calibration locates a part of the distribution close to the borrowing constraint with high MPCs, while another part of the distribution is made up of high-income patient agents with very high target stocks of assets.

3.3 Policy transmission with heterogeneous households

We now revisit the transmission of monetary and fiscal policy with heterogeneous households.

Monetary policy. Figure 3 plots the impulse responses to the monetary policy shock, analogously to Figure 1 from the representative agent benchmark. The aggregate impulse responses are nearly identical, by construction: we re-calibrated the EIS σ^{-1} so that the impact effect is exactly the same, and the heterogeneous-household model does not have very different dynamics after that. Given that all other aggregate paths are the same, so is the path of aggregate investment.

However, the decomposition of consumption into direct and indirect effects is starkly different: the indirect effects of monetary policy now represent about 40% of the consumption response, compared to near 0% in the representative agent benchmark.[19] The reason is that constrained households have high MPCs: the initial increase in consumption and investment from monetary policy generates additional income for these households, and their high MPCs imply that this additional increase in income translates into higher consumption. This is the key result in Kaplan et al. (2018). As they emphasize, the fact that the transmission channels are different implies monetary policymakers and researchers need to consider a host of new mechanisms more carefully than in the representative agent benchmark.

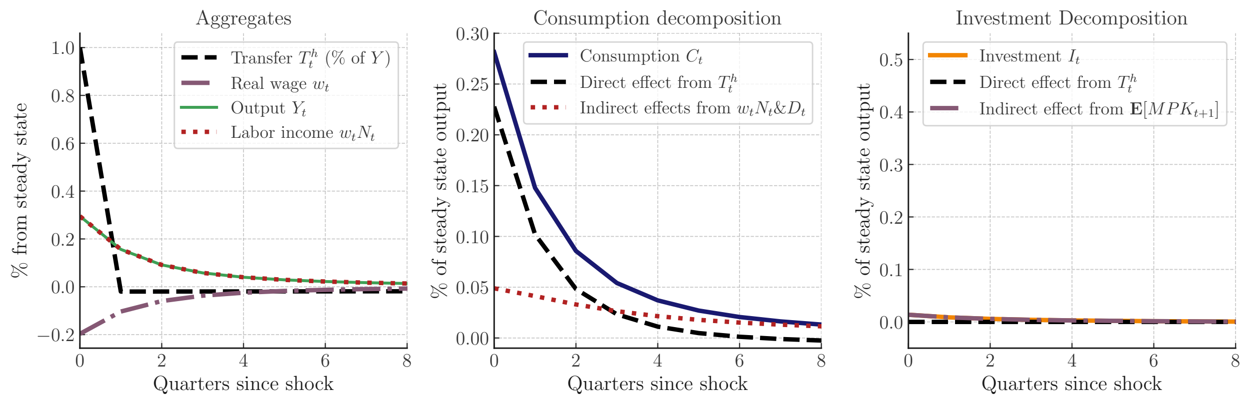
Fiscal policy. Figure 4 plots the impulse responses to a lump-sum transfer to households T_t^h. Upon impact, the aggregate size of the transfer is equal to 1% of GDP and is financed by government borrowing. Over time, the government pays off the debt through lower future taxes; the entire implied path of transfers is in the left panel of Figure 4.[20] As we described in Section 2, this type of "NPV-0" transfer has no effect on consumption in the representative agent benchmark due to Ricardian equivalence.

Our heterogeneous household model breaks Ricardian equivalence because borrowing-constrained households' decisions are dictated by their current budget constraint only, not their intertemporal budget constraint. In other words, constrained households' MPCs out of the transfer are 1. Since the transfer is uniformly distributed across households, its direct effect is determined by the *unweighted MPC* averaged across households, and is close to 0.25 upon impact. The higher consumption demand then increases aggregate income $w_t N_t$, which increases individual income $e w_t N_t$ in proportion to labor productivity e. These additional rounds of spending from higher incomes are controlled by the *income-weighted MPC*, which

[19] In doing this decomposition, we count the component of the capital gain that comes from rising dividends as part of the indirect effect, and the component that comes from falling real interest rates as part of the direct effect, in order to keep it comparable to the representative-household decomposition.

[20] Specifically, we parameterize the path of transfers such that government debt follows an AR(1) path with quarterly persistence $\rho_B = 0.99$.

Figure 4: Fiscal Policy Transmission in the Heterogeneous Household Model



Notes: impulse responses of aggregates (left), consumption (middle, blue line) and investment (right, orange line) to a transfer shock of 1% of GDP, financed by an increase in debt that reverts back to steady state with an autocorrelation of 0.99. Variables in the left panel are in percent deviations from their steady state values except for the transfer, which is in percent of GDP. Variables in the middle and right panels are in percent of GDP. The middle panel also displays the decomposition of consumption into a direct and indirect effect in the spirit of equation (15). The right panel also displays the decomposition of investment into a direct and indirect effect by considering the first order effects from equilibrium movements in r_t and MPK_{t+1} on I_t implied by equations (18)–(19).

is equal to 0.19 in our calibration. The overall transfer multiplier, which combines both the direct and indirect effects, is nearly 0.3 upon impact.

Over time, consumption remains elevated for approximately two years, reflecting endogenous propagation of this one-time transfer. The reason is that households near the borrowing constraint, but not actually constrained, have an MPC which is significantly above $1 - \beta$—so they consume out of the shock upon impact—but is also significantly below 1—so they also save part of the transfer to consume in future periods. Auclert et al. (2024) call these dynamic responses the *intertemporal MPCs* or "iMPCs."

4 Introducing Heterogeneous Firms

The heterogeneous household model unlocked new policy transmission channels for consumption by matching data on MPCs. We now turn to performing a similar analysis for the role of firm heterogeneity in investment. While the consumption block was built on the foundation of the standard incomplete markets model, no widely-used benchmark exists on the firm side. One objective of this section is therefore to build a canonical model capturing the core elements of both firm dynamics models (as in Hopenhayn, 1992) and financial frictions models (as in Kiyotaki and Moore, 1997 and Bernanke et al., 1999). Specifically, our model is a version of Ottonello and Winberry (2020), with collateral constraints replacing endogenous debt default, and extended to include firm-level capital adjustment costs.

4.1 Model

A mass 1 of ex-ante identical, but ex-post heterogeneous firms are subject to idiosyncratic risk to their productivity z and their ability to operate (exit shocks). The evolution of z is governed by a Markov transition matrix Π^z, and we assume that new entrants always draw from the stationary distribution of Π^z, such that the cross–sectional distribution of z is always equal to its stationary distribution. Given its observed productivity z_{jt}, firm j produces output using the production function $y_{jt} = z_{jt} k_{jt-1}^{\alpha} n_{jt}^{\nu}$, where k_{jt-1} is its capital stock chosen at $t-1$ and n_{jt} is its labor demand at t. Decreasing returns to scale $\alpha + \nu < 1$ ensure that there is an optimal scale of the firm and therefore a non-degenerate distribution of firm sizes.

After production, firms receive an idiosyncratic exit shock and make investment decisions. With probability θ, the firm continues into the next period, in which case it chooses its investment i_{jt} in order to accumulate capital $k_{jt} = (1-\delta)k_{jt-1} + i_{jt}$. Investment incurs adjustment costs $\varphi\left(\frac{k_{jt}-k_{jt-1}}{k_{jt-1}}\right) k_{jt-1}$ in units of output, as in Section 2. With complementary probability $1-\theta$, the firm is forced to exit the economy, in which case it sells its undepreciated capital, pays back its debt (described below), and permanently exits the economy. The capital sale of these exiting firms is not subject to adjustment costs.

Firms have two sources of finance for their investment decisions, both of which are subject to frictions. First, the firm can issue debt in the form of coupon bonds, receiving $\frac{b_{jt}}{1+r_t}$ units of output in period t for each promise to repay b_{jt} units of output in period $t+1$ (with $b_{jt} < 0$ interpreted as financial savings).[21] This borrowing is subject to the collateral constraint

[21]Therefore, debt in our model is real and short term. Jungherr, Meier, Reinelt and Schott (2024) incorporate both short- and long-term nominal debt to study how maturity structure affects the transmission of monetary shocks.

$b_{jt} \leq \chi k_{jt}$ with $\chi \leq 1-\delta$.[22,23] The second source of finance is the firm's equity, i.e., using internal resources and therefore reducing dividend payments d_{jt} to shareholders. However, firms cannot issue new equity, implying that[24]

$$d_{jt} = y_{jt} - w_t n_{jt} - b_{jt} - (k_{jt} - (1-\delta)k_{jt-1}) - \varphi\left(\frac{k_{jt} - k_{jt-1}}{k_{jt-1}}\right) k_{jt-1} + \frac{b_{jt}}{1+r_t} + T_t^f \geq 0.$$

One can view our representative agent benchmark as an all-equity financed firm in which we implicitly removed this non-negativity constraint on dividends.

We employ a simple "one out, one in" entry process to keep the mass of firms fixed over time. As described above, exiting firms repay their debt and sell their undepreciated capital before exiting. Entering firms are endowed with k_0 units of capital, b_0 units of debt, and draw idiosyncratic productivity z from the stationary distribution. One can view $\frac{b_0}{1+r_t} - k_0$ as the initial equity injection into the firm.

4.2 Solving the Firm's Problem

We now characterize the solution to the firm's problem and show that financial frictions generate positive MPIs. The dynamic problem of a continuing firm with exogenous productivity

[22] Collateral constraints can be derived from an environment in which firms have limited commitment to repay debt and lenders are only willing to lend up to the amount which firms would repay with probability one. Ottonello and Winberry (2020) develop a model in which firms can default, and lenders compensate for this risk by charging higher credit spreads. An alternative to collateral constraints is earnings-based constraints, with firms' borrowing limited by their cash flows rather than their capital assets. Lian and Ma (2021) suggest that these constraints are quite prevalent empirically, and Caglio, Darst and Kalemli-Özcan (2021), Drechsel (2023), and He (2025) explore their macroeconomic implications.

[23] Our model features a "financial accelerator" in the sense that the investment of constrained firms endogenously declines with aggregate economic activity. However, since we have a one-good model, the relative price of capital is always 1, ruling out changes in the market value of capital directly relaxing the constraint as in Bernanke et al. (1999). One could endogenize the relative price of capital by incorporating "external adjustment costs," i.e. curvature in the transformation of output into investment goods. Disciplining these costs is a matter of some debate given the relative acyclicality of the relative price of capital in the data (e.g. House and Shapiro, 2008, Schmitt-Grohé and Uribe, 2012). Therefore, as our baseline, we keep all adjustment costs internal to the firm.

[24] One could relax this assumption by incorporating costly equity issuance as in Chen (2025).

state z, stock of capital k, and outstanding debt b can be written as follows:

$$v_t(z,k,b) = \max_{n,k',b',d} d + \frac{1}{1+r_t}\mathbb{E}\left[(1-\theta)v_{t+1}(z',k',b') + \theta v_{t+1}^{\text{exit}}(z',k',b')|z\right]$$
$$\text{s.t. } d = zk^\alpha n^\nu - w_t n - b - (k'-(1-\delta)k) - \varphi\left(\frac{k'-k}{k}\right)k + \frac{b'}{1+r_t} + T_t^f$$
$$d \geq 0 \text{ and } b' \leq \chi k', \quad \text{ and where } v_t^{\text{exit}}(z,k,b) = zk^\alpha n^\nu - w_t n - b + (1-\delta)k.$$

As in Section 2, labor demand equalizes the wage to the marginal product of labor (see equation (10)).[25] We define variable profits as

$$\pi_t(z,k) \equiv \max_n zk^\alpha n^\nu - w_t n = \widetilde{\nu} z^{\frac{1}{1-\nu}} w_t^{-\frac{\nu}{1-\nu}} k^{\frac{\alpha}{1-\nu}}$$

and refer to $\pi_t(z,k)$ as the firm's *cash flows*. The dynamic investment and financing problem is more complicated than the households' consumption/savings decisions due to the two occasionally-binding constraints $d \geq 0$ and $b' \leq \chi k'$. Nevertheless, the solution is still tractable, following the approach in Khan and Thomas (2013). We next summarize the properties of this solution, with details provided in Appendix A.

The multiplier on the non-negativity constraint for dividends, which we denote by $\lambda_t(z,k,b)$, summarizes the tightness of financial constraints. In particular, the marginal value of an additional unit of cash inside the firm is $-\frac{\partial v_t(z,k,b)}{\partial b} = 1 + \lambda_t(z,k,b)$, the value of the cash itself, 1, plus the shadow value of relaxing the non-negativity constraint, $\lambda_t(z,k,b)$. In other words, $\lambda_t(z,k,b)$ captures the additional value of liquidity inside the firm compared to outside the firm (Holmstrom and Tirole, 2011). At the optimum, firms equate $\lambda_t(z,k,b)$ to the present value of multipliers on the borrowing constraint: $\lambda_t(z,k,b) = \mu_t(z,k,b)(1+r_t) + (1-\theta)\mathbb{E}_t[\lambda_{t+1}(z',k',b')]$. Hence, resources inside the firm are more valuable than resources outside the firm to the extent that they relax current or future borrowing constraints; if there is any probability of being borrowing constrained, then $\lambda_t(z,k,b) > 0$ and the firm does not issue dividends.

With these objects in hand, we can characterize firms' investment decisions as follows.

[25] In particular, the marginal propensity to hire (MPH)—the response of employment to a firm transfer, holding the wage fixed—is zero for all firms in this model. See Melcangi (2022) for evidence of positive MPH for some firms and a heterogeneous-firm model with a working capital constraint that generates this fact.

***Proposition* 1.** *Consider a firm that will continue into period $t+1$ and has initial state variable (z,k,b). Its optimal decisions are characterized by three regions:*

(i) ***Unconstrained:*** *If $b \leq \underline{b}_t(z,k)$, then $\lambda_t(z,k,b)=0$. In this case, the capital accumulation policy k' is independent of initial debt b and solves*

$$\frac{k'-k}{k} = \frac{1}{\phi}\left(\frac{1}{1+r_t}\mathbb{E}_t\left[\pi_{2,t+1}(z',k') + 1 - \delta + (1-\theta)\frac{\partial}{\partial k'}\varphi\left(\frac{k''-k'}{k'}\right)k'\right] - 1\right). \tag{21}$$

(ii) ***Actively constrained:*** *If $b \geq \overline{b}_t(z,k,b)$, then both $d \geq 0$ and $b' \leq \chi k'$ are binding. In this case, the capital accumulation policy is $k' = \overline{k}_t(z,k,b)$, where $\overline{k}_t(z,k,b)$ solves*

$$\overline{k}_t(z,k) - (1-\delta)k + \varphi\left(\frac{\overline{k}_t(z,k)-k}{k}\right)k = \pi_t(z,k) + T_t^f - b + \frac{\chi\overline{k}_t(z,k)}{1+r_t}. \tag{22}$$

(iii) ***Potentially constrained:*** *If $b \in (\underline{b}_t(z,k), \overline{b}_t(z,k,b))$, then $d \geq 0$ is binding but $b' \leq \chi k'$ is slack. In this case, the capital accumulation policy solves*

$$\frac{k'-k}{k} = \frac{1}{\phi}\left(\frac{1}{1+r_t}\mathbb{E}_t\left[\begin{array}{l}\frac{\theta+(1-\theta)(1+\lambda_{t+1}(z',k',b'))}{1+\lambda_t(z,k,b)}\pi_{2,t+1}(z',k') + 1 - \delta \\ + \frac{(1-\theta)(1+\lambda_{t+1}(z',k',b'))}{1+\lambda_t(z,k,b)}\frac{\partial}{\partial k'}\varphi\left(\frac{k''-k'}{k'}\right)k'\end{array}\right] - 1\right). \tag{23}$$

Proof. See Appendix A. ∎

Firms with low enough debt (or high enough financial savings) $b \leq \underline{b}_t(z,k)$ are *unconstrained* in the sense that there is zero probability of being borrowing constrained at any point in the future. In this case, $\lambda_t(z,k,b)=0$, so the value of liquidity inside the firm is the same as outside the firm. Here, the firm's financial structure is irrelevant for investment, and investment simply satisfies the first-order condition (21). As in equation (18) from the representative-firm benchmark, the RHS of this expression is the firm's marginal Q (adjusted for the exit shock), so unconstrained firms' investment follows the Q theory.[26]

[26]Unconstrained firms' mix of debt and equity is indeterminate in the sense that any pair (d,b') which satisfies $b' \leq \chi k'$ and $d \geq 0$ is a solution to the firm's problem—the Modigliani-Miller theorem. Again following Khan and Thomas (2013), we resolve this indeterminacy by assuming unconstrained firms issue the maximum amount of debt b' which leaves them unconstrained. One motivation for this choice is that our model does not incorporate the tax advantage of debt, which would break Modigliani-Miller and incentivize firms to issue debt.

Firms with sufficiently high debt $b \geq \overline{b}_t(z,k,b)$ are *actively constrained* in the sense that the borrowing constraint $b' \leq \chi k'$ currently binds, which implies that $d \geq 0$ binds as well. Together, these two constraints define the maximum level of capital accumulation $\overline{k}_t(z,k)$ that the firm can afford in (22). To see this, consider the case with no adjustment costs to $\varphi(\cdot) = 0$, which yields

$$\overline{k}_t(z,k,b) = \frac{\pi_t(z,k) + T_t^f + (1-\delta)k - b}{1 - \frac{\chi}{1+r_t}}. \tag{24}$$

In this case, the maximal investment levers up the firm's current cash on hand $\pi_t(z,k) + T_t^f + (1-\delta)k - b$ according to the collateralizability of new investment χ.[27]

Finally, firms with intermediate levels of debt $b \in (\underline{b}_t(z,k), \overline{b}_t(z,k,b))$ are *potentially constrained* in the sense that their borrowing constraint $b' \leq \chi k'$ is currently slack, but there is a positive probability of facing a binding borrowing constraint in the future, so $\lambda_t(z,k,b) > 0$ and $d \geq 0$ currently binds. For these firms, financial frictions affect investment through the multipliers $\lambda_t(z,k,b)$ in (23)—essentially, concerns about hitting future borrowing constraints (encoded in the multipliers) create risk aversion for the firm. This effective risk aversion generates behavior akin to that generated by household's precautionary savings motives.

Marginal Propensities to Invest. We define the MPI as $\frac{-\partial k_t'(z,k,b)}{\partial b}$, i.e. the response of investment to a marginal infusion of liquidity (which is isomorphic to reducing their net debt position b). In our model, increases in aggregate demand raise cash flows $\pi_t(z,k)$ by reducing the real wage w_t firms must pay to labor.

Since unconstrained firms' investment follows the Q theory relationship (21), their MPI is $-\frac{\partial k_t'(z,k,b)}{\partial b} = 0$, as in the representative firm benchmark. Currently constrained firms' MPI is $-\frac{\partial \overline{k}_t(z,k)}{\partial b}$, which (applying the implicit function theorem) is strictly positive. In fact, using the special case without adjustment costs (24), we see actively constrained firms' $\frac{\partial \overline{k}_t(z,k)}{\partial b} = \frac{1}{1-\frac{\chi}{1+r_t}}$, which is generally larger than 1 because firms can lever up. Finally, we will find numerically that potentially constrained firms have an MPI $\frac{-\partial k_t'(z,k,b)}{\partial b} > 0$ because higher cash flows reduce effective risk aversion. Hence, putting all these cases together, we

[27] In the general case, $\overline{k}_t(z,k,b)$ solves a quadratic equation, but the logic is similar to (24).

get that MPIs are highest for constrained firms and then converge to zero as firms become unconstrained—similar to the finding that MPCs were highest for constrained households and then converged to the PIH level as households became unconstrained.

Aggregate Investment. As in the heterogeneous household model, we construct the sequence-space representation of the aggregate investment block $\mathcal{I}_t(\{r_s, w_s\})$ in two steps. First, given the path of prices $\{r_s, w_s\}$, we compute the decision rules as a function of individual states (z, k, b). Second, given a path of these decisions, we compute the path of distributions by iterating forward on the law of motion for the distribution. Aggregating up investment decisions gives the path of aggregate investment $\mathcal{I}_t(\{r_s, w_s\})$.

5 Disciplining MPIs From the Data

We now calibrate the new parameters of the heterogeneous firm version of the investment block, paying particular attention to those governing the MPIs.

5.1 The Empirical Literature Since the Q theory

As described in Section 2, the empirical literature on the Q theory finds a positive correlation between investment and cash flows, suggesting that the average MPI is positive. However, variation in cash flows is endogenous, which may bias these estimates relative to the true MPI. For example, suppose that higher cash flows are driven by a higher realization of idiosyncratic productivity z. Since productivity is persistent, the expected marginal product of capital in the future also increases, which would increase investment even for unconstrained firms. If we observed the true marginal Q, then we could control for the effect of higher future marginal product of capital; however, we only observe average Q, which is only an imperfect proxy for marginal Q. As a result, the estimated investment-cash flow sensitivity may still be positive even if firms are unconstrained and have a true MPI of 0.[28]

[28]Furthermore, even among constrained firms, the estimated investment-cash flow sensitivity may not resemble the true MPI; see Gomes (2001) for an example, and Kaplan and Zingales (1997), Cleary (1999), Chen and Chen (2012) for an earlier literature arguing that investment-cash flow sensitivities are poor measures of financial constraints. For the same reason, strategies that regress investment on unexpected revisions to sales growth, such as proposed in Alati et al. (2024), will not recover the MPI as defined in this

Unlike with MPCs, researchers have not found exogenous variation in cash flows to directly estimate MPIs for a broad enough cross-section of firms to be useful for an aggregate model.[29] Instead, the investment literature has come to rely on structural models that are calibrated to match other moments in the data, and implicitly trust in the structure of the model to infer the distribution of MPIs.

That said, the empirical literature has provided evidence that financial constraints affect investment decisions, which is a pre-requisite for positive MPIs to exist in the first place. For example, Chaney, Sraer and Thesmar (2012) argue that increases in local real estate prices lead to an increase in investment by increasing the value of the firm's corporate real estate. These types of papers are consistent with the idea that variation in credit supply—in our model, variation in the collateral constraint parameter χ—impacts firms' investment decisions, rejecting the hypothesis that all firms are unconstrained.[30] However, these papers do not directly estimate the MPI, which captures the effects of changes in cash flows rather than changes in credit supply.

A related literature documents heterogeneous responses of investment to price changes, also consistent with financial frictions impacting investment. For example, Zwick and Mahon (2017) estimate the response of investment to the Bonus Depreciation Allowance, a temporary tax incentive that reduces the after-tax relative price of investment. They find that small, non-dividend paying firms are more responsive to the Bonus. More related to monetary policy, a number of papers document heterogeneous responses to monetary policy shocks (for example, Gertler and Gilchrist, 1994 by size, Ottonello and Winberry, 2020 by default risk, Jeenas, 2025 by cash holdings, Cloyne, Ferreira, Froemel and Surico, 2023 by age, Cao, Hegna, Holm, Juelsrud, König and Riiser, 2023 across firms in Norway, etc.). Best, Born and Menkhoff (2025) provide similar empirical evidence using a high-quality firm survey of CEOs. Broadly speaking, these papers show that the response of investment to price

paper: this regression would yield a non-zero regression coefficient even in the version of our model where all firms are unconstrained and $MPI_j = 0$ for all firms j.

[29] Blanchard, Lopez-de Silanes and Shleifer (1994) estimates MPIs out of windfall cash winnings from court cases, but the sample only contains eleven firms. Rauh (2006) estimates MPIs using variation in pension funding requirements, but these are based on discontinuities that affect a very small subset of firms (see Bakke and Whited, 2012).

[30] Indeed, Catherine, Chaney, Huang, Sraer and Thesmar (2022) use the Chaney et al. (2012) estimates to calibrate a model with collateral constraints and shows that the constraints are tight.

changes depends on the value of inside liquidity $\lambda_t(z, k, b)$, but again that does not directly estimate the MPI.[31]

We believe that there is scope for a literature to estimate MPIs out of transitory cash flow shocks using the same credibility-revolution methods that have been used over the past decade to estimate MPCs. Structural models with financial frictions such as the one in this paper require not just estimates of average MPIs, but also of the distribution and firm correlates of MPIs in the data, as well as the time path of the investment response to a transitory cash flow shock (the *intertemporal Marginal Propensity to Invest*, or "iMPI"). A promising step in this direction is the work of Martin-Baillon (2021), who uses the Blundell, Pistaferri and Preston (2008) semi-structural method to estimate average and sectoral MPIs.

5.2 Calibration

Without direct estimates of the MPI, we follow the approach of the structural investment literature: calibrate the heterogeneous-firm block to match other moments in the data, and infer the distribution of MPIs using the structure of the model. The left panel of Table 3 contains all the parameters for the heterogeneous firm block, which we choose to match the moments in the right panel. While all these parameters are jointly chosen to match all the moments, there is a clear mapping between certain parameters and moments.[32]

The first three parameters—the coefficients of the production function and the capital depreciation rate—are the same as in the representative agent benchmark in order to match the same moments. We also calibrate the adjustment cost parameter ϕ to match the peak impulse response of aggregate investment to a monetary shock, as in the representative agent benchmark. We find that the required adjustment cost parameter is $\phi = 13.5$, somewhat higher than its value in the representative agent benchmark.

[31] See Koby and Wolf (2020) for a structural interpretation of the empirical results in Zwick and Mahon (2017) in the context of a model with fixed adjustment costs.

[32] Some of the moments which are influenced by the household side through general equilibrium; we continue to use the heterogeneous household specification for now, and discuss their interaction in Section 7 below. We again re-calibrate the EIS σ^{-1} on the household side to ensure that we continue to match the same impact response of consumption to a monetary policy shock, and re-calibrate the β process to match the level of aggregate wealth. A small recalibration is needed relative to that in the model of Section 3 since the level of aggregate assets relative to labor income now depends on the endogenous distribution of firm rents and firm debt issuance decisions.

Table 3: Parameters and Targets for the Heterogeneous Firm Block

Parameter	Description	Value	Target (all joint)	Data	Model
Production function and depreciation rate					
ν	Labor exponent	0.60	Labor share	0.60	0.60
α	Capital exponent	0.32	Returns to scale	0.92	0.92
δ	Capital depreciation rate	0.025	Aggregate I/K	0.025	0.025
Adjustment costs					
ϕ	Quadratic ACs	13.50	IRF to monetary shock	1%	1%
Collateral constraint					
χ	Collateral constraint	0.35	Kermani and Ma (2023)	Direct measurement	
Lifecycle					
k_0	Initial capital	0.50	New entrant size	4%	4%
b_0	Initial debt	0.175	χk_0	Imposed	
Idiosyncratic shocks: $\log z' = \rho_z \log z + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_z^2)$					
ρ_z	Persistence	0.94	Syverson (2011)	Direct measurement	
σ_z	SD of innovations	0.085	Syverson (2011)	Direct measurement	
θ	Exit probability	0.02	BDS data	0.02	0.02

Notes: parameters in the left panel chosen to match moments in the right panel. "Production function and depreciation rate" parameters chosen as in representative agent benchmark from Section 2. Adjustment costs are quadratic $\varphi(x) = \frac{\phi}{2}x^2$ and ϕ chosen to match 1% peak response of aggregate investment to a 50bps annualized decline of real rate (also following procedure from Section 2). Collateral constraint $\chi = 0.35$ chosen to match average capital recovery ratios estimated by Kermani and Ma (2023). Initial capital stock k_0 chosen to be 4% of average capital and initial debt chosen to be $b_0 = \chi k_0$ so that new entrants are maximally leveraged. Persistence ρ_z and standard deviation σ_z of idiosyncratic productivity shocks chosen to match stylized facts from Syverson (2011) described in main text. Exit probability θ chosen to match average exit rate of 8% from Business Dynamic Statistics (BDS).

The collateral constraint χ controls the MPI for constrained firms through $\overline{k}_t(z,k)$. We choose $\chi = 0.35$ to match the average liquidation value of capital during bankruptcy estimated by Kermani and Ma (2023). This value is directly informative about the ex post value of collateral to lenders during default; Kermani and Ma (2023) also show it is consistent with surveys of lenders' ex ante collateral requirements when making loans, which is more relevant for our model (which does not have default).[33]

The initial conditions for new entrants (k_0, b_0) control how much young firms need to

[33] An alternative approach would be to choose χ to match features of the distribution of leverage. We do not follow that approach because our model is an extremely stylized representative of leverage among constrained firms, is driven by arbitrary choices about the borrowing decisions of unconstrained firms, and doesn't have features like the tax advantage of debt that are likely important in the data. That said, average leverage in our model is 0.29, close to the 0.34 reported by Crouzet and Mehrotra (2020) from the microdata underlying the Quarterly Financial Reports (QFR).

grow and, therefore, how likely they are to be constrained. Ideally, we would choose (k_0, b_0) to match firm dynamics over their lifecycle; we make a simple choice in this spirit. First, to cut down on free parameters, we impose $b_0 = \chi k_0$, i.e that new entrants are maximally leveraged. Second, we choose k_0 to be 4% of the capital of the average incumbent firm as in Li, Lian, Ma and Martell (2025). As a result, small and young firms grow faster than the average firm in our model, as in the data.

Finally, we take the idiosyncratic shocks—both idiosyncratic productivity z and the exit shocks—from the data. There is a large literature estimating idiosyncratic productivity differences across firms.[34] Syverson (2011) summarizes this literature with two stylized facts: productivity is persistent (with an annual autocorrelation between 0.6 and 0.8) and highly dispersed (with a within-sector productivity gap between the 90th and 10th percentiles around 0.65 log points). We assume that z follows an AR(1) process, $\log z' = \rho_z \log z + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_z^2)$, and choose the parameters ρ_z and σ_z to match Syverson (2011)'s stylized facts.[35] For the exit shocks, we choose the probability of exit $\theta = 0.02$ quarterly to match the roughly 8% annual average exit rate reported in the Business Dynamics Statistics (BDS).

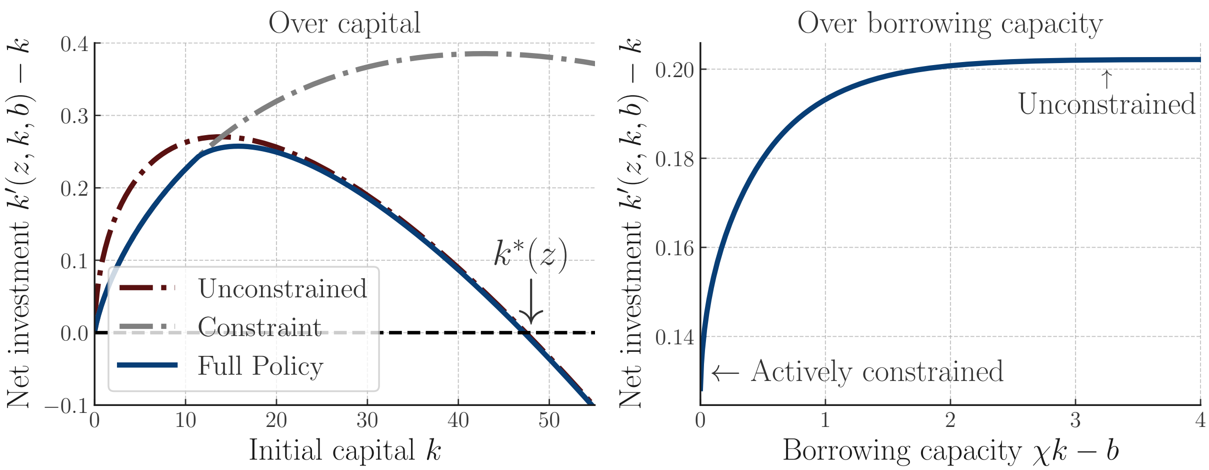
5.3 Steady State Decision Rules and MPI Distribution

We now inspect the firm's decision rules and implied MPIs in the steady state of our calibrated model. The left panel of Figure 5 plots the net investment policy, $k'(z, k, b) - k$, as a function of the firm's initial capital k (for a given level of productivity z). The "unconstrained" curve in purple plots the net investment rate for a firm with low enough b to be unconstrained for all levels of k. The unconstrained policy has a target level of capital, $k^*(z)$, to which the firm would eventually converge if its productivity were to remain constant—similar to the buffer-stock target asset level $a^*(e, \beta)$ in the household model. The

[34] A key issue in this literature is the distinction between physical productivity, which is quantity produced per bundle of inputs, and revenue productivity, which is revenue per bundle of inputs. Typically, studies are only able to estimate revenue productivity because most datasets do not record prices and quantities separately. Given this challenge, some researchers calibrate the productivity process by matching another set of statistics, such as the implied distribution of investment rates. While we do not follow that approach here, we note that our cross-sectional dispersion of investment rates is 0.11 annualized, compared to 0.14 in Zwick and Mahon (2017).

[35] In particular, we set our quarterly autocorrelation ρ_z to match an annual autocorrelation of 0.8 and then set the dispersion of idiosyncratic shocks σ_z so that the cross sectional 90-10 gap in log productivity is 0.65.

Figure 5: Investment Decision Rules with Heterogeneous Firms



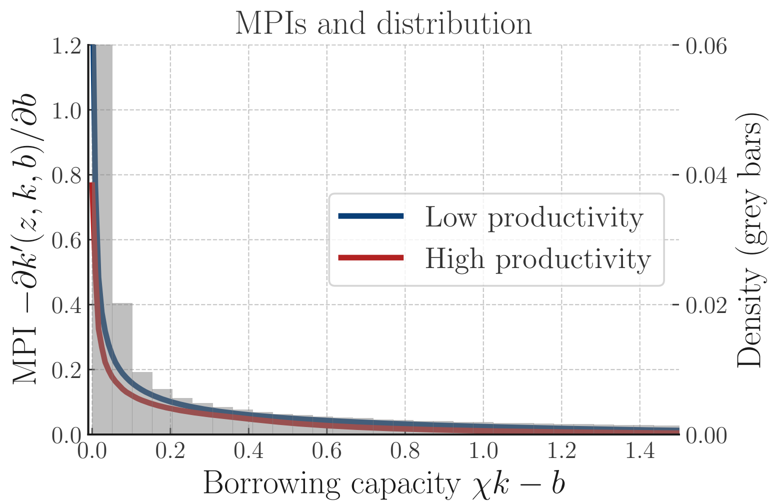
Notes: firms' decision rules in steady state of calibrated model. Left panel plots net investment $k'(z,k,b)-k$ for a firm with given level of productivity z and borrowing capacity $\chi k - b$ as a function of initial capital k. "Unconstrained" line refers to $k^*(z,k)$ which solves (21). "Constraint" refers to $\overline{k}(z,k,b)$ which solves (22). "Full policy" refers to the true policy function $k'(z,k,b)$. Right panel plots net investment $k'(z,k,b)$ as a function of its borrowing capacity $\chi k - b$, holding fixed productivity z and capital k.

grey "constraint" curve plots the maximal net investment level $\overline{k}(z,k,b)-k$ for a firm with a positive amount of initial debt $b > 0$. The constraint is increasing in initial capital k because, all else equal, firms with higher k have higher cash on hand $\pi(z,k) + (1-\delta)k - b + T_t^f$.

The firm's full policy function in blue pastes together the constraint (which is binding for low levels of capital k) and the unconstrained policy (which is optimal for high enough levels of k). For intermediate levels of k, the firm is potentially constrained and invests less than the unconstrained choice even though the unconstrained choice is feasible. The reason is that the marginal product of capital is lower when productivity z' falls, and these are also the states in which the constraint becomes tighter. As a result, the covariance between the multipliers $\lambda_{t+1}(z',k',b')$ and the return to capital $\pi_{t+1}(z',k')$ in (23) is negative, inducing the firm to invest less than it would if it were unconstrained.

The right panel of Figure 5 plots net investment $k'(z,k,b)-k$ as a function of its financial position. To make the units interpretable, we measure financial position using the firm's initial *borrowing capacity* $\chi k - b$ inherited from last period's choices and vary b holding k fixed. If the firm were actively constrained in the last period, it would have set $b = \chi k$

Figure 6: MPIs and Distribution

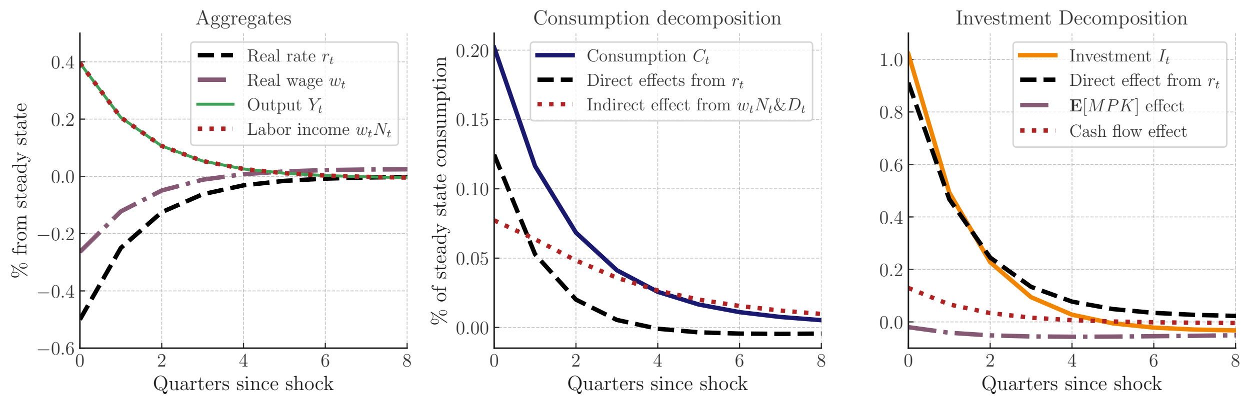


Notes: MPIs $-\frac{\partial k'(z,k,b)}{\partial b}$ (measured on the left y-axis) and stationary distribution of firms (measured on the right y-axis) as a function of borrowing capacity $\chi k - b$. MPIs are plotted for firms with low and high values of productivity z, fixing a value of k.

and therefore have zero borrowing capacity. These firms are also constrained in the current period, so their investment policy is given by $k'(z,k,b) = \overline{k}(z,k,b)$. As initial borrowing capacity increases, the firm has more cash on hand and can therefore invest more. Eventually, for high initial borrowing capacity, the firm becomes unconstrained and its investment is independent of borrowing capacity. Hence, the investment policy is a concave function of borrowing capacity, similar to how the consumption function is concave out of wealth.

Figure 6 plots the MPI $-\frac{\partial k'(z,k,b)}{\partial b}$, which equals the slope of the investment policy with respect to initial borrowing capacity. Firms with low levels of initial borrowing capacity are actively constrained and have the highest MPIs (in some cases above 1 due to the ability to leverage described above); in our calibration, 66% of firms are actively constrained. As initial borrowing capacity increases, firms become potentially constrained, and therefore have smaller but still positive MPIs; in our calibration, 21% of firms are potentially constrained. Finally, for high enough levels of initial borrowing capacity, firms become unconstrained and have MPI= 0; in our calibration, 13% of firms are unconstrained. Averaging across all firms, the unweighted average MPI is 0.50.

FIGURE 7: Monetary Transmission with Heterogeneous Firms



Notes: impulse responses of aggregates (left), consumption (middle, blue line) and investment (right, orange line) to a monetary policy shock modeled as a 50bp annualized decline in the real interest rate with quarterly persistence of 0.5, similar to Christiano et al. (2005). All variables are in percent deviations from their steady state values, except the real interest rate which is in annualized percentage points. The middle panel also displays the decomposition of consumption into a direct and indirect effect in the spirit of equation (15). The right panel also displays the decomposition of investment into a direct and indirect effect using the sequence-space Jacobians of the firm model. The $\mathbf{E}[MPK]$ effect is the effect of w_t via the expected marginal product of capital. The new cash flow effect is the indirect effect resulting from the effect that changes in w_t have through constraints.

6 Policy Transmission with Heterogeneous Firms

We now show that the presence of financially constrained firms with high MPIs creates new transmission channels for investment, just like the presence of constrained, high-MPC households did for consumption.

6.1 Monetary Policy

Figure 7 plots the impulse responses to the expansionary monetary shock. As in Sections 2 and 3, the shock reduces the real interest rate by 50 basis points upon impact and reverts back to steady state with quarterly autocorrelation 0.5. Given our calibration strategy, the aggregate responses (left panel) and the responses of consumption (middle panel) are nearly identical to Section 3.

The right panel of Figure 7 decomposes the response of aggregate investment now that we have heterogeneous firms. The existence of high-MPI firms generates new cash flow effects of monetary policy: higher aggregate demand lowers the real wage w_t, which increases firms' cash flows and therefore increases investment among constrained firms. However, these new indirect effects are relatively modest, accounting for 13% of the total response. The majority of the response of investment continues to be driven by the direct effects of monetary policy; the effect of changes in future real wages on the future MPK continues to play a minor role.[36]

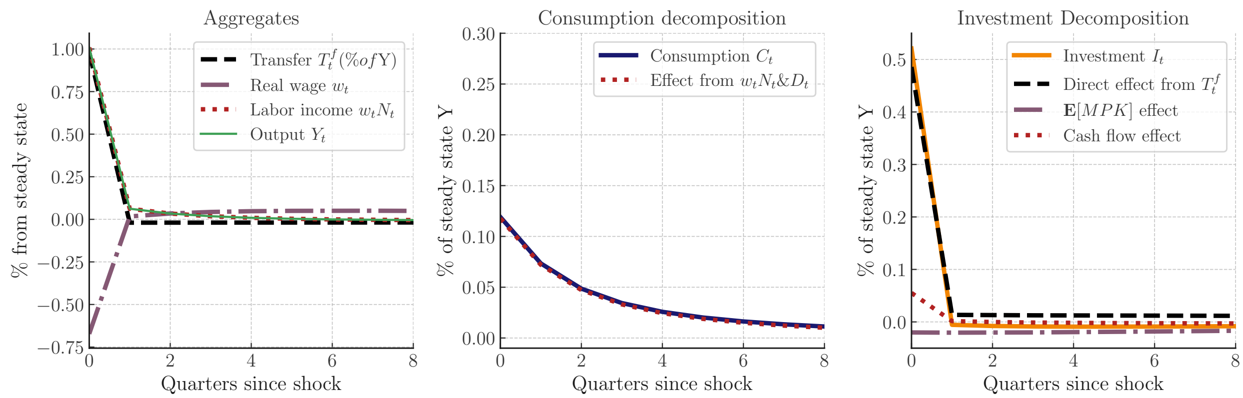
The indirect effects of monetary policy are modest because they are governed by the *size-weighted MPI*, which downweights the contribution of small, high MPI firms. To see why, note that our Cobb-Douglas production function implies that cash flows are proportional to output, $\pi_t(z,k) = y_{jt} - w_t n_{jt} = (1-\nu)y_{jt}$. Since the real wage w_t is the same for all firms, changes in aggregate demand dY_t affect individual firms' cash flows in equal proportion, $\frac{d\pi_{jt}}{\pi_{jt}} = \frac{dY_t}{Y_t}$, implying that the change in an individual firm's cash flow is proportional to its initial size $d\pi_{jt} = \pi_{jt}\frac{dY_t}{Y_t}$. The aggregate MPI, which we denote $\overline{\text{MPI}}$, is then the average MPI weighted by initial size: $\overline{\text{MPI}} = (1-\nu)\mathbb{E}\left[\left(\frac{\pi_{jt}}{\mathbb{E}[\pi_{jt}]}\right)\text{MPI}_{jt}\right]$. In our calibration, the size-weighted MPI is 0.05, an order of magnitude smaller than the unweighted MPI of 0.50.

The fact that the size-weighted MPI is an order of magnitude smaller than the unweighted MPI reflects an extreme negative correlation between size and MPI in our particular calibration. However, there are good reasons to be suspicious of this result; for instance, the correlation between size and leverage is extremely negative in our model but somewhat positively correlated in the data, suggesting that large firms may also be affected by borrowing constraints but have a higher borrowing capacity (see Chatterjee and Eyigungor, 2023). As we discussed in the introduction, it may well be that the true empirical correlation between size and MPI is weaker than in our model, just like the empirical correlation between income and MPC is weaker than in the first generation of heterogeneous household models.

There are at least three natural candidates to break the extreme correlation between size and MPI that arises in our model. First, incorporating the tax advantage of debt would incentivize large firms to borrow and get closer to the constraint. Second, permanent heterogeneity in returns to scale would imply that many large firms have high returns to

[36] Appendix B studies the heterogeneity in the effects of changes in interest rates on households and firms in our model. This effect, aggregated across households and firms, kick-starts a transmission mechanism to aggregate wage and dividend income and firm cash flows, and back to consumption and investment, which our paper studies in detail.

FIGURE 8: Fiscal Policy Transmission with Heterogeneous Firms



Notes: impulse responses of aggregates (left), consumption (middle, blue line) and investment (right, orange line) to a transfer shock of 1% of GDP, financed by an increase in debt that reverts back to steady state with an autocorrelation of 0.99. Variables in the left panel are in percent deviations from their steady state values except for the transfer, which is in percent of GDP. Variables in the middle and right panels are in percent of GDP. The middle panel also displays the decomposition of consumption into a direct and indirect effect in the spirit of equation (15). The right panel also displays the decomposition of investment into a direct and indirect effect using the sequence-space Jacobians of the firm model. The $\mathbf{E}[MPK]$ effect is the effect of w_t via the expected marginal product of capital. The new cash flow effect is the indirect effect resulting from the effect that changes in w_t have through constraints.

scale and would want to grow by even more, making them constrained. Third, increases in aggregate demand could affect small firms by more than large firms, upweighting their contribution to the aggregate MPI.[37]

6.2 Fiscal Policy

We now study the effects of a one-time fiscal transfer T_t^f to firms, equivalent in size to 1% of steady state GDP. These transfers are similar in spirit to the debt forgiveness programs enacted during COVID because an increase in the transfer T_t^f has the same effect on cash on hand $\pi_t(z,k) + (1-\delta)k + T_t^f - b$ as a decline in outstanding debt b. As with the household transfer in Section 3, we assume that the transfer is deficit-financed upon impact, and that the deficits are gradually paid back using lump-sum taxes on firms (such that the quarterly persistence of government debt is again 0.99).

[37]More broadly, just like a literature estimates the incidence of demand shocks across the household distribution and its correlation with MPCs (e.g. Patterson, 2023), it would be natural to perform a similar

The right panel of Figure 8 shows that the transfer increases investment by 3% upon impact, which is equivalent to more than 0.5 percent of steady state GDP. Most of this increase is driven by the direct effect of the transfer, i.e. higher T_t^f increases investment of constrained firms according to their MPI. Since the transfer is uniformly distributed across firms, the magnitude of this direct effect is governed by the unweighted average MPI, which is 0.50 in our calibration. Since the unweighted MPI is larger than the unweighted MPC, the transfer has a larger effect on aggregate spending than the same-size transfer to households.

The transfer also generates indirect effects on investment, but Figure 8 shows that those effects are small. As with monetary policy, the increase in aggregate demand lowers the real wage w_t, increasing cash flows $\pi_t(z, k)$. But, as discussed above, the aggregate effects of these cash flows are determined by the size-weighted MPI, which is an order of magnitude smaller than the unweighted MPI; as a result, the indirect effects out of higher cash flows are an order of magnitude smaller than the direct effects of the transfer.

The right panel of Figure 8 also shows that the dynamic effects of the transfer are not persistent; investment essentially reverts back to steady state after impact. This finding contrasts with the household model, where the iMPCs generated endogenous persistence out of the transfer; apparently, the heterogeneous-firm model does not have very persistent iMPIs. Characterizing these iMPIs and relating them to aggregate demand propagation, similar to how Auclert et al. (2024) did for iMPCs on the household side, would be an interesting avenue for future research.

The middle panel of Figure 8 illustrates a new spillover mechanism in our model from combining household and firm heterogeneity: the transfer to firms also increases consumption spending (by 0.12 percent of steady state GDP). This spillover occurs because higher investment demand increases aggregate demand and therefore household incomes, which then stimulates consumption according to the income-weighted MPC. We turn to the study of this type of complementarity between household and firm heterogeneity next.

exercise for the firm distribution and MPIs.

7 Aggregate Demand Complementarities

So far, we have mostly studied the heterogeneous-household and heterogeneous-firm blocks separately from each other. Here, we show that having these two blocks together creates *aggregate demand complementarities*: having both high MPCs and high MPIs generates larger effects than either does in isolation.

We illustrate aggregate demand complementarities in the context of uniform, lump-sum fiscal transfers. Since monetary policy holds the real interest rate constant, goods market clearing is

$$\mathcal{C}_t(r^0, \{T_s^h, w_s N_s\}) + \widetilde{\mathcal{I}}_t(\{T_s^f, w_s\}) = \mathcal{Y}_t(\{T_s^f, w_s\}),$$

where r^0 is the implied capital gain for households and $\widetilde{\mathcal{I}}_t(\{w_s, r_s\})$ includes both aggregate investment and adjustment costs. To build intuition, consider the impact effect of a given combination of transfers dT_0^h and dT_0^f, ignoring the contribution of changes in r^0 or $\{w_s\}_{s\geq 1}$.[38] In this case, one can show that the change in aggregate income is

$$\text{MPC} \times dT_0^h + \text{MPI} \times dT_0^f + \overline{\text{MPC}} \times dY_0 + \overline{\text{MPI}} \times dY_0 = dY_0. \tag{25}$$

The first two terms in (25) capture the direct impacts of the transfer on consumption and investment spending; as we showed in Sections 3 and 4, these direct effects are controlled by the unweighted MPC and MPI. The second two terms in (25) then capture the indirect impacts through changes in household income and firms' cash flows; these effects are controlled by the income-weighted $\overline{\text{MPC}}$ and size-weighted $\overline{\text{MPI}}$. Solving this equation yields

$$dY_0 = \frac{\text{MPC} \times dT_0^h + \text{MPI} \times dT_0^f}{1 - \overline{\text{MPC}} - \overline{\text{MPI}}}. \tag{26}$$

We say that household and firm heterogeneity generates aggregate demand complementarities because the implied $\overline{\text{MPC}}$ and $\overline{\text{MPI}}$ enter nonlinearly in (26). This nonlinearity reflects the addition of investment to the Keynesian cross: a given shock to spending (here, the

[38] Properly accounting for these future changes would require working with iMPCs, as in Auclert et al. (2024), or the analogous iMPIs on the firm side. We leave this analysis for future research; our goal here is simply to motivate and interpret the numerical results in this section.

Table 4: Policy effects by household/firm heterogeneity

(A) To firms

	Rep firm	Het firm
Rep household	0	0.52
Het household	0	0.64

(B) To households

	Rep firm	Het firm
Rep household	0	0
Het household	0.29	0.31

Notes: percentage change in GDP in response to fiscal transfers to households (panel A) or firms (panel B). Transfers are 1% of GDP upon impact, which are then deficit-financed through higher future taxes, as described in Sections 3 and 4. Effects computed in the four models described in the main text.

transfers) raises aggregate demand, which stimulates both consumption through constrained households and investment through constrained firms, etc. To quantify their effects, we compute the effects of transfers in four models:[39]

(i) Representative household, representative firm. On the household side, the direct effect of the transfer is $\text{MPC} = 0$ and the indirect effect of aggregate demand is $\overline{\text{MPC}} \approx 0$. On the firm side, both $\text{MPI} = \overline{\text{MPI}} = 0$.

(ii) Heterogeneous household, representative firm. The presence of household heterogeneity implies that $\text{MPC} > 0$ and $\overline{\text{MPC}} > 0$, but the representative firm still implies $\text{MPI} = \overline{\text{MPI}} = 0$.

(iii) Representative household, heterogeneous firm: the representative household has $\text{MPC} = 0$ and $\overline{\text{MPC}} \approx 0$, but the presence of firm heterogeneity implies that $\text{MPI} > 0$ and $\overline{\text{MPI}} > 0$.

(iv) Heterogeneous household, heterogeneous firm: combining both gives $\text{MPC} > 0$, $\overline{\text{MPC}} > 0$, $\text{MPI} > 0$, $\overline{\text{MPI}} > 0$.

For each model, we separately compute the effects of a lump-sum transfer to households and to firms, each of which are equivalent to 1% of steady state GDP.

The left panel of Table 4 shows that the aggregate demand complementarities are strong in the case of the firm level transfer. With heterogeneous firms but only a representative household, the multiplier in (26) simply becomes $\frac{\text{MPI}}{1-\overline{\text{MPI}}}$; since $\text{MPI} = 0.50$, the direct effect

[39]In each case, we re-calibrate the EIS σ^{-1} and adjustment cost parameter ϕ to match the same impact effects of monetary policy on aggregate consumption and investment.

of the transfer is large. However, since $\overline{\text{MPI}} = 0.05$, the indirect effects are small. Once we incorporate household heterogeneity, the multiplier becomes $\frac{\text{MPI}}{1-\overline{\text{MPI}}-\overline{\text{MPC}}}$; since $\overline{\text{MPC}} = 0.2$, the indirect effects are significantly strengthened through consumption, leading to a 20% larger total response to the transfer.

In contrast, the right panel of Table 4 shows that the complementarities for the household transfer are modest. This occurs because adding firm heterogeneity increases the multiplier from $\frac{\text{MPC}}{1-\overline{\text{MPC}}}$ to $\frac{\text{MPC}}{1-\overline{\text{MPC}}-\overline{\text{MPI}}}$, which is small given that $\overline{\text{MPI}}$ is small.

8 Analogies and Directions for Future Research

To conclude our analysis, we summarize the various analogies between heterogeneous household and heterogeneous firm models that arose in our sufty. We then use these analogies as a framework to guide possible directions for future research.

Analogies. Table 5 collects the key analogies. In steady state, households with idiosyncratic shock (e, β) eventually converge to a target level of assets $a^*(e, \beta)$, just like firms with productivity z converge to a target capital stock $k^*(z)$. The speed of convergence is controlled by the EIS σ^{-1} for unconstrained households and by the adjustment costs ϕ for unconstrained firms. However, these decisions are distorted by occasionally binding borrowing constraint $a' \geq 0$ for households and the financial frictions $d \geq 0$ and $b' \leq \chi k'$ for firms. Even for agents not currently constrained, idiosyncratic risk still affects their behavior through prudence for households and the possibility of facing a binding constraint in the future for firms. All together, these forces give rise to a concave consumption function out of wealth and a concave investment function out of initial borrowing capacity.

The bottom panel of Table 5 summarizes how these mechanisms shape the response to monetary and fiscal shocks. Relative to the representative agent benchmark, the new channels operate through having high MPCs on the household side and high MPIs on the firm side. As a result, monetary policy generates indirect effects by increasing incomes among constrained households and cash flows among constrained firms. In addition, high MPCs and MPIs break Ricardian equivalence on the household side and the fact that the

Table 5: Analogies Between the Heterogeneous Agent Blocks

	Households	**Firms**
Steady state	Target buffer stock of assets	Target stock of capital
	EIS	Adjustment costs
	Borrowing constraint	Financial frictions
	Prudence	Concern about hitting constraints
	Concave consumption function	Concave investment function
Response to shocks	MPCs/iMPCs	MPIs/iMPIs
	Indirect effects of monetary policy	Cash flow effects of monetary policy
	Breaking Ricardian equivalence	Breaking inside = outside liquidity
	Income-weighted MPC governs multiplier	Size-weighted MPI governs multiplier

Notes: summary of conceptual analogies between heterogeneous household and heterogeneous firm models described in the main text.

MPI = 0 on the firm side, allowing deficit-financed fiscal transfers to stimulate spending. Finally, high MPCs and MPIs generate an aggregate demand multiplier in the spirit of the Keynesian cross, governed by the income-weighted MPC for consumption and size-weighted MPI for investment.

Possible Directions for Future Research. Over the last ten years, the HANK literature has done a lot of work to characterize the new transmission mechanisms in the bottom left panel of Table 5, but the bottom right panel remains relatively underexplored. For example, a full characterization of the iMPIs may provide a useful sufficient statistic for the full dynamic indirect effects of changes in aggregate demand on investment. The indirect effects of monetary policy likely differ under different monetary policy rules, specifications of debt (real vs. nominal, short vs. long maturity), or other extensions such as those described in Section 6. Similarly, we have only studied a very simple fiscal transfer, but the effects of other fiscal policies like investment tax credits, debt forgiveness, or loan guarantees may be different. In addition, government debt may crowd out investment and generate interesting interactions between monetary and fiscal policy.

At the same time, the heterogeneous firm literature has done a lot of work to understand investment behavior outside of monetary and fiscal policy which may prove useful for the consumption side as well. In particular, the durability of capital has two implications missing from models of nondurable consumption: (i) the interest sensitivity of investment is much higher than nondurable consumption and (ii) capital is collateralizable, changing the behavior of MPIs. These implications also arise in models of durable consumption, though these models are not as well studied as those of nondurable consumption.[40]

Finally, households and firms are linked together through many rich channels that we have excluded from our benchmark models. For example, our mutual fund was a simple, frictionless financial intermediary which transformed households' savings into firms' investment. However, financial intermediaries are often constrained by various balance sheet constraints or capital requirements. Changes in the tightness of these constraints—such as during financial crises—may affect the ability of monetary policy to impact consumption and investment. As another example, in the labor market, households are matched to individual firms, so the performance of those firms directly impacts their labor income. In this type of setting, policies that impact firms may pass through differentially to the workers in the economy.

9 Conclusion

The New Keynesian framework has undergone substantial development since the simplest versions were first introduced. The first wave of research, working in the representative agent version of the model, incorporated various elements to help the model quantitatively match the data. The second wave of research introduced HANK, which broke the representative household assumption. As we reviewed, the presence of high-MPC households unlocked new indirect policy transmission mechanisms, spawning a new literature to fully unpack them.

Our hope is that we are now on the cusp of a third wave of research which will break the representative firm assumption and incorporate heterogeneous firms into HANK. In this paper, we proposed a canonical heterogeneous-firm model with financial frictions from which others can build. Within this benchmark, we showed how high-MPI firms generate new indirect channels of policy transmission. However, given the constraints of space, our analysis can only offer a preview of what may still be to come.

[40] For notable exceptions, see Aaronson, Agarwal and French (2012), McKay and Wieland (2021), Laibson, Maxted and Moll (2022), and Beraja and Zorzi (2025).

References

Aaronson, Daniel, Sumit Agarwal, and Eric French, "The Spending and Debt Response to Minimum Wage Hikes," *American Economic Review*, December 2012, *102* (7), 3111–3139.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, August 1994, *109* (3), 659–684.

Alati, Andrea, Johannes Fischer, Maren Froemel, and Ozgen Ozturk, "Firms' Sales Expectations and Marginal Propensity to Invest," *SSRN Working Paper*, 2024.

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

Auclert, Adrien, Bence Bardóczy, and Matthew Rognlie, "MPCs, MPEs, and Multipliers: A Trilemma for New Keynesian Models," *The Review of Economics and Statistics*, May 2023, *105* (3), 700–712.

_ , _ , _ , and Ludwig Straub, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, September 2021, *89* (5), 2375–2408.

_ , Hannes Malmberg, Matthew Rognlie, and Ludwig Straub, "The Race Between Asset Supply and Asset Demand," *Proceedings of the Jackson Hole Economic Policy Symposium*, November 2025.

_ , Matthew Rognlie, and Ludwig Straub, "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model," Working Paper 26647, National Bureau of Economic Research, January 2020.

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

_ , _ , and _ , "Fiscal and Monetary Policy with Heterogeneous Agents," *Annual Review of Economics*, August 2025, *17* (Volume 17, 2025), 539–562.

_ , Rodolfo Rigato, Matthew Rognlie, and Ludwig Straub, "Beyond Certainty Equivalence: Second-Order Solutions in the Sequence Space," *Manuscript*, December 2025.

Bakke, Tor-Erik and Toni M. Whited, "Threshold Events and Identification: A Study of Cash Shortfalls," *The Journal of Finance*, 2012, *67* (3), 1083–1111.

Beraja, Martin and Nathan Zorzi, "Durables and the Marginal Propensity to Spend," *Manuscript*, 2025.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist, "Chapter 21 – The Financial Accelerator in a Quantitative Business Cycle Framework," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1, Part C, Elsevier, 1999, pp. 1341–1393.

Best, Lea, Benjamin Born, and Manuel Menkhoff, "The Impact of Interest: Firms' Investment Sensitivity to Interest Rates," *Manuscript*, 2025.

Bewley, Truman, "The Permanent Income Hypothesis: A Theoretical Formulation," *Journal of Economic Theory*, December 1977, *16* (2), 252–292.

Bilbiie, Florin O, "Monetary Policy and Heterogeneity: An Analytical Framework," *The Review of Economic Studies*, July 2025, *92* (4), 2398–2436.

Bilbiie, Florin O., Diego R. Känzig, and Paolo Surico, "Capital and Income Inequality: An Aggregate-Demand Complementarity," *Journal of Monetary Economics*, March 2022, *126*, 154–169.

Bils, Mark and Peter J. Klenow, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, October 2004, *112* (5), 947–985.

Blanchard, Olivier J., Florencio Lopez de Silanes, and Andrei Shleifer, "What Do Firms Do with Cash Windfalls?," *Journal of Financial Economics*, December 1994, *36* (3), 337–360.

Blundell, Richard, Luigi Pistaferri, and Ian Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, December 2008, *98* (5), 1887–1921.

Broer, Tobias, Niels-Jakob Harbo Hansen, Per Krusell, and Erik Öberg, "The New Keynesian Transmission Mechanism: A Heterogeneous-Agent Perspective," *Review of Economic Studies*, January 2020, *87* (1), 77–101.

Caballero, Ricardo J. and Eduardo M. R. A. Engel, "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach," *Econometrica*, 1999, *67* (4), 783–826.

Caglio, Cecilia R., R. Matthew Darst, and Ṣebnem Kalemli-Özcan, "Collateral Heterogeneity and Monetary Policy Transmission: Evidence from Loans to SMEs and Large Firms," Working Paper 28685, National Bureau of Economic Research, April 2021.

Calvo, Guillermo A., "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, September 1983, *12* (3), 383–398.

Campbell, John Y. and N. Gregory Mankiw, "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," *NBER Macroeconomics Annual*, January 1989, *4*, 185–216.

Cao, Jin, Torje Hegna, Martin Blomhoff Holm, Ragnar Enger Juelsrud, Tobias König, and Mikkel Riiser, “The Investment Channel of Monetary Policy: Evidence from Norway,” *Manuscript*, 2023.

Carroll, Christopher D., “Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis,” *Quarterly Journal of Economics*, February 1997, *112* (1), 1–55.

— **and Miles S. Kimball**, “On the Concavity of the Consumption Function,” *Econometrica*, July 1996, *64* (4), 981–992.

Catherine, Sylvain, Thomas Chaney, Zongbo Huang, David Sraer, and David Thesmar, “Quantifying Reduced-Form Evidence on Collateral Constraints,” *The Journal of Finance*, 2022, *77* (4), 2143–2181.

Chaney, Thomas, David Sraer, and David Thesmar, “The Collateral Channel: How Real Estate Shocks Affect Corporate Investment,” *American Economic Review*, October 2012, *102* (6), 2381–2409.

Chatterjee, Satyajit and Burcu Eyigungor, “The Firm Size-Leverage Relationship and Its Implications for Entry and Business Concentration,” *Review of Economic Dynamics*, April 2023, *48*, 132–157.

Chen, Huafeng (Jason) and Shaojun (Jenny) Chen, “Investment-Cash Flow Sensitivity Cannot Be a Good Measure of Financial Constraints: Evidence from the Time Series,” *Journal of Financial Economics*, February 2012, *103* (2), 393–410.

Chen, Yuchen, “Monetary Policy, Debt Structure, and Credit Reallocation,” *Manuscript*, February 2025.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, February 2005, *113* (1), 1–45.

Clarida, Richard, Jordi Galí, and Mark Gertler, “The Science of Monetary Policy: A New Keynesian Perspective,” *Journal of Economic Literature*, December 1999, *37* (4), 1661–1707.

Cleary, Sean, “The Relationship between Firm Investment and Financial Status,” *The Journal of Finance*, 1999, *54* (2), 673–692.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico, “Monetary Policy, Corporate Finance, and Investment,” *Journal of the European Economic Association*, December 2023, *21* (6), 2586–2634.

Cooper, Russell W. and John C. Haltiwanger, “On the Nature of Capital Adjustment Costs,” *The Review of Economic Studies*, July 2006, *73* (3), 611–633.

Crouzet, Nicolas and Neil R. Mehrotra, "Small and Large Firms over the Business Cycle," *American Economic Review*, November 2020, *110* (11), 3549–3601.

Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica*, 1991, *59* (5), 1221–1248.

Drechsel, Thomas, "Earnings-Based Borrowing Constraints and Macroeconomic Fluctuations," *American Economic Journal: Macroeconomics*, April 2023, *15* (2), 1–34.

Dupor, Bill, "Investment and Interest Rate Policy," *Journal of Economic Theory*, May 2001, *98* (1), 85–113.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin, "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics*, October 2000, *46* (2), 281–313.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik, "MPC Heterogeneity and Household Balance Sheets," *American Economic Journal: Macroeconomics*, October 2021, *13* (4), 1–54.

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen, "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, January 1988, *1988* (1), 141–206.

Flavin, Marjorie A., "The Adjustment of Consumption to Changing Expectations About Future Income," *Journal of Political Economy*, October 1981, *89* (5), 974–1009.

Friedman, Milton, "The Permanent Income Hypothesis," in "A Theory of the Consumption Function," Princeton, N.J: Princeton University Press, January 1957.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, February 2008.

Gertler, Mark and Simon Gilchrist, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *Quarterly Journal of Economics*, 1994, *109* (2), 309–340.

Gomes, Joao F., "Financing Investment," *American Economic Review*, December 2001, *91* (5), 1263–1285.

Grigsby, John, Erik Hurst, and Ahu Yildirmaz, "Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data," *American Economic Review*, February 2021, *111* (2), 428–471.

Hayashi, Fumio, "Tobin's Marginal q and Average q: A Neoclassical Interpretation," *Econometrica*, January 1982, *50* (1), 213–224.

He, Calvin, "Monetary Policy, Capital Reallocation, and Financial Frictions," *Work in Progress*, June 2025.

Hennessy, Christopher A. and Toni M. Whited, "How Costly Is External Financing? Evidence from a Structural Estimation," *The Journal of Finance*, 2007, *62* (4), 1705–1745.

Holmstrom, Bengt and Jean Tirole, *Inside and Outside Liquidity*, MIT Press, January 2011.

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

House, Christopher L. and Matthew D. Shapiro, "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation," *American Economic Review*, June 2008, *98* (3), 737–68.

Huggett, Mark, "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics*, December 1996, *38* (3), 469–494.

Jappelli, Tullio and Luigi Pistaferri, "The Consumption Response to Income Changes," *Annual Review of Economics*, 2010, *2* (1), 479–506.

Jeenas, Priit, "Firm Balance Sheet Liquidity, Monetary Policy Shocks, and Investment Dynamics," *Manuscript*, October 2025.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, December 2006, *96* (5), 1589–1610.

Jungherr, Joachim, Matthias Meier, Timo Reinelt, and Immo Schott, "Corporate Debt Maturity Matters for Monetary Policy," *Manuscript*, December 2024.

Kaplan, Greg and Giovanni L. Violante, "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, July 2014, *82* (4), 1199–1239.

— and —, "Microeconomic Heterogeneity and Macroeconomic Shocks," *Journal of Economic Perspectives*, August 2018, *32* (3), 167–194.

— and —, "The Marginal Propensity to Consume in Heterogeneous Agent Models," *Annual Review of Economics*, August 2022, *14* (1), 747–775.

—, Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," *American Economic Review*, March 2018, *108* (3), 697–743.

Kaplan, Steven N. and Luigi Zingales, "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?," *Quarterly Journal of Economics*, February 1997, *112* (1), 169 –215.

Kermani, Amir and Yueran Ma, "Asset Specificity of Nonfinancial Firms," *Quarterly Journal of Economics*, February 2023, *138* (1), 205–264.

Khan, Aubhik and Julia K. Thomas, "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, March 2008, *76* (2), 395–436.

— **and** — , "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, December 2013, *121* (6), 1055–1107.

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

Koby, Yann and Christian K. Wolf, "Aggregation in Heterogeneous-Firm Models: Theory and Measurement," *Manuscript*, July 2020.

Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896.

Laibson, David, Peter Maxted, and Benjamin Moll, "A Simple Mapping from MPCs to MPXs," Working Paper 29664, National Bureau of Economic Research, January 2022.

Li, Huiyu, Chen Lian, Yueran Ma, and Emily Martell, "Borrowing Constraints, Markups, and Misallocation," Working Paper 33960, National Bureau of Economic Research, 2025.

Lian, Chen and Yueran Ma, "Anatomy of Corporate Borrowing Constraints," *Quarterly Journal of Economics*, February 2021, *136* (1), 229–291.

Lucas, Robert E., "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 1978, *9* (2), 508–523.

Luetticke, Ralph, "Transmission of Monetary Policy with Heterogeneity in Household Portfolios," *American Economic Journal: Macroeconomics*, April 2021, *13* (2), 1–25.

Martin-Baillon, Alaïs, "Firms' Marginal Propensities to Invest," *Manuscript*, September 2021.

McKay, Alisdair and Christian K. Wolf, "Monetary Policy and Inequality," *Journal of Economic Perspectives*, February 2023, *37* (1), 121–144.

— **and Johannes F. Wieland**, "Lumpy Durable Consumption Demand and the Limited Ammunition of Monetary Policy," *Econometrica*, November 2021, *89* (6), 2717–2749.

— **, Emi Nakamura, and Jón Steinsson**, "The Power of Forward Guidance Revisited," *American Economic Review*, October 2016, *106* (10), 3133–3158.

Melcangi, Davide, "The Marginal Propensity to Hire," *Manuscript*, May 2022.

Nakamura, Emi and Jón Steinsson, "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, November 2008, *123* (4), 1415–1464.

Ottonello, Pablo and Thomas Winberry, "Financial Heterogeneity and the Investment Channel of Monetary Policy," *Econometrica*, 2020, *88* (6), 2473–2502.

Patterson, Christina, "The Matching Multiplier and the Amplification of Recessions," *American Economic Review*, April 2023, *113* (4), 982–1012.

Rauh, Joshua D., "Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans," *Journal of Finance*, February 2006, *61* (1), 33–71.

Rupert, Peter and Roman Šustek, "On the Mechanics of New-Keynesian Models," *Journal of Monetary Economics*, April 2019, *102*, 53–69.

Schmitt-Grohé, Stephanie and Martín Uribe, "What's News in Business Cycles," *Econometrica*, November 2012, *80* (6), 2733–2764.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, June 2007, *97* (3), 586–606.

Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–365.

Werning, Iván, "Incomplete Markets and Aggregate Demand," Working Paper 21448, National Bureau of Economic Research, August 2015.

Woodford, Michael, *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, August 2003.

_ , "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 2011, *3* (1), 1–35.

Zeldes, Stephen P., "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, April 1989, *97* (2), 305–346.

Zwick, Eric and James Mahon, "Tax Policy and Heterogeneous Investment Behavior," *American Economic Review*, January 2017, *107* (1), 217–48.

A Proof of Proposition 1

From the main text, the firm's dynamic problem is

$$v_t(z,k,b) = \max_{k',b',d} d + \frac{1}{1+r_t}\mathbb{E}\left[(1-\theta)v_{t+1}(z',k',b') + \theta v_{t+1}^{\text{exit}}(z',k',b')|z\right] \text{ s.t.} \tag{27}$$

$$d = \pi_t(z,k) - b - (k' - (1-\delta)k) - \varphi\left(\frac{k'-k}{k}\right)k + \frac{b'}{1+r_t} + T_t^f$$

$$d \geq 0 \text{ and } b' \leq \chi k', \quad \text{and where } v_t^{\text{exit}}(z,k,b) = zk^\alpha n^\nu - w_t n - b + (1-\delta)k.$$

We begin by taking the first-order conditions and then simplifying them to arrive at the characterization in Proposition 1.

A.1 First-Order Conditions

We solve the firm's problem using Lagrangian methods. We plug in the definition of dividends d to the objective function and attach multipliers $\lambda_t(z,k,b)$ and $\mu_t(z,k,b)$ to the $d \geq 0$ and $b' \leq \chi k'$ constraints.

Borrowing The first-order condition for b' is

$$0 = \frac{1+\lambda_t(z,k,b)}{1+r_t} - \mu_t(z,k,b) - \frac{\theta}{1+r_t} + \frac{1-\theta}{1+r_t}\mathbb{E}_t\left[v_{3,t+1}(z',k',b')\right]$$

$$\implies 1 + \lambda_t(z,k,b) = \mu_t(z,k,b)(1+r_t) + \theta - (1-\theta)\mathbb{E}_t\left[v_{3,t+1}(z',k',b')\right]. \tag{28}$$

To get $v_{3,t+1}(z',k',b')$, we use the envelope condition to see

$$v_{3,t}(z,k,b) = -(1+\lambda_t(z,k,b)) \implies v_{3,t+1}(z',k',b') = -(1+\lambda_{t+1}(z',k',b').$$

Plug this into (28) to get

$$1 + \lambda_t(z,k,b) = \mu_t(z,k,b)(1+r_t) + \theta + (1-\theta)\mathbb{E}_t\left[(1+\lambda_{t+1}(z',k',b')\right]$$

To simplify, note that $\theta + (1-\theta)\mathbb{E}_t[1+\lambda_{t+1}(z',k',b')] = 1 + (1-\theta)\mathbb{E}_t[\lambda_{t+1}(z',k',b')]$. Plugging this in and subtracting 1 from both sides, we have the final expression

$$\lambda_t(z,k,b) = \mu_t(z,k,b)(1+r_t) + (1-\theta)\mathbb{E}_t[\lambda_{t+1}(z',k',b')]. \tag{29}$$

The expression (29) says that the shadow price of issuing equity today is related to the shadow price of issuing debt plus a continuation value. We can solve (29) in sequence form for firm j to see

$$\lambda_{jt} = \mathbb{E}_t\left[\sum_{s=0}^{\infty}(1-\theta)^s(1+r_{t+s})\mu_{jt+s}\right].$$

In other words, firms equate the shadow price of issuing equity to the shadow price of issuing debt in all future states of the world in which the firm may be borrowing constrained. Therefore, if there is a positive probability of being borrowing constrained in some future state, then $\lambda_{jt} > 0$, so the firm will not pay dividends. Another way to say that is that funds are more valuable inside the firm if there is ever a future state in which the firm may be borrowing constrained.

Investment The first-order condition for capital investment k' is

$$-(1+\lambda_t(z,k,b))\left[1+\varphi'\left(\frac{k'}{k}-1\right)\right] + \mu_t(z,k,b)\chi + \frac{\theta}{1+r_t}\mathbb{E}_t\left[\pi_{2,t+1}(z',k') + (1-\delta)\right] + \frac{1-\theta}{1+r_t}\mathbb{E}_t\left[v_{2,t+1}(z',k',b')\right] = 0. \tag{30}$$

To get the marginal value function, we use the envelope condition

$$v_{2,t}(z,k,b) = (1+\lambda_t(z,k,b))\left(\pi_{2,t}(z,k) + (1-\delta) - \varphi\left(\frac{k'}{k}-1\right) + \varphi'\left(\frac{k'}{k}-1\right)\frac{k'}{k}\right).$$

For ease of notation, define $\widehat{\varphi}_{k,k'} = -\varphi\left(\frac{k'}{k}-1\right) + \varphi'\left(\frac{k'}{k}-1\right)\frac{k'}{k}$, so that $v_{2,t}(z,k,b) = (1+\lambda_t(z,k,b))\left(\pi_{2,t}(z,k) + (1-\delta) + \widehat{\varphi}_{k,k'}\right)$. Now plug this expression for the marginal value function into the first-order condition for investment (30) and rearrange to get

$$
\begin{aligned}
(1+\lambda_t(z,k,b))&\left[1+\varphi'\left(\frac{k'}{k}-1\right)\right] \\
&= \mu_t(z,k,b)\chi + \frac{\theta}{1+r_t}\mathbb{E}_t\left[\pi_{2,t+1}(z',k') + (1-\delta)\right] \\
&+ \frac{1-\theta}{1+r_t}\mathbb{E}_t\left[(1+\lambda_{t+1}(z',k',b'))\left(\pi_{2,t+1}(z',k')+(1-\delta)+\widehat{\varphi}_{k',k''}\right)\right].
\end{aligned} \tag{31}
$$

Summary To summarize, a solution to the firm's problem is a set of policy functions $k'_t(z,k,b)$ and $b'_t(z,k,b)$ together with multiplier functions $\mu_t(z,k,b)$ and $\lambda_t(z,k,b)$ such that the following four conditions hold:

(i) Borrowing FOC:

$$
\lambda_t(z,k,b) = \mu_t(z,k,b)(1+r_t) + (1-\theta)\mathbb{E}_t[\lambda_{t+1}(z',k',b')]. \tag{32}
$$

(ii) Investment FOC:

$$
\begin{aligned}
(1+\lambda_t(z,k,b))&\left[1+\varphi'\left(\frac{k'}{k}-1\right)\right] \\
&= \mu_t(z,k,b)\chi + \frac{\theta}{1+r_t}\mathbb{E}_t\left[\pi_{2,t+1}(z',k') + (1-\delta)\right] \\
&+ \frac{1-\theta}{1+r_t}\mathbb{E}_t\left[(1+\lambda_{t+1}(z',k',b'))\left(\pi_{2,t+1}(z',k')+(1-\delta)+\widehat{\varphi}_{k',k''}\right)\right].
\end{aligned} \tag{33}
$$

(iii) No equity issuance constraint: $\lambda_t(z,k,b)d = 0$ with $\lambda_t(z,k,b) \geq 0$.

(iv) Borrowing constraint: $\mu_t(z,k,b)(\chi k' - b') = 0$ with $\mu_t(z,k,b) \geq 0$.

A.2 Analysis

Following Khan and Thomas (2013), we will now show that the solution to this problem falls into three regions of the (k,b) state space:

(i) Unconstrained: $\lambda_t(z,k,b) = 0$.

(ii) Potentially constrained: $\lambda_t(z,k,b) > 0$ but $\mu_t(z,k,b) = 0$.

(iii) Actively constrained: $\mu_t(z,k,b) > 0$ (which implies that $\lambda_t(z,k,b) > 0$).

We will sometimes collapse categories (ii) and (iii) into one category of "constrained."

Unconstrained To characterize decisions of the unconstrained firms, we will presume that $\lambda_t(z,k,b) = 0$ and then derive the optimal decisions under this presumption. If these decisions are feasible, then the firm will want to choose them because they maximize the firm's objective subject to fewer constraints than the full problem.

We know that from (32) that, if $\lambda_t(z,k,b) = 0$, then $\mu_t(z,k,b) = 0$ and $\lambda_{t+1}(z',k',b') = 0$ with probability 1. Plugging this into the capital FOC and simplifying gives

$$1 + \varphi'\left(\frac{k'}{k} - 1\right) = \frac{1}{1+r_t}\mathbb{E}_t\left[\pi_{2,t+1}(z',k') + (1-\delta) + (1-\theta)\widehat{\varphi}_{k',k''}\right]. \tag{34}$$

Call the policy function which emerges from solving this problem $k_t^*(z,k)$. Note that it only depends on current productivity z and capital k.

For unconstrained firms, the mix of financing between dividends d and borrowing b' is not pinned down (a version of the Modigliani-Miller theorem). Following Khan and Thomas (2013), we assume these firms follow the "maximum borrowing policy" $b_t^*(z,k)$, i.e. the choice of borrowing b' such that borrowing is set to its highest level such that the firm remains unconstrained. To solve for the maximum borrowing policy, we conduct the following thought experiment: suppose the firm adopts the unconstrained capital accumulation policy starting today and the minimum savings policy starting tomorrow — what is the highest level of borrowing today b' such that the firm can pay positive dividends with probability one?

If the firm receives the exit shock, then dividends in the next period, conditional on idiosyncratic productivity level z', is

$$d_{t+1}(z,k,z') = \pi_{t+1}(z', k_t^*(z,k)) + (1-\delta)k_t^*(z,k) + T_{t+1}^f - b'.$$

For these dividends to be positive with probability one, we need

$$b' \leq \pi_{t+1}(z', k_t^*(z,k)) + (1-\delta)k_t^*(z,k) + T_{t+1}^f \equiv b_t^1(z,k,z')$$

for all z' with $p(z'|z) > 0$.

If the firm does not receive the exit shock in the next period, then dividends conditional on idiosyncratic productivity z' are

$$\begin{aligned} d_{t+1}(z,k,z') = \pi_{t+1}(z', k_t^*(z,k)) + (1-\delta)k_t^*(z,k) + T_{t+1}^f - b' - k_{t+1}(z', k_t^*(z,k)) \\ - \varphi\left(\frac{k_{t+1}(z', k_t^*(z,k))}{k_t^*(z,k)} - (1-\delta)\right) k_t^*(z,k) + \frac{b_{t+1}(z', k_t^*(z,k))}{1+r_t} \end{aligned}$$

For these dividends to be positive with probability one, we need

$$\begin{aligned} b' \leq & \pi_{t+1}(z', k_t^*(z,k)) + (1-\delta)k_t^*(z,k) + T_{t+1}^f - k_{t+1}(z', k_t^*(z,k)) \\ & - \varphi\left(\frac{k_{t+1}(z', k_t^*(z,k))}{k_t^*(z,k)} - (1-\delta)\right) k_t^*(z,k) + \frac{b_{t+1}(z', k_t^*(z,k))}{1+r_t} \equiv b_t^2(z,k,z') \end{aligned}$$

for all z' with $p(z'|z) > 0$.

In addition, the minimum savings policy must satisfy the borrowing constraint in order to be feasible. Putting all these cases together, we have

$$b_t^*(z,k) = \min\{\min_{z'} b_t^1(z,k,z'), \min_{z'} b_t^2(z,k,z'), \chi k\}. \tag{35}$$

If the unconstrained choices are feasible, then the firm will want to choose them. These choices are feasible if adopting them starting in the current period does not violate the no-equity issuance constraint:

$$d = \pi_t(z,k) + (1-\delta)k + T_t^f - b - k_t^*(z,k) - \varphi\left(\frac{k_t^*(z,k)}{k} - 1\right) k + \frac{b_t^*(z,k)}{1+r_t} \geq 0.$$

This condition is satisfied as long as current debt b is not high enough:

$$b \leq \pi_t(z,k) + (1-\delta)k + T_t^f - k_t^*(z,k) - \varphi\left(\frac{k_t^*(z,k)}{k} - 1\right) k + \frac{b_t^*(z,k)}{1+r_t} \equiv \underline{b}_t(z,k). \tag{36}$$

All firms with state variable (z,k,b) such that $b \leq \underline{b}_t(z,k)$ will be unconstrained.

Potentially Constrained Firms Constrained firms have $\lambda_t(z,k,b) > 0$ and therefore set $d = 0$. To derive the cutoff $\overline{b}_t(z,k)$, suppose that the collateral constraint were slack so that $\mu_t(z,k) = 0$. Denote the associated policy functions $k_t^{\mathrm{p}}(z,k,b)$ and $b_t^{\mathrm{p}}(z,k,b)$. Since $d = 0$ is binding, we must have that

$$b_t^{\mathrm{p}}(z,k,b) = -(1+r_t)\left(\pi_t(z,k) + (1-\delta)k + T_t^f - b - k_t^{\mathrm{p}}(z,k,b) - \varphi\left(\frac{k_t^{\mathrm{p}}(z,k,b)}{k} - 1\right)k\right).$$

If the firm could adopt these policies, then it would be optimal to do so because they solve a relaxed version of the full problem. The policies are feasible as long as $b_t^{\mathrm{p}}(z,k,b) \leq \chi k$. Using the expression above, this occurs if

$$b \leq \frac{\chi k}{1+r_t} + \pi_t(z,k) + (1-\delta)k + T_t^f - k_t^{\mathrm{p}}(z,k,b) - \varphi\left(\frac{k_t^{\mathrm{p}}(z,k,b)}{k} - 1\right)k \equiv \underline{b}_t(z,k,b). \quad (37)$$

Actively Constrained Firms Given the discussion above, firms with $b > \underline{b}_t(z,k,b)$ will have both constraints binding. Plugging $b' = \chi k'$ into $d = 0$ gives

$$\pi_t(z,k) + (1-\delta)k + T_t^f - b - k' - \varphi\left(\frac{k'}{k} - 1\right)k + \frac{\chi k'}{1+r_t} = 0.$$

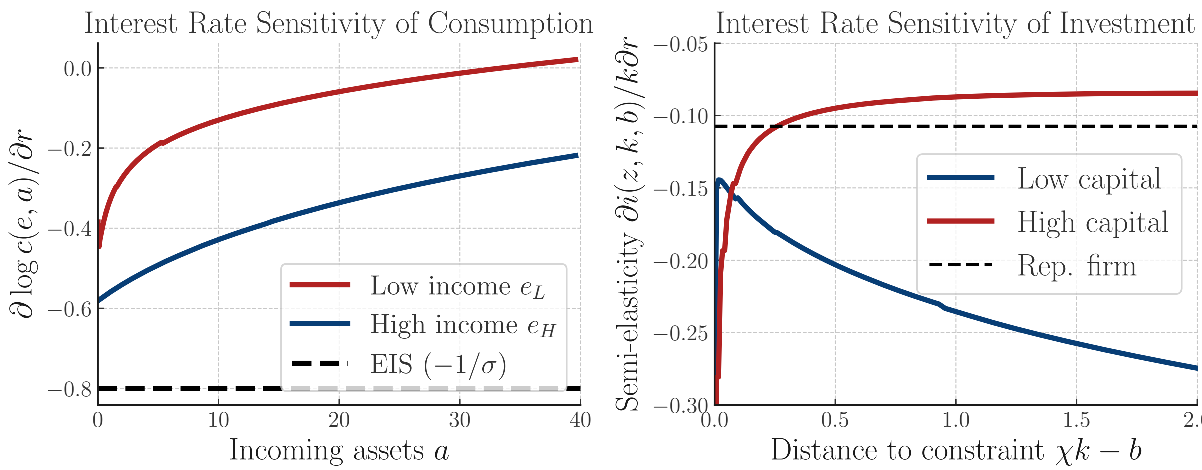
The highest level of k' that solves this equation is the maximal investment level $\overline{k}_t(z,k,b)$ from the main text.

B Interest rate sensitivities

Figure B.1 compares the interest rate sensitivities of households and firms, defined as the sensitivity of their policy functions to a one-time change in the real interest rate. In the representative-agent model, $\partial \log c/\partial r$ is equal to the EIS σ^{-1}, while in the representative firm model $\partial i/k\partial r$ is equal to $1/\phi$.

The heterogeneous household model features heterogeneity in the consumption sensitivity to interest rates, for two reasons. First, households have higher marginal propensities to consume, dampening their response to interest rates. Second, households with higher assets experience a positive income effect from a rise in interest rates, which offsets their

FIGURE B.1: Interest Rate Sensitivies of Households and Firms



Notes: steady-state interest rate sensitivities of households' decision rules (left panel) and firms (right panel). Household model calibrated as in Table 2, firm model calibrated as in Table 3. Interest rates have heterogeneous income and substitution effects on households and firms, leading them to be affected differentially.

substitution effect. This effect is larger for richer agents, explaining why richer agents have lower sensitivity of consumption to interest rates. Eventually, for rich enough agents, this income effect turns positive.

Likewise, the heterogeneous firm model features heterogeneity in the investment sensitivity to interest rates. This heterogeneity has been the focus of much of the literature following Ottonello and Winberry (2020). Firms close to the borrowing constraint are affected by a cash flow effect from movements in the rates, explaining relatively high sensitivity at the constraint. Just like for households, there are income effects from changes in rates. Overall, it is ambiguous whether firms at the constraint respond more or less than firms far from the constraint: in our calibration, unconstrained low-capital firms respond more, while unconstrained high-capital firms respond less than constrained firms.