Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Cyclical Fluctuations, Financial Frictions, and Productivity Differences across Firms

Luca Guerrieri, Jinill Kim, Arsenii Mishin

2026-047

Please cite this paper as:
Guerrieri, Luca, Jinill Kim, and Arsenii Mishin (2026). "Cyclical Fluctuations, Financial Frictions, and Productivity Differences across Firms," Finance and Economics Discussion Series 2026-047. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2026.047.

Cyclical Fluctuations, Financial Frictions, and Productivity Differences across Firms*

Luca Guerrieri
Federal Reserve Board

Jinill Kim
Korea University
Bank of Korea

Arsenii Mishin
HSE University

June 23, 2026

Abstract

Within narrowly defined industries, the most productive firms produce far more than the least productive from the same inputs, and this dispersion widens in downturns. We build a tractable representative-agent model in which financial frictions—adverse selection and moral hazard—make firms sort endogenously into lenders, strategic defaulters, and producers. As credit conditions vary, the resulting misallocation gives aggregate total factor productivity (TFP) an endogenous component that accounts for about 30 percent of the variance of TFP at business-cycle frequencies, a third of it from strategic default. We show that our tractable model can match key features of the observed distribution of productivity across firms and its co-movement with output growth and credit conditions in the data.

JEL Classification: E23, E32, E44
Keywords: productivity dispersion, endogenous productivity, financial intermediation, business cycles

*The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, of any other person associated with the Federal Reserve System, of the Bank of Korea, or of any other person associated with the Bank of Korea. An online appendix is available at the following link: https://www.lguerrieri.com/downloads/Appendix_GKM.pdf

1 Introduction

Within narrowly defined industries, a firm near the top of the productivity distribution produces about twice as much as a firm near the bottom from the same measured inputs—and that gap widens in the years when output growth runs below average. Productivity differences this large are pervasive across the economy, and a substantial literature traces them to financial frictions that keep capital from flowing to its most productive uses.[1] What matters for the aggregate economy is that this dispersion is not static but moves with the business cycle: if the spread of productivity across firms widens and narrows over the cycle, then so does the efficiency with which the economy uses its inputs, and so does aggregate total factor productivity (TFP).[2] Capturing this link in full generality, however, requires tracking distributions of firm-level state variables—models that are powerful but difficult to embed in the representative-agent frameworks used to study aggregate fluctuations.

This paper resolves that tension. Our central contribution is a representative-agent model that delivers endogenous productivity dispersion driven by financial frictions while remaining tractable enough to be estimated and embedded in standard business-cycle frameworks. In our model, each firm is endowed with two technologies: a production technology, whose idiosyncratic efficiency is private information, and a financial intermediation technology, which is available to all firms on equal terms. Because firm-level production efficiency is not directly observed by investors and the claims of firms are not credible, this informational friction results in all firms being financed with an aliquot share of household savings, regardless of productivity. In a secondary market for funds, where the intermediation technology of a firm allows a little more screening of productivity to be performed, firms compare the returns from their two technologies and sort themselves into lenders, firms deploying their intermediation technology; borrowers that default strategically without producing; and producers. This sorting depends on the realization of the production-technology draw, and characterizing the equilibrium reduces to locating the productivity cutoffs that demarcate

[1] Foundational contributions include Foster et al. (2008), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Khan and Thomas (2013), Moll (2014), Gopinath et al. (2017), and David and Venkateswaran (2019).

[2] On the decomposition of aggregate TFP into the average of firm-level productivities and the efficiency of input allocation across firms, see Hsieh and Klenow (2010).

these groups. As these cutoff points respond to aggregate macroeconomic conditions, the resulting model delivers endogenous dynamics for aggregate TFP akin to those in the endogenous growth literature: Roughly 30 percent of the variance of TFP at business cycle frequencies is endogenous in our model.

The appeal of this structure is its tractability. The aggregate consequences of capital misallocation turn out to be captured by just a few objects: the two cutoffs that separate lenders, defaulting borrowers, and producers, and the interest rate at which firms lend to one another. As aggregate conditions change, these boundaries shift, and the changing composition of producing firms is all that the aggregate economy needs to track. The model therefore adds only a handful of equations to a standard real business cycle framework, rather than requiring us to follow an evolving distribution of firms over time. Accordingly, we can embed the mechanism in an otherwise standard business-cycle model, solve and estimate it with off-the-shelf methods. We view the model as a building block that others can readily reuse or extend.

But our contribution extends beyond tractability. The framework captures strategic default as an equilibrium outcome, not a knife-edge case ruled out by contract design, and the equilibrium variation in the mass of defaulting firms is quantitatively important: it accounts for roughly one-third of the endogenous TFP component, over and above what a single-cutoff model without strategic default can generate. Strategic default thus serves a dual purpose: it amplifies the endogenous TFP component by concentrating production among high-productivity firms, and in our simulations also stabilizes the dynamic path of that component—when aggregate conditions shift, firms near the default threshold adjust first, distributing the reallocation burden across two margins rather than concentrating it in a single threshold, which yields a smoother and more persistent endogenous TFP response than the no-default variant produces.

As in Stiglitz and Weiss (1981), SW henceforth, credit rationing is at the center of our model, but it arises from some additional incentives built into our model. In SW, default arises involuntarily when project returns fall short of the promised repayment; the interest rate affects bank returns by changing the composition of applicants (adverse selection) and the riskiness of projects chosen (incentive effect/moral hazard). Our model adds a different channel: firms with sufficiently low productivity find it profitable to divert borrowed funds to an outside option rather than produce, making default a strategic, equilibrium choice.[3] This moral hazard margin, different from the one in SW, generates a two-tier sorting structure. In our model interest rates influence both the strategic default rate and the average productivity of firms that choose to borrow, produce, and repay their loans.[4]

In our model, misallocation is not static. As aggregate conditions change, the thresholds governing which firms lend, borrow, and default fluctuate, and with them the dispersion of productivity across firms and aggregate TFP. The model thus encompasses two sources of aggregate TFP: a standard exogenous component and an endogenous component driven by changes in the distribution of productive activity across firms. Under our baseline calibration, this endogenous channel generates meaningful differences in impulse responses to TFP shocks relative to models without financial frictions and productivity dispersion. In the efficient limit of our model, all credit flows to the most productive firm and the economy collapses to a standard RBC model.

Despite its relative simplicity, the model is sufficiently flexible to match key features of the data. We calibrate the model to standard targets for credit market conditions and productivity dispersion, and find that it comes close to replicating several untargeted moments. In particular, OLS regressions estimated on model-simulated data reproduce the signs of their counterparts in U.S. data: within-industry productivity dispersion narrows when GDP growth runs above average, and delinquencies on business loans rise when growth is depressed and when real interest rates are high.

[3] There are other important differences between our analysis and the one in Stiglitz and Weiss (1981). We develop a general equilibrium model whereas theirs is a static partial equilibrium model. The model in Stiglitz and Weiss (1981) has no capital accumulation, no goods market, no aggregate TFP, and no business cycle dynamics. The bank's interest rate is a parameter, not an endogenous price that co-moves with output, investment, and consumption as in our model.

[4] In an entrepreneurship model with multiple sources of financial frictions, Paulson et al. (2006) showed—via structural estimation on Thai data—that moral hazard due to information asymmetry is the dominant source.

2 Literature Review

Our work has many antecedents. It is related first to papers that showcase models with costly state verification, such as Carlstrom and Fuerst (1997), Bernanke et al. (1999), and Christiano et al. (2014). These papers established that the agency costs of financial intermediation can substantially amplify and propagate aggregate shocks, a foundational insight that allows us to go further. A common modelling device for these papers is the assumption that credit needs to be allocated to firms with heterogeneous productivity. However, the productivity dispersion is static and exogenously given—it cannot respond to economic conditions because the distribution of idiosyncratic productivity is assumed fixed and independent of aggregate developments. Default in those models is also not strategic: it stems from the inability to repay after uncertainty resolves unfavorably. A further distinction concerns the source of heterogeneity: in those frameworks it is idiosyncratic project-return or entrepreneurial risk, rather than the producing-firm productivity-selection margin that is central to our model. In Christiano et al. (2014) in particular, idiosyncratic risk varies over time, but exogenously, rather than responding to the equilibrium sorting of firms. By contrast, in our model default is strategic, varies with economic conditions, and feeds back into productivity dispersion and aggregate TFP.

Our work is also related to papers that include endogenously evolving productivity dispersion. Khan and Thomas (2013) were the first to explore the endogenous TFP channel in a quantitative DSGE setting where reallocation of capital across heterogeneous firms determines aggregate TFP. Their model works through the interaction of two frictions—collateralized borrowing constraints and partial capital irreversibility—that together distort capital allocation and generate persistent endogenous TFP movements. Our model captures the same channel through a more parsimonious mechanism: strategic default sorts firms by productivity and generates credit rationing without requiring capital-level heterogeneity or (S,s) adjustment rules. And unlike their model, which is designed to rule out default in equilibrium, default is an equilibrium outcome in ours. Khan and Thomas's rich framework allows them to match detailed microeconomic moments on firm investment behavior that our model does not target; the tradeoff is that their solution algorithm requires tracking a three-dimensional firm distribution. By contrast, the simplicity of our framework allows us to embed it in many variants of an RBC model, and we view it as a building block that others can readily reuse or extend.

Our emphasis on tractability connects us to Buera and Moll (2015), who developed a model where heterogeneous firms face collateral constraints and shocks to those constraints open productivity wedges. In their benchmark model, idiosyncratic firm productivity is independently and identically distributed (i.i.d.) across entrepreneurs and over time, so the cutoff governing which firms produce is determined exogenously by the collateral constraint rather than by equilibrium sorting. When they allow for persistent idiosyncratic shocks—treated in their online appendix under logarithmic utility to preserve tractability—the conclusions were the same as in the i.i.d. case. Under logarithmic preferences, moreover, a tightening of collateral constraints is exactly isomorphic to an exogenous TFP shock, and no other shocks are considered; the endogenous TFP component therefore cannot be separately identified from the exogenous one. Our model, by contrast, delivers an endogenous TFP component that is not isomorphic to an exogenous technology shock and can therefore be separately identified, even in the i.i.d. limit.

Our work is also related to Liu and Wang (2014) and Dong et al. (2025), who studied heterogeneous firms under collateral constraints in similar settings. In Liu and Wang (2014), borrowing constraints are governed by an exogenous parameter rather than by the equilibrium sorting of firms across lending, production, and default. By design, no firm defaults in equilibrium in their framework; default is instead an equilibrium outcome in ours, and it is precisely the endogenous variation in the mass of defaulting firms that drives an important part of the endogenous TFP response. A further difference is that the endogenous TFP channel in our model operates through equilibrium sorting at the default margin, without relying on entry and exit costs. Dong et al. (2025) is a targeted extension of Liu and Wang (2014) that focuses on the effects of turbulence shocks. In their model, misallocation arises from the joint presence of credit constraints and production distortions; turbulence operates by tightening the borrowing constraints of high-productivity firms and amplifying pre-existing misallocation. In our model, misallocation arises endogenously from financial frictions alone, without requiring pre-existing production distortions.

Our paper also contributes to a body of work exploring heterogeneous firms, financial frictions, and default. Gilchrist et al. (2014) studied credit spreads under uncertainty shocks in a model with default. Arellano et al. (2019) examined the role of uncertainty shocks in a model with noncontingent debt and equilibrium default. Gomes and Schmid (2021) developed a model with endogenous default where firms vary in leverage and studied the implications for credit spreads.[5] Closest to our treatment of default is the limited-enforcement friction in Cao et al. (2019), in which an entrepreneur who defaults can divert a fraction of the firm's capital—much as in our diversion technology—generating an endogenous borrowing constraint. Their analysis is set in partial equilibrium at the firm level and targets the relationship between investment, Tobin's q, and cash flow; default influences an always-binding constraint but does not occur in equilibrium. In our model the friction operates in general equilibrium, a positive mass of firms defaults each period, and the object of interest is the endogenous component of aggregate TFP.

The misallocation literature more broadly provides important context for our work. Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) established that policy distortions and wedges between marginal products account for large aggregate productivity losses. Midrigan and Xu (2014) and Moll (2014) emphasized financial frictions as a central driver of misallocation and productivity losses. Moll (2014) established an important benchmark: With serially persistent idiosyncratic productivity, entrepreneurs can gradually self-finance out of credit constraints, which reduces misallocation and allows TFP to recover over time. Without serial correlation, however, that particular self-financing channel disappears and TFP is fully driven by exogenous factors. Our paper goes further, showing that endogenous TFP dynamics arise generically even in this i.i.d. case. First, an equilibrium-selection channel makes the cutoffs for lending, borrowing, and default responsive to aggregate conditions, producing endogenous TFP variation independent of wealth dynamics. Second, explicitly modeling strategic default creates an additional, quantitatively relevant source of TFP fluctuations through firms' default choices and their general-equilibrium effects. Together these

[5] Jo et al. (2021) and Ottonello and Winberry (2020) also introduced financial frictions and firm default, but our model is more tractable and intuitive by incorporating moral hazard and strategic default; they focused on size and leverage distributions, whereas our model emphasizes consistency with productivity dispersion. Guo et al. (2025) discussed business cycles with firms' private information and equity financing.

mechanisms substantially broaden the range of financial-contracting features that can drive aggregate misallocation and TFP dynamics beyond the scope of Moll (2014). David and Venkateswaran (2019) and Gopinath et al. (2017) documented the importance of capital misallocation both in the United States and in Southern Europe. On the firm dynamics side, Clementi and Hopenhayn (2006) and Cooley and Quadrini (2001) developed theories of financing constraints and firm dynamics that inform our modeling of credit frictions.

Beyond the financial-friction and misallocation literatures, the endogenous component of TFP in our model connects to a long tradition in the endogenous growth literature, which uses representative-agent frameworks to study the evolution of aggregate productivity, building on foundational contributions by Lucas (1988), Romer (1990), and Aghion and Howitt (1992).[6] Our approach, however, avoids a classic criticism of that literature. Models of endogenous growth can lead to strong predictions and policy recommendations based on long-run growth effects that are difficult to ascertain empirically. In our model, the endogenous effects are not permanent, and the tractable structure allows us to deploy standard methods for an empirical verification. While the endogenous growth literature has generated deep insights about long-run forces shaping aggregate TFP, empirically isolating those forces has proven difficult: aggregate time series data cannot readily distinguish neoclassical from endogenous growth models (Pack, 1994), and the low-frequency statistical models underlying many empirical exercises are nearly indistinguishable from one another at typical macroeconomic sample lengths (Müller & Watson, 2008). By contrast, our mechanism leaves a testable footprint in the cross section: we can use data on the dispersion of productivity for firms within narrowly defined industries in support of the micro-foundations of our model.

A more recent strand of this literature generates endogenous aggregate productivity through innovation and the reallocation of activity across heterogeneous firms. Akcigit et al. (2016) study a market for ideas in which search frictions and technological mismatch deter-

[6] Lucas (1988), Romer (1990), and Aghion and Howitt (1992) established how human capital accumulation, R&D, and innovation endogenously determine the level and growth rate of productivity. King and Rebelo (1990) and Rebelo (1991) embedded these mechanisms in calibrated representative-agent frameworks, showing how endogenous growth alters the predictions of RBC models with implications for policy analysis. Comin and Gertler (2006) showed how endogenous technology adoption and R&D generate medium-frequency TFP dynamics that persist well beyond the business cycle. For further examples in this tradition, see Jones et al. (2005) and Dang et al. (2012) on endogenous growth with business cycle implications, and Gornemann et al. (2025) for a quantitative open-economy application with endogenous productivity.

mine how efficiently innovations are allocated across firms; Acemoglu et al. (2018) emphasize selection in which firms innovate and survive, showing that policies that reallocate research resources toward more innovative firms raise growth; and Peters (2020) derives misallocation endogenously from heterogeneous mark-ups, with a two-way feedback between growth and the dispersion of distortions. As in our model, aggregate productivity in these frameworks is an endogenous object that depends on how a key input is allocated across firms, and the models are disciplined by the observed dispersion of revenue productivity. They differ from ours in three respects that together delineate our contribution. First, the endogenous component there is a permanent feature of long-run growth, whereas in our model it is a transitory component of TFP at business-cycle frequencies, with long-run growth left exogenous. Second, the misallocation is technological or driven by market power rather than by financial frictions. Third, these are heterogeneous-firm growth models that track an evolving firm distribution along a balanced growth path, whereas our mechanism reduces to two cutoffs and a price and embeds in a standard representative-agent business cycle model.

Finally, our paper is related to an extensive literature explaining productivity differences across firms. Beyond financial frictions, alternative explanations include slow learning of average productivity in the face of noise shocks, as in Jovanovic (1982); the coexistence of older and newer technology vintages, as in Caballero and Hammour (1994), Caballero and Hammour (1996), and Caballero and Hammour (1998); search and matching frictions between workers and firms, as in Barlevy (2002); and imperfect product substitutability, as in Bernard et al. (2003) and Melitz (2003).

3 Data Motivation

We document that within-industry productivity differences in the U.S. manufacturing sector are sizable and pervasive, and that they tend to widen when GDP growth runs below average. The summary statistics constructed here are the same ones that we use to assess our theoretical model.

We draw dispersion in productivity across 86 four-digit NAICS manufacturing industries from experimental productivity dispersion statistics, Dispersion Statistics on Productivity

(DiSP), derived from Census Bureau microdata.[7] The most recent release of DiSP covers the years 1987–2021 on an annual basis. Our analysis focuses on the second-moment measure of establishment-level total factor productivity.

Figure 1 shows the evolution of TFP dispersion over time. The top panel considers the percentage difference in productivity between the establishment at the 90^{th} percentile of the TFP distribution and the 10^{th} percentile establishment in the same industry. The solid line denotes the average of these within-industry 90–10 gaps, weighted by each industry's contribution to GDP.[8] The units on the vertical axis are percentage points: a reading of one hundred means that the plant at the 90^{th} percentile makes twice as much output as the 10^{th} percentile plant with the same measured inputs. By that yardstick, the average 90–10 gap is sizable throughout the sample, fluctuating around 100 percent without an evident trend.

The dot-dashed bands describe how that gap varies across industries. Year by year, we rank the 86 industries by their own 90–10 gap and plot the gap of the industry at the 15^{th} percentile and at the 85^{th} percentile of the cross-industry distribution. The band is fairly symmetric around the solid line, and even the industry at the 15^{th} percentile across industries still has a sizable 90–10 gap. The average TFP gap is therefore not driven by a handful of outliers: large within-industry productivity differences are pervasive across the manufacturing sector.

For robustness, the bottom panel repeats the same analysis for the interquartile range (IQR). Within-industry productivity differences remain important quantitatively when we focus on segments closer to the center of the distribution, not just the tails.

Turning to cyclical variation, the shaded areas in Figure 1 mark recessions as dated by the National Bureau of Economic Research. Reading the solid line in the top panel against those shaded windows—and against the lower-frequency variation in GDP growth that they highlight—the 90–10 gap tends to widen in the years when GDP growth runs below its sample average and, most visibly, during NBER recessions. The same broad pattern is apparent in the bottom panel for the interquartile range.

[7] This dataset is described in Cunningham et al. (2023).

[8] The industry value-added data used to construct the weighted average TFP dispersion shown in Figure 1 is from the Bureau of Economic Analysis. To construct the industry weights, we aggregate the value added of multiple NAICS manufacturing industries that correspond to the same BLS code.

4 Model

Our analysis focuses on the key role of credit markets in reallocating capital among heterogeneous firms. Each firm in the goods sector is endowed with two technologies. The first is a production technology that combines capital and labor using an idiosyncratic efficiency parameter ω drawn from a distribution on the unit interval; this draw is transitory and privately observed by the firm. The second is a financial intermediation technology, available on equal terms to every firm, which allows it to lend funds to other firms at the prevailing inter-firm rate. Firms also experience persistent aggregate technology shocks that shift the productivity of the production technology for all firms simultaneously. Because the idiosyncratic efficiency of each firm's production technology is private information, households allocate an aliquot share of their savings to each firm without conditioning on ω. We will show that firms with sufficiently high realizations of ω borrow from firms with low realizations to purchase additional capital, while the latter find it optimal to deploy their financial intermediation technology instead of producing.

Unlike in a frictionless credit market, two distinct informational frictions interact to generate misallocation of capital and reduced productivity in our framework. The first is adverse selection: the production-technology efficiency ω above some threshold level is private and unobservable to outside investors, so the composition of borrowers cannot be screened perfectly. Because all eligible firms could credibly claim high productivity, offering to pay a higher rate to borrow more funds lacks credibility, and credit is rationed as in Stiglitz and Weiss (1981). The second friction is moral hazard: firms with sufficiently low realizations of ω find it profitable to divert borrowed funds to an outside option rather than produce and repay, making default a strategic, equilibrium choice. It is this moral hazard channel—not simply the adverse-selection problem—that limits the borrowing capacity of even the most efficient firm, because lenders must price in the possibility that some borrowers will divert.

Our interpretation of inter-firm lending reflects the broader role of credit markets in allocating credit. As Bernanke and Gertler (1990) note, this lending can be viewed as intermediated by competitive financial institutions that neither use resources nor earn profits in equilibrium. Variation in the cutoff points we just discussed will also lead to endogenous changes in the size of the intermediation sector, as the boundary between firms deploying their production technology and those deploying their financial intermediation technology shifts with aggregate conditions.

Aside from the endogenous financial intermediation, the other main innovation of the model consists of tracking the dispersion of productivity across goods-producing firms in a tractable way. Relative to a standard model, the complications will consist of characterizing three additional equilibrium objects. The first two objects are a couple of cutoff points within the distribution of firm productivity. A lower cutoff point will separate firms with such low productivity that lending to more productive firms will be more profitable than producing. Just on the other side of this cutoff will be more-productive firms that borrow in the inter-firm market. An upper cutoff point will demarcate firms that borrow and default from those that do not default. The third additional object is the price at which funds are lent across firms.

As these cutoff points and the cost of credit fluctuate, productivity differences across firms will also fluctuate, as will default rates, together with the overall efficiency of the economy. We will show that despite the simplicity of our model, key characteristics of these fluctuations are well aligned with U.S. data.

Firms are at the center of our model, but we start with households for ease of exposition.

4.1 Households

There is an infinitely-lived representative household that has preferences over consumption and labor, respectively C_t and H_t, in line with King et al. (1988). The household solves the following problem:

$$\max_{\left\{A_{t+\tau},C_{t+\tau},H_{t+\tau},B^H_{t+\tau}\right\}_{\tau=0}^{\infty}} E_t \sum_{\tau=0}^{\infty} \beta^{\tau} \left[\ln(C_{t+\tau} - \nu_{ct+\tau}) - \frac{\vartheta}{1+\nu} (H_{t+\tau})^{1+\nu} \right], \tag{1}$$

subject to

$$C_{t+\tau} + A_{t+\tau} + B^H_{t+\tau} = R^A_{t+\tau} A_{t+\tau-1} + W_{t+\tau} H_{t+\tau} + R^B_{t+\tau-1} B^H_{t+\tau-1} + \Pi_{t+\tau} + T_{t+\tau} + \Xi_{t+\tau}. \tag{2}$$

In period $t+\tau$, the household chooses how much to consume, $C_{t+\tau}$, and work, $H_{t+\tau}$. The term $\nu_{ct+\tau}$ is an exogenous shock to consumption, which follows an autoregressive process of order one:

$$\nu_{ct} = \rho_\nu \nu_{ct-1} + \varepsilon_{\nu t}, \qquad \varepsilon_{\nu t} \sim \mathcal{N}(0, \sigma_\nu^2). \tag{3}$$

The parameter β discounts future utility. The parameter ϑ captures the disutility from working, and the parameter ν governs the Frisch elasticity of labor supply. The household also chooses the amount of assets $A_{t+\tau}$ and a government bond $B^H_{t+\tau}$. The household enters period $t+\tau$ with assets $A_{t+\tau-1}$ which carry a state-contingent return $R^A_{t+\tau}$ from equity ownership of the goods-producing firms (which act as both producers and financial intermediaries) and receives a non-state contingent return $R^B_{t+\tau-1}$ from its holdings of the government bond $B^H_{t+\tau-1}$. It also supplies $H_{t+\tau}$ units of labor at the wage rate $W_{t+\tau}$. The term $\Pi_{t+\tau}$ captures profits from ownership of capital-producing firms; goods-producing firms earn zero profits in equilibrium due to perfect competition, while capital-producing firms can earn non-zero profits in equilibrium. The term $T_{t+\tau}$ represents a lump-sum transfer from the government, which runs a balanced budget, period by period (defined in equation (29) of Section 4.4). Finally, the term Ξ_t, defined in equation (30) of Section 4.4, refers to transfers that the household receives because firms that take the outside option, which is described below, are subject to a haircut on the returns from the outside option.

The first-order conditions for assets, consumption, labor, and government bonds are derived in Section C.1 of the online appendix; see equations (C.1)–(C.4), where λ_{ct} is the Lagrange multiplier attached to the budget constraint (2).

4.2 Production and Financial Intermediation

Firms in the goods sector are at the heart of our model. The firms in this sector follow a two-period overlapping structure from period t to $t+1$. This device is familiar from decentralizations of the RBC model that rely on equity contracts to allocate household savings to firms. Table 1 provides a roadmap to the sequence of choices made and actions taken by firms within each period. We highlight with a **boldface** font the choices or actions that are specific to our model. Lowercase letters will denote variables of individual firms; we shall reserve uppercase letters for aggregate variables.

Firms operate in a perfectly competitive market, producing a homogeneous good. Ex ante, each firm is endowed with two technologies: a production technology and a financial intermediation technology. An idiosyncratic efficiency parameter $\omega \in [0,1]$, drawn from an identical distribution at the start of each period, governs how productive the firm would be if it chose to deploy its production technology. The financial intermediation technology, which consists of lending funds in the inter-firm market, is identical across firms. Depending on the realization of ω, firms endogenously choose which technology to deploy: those with high ω find production more profitable and borrow to expand their capital stock, while those with low ω find it optimal to act as financial intermediaries and lend their funds to higher-productivity firms.

In a slight abuse of notation, the term ω will do double duty by denoting the idiosyncratic production efficiency and by serving as an index for other firm-specific variables. Since households have no visibility into the idiosyncratic production efficiency of firms, each firm receives an aliquot share of the households' savings, a_t. So, notice that a_t does not depend on ω in equilibrium.

Starting from the production function, the output of an individual firm in period $t+1$, $y_{t+1}(\omega)$, is governed by

$$y_{t+1}(\omega) = \omega Z_{t+1} k_t(\omega)^{\alpha} h_{t+1}(\omega)^{1-\alpha}. \tag{4}$$

The terms $k_t(\omega)$ and $h_{t+1}(\omega)$ denote the levels of capital and labor inputs used by the firm. The idiosyncratic productivity ω follows the cumulative distribution function $\mu(\omega)$ on the interval $[0,1]$, satisfying $\mu(0)=0$, $\mu(1)=1$, and $\mu'(\omega)>0$. The aggregate technology shock Z_{t+1} evolves according to

$$\log Z_{t+1} = \rho_z \log Z_t + \varepsilon^z_{t+1}, \qquad \varepsilon_{zt} \sim \mathcal{N}(0, \sigma_z^2), \tag{5}$$

where the parameter ρ_z governs the persistence of the shock process.

Firms make plans in period t to produce in period $t+1$. After their idiosyncratic productivity ω is known, the intermediation market opens. Depending on their productivity, some firms borrow while others lend at the predetermined rate ρ_t. Firms that borrow decide whether to use an outside option by purchasing the government bond or produce.[9] If they choose to produce, they purchase physical capital.

Firms can walk away from loans and default. If a firm decides to default, it can retain a fraction $\Theta_t(\omega)$ of the funds borrowed from other firms $b_t(\omega)$. We assume that $\Theta_t(\omega)$ increases in ω and is concave. Section B of the online appendix provides a structural interpretation for our modeling of $\Theta_t(\omega)$. For simplicity, we make the average fraction of funds that can be diverted equal to θ, so

$$\theta = \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \Theta_t(\omega) \frac{\mu'(\omega)}{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)} \mathrm{d}\omega, \tag{6}$$

where, as will be shown later, $\overline{\omega}_t$ and $\overline{\overline{\omega}}_t$ define the boundaries of defaulting firms, and $\frac{\mu'(\omega)}{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)}$ gives the corresponding probability density over that interval.[10]

Creditor firms are assumed not to be able to go after any funds set aside for the outside option. The diverted assets are placed with the government bond; taking this outside option incurs a haircut ξ on the bond return, which is rebated in lump-sum fashion to households. The precise outside-option payoff and the role of ξ are set out below when we characterize the optimization problem of defaulting firms.

Whereas the idiosyncratic productivity of firms is not directly observed by households, we assume that firms have a bit more information. While still far from perfect, a screening technology allows lending firms to tell whether borrowers can be expected to make more than the outside option by producing. This technology prevents firms with the lowest levels of production efficiency from borrowing and defaulting, thereby supporting a mass of lenders among firms with $\omega \leq \overline{\omega}_t$.[11] Firms whose private productivity ω falls between $\overline{\omega}_t$ and $\overline{\overline{\omega}}_t$ find it advantageous to divert all available funds towards the outside option. Only firms that can

[9]The inter-firm loan rate ρ_t is pinned in equilibrium by the zero-profit condition of lending firms, equation (32): lenders must be indifferent between lending at rate ρ_t and taking the outside option themselves. In other words, ρ_t is not an exogenous constant but an endogenous price determined jointly with the two cutoffs by the equilibrium conditions characterized in Section 5.1.

[10]In Section C.5 of the online appendix, we show that, as long as $\mu'(\overline{\omega}_t) \neq 0$ (a condition satisfied for our baseline calibration), the integral in equation (6) exists in the limiting case in which $\overline{\overline{\omega}}_t$ tends to $\overline{\omega}_t$.

[11]A fully accurate screening technology would deliver the first-best allocation with only the most productive firm producing. The imperfect screening we assume guarantees instead that low-efficiency firms choose to lend rather than switch to borrowing and defaulting. Formally, the cutoff $\overline{\omega}_t$ is not set directly by the screening technology but is determined endogenously by the equilibrium indifference condition (31) and the zero-profit condition (32); the screening technology's role is to support this equilibrium by preventing deviations in which low-ω firms would mimic borrowers.

make higher returns by producing will produce.

With constant returns to scale production, the economy would attain the first-best allocation if the most efficient firm, the one with $\omega = 1$, could borrow all financing from other firms, allowing it to be the only firm to produce. This allocation is hindered by the presence of moral hazard due to asymmetric information. By limiting the borrowing capacity of the most efficient firm, moral hazard gives less efficient firms room to borrow from financial intermediaries. Asymmetric information refers to uncertainty about the quality of borrowers. The productivity level above $\overline{\omega}_t$ is private information that is not observed by other firms. Producing firms may seek more funds by offering to borrow above the prevailing lending rate. However this offer lacks credibility as all firms above a certain threshold could claim high productivity. The result is credit rationing, as in Stiglitz and Weiss (1981). Consequently, all inter-firm financial contracts are identical and do not depend on ω in equilibrium.

In the next three sections, we will consider three distinct optimization problems spanning the possible combinations of actions that firms can take. Firms can borrow from other firms and produce, borrow and default, or lend. We will show that firms sort themselves in each group depending on their idiosyncratic productivity, ω. For now, take this result as a hypothesis. The proof of the hypothesis is set up in Section 5 with details pushed to Section A of the online appendix.

4.2.1 Firms Choosing to Produce ($\omega \geq \overline{\overline{\omega}}_t$)

Firms acquire the capital stock required for production in period $t + 1$, by combining funds from households $a_t(\omega)$ with a loan $b_t(\omega)$ from the mass of lending firms; consequently,

$$b_t^{tot}(\omega) = a_t(\omega) + b_t(\omega), \tag{7}$$

where $b_t^{tot}(\omega)$ denotes the total borrowing of the firm. By contractual agreement, the funds borrowed can only be used to purchase capital; so,

$$b_t^{tot}(\omega) = Q_t k_t(\omega), \tag{8}$$

where Q_t is the price of capital in terms of consumption. Let $R_{t+1}(\omega)$ denote the rate of return on capital ownership and $\pi_{t+1}(\omega)$ describe the profit of the firm. The revenue of the firm includes the proceeds from the sale of output as well as from the sale of the undepreciated fraction of capital. The expenses encompass commitments associated with loan servicing as well as remuneration for labor services. Thus,

$$\pi_{t+1}(\omega) = z_{t+1}\omega k_t(\omega)^{\alpha} h_{t+1}(\omega)^{1-\alpha} + (1-\delta)Q_{t+1}k_t(\omega) - R_{t+1}(\omega)b_t^{tot}(\omega) - W_{t+1}h_{t+1}(\omega), \quad (9)$$

where δ is the depreciation rate.

At time t, the problem of this producer of goods is to choose $b_t^{tot}(\omega)$, $k_t(\omega)$, and $h_{t+1}(\omega)$ to maximize expected profits in period $t+1$. Both capital and labor are committed in period t, before the firm-specific productivity ω is realized: capital is predetermined through credit rationing, and labor is hired one period in advance, reflecting the time required to contract, train, and match workers to firms. This maximization problem can be expressed as:

$$\max_{k_t(\omega),\, h_{t+1}(\omega),\, b_t^{tot}(\omega)} E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \pi_{t+1}(\omega) \right\}, \quad (10)$$

where $\beta \frac{\lambda_{ct+1}}{\lambda_{ct}}$ is the stochastic discount factor coming from the household's problem. This maximization is subject to (8).

Because both capital and labor are committed before ω is observed, and all producing firms are ex-ante identical, the optimal labor allocation is uniform across producing firms,

$$h_{t+1}(\omega) = \tilde{h}_{t+1}, \quad (11)$$

so the labor–capital ratio $h_{t+1}(\omega)/k_t(\omega) = \tilde{h}_{t+1}/\tilde{k}_t$ is equalized across producing firms (see Section C.3 of the online appendix). Aggregating over the producing segment yields the familiar aggregate labor-demand condition

$$W_t = (1-\alpha)\,\frac{Y_t}{H_t}. \quad (12)$$

Using zero-profit pricing, the return on capital earned by a firm with productivity ω is

$$R_{t+1}(\omega) = \frac{\kappa_{t+1}}{Q_t}\,\omega + \frac{(1-\delta)}{Q_t}Q_{t+1}, \tag{13}$$

where $\kappa_{t+1} \equiv \alpha Z_{t+1}\Big(\dfrac{H_{t+1}}{K_t}\Big)^{1-\alpha}$ is a common factor across all producing firms. Because ω is strictly increasing, $R_{t+1}(\omega)$ is strictly increasing in ω: more productive firms earn higher returns. The factor $\frac{1}{Q_t}$ converts one unit of consumption into $\frac{1}{Q_t}$ units of capital; the term $\kappa_{t+1}\,\omega$ is the marginal product of capital at productivity ω; and $(1-\delta)Q_{t+1}/Q_t$ captures the capital-gain component.

Upon observing ω, the producing firm determines the demand for inter-firm loans by solving the following problem:

$$\max_{b_t(\omega)|\omega} E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\Big(R_{t+1}(\omega)\left(a_t(\omega)+b_t(\omega)\right) - R^A_{t+1}(\omega)a_t(\omega) - \rho_t b_t(\omega)\Big)\right]. \tag{14}$$

It generates returns from owning capital by utilizing borrowed funds and distributes these returns to meet equity commitments of households and repay loans from other firms.

Note that, given constant returns to scale, a firm choosing to produce would not borrow if the cost of inter-firm funds exceeded the expected return from production. A sufficient condition for the corner solution in which such a firm borrows as much as possible is therefore $\rho_t \leq E_t R_{t+1}(\omega)$ for all $\omega \geq \overline{\overline{\omega}}_t$; this inequality is verified in equilibrium for the calibrated model. Accordingly, supply conditions will have to determine how much these firms can borrow.

We use the zero-profit condition that holds under every state of nature to size the return paid to households by firms in this segment, i.e.

$$R^A_{t+1}(\omega) = \frac{R_{t+1}(\omega)\left(a_t(\omega)+b_t(\omega)\right) - \rho_t b_t(\omega)}{a_t(\omega)}. \tag{15}$$

4.2.2 Firms Choosing to Borrow and Default ($\overline{\omega}_t < \omega < \overline{\overline{\omega}}_t$)

The problem of the firm that diverts the borrowed funds by choosing the outside option can be described as follows:

$$\max_{b_t(\omega)|\omega} E_t \left[\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left(\left(R_t^B - \xi \right) \left(a_t(\omega) + \Theta_t(\omega) b_t(\omega) \right) - R_{t+1}^A(\omega) a_t(\omega) \right) \right]. \tag{16}$$

The firm earns returns by investing borrowed funds in the government bond and uses these to pay the returns on equity to households. When taking the outside option, a haircut ξ is applied to the yield R_t^B on the government bond. The term $\Theta_t(\omega)$ reflects that only a fraction $\Theta_t(\omega)$ of the funds borrowed in the inter-firm market can be retained by the borrower when diverting the funds. Notice that this firm will borrow as much as possible, just as long as government bonds minus a haircut ξ pay a positive return. Therefore, similar to the previous situation, the supply conditions dictate the borrowing capacity of the firms in this segment.

We use the zero-profit condition that holds under every state of nature to size the return paid to households by firms in this segment, i.e.

$$R_{t+1}^A(\omega) = \frac{\left(R_t^B - \xi \right) \left(a_t(\omega) + \Theta_t(\omega) b_t(\omega) \right)}{a_t(\omega)}. \tag{17}$$

4.2.3 Firms Choosing to Lend ($\omega \leq \overline{\omega}_t$)

The problem of the firm that lends in the inter-firm market collapses to

$$\max_{l_t(\omega)|\omega} E_t \left[\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left(\rho_t \frac{1 - \mu(\overline{\overline{\omega}}_t)}{1 - \mu(\overline{\omega}_t)} l_t(\omega) + \frac{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)}{1 - \mu(\overline{\omega}_t)} (1 - \theta) l_t(\omega) - R_{t+1}^A(\omega) a_t(\omega) \right) \right], \tag{18}$$

subject to the constraint that $l_t(\omega) \leq a_t(\omega)$.

The term $\frac{1 - \mu(\overline{\overline{\omega}}_t)}{1 - \mu(\overline{\omega}_t)}$ represents the share of loans to firms that choose to produce, yielding the return ρ_t. The term $\frac{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)}{1 - \mu(\overline{\omega}_t)}$ represents the share of loans to firms that choose to divert the borrowed funds, yielding the average fraction that can be recovered, $1 - \theta$.[12] The firm pays the household the return on assets $R_{t+1}^A(\omega)$.

[12] Proposition 2 establishes that $\overline{\overline{\omega}}_t \geq \overline{\omega}_t$, ensuring that the share of loans to firms that choose to produce and the share of loans to firms that choose to divert the borrowed funds are non-negative and not greater than 1.

Notice that the revenues of the firm are increasing in the amount of loans it supplies. Thus, the budget constraint will hold with equality:

$$l_t(\omega) = a_t(\omega). \tag{19}$$

We use the zero-profit condition that holds under every state of nature to size the return paid to households by firms in this segment, i.e.

$$R^A_{t+1}(\omega) = \rho_t \frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)} + \frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}(1-\theta). \tag{20}$$

4.3 Capital-Producing Firms

In period t, competitive capital-producing firms buy capital from the goods-producing firms, repair depreciated capital and build new capital. They sell both the new and refurbished capital next period.

Let I^g_t denote aggregate gross investment expenditures. We introduce quadratic adjustment costs measured in units of investment such that the supply of investment goods is given by:

$$I^n_t = Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I^g_t}{I^g_{t-1}} - 1\right)^2\right] I^g_t. \tag{21}$$

The parameter ϕ governs investment adjustment costs for current production relative to past production. The term Z_{it} is a shock process driven by

$$\ln(Z_{It}) = \rho_I \ln(Z_{It-1}) + \varepsilon_{It}, \qquad \varepsilon_{It} \sim \mathcal{N}(0, \sigma^2_I). \tag{22}$$

The aggregate capital stock evolves according to:

$$K_t = I^n_t + (1-\delta)K_{t-1}, \tag{23}$$

where K_t is the amount of capital allocated to the goods-producing firms.

The capital producing firms are owned by households, and solve the problem

$$\max_{I^g_{t+i}} E_t \sum_{i=0}^{\infty} \beta^i \frac{\lambda_{ct+i}}{\lambda_{ct}} \left\{ Q_{t+i} Z_{It+i} \left[1 - \frac{\phi}{2} \left(\frac{I^g_{t+i}}{I^g_{t+i-1}} - 1 \right)^2 \right] I^g_{t+i} - I^g_{t+i} \right\}. \tag{24}$$

The first-order condition, derived in Section C.2 of the online appendix (equation (C.5)), determines the equilibrium price of capital Q_t via a standard Tobin's q condition.

Competition will ensure zero profits for the goods-producing firms. Accordingly, the profits rebated to households as aliquot shares will be given by

$$\Pi_t = Q_t Z_{It} \left[1 - \frac{\phi}{2} \left(\frac{I^g_t}{I^g_{t-1}} - 1 \right)^2 \right] I^g_t - I^g_t. \tag{25}$$

4.4 The Government

The government finances its transfers T_t by issuing government bonds B^G_t to balance its budget period by period:

$$T_t = B^G_t - R^B_{t-1} B^G_{t-1}. \tag{26}$$

The household and firms can buy government bonds, so

$$B^G_t = B^H_t + D_t, \tag{27}$$

where D_t denotes firms' holdings of government bonds. We assume that the government does not make bonds available to households but sells them only to firms, so that

$$B^H_t = 0 \tag{28}$$

and price the government bond using the households' stochastic discount factor. Although firms hold the bonds in equilibrium, the household's Euler equation for bonds (equation (C.4) in Section C.1 of the online appendix) still pins down the risk-free rate R^B_t because households are the ultimate owners of all assets in this economy and must be indifferent at the margin.

It implies that the government budget constraint can be written as:

$$T_t = D_t - R^B_{t-1} D_{t-1}. \tag{29}$$

Since firms that take the outside option are subject to a haircut cost ξ on their investment in government bonds in the previous period, the amount of transfers rebated to the household to ensure that there are no deadweight losses in the economy is equal to

$$\Xi_t = \xi D_{t-1}. \tag{30}$$

5 Analytical Characterization of the Equilibrium

This section characterizes the equilibrium analytically and derives its key implications for aggregate output and measured TFP. Section 5.1 provides conditions that pin down the loan rate ρ_t and two cutoff points $\overline{\omega}_t$ and $\overline{\overline{\omega}}_t$. These cutoff points sort firms into three segments depending on the realization of the idiosyncratic productivity shock: 1) firms with $\omega \leq \overline{\omega}_t$ lend; 2) firms with $\overline{\omega}_t < \omega < \overline{\overline{\omega}}_t$ borrow and default; 3) firms with $\omega \geq \overline{\overline{\omega}}_t$ borrow and produce. Our strategy is to posit a solution that satisfies the first-order conditions of the intermediate goods sector and then verify that no firm can or has an incentive to switch to a different segment. To this purpose, we formulate three propositions. The aggregation of individual firm decisions and the formal definition of competitive equilibrium are left for Section C.3 of the online appendix. Section 5.2 then integrates individual firm outputs across the producing segment to obtain aggregate output, and shows how the equilibrium cutoff $\overline{\overline{\omega}}_t$ generates an endogenous component of the measured Solow residual.

5.1 The Loan Rate and Cutoff Points

Here we clarify the importance of the screening technology that allows lenders to tell whether borrowers can be expected to make more than the outside option. This technology prevents the firms with the lowest firm-specific technology levels from borrowing.

Assume (and then verify) that if the firms that have a firm-specific productivity $\omega < \overline{\omega}_t$

will choose to lend, then the following condition holds:

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{\kappa_{t+1}}{Q_t}\,\overline{\omega}_t+\frac{(1-\delta)}{Q_t}Q_{t+1}\right)\right]=E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\right]. \tag{31}$$

This condition imposes that for a firm with productivity right at the cutoff point between borrowing and lending, $\omega=\overline{\omega}_t$, the expected return from producing $R_{t+1}(\overline{\omega}_t)$ equals the expected return of the outside option $R_t^B-\xi$.

Additionally, the following condition applies to firms that lend. We need to compare the expected return from lending to the expected return from diverting funds. For a firm whose productivity lies exactly at the cutoff between borrowing and lending, $\omega=\overline{\omega}_t$, the expected return from lending, i.e., the lending rate net of the losses expected from firms diverting funds, matches the expected return from borrowing and defaulting:

$$\begin{aligned}&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\rho_t\frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)}+(1-\theta)\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}\right)a_t\right]=\\&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\left(a_t+\Theta_t(\overline{\omega}_t)b_t\right)\right].\end{aligned} \tag{32}$$

Notice there is no dependence of b_t on ω because $b_t=b_t(\omega)$ for all ω due to asymmetric information. At this point, we have spelled out all the elements that are needed for our first proposition.

Proposition 1. *Given that $\Theta_t(\omega)$ is non-negative and increasing in ω, firms with idiosyncratic productivity ω less than the cutoff point $\overline{\omega}_t$ will have no incentive to deviate from lending.*

A proof of this proposition is offered in Section A of the online appendix. The proof relies primarily on equations (31) and (32).

For the next step of pinning down $\overline{\overline{\omega}}_t$, consider the following condition:

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\left(a_t+\Theta_t(\overline{\overline{\omega}}_t^*)b_t\right)\right]=E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\overline{\omega}}_t^*)(a_t+b_t)-\rho_t b_t\right)\right]. \tag{33}$$

It imposes that the marginal firm with productivity level $\overline{\overline{\omega}}_t^*$ will be indifferent between diverting funds and producing.

Taking the first derivative with respect to $\overline{\overline{\omega}}_t^*$ of both sides of equation (33), we will also need to impose the following condition on these derivatives:

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\overline{\overline{\omega}}_t^*)b_t\right]<E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}'(\overline{\overline{\omega}}_t^*)(a_t+b_t)\right)\right]. \tag{34}$$

In other words, for a marginal firm with the productivity level $\overline{\overline{\omega}}_t^*$ the expected returns from diverting funds grow less steeply than the expected returns from producing. This condition is really an additional condition on the choice of the function $\Theta_t(\omega)$. It implies that a firm with a slightly higher productivity than $\overline{\overline{\omega}}_t^*$ will prefer producing to diverting funds, while a firm with a slightly lower productivity than $\overline{\overline{\omega}}_t^*$ will prefer diverting funds to producing. Note that it is a local condition in the sense that it applies to a marginal firm with the productivity level $\overline{\overline{\omega}}_t^*$.

We define $\overline{\overline{\omega}}_t=\max(\overline{\omega}_t,\overline{\overline{\omega}}_t^*)$. This definition ensures that terms that represent shares in the expression of the return from lending, for instance, $\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}$ in equation (32), are well-defined. But we can do more and establish that $\overline{\overline{\omega}}_t^*$ is never less than $\overline{\omega}_t$, in line with the following proposition.

Proposition 2. *Take $\overline{\overline{\omega}}_t^*$ to be the productivity of a firm indifferent between diverting funds and producing. Given that $\Theta_t(\omega)$ is non-negative, then it must be that $\overline{\overline{\omega}}_t^*\geq\overline{\omega}_t$.*

Notice that, by definition of $\overline{\overline{\omega}}_t$, a corollary of Proposition 2 is that $\overline{\overline{\omega}}_t=\overline{\overline{\omega}}_t^*$, so $\overline{\overline{\omega}}_t\geq\overline{\omega}_t$.

Before moving to our final proposition, we also need to choose a function such that

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\overline{\omega}_t)b_t\right]<E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\frac{\kappa_{t+1}}{Q_t}(a_t+b_t)\right]. \tag{35}$$

The left-hand side is the maximum expected marginal gain from diversion, evaluated at the marginal firm $\omega=\overline{\omega}_t$ that is indifferent between diverting funds and lending. Because of the concavity of the function, $\Theta_t(\overline{\omega}_t)$ is at least as large as for any other firm with $\omega>\overline{\omega}_t$. The right-hand side can also be written as $E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}'(\omega)(a_t+b_t)\right]$, the expected marginal gain from producing; since $R_{t+1}(\omega)=(\kappa_{t+1}/Q_t)\,\omega+(1-\delta)Q_{t+1}/Q_t$, its derivative at ω equals κ_{t+1}/Q_t which does not depend on ω. Because the expected marginal benefit from diversion is decreasing in ω (as Θ_t is concave), if this global condition holds at $\omega=\overline{\omega}_t$, then it implies the local condition (34). We can use this global condition to prove that no diverting firm has an incentive to deviate by producing, and no producing firm has an incentive to switch to diverting. With this condition specified, we can state our final proposition.

Proposition 3. *Given that $\Theta_t(\omega)$ is non-negative, concave and increasing in ω, equations (31), (32), and (33), together with the slope condition under (35) are sufficient to ensure that, depending on the realization of their idiosyncratic productivity ω, firms sort themselves into three groups:*

1. *firms with $\omega \leq \overline{\omega}_t$ lend;*
2. *firms with $\overline{\omega}_t < \omega < \overline{\overline{\omega}}_t$ borrow and default;*
3. *firms with $\omega \geq \overline{\overline{\omega}}_t$ borrow and produce.*

A proof of this proposition is in Section A of the online appendix.

5.2 Aggregate Output and the TFP Decomposition

Once the propositions above pin down the equilibrium cutoffs, individual producing-firm outputs can be integrated across the mass of producing firms. The full derivation is in Section C of the online appendix. The key step combines two facts: (i) all producing firms hold the same capital $\tilde{k}_t$ (credit rationing); and (ii) because the labor commitment is also made before ω is observed, each firm's optimal labor satisfies $h_{t+1}(\omega) = \tilde{h}_{t+1}$ from (11). Individual output is therefore $y_{t+1}(\omega) = Z_{t+1}\,\omega\,\tilde{k}_t^{\alpha}\tilde{h}_{t+1}^{1-\alpha}$, which is linear in ω. Integrating across the producing segment yields:

$$Y_t = Z_t\,K_{t-1}^{\alpha}\,H_t^{1-\alpha}\,\frac{\int_{\overline{\overline{\omega}}_{t-1}}^{1} \omega\,\mu'(\omega)\,\mathrm{d}\omega}{1-\mu(\overline{\overline{\omega}}_{t-1})}. \tag{36}$$

This equation reveals the central insight: measured TFP—the Solow residual defined by $Y_t/(K_{t-1}^{\alpha}H_t^{1-\alpha})$—equals the exogenous shock Z_t multiplied by $E[\omega \mid \omega \geq \overline{\overline{\omega}}_{t-1}]$, the conditional mean of idiosyncratic productivity among producing firms. Because the cutoff $\overline{\overline{\omega}}_{t-1}$ responds to financial conditions, the Solow residual contains an endogenous component even when Z_t is held fixed: a tightening that raises $\overline{\overline{\omega}}_{t-1}$ selects more productive firms into operation, lifting $E[\omega \mid \omega \geq \overline{\overline{\omega}}_{t-1}]$. Evaluating the integral in closed form requires a distributional assumption on μ.

5.2.1 Functional Form for the Distribution of Idiosyncratic Productivity

We choose the Beta distribution to govern draws of the idiosyncratic productivity $\omega \sim \text{Beta}(\eta_1, \eta_2)$ on $[0, 1]$. With shape parameters η_1 and η_2, the probability density function of ω is

$$\mu'_{\eta_1,\eta_2}(\omega) = \frac{\omega^{\eta_1-1}(1-\omega)^{\eta_2-1}}{B(\eta_1, \eta_2)}, \tag{37}$$

where $B(\eta_1, \eta_2) = \Gamma(\eta_1)\Gamma(\eta_2)/\Gamma(\eta_1 + \eta_2)$ and Γ is the Gamma function.

We also need to specify the diversion function $\Theta_t(\omega)$, which governs the fraction of borrowed funds $b_t(\omega)$ that a firm can divert. To align our functional choice with the structural justification for this function derived in Section B of the online appendix, this function must be non-negative, increasing, and concave in ω. Among the many functional forms consistent with these requirements, we choose the power specification

$$\Theta_t(\omega) = \omega^{\psi} F_t, \tag{38}$$

where ψ is a parameter and F_t is determined endogenously to satisfy

$$\theta = \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \omega^{\psi} F_t \frac{\mu'(\omega)}{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)} \mathrm{d}\omega. \tag{39}$$

Section C.9 of the online appendix provides the functional form of F_t.[13] We have verified numerically that the choices of ψ and other parameters that influence F_t satisfy the slope condition (35), which is used in the proof that ω determines each firm's segment. [14]

[13]Section C.9 of the online appendix also shows that $F_t = \dfrac{\theta}{\overline{\omega}_t^{\psi}}$ in the limiting case as $\overline{\overline{\omega}}_t \to \overline{\omega}_t$.

[14]The equation for F_t makes $\Theta(\omega)$ depend on the equilibrium cutoffs, so a firm with a given ω can have a different diversion share when the defaulting interval changes. Individual firms do not internalize this general-equilibrium effect. For the model developed in Section B of the online appendix, we show that a general equilibrium mechanism compresses the diversion rate when more firms attempt to divert resources. Numerically, we verify that the chosen diversion function shares the same property in the steady state of the

5.2.2 Closed-form Decomposition of TFP into Endogenous and Exogenous Components

With the Beta distribution, the integral in equation (36) evaluates in closed form. The key step rewrites $\omega \cdot \mu'_{\eta_1,\eta_2}(\omega) \propto \omega^{\eta_1}(1-\omega)^{\eta_2-1}$ as a scalar multiple of the Beta(η_1+1, η_2) density kernel, converting the integral via the ratio of Beta functions $B(\eta_1+1, \eta_2)/B(\eta_1, \eta_2) = \eta_1/(\eta_1+\eta_2)$. Define

$$c_\alpha \;\equiv\; \frac{B(\eta_1+1,\, \eta_2)}{B(\eta_1, \eta_2)} \;=\; \frac{\eta_1}{\eta_1+\eta_2}, \tag{40}$$

which is the unconditional mean of $\omega \sim \text{Beta}(\eta_1, \eta_2)$. The full derivation is provided in Section C of the online appendix, equation (C.53). Aggregate output then simplifies to

$$Y_t = \Phi_t \cdot Z_t \cdot K_{t-1}^{\alpha} H_t^{1-\alpha}, \tag{41}$$

where the endogenous TFP prefactor is

$$\Phi_t \;\equiv\; c_\alpha \, \frac{1-\mu_{\eta_1+1,\,\eta_2}(\overline{\overline{\omega}}_{t-1})}{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_{t-1})}. \tag{42}$$

The right-hand side equals $E[\omega \mid \omega \geq \overline{\overline{\omega}}_{t-1}]$, the conditional expectation of ω among producing firms. We retain the notation c_α for consistency with the appendix; here it collapses to the unconditional Beta mean $\eta_1/(\eta_1+\eta_2)$.

Equations (41)–(42) give the TFP decomposition a precise quantitative form. The measured Solow residual is

$$\text{TFP}_t \;\equiv\; \frac{Y_t}{K_{t-1}^{\alpha} H_t^{1-\alpha}} \;=\; \underbrace{\Phi_t}_{\text{endogenous}} \cdot \underbrace{Z_t}_{\text{exogenous}}.$$

As the economy approaches the frictionless RBC limit, $\overline{\overline{\omega}}_t \to 1$ for all t; consequently $\Phi_t \to 1$ (see equation F.28 in Section F of the online appendix), and the Solow residual moves one-for-one with Z_t. In our model, $\overline{\overline{\omega}}_{t-1}$ responds endogenously to financial conditions, causing Φ_t to fluctuate. A tightening that raises the default cutoff concentrates production among higher-productivity firms and raises Φ_t; a loosening has the opposite effect. The Solow residual

model.

therefore inherits variation from both the exogenous technology process and the endogenous reallocation of production across firms. Section 7.1 and the quantitative analysis document the empirical importance of this channel.

6 Parameter Choices and Model Solution

We now turn to the quantitative implementation of the model. This section is organized as follows. Section 6.1 describes the choice of parameter values, combining steady-state calibration targets with a simulated method of moments estimation of the shock processes and adjustment-cost parameter. Section 6.2 outlines the second-order perturbation solution method. Section 6.3 reports the targeted moments to assess the overall fit of the model. We leave a discussion of untargeted moments for Section 7.3.

6.1 Parameters

We choose model parameters with a mix of estimation and calibration. Specifically, we estimate model-specific parameters using the simulated method of moments (SMM) on second moments. We pin down other parameters by matching first moments. Finally, we set a few parameters to standard values from the literature.

We use the simulated method of moments (SMM) to size the shock processes, total factor productivity, Z_t, investment-specific technology, Z_{It}, and the consumption shock ν_{ct}. We model each shock as an AR(1) process. We size the persistence parameters and standard deviations of the innovations. Along with sizing these shocks, we also estimate by SMM the investment adjustment cost parameter, ϕ.

For the implementation of the SMM estimation, we minimize a quadratic objective function based on variances, correlations, and autocorrelations of real GDP, real consumption, the relative price of investment, and the average delinquency rate on business loans at commercial banks.[15]

[15] For GDP and relative prices we used data from the NIPA Release of the U.S. Bureau of Economic Analysis. We use chain-type indexes. The relative price of investment is the ratio of the price index for gross private domestic investment to the price index for personal consumption excluding food and energy. The delinquency rate on business loans at commercial banks is from the data release Charge-Off and Delinquency

Our data run from the first quarter of 1987 through the fourth quarter of 2021, matching the sample period in Section 3, a choice dictated by the availability of data on the dispersion of productivity. The data moments for the SMM exercise are computed after HP-filtering the data (using a value for the smoothing parameter of 1,600 as standard for quarterly data). We take the same approach for observed and simulated data. For the SMM objective function, we employ a modified optimal weighting matrix with model moments from the average of 10 simulated samples of the same length as the observed data.[16]

We choose selected parameters to match steady-state targets for key variables, using first moments from the same sample period as the SMM estimation. We size the two parameters of the Beta distribution η_1 and η_2, the haircut ξ, and the diversion share θ at values that jointly match the 1) within-industry dispersion of productivity as captured by the weighted average gap in TFP between the 90^{th} and 10^{th} percentile firm, 2) the average spread between the return on assets for households and the risk-free rate of the outside option of firms, 3) the average share of bank credit, and 4) the average delinquency rate on business loans.[17] The idea behind these choices of the parameters is the following. The shape of the productivity distribution is important to match that the establishment at the 90^{th} percentile is 2.08 as productive as the one at the 10^{th} percentile. We compute the target ratio of weighted-average within-industry dispersion over the sample period by taking the mean of the data points underlying the solid line in the top panel of Figure 1. Section E of the online appendix provides formulas to calculate this target ratio from the model. We map lending firms into stylized banks, which provide about 47% of private credit in the United States.[18] We match

Rates on Loans and Leases at Commercial Banks of the Board of Governors of the Federal Reserve System. The series for real GDP and delinquency rates are the same as those used in Section 3. In our simple model, there is no distinction between delinquency rates, charge-off rates, and the outright default rate of borrowers.

[16]The standard optimal weighting matrix assigns a weight proportional to the inverse of the long-run variance of each sample moment. The variance of consumption is estimated very tightly from the data; so, the SMM optimal weighting matrix assigns a disproportionate weight to that moment to the detriment of pinning down other important moments —the variance of the investment price, the variance of the default rate, and the correlation between GDP and the default rate. To prevent this single-moment dominance, we increase the weight on the GDP–default-rate correlation to the full weight of the consumption variance, while while the other two moments receive half the weight of the consumption variance.

[17]The shape of the productivity distribution affects all the described targets, so the word "*jointly*" follows.

[18]We construct our measure using data from the Z.1 Financial Accounts of the United States. We sum non-financial corporate and non-financial non-corporate business loans and subtract non-financial business other loans and advances to form the numerator. We divide by total credit to the non-financial private sector from the Bank for International Settlements and take the average. We note that the BIS denominator covers a broader set of borrowers than the non-financial non-corporate sector included in the Z.1 numerator, so our

this share to our model counterpart expression $\frac{L_t}{B_t^{tot}}$ evaluated in the steady state. Because all external funds in the model are intermediated through lending firms, $\frac{L_t}{B_t^{tot}}$ is the model analogue of the share of firm financing provided through the banking system. We use the same data on delinquent loans as in Section 3 to size the average delinquency rate at 2.6%. We map it to χ_t, the share of borrowing firms diverting funds,[19]

$$\chi_t \equiv \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)}, \tag{43}$$

and calibrate θ so that its steady-state value matches the 2.6% target. We match the average spread between the bank prime loan rate and a one-year Treasury bill to the haircut parameter ξ. At the deterministic steady state, $R^A = R^B$ from the household Euler equations, so the spread between the return to assets for households and the outside-option return $R^B - \xi$ equals exactly ξ; equivalently, ξ is the model counterpart of the observed prime-rate–Treasury-bill spread. The prime rate is a good match for R_t^A, since it is an unsecured (defaultable) rate.[20]

Before discussing SMM-estimated parameters, we first outline the calibration of the remaining parameters. Calibrated values appear at the top of Table 2 and follow guidance from prior literature: the discount rate $\beta = 0.9925$, the capital share $\alpha = 0.3$, the depreciation rate $\delta = 0.01$, the inverse Frisch elasticity of labor supply $\nu = 2$. Under KPR preferences, the utility function features log utility in consumption, so $\sigma = 1$ is imposed implicitly and is not a free parameter.

We set ϑ at 0.84, a value that ensures that aggregate labor takes a value of 1 in the steady state.

measure should be interpreted as an approximation.

[19] Our model does not distinguish between delinquency, charge-offs, and outright default: a diverting firm fails to repay in full, which we treat as a single default event. We therefore take the observed delinquency rate on business loans as a proxy for the model's default rate χ_t.

[20] The data for the bank prime loan rate and the return on one-year Treasury bills are from the Federal Reserve Board's H.15 Release.

6.2 Model Solution

We solve the model with an exact second-order perturbation method using the pruning algorithm of Kim et al. (2008). When drawing data samples from the model, we found that a first-order perturbation method was not sufficiently accurate to avoid the upper cutoff point $\overline{\overline{\omega}}_t$ falling below the lower cutoff point $\overline{\omega}_t$, a theoretical impossibility. By contrast, a second-order perturbation solution is sufficiently accurate to avoid this problem. We provide further accuracy checks in Section H of the online appendix.

6.3 Simulated Method of Moments, Results

Table 3 reports confidence intervals for both the data moments and the model moments alongside the point estimates. The model matches most of the targeted second moments well: the data and model confidence intervals overlap for 10 of the 14 moments. The four variances are all matched within sampling uncertainty, as are the correlations involving GDP, consumption, and the default rate, and the autocorrelations of GDP and consumption. For four moments—the correlation between GDP and the investment price, the correlation between the investment price and the default rate, and the first-order autocorrelations of the investment price and the default rate—the confidence intervals do not overlap, indicating a quantitatively significant discrepancy. Nonetheless, for every moment in the table, including these four, the model correctly matches the sign.

Table 4 reports the relative contributions of the shocks in our calibration to fluctuations in variables targeted by the SMM procedure and other key variables. Aggregate TFP shocks are the primary drivers of fluctuations in GDP, but we find that consumption shocks are the most important driver of TFP dispersion. The table also shows that IST shocks are less important for our model to hit the moments targeted in the SMM estimation procedure.

7 Dynamic Responses to Shocks and Model Assessment

Before turning to an assessment of the model by checking untargeted moments, Section 7.1 discusses the effects of an aggregate TFP shock, and Section 7.2 considers consumption and investment specific shocks. We compare our model to a frictionless RBC model, a special case of our model in which the most efficient firm is the only one to produce and attracts all household savings.[21] Whereas the first two parts of this section focus on moments conditional on specific shocks, the last part turns to unconditional moments. We compare moments untargeted by our calibration using observed and model synthetic data.

7.1 Effects of a TFP Shock

Figure 2 shows the responses of key aggregate variables to a one-standard-deviation expansionary TFP shock, ε_t^z. The size of the shock process and its persistence follow the estimates from Section 6. In each panel, the solid lines show responses from the baseline model with financial frictions and endogenous TFP dispersion. The dashed lines show responses in a frictionless RBC model, a special case of our model in which all production is carried out by the most efficient firm.

On impact (period 1 in the figure), the positive productivity shock increases the marginal product of capital, which in turn raises the average return on capital. It is this standard exogenous productivity that pushes up aggregate output on impact. An endogenous TFP channel kicks in from the second period.

Since the shock is autocorrelated but decaying ($\rho_z < 1$), the expected marginal product of capital for the next period is lower relative to the impact response. In line with this effect, the demand for capital after initially expanding starts declining. This expected decline brings down the expected price of capital. Accordingly, the average return to capital falls. This return is averaged across firms with different idiosyncratic productivity, and it includes a term that reflects the marginal product of capital and another term that captures gains or

[21] Section F of the online appendix proves that when firms can credibly disclose their productivity and $\xi = 0$, our model collapses to a standard RBC model.

losses from the resale of undepreciated capital. It is this second term that pushes down the average return on capital.

As a reminder, Equation (33) pins down $\overline{\overline{\omega}}_t$.[22] From this equation, we can see that, all else equal, a decrease in the average return of capital pushes up $\overline{\overline{\omega}}_t$. While many of the variables entering equation (33) do change, the decrease in the average return of capital is the most important quantitatively. Intuitively, more productive firms are better able to withstand a decrease in the average return to capital. As $\overline{\overline{\omega}}_t$ rises, the productivity dispersion declines, as captured by a reduction in the gap between the productivity of the 90^{th} and 10^{th} percentile firms.

As the aggregate technology shock is expected to gradually wane, the expected return to producing falls for the marginal firm—chiefly because the expected resale price of capital declines, the same force that lowers the average return on capital and raises $\overline{\overline{\omega}}_t$. The risk-free outside-option return R_t^B declines as well, but by less, so the firm indifferent between lending and borrowing tips toward lending and the lower cutoff $\overline{\omega}_t$ rises.

Notice that as $\overline{\overline{\omega}}_t$ does not rise quite as much as $\overline{\omega}_t$, the default rate χ_t defined in equation (43) goes down. Accordingly, default rates move countercyclically. More generally, expansionary shocks can lead to an increase or a decrease in the default rate depending on whether the effects on the expected returns dominate or the effects through the outside option dominate.

It is useful to keep the structural roles of the two cutoffs distinct, in anticipation of the comparison developed in Section 8. The upper cutoff $\overline{\overline{\omega}}_{t-1}$ enters aggregate production directly: it is the threshold above which firms choose to produce and repay; firms with $\overline{\omega}_{t-1} < \omega < \overline{\overline{\omega}}_{t-1}$ borrow and default, so only firms with $\omega \geq \overline{\overline{\omega}}_{t-1}$ contribute to output in period t. It is therefore $\overline{\overline{\omega}}_{t-1}$ that governs the endogenous prefactor Φ_t and shapes measured TFP. The lower cutoff $\overline{\omega}_t$, by contrast, is purely forward-looking: it marks the boundary between lenders and all borrowers—producers and defaulters alike are on the borrowing side of this threshold—and responds to current and expected-future conditions, but it does not appear directly in the aggregate production function. The two cutoffs co-move in equilibrium, yet their separate determination allows the production-relevant margin ($\overline{\overline{\omega}}$) and the lending-

[22] The aggregate counterpart of equation (33) is given in Section C of the online appendix.

incentive margin ($\overline{\omega}$) to absorb shocks with some degree of independence. This independence is precisely what is lost in the no-default model, where a single cutoff must play both roles simultaneously.

Due to a fall in the productivity dispersion, investment rises more than in a frictionless RBC model. Total output also increases by more. On impact, consumption rises relatively less compared to a frictionless RBC model as consumption is crowded out by more productive investments. Subsequently, starting from the next period, consumption rises relatively more due to higher total output.

Summing up, financial frictions and firm dispersion lead to persistent output and interest rate dynamics.

7.2 Effects of Other Shocks

Figure 3 shows the responses of key aggregate variables to a one-standard-deviation shock to consumption preferences, $\varepsilon_{\nu t}$, leading to an expansion in consumption. Again, the size of the shock process and its persistence follow the estimates from Section 6.

The differences between our model and the canonical RBC model are more dramatic for this shock. Starting from the risk-free rate, its rise is supercharged relative to the rise in the RBC model. Two forces cumulate to produce this difference, a substitution and a wealth effect. By construction, the shock has a pronounced substitution effect, which results in a rise in the risk-free rate in both models. In a closed-economy setting which precludes the possibility of borrowing from abroad, the rise is necessary to push down investment leaving room to consume more today for any level of labor input. But in our model with financial frictions, this rise in the risk free rate also affects the outside option for defaulting firms. There is a rise in borrowing costs and an increase in defaults. The rise in borrowing costs also contributes to concentrating production in higher productivity firms—the ones more able to withstand increases in borrowing costs. As a result, the productivity dispersion falls. This fall in dispersion is the extra endogenous kick to productivity which is the main determinant of the difference between the evolution of output across the two models. But consider that this effect generates a wealth effect that boosts consumption by more, hence the need for a much larger rise in the risk-free interest rate in our model relative to the frictionless model.

Although the differences between the two models are quite persistent, we have confirmed that both models are stationary. The model with financial frictions does return to its steady state at horizons well beyond the one for the figure.

Figure 4 shows the responses of key aggregate variables to an expansionary one-standard-deviation shock to investment technology ε_{It}. And once again, the size of the shock process and its persistence follow the estimates from Section 6.

Investment-specific technology shocks generate strong substitution effects as the benefits of higher productivity can only be accrued by boosting investment. Accordingly very similar considerations apply as for consumption shocks. The real interest rate needs to rise to curb consumption in this case. Default rates rise, as the rise in borrowing costs is the dominant force, dispersion decreases, and total factor productivity rises endogenously, giving rise to a sizable wealth effect (relative to the size of the shock). As Section 8 makes clear, the strategic-default margin plays a dual role throughout: it acts as an amplifier by concentrating production among high-productivity firms, and simultaneously as a stabilizer that prevents the oscillatory TFP dynamics that emerge when a single cutoff must govern both the selection of producers and the lending incentive.

7.3 Untargeted Moments

Having traced the model's response to each shock, we now turn to a set of regression-based untargeted moments that ask whether the same conditional correlations apparent in the U.S. data also arise in the model.

Table 5 reports estimates from four ordinary least squares regressions on annual U.S. data from 1987 to 2021. In columns 1 and 2, the dependent variable is the weighted-average within-industry 90–10 TFP gap from Figure 1, expressed in percentage points.[23] Column 1 regresses that gap on the contemporaneous growth rate of real GDP per capita (in annual percent) and the delinquency rate on business loans at commercial banks (in annual percent); column 2 adds one lag of the growth rate.[24] In columns 3 and 4, the dependent variable is the

[23]This is the value-added-weighted 90–10 measure plotted by the solid line in the top panel of Figure 1, built from the BLS Dispersion Statistics on Productivity (Cunningham et al., 2023) together with industry value-added weights from the Bureau of Economic Analysis.

[24]Real GDP per capita is the ratio of real gross domestic product (in chained 2017 dollars) to mid-

delinquency rate itself, regressed on its own one-period lag, contemporaneous GDP per capita growth, and—in column 4—an ex-ante one-year real interest rate (in annual percent).[25] We are deliberately humble about what these regressions deliver: they are partial correlations, not causal estimates, and we use them as additional untargeted moments against which to judge the calibration.

The point estimates in columns 1 and 2 of Table 5 confirm the pattern visible in Figure 1. A one-percentage-point higher annual growth rate is associated with a 2.86 pp narrower 90–10 gap in column 1, and the lagged-growth coefficient in column 2 is also negative and significantly different from zero. Both columns also show a sizable negative coefficient on the delinquency rate (about -4.2 pp). For the results shown in columns 3 and 4, consider that the delinquency rate is persistent (lagged coefficient around 0.8), it falls when growth is high (by about 0.22 pp per percentage point of growth), and—once the real rate is added in column 4—it rises with the real rate at a coefficient of 0.14 pp. Read together, these last two columns simply summarise the empirical fact that delinquencies are higher when growth is depressed and when real rates are high.

To assess the model along these dimensions, we run exactly the same four regressions on simulated data. For each of 1,000 replications, we simulate the model at quarterly frequency over a span that matches the length of the observed sample and then aggregate to annual frequency by averaging within each year, mirroring the construction in the data. The model's analogue of the dispersion measure is the time series of $100 \cdot \omega_{90}/\omega_{10}$ built from the upper cutoff $\overline{\overline{\omega}}_t$; the analogue of the delinquency rate is the model default rate times one hundred; growth is the annual log-difference of yearly model GDP per capita, expressed in percent; and the analogue of the real rate is the net annualised return on bonds, $400 \cdot (R^B - 1)$, also in percent. Lags are constructed exactly as in the data and we estimate the four specifications

year population, both from the U.S. Bureau of Economic Analysis Gross Domestic Product release within the National Income and Product Accounts; the population concept covers resident population plus armed forces overseas. The delinquency rate on business loans at commercial banks comes from the Federal Reserve Board's Charge-Off and Delinquency Rates on Loans and Leases at Commercial Banks release. We aggregate the underlying quarterly delinquency series to annual frequency by averaging within the year.

[25]The ex-ante one-year real interest rate is from the Federal Reserve Bank of Cleveland's Inflation Expectations release; the underlying model uses Treasury yields, inflation data, inflation swaps, and survey-based measures of inflation expectations. The original monthly series is aggregated to annual frequency by averaging within the year.

by OLS. Table 6 reports the average slope coefficient across the 1,000 replications, with the 5^{th} and 95^{th} percentiles in brackets.

In the dispersion regressions (columns 1 and 2 of Table 6), the model produces a negative and significant coefficient on current GDP per capita growth and on its lag, matching the sign of the data estimates in Table 5. The contemporaneous delinquency coefficient is also negative in the model, and its point estimate (about -4.2) is essentially the same as in the data, although the model's 90% range is wide enough to include zero. The dynamics behind these signs are visible in the impulse responses of Section 7.1: in Figure 2, the same TFP shock that raises output narrows the 90–10 gap and lowers the default rate, so dispersion, growth, and delinquencies co-move in a way that produces negative partial correlations period by period. The wide range on the delinquency coefficient in columns 1 and 2 is a signal that the model captures the sign of this conditional correlation more sharply than its magnitude across small simulated samples.

The delinquency regressions (columns 3 and 4 of Table 6) provide a sanity check on a key contribution of the paper. The lagged delinquency coefficient is positive and significant, as in the data; the contemporaneous growth coefficient in column 3 is negative and significant (-0.11 in the model versus -0.22 in the data); and the real-interest-rate coefficient in column 4 is positive and tightly estimated (0.26 in the model versus 0.14 in the data). These signs line up with the dynamics shown in Sections 7.1 and 7.2: the consumption and investment-specific shocks of Figures 3 and 4 push the real rate up and the default rate up at the same time; the TFP shock of Figure 2 raises growth and pulls delinquencies down. The one cell that does not line up is the growth coefficient in column 4, which is essentially zero in the model: once the real rate enters the regression, it absorbs the cyclical variation in delinquencies that growth carries on its own in column 3.

Taken together, these untargeted regressions reproduce, in sign, four of the partial correlations that we read off the data, including the central one that delinquencies rise when growth is depressed and when real rates are high. Section 8 now asks how much of the endogenous TFP behaviour underlying these regressions depends specifically on the strategic-default margin.

8 Endogenous TFP Dynamics With and Without Strategic Default

To sharpen the intuition for how financial frictions shape measured TFP, we contrast the baseline model with a simplified variant in which strategic default is ruled out by assumption. The comparison between these two model variants allows us to showcase the important role of default in driving the endogenous component of TFP.

In the *no-default model*, the outside option is unavailable, so no firm has an incentive to divert borrowed funds; the two-cutoff structure of the baseline collapses to a single cutoff $\overline{\omega}_t$ that separates lenders from producers. All other model blocks—the household problem, capital-producing firms, and the shock processes—are unchanged. Because no funds are diverted, the government sector drops out entirely. Section G of the online appendix develops this variant in full detail.

Figure 5 presents a side-by-side comparison of the TFP decomposition for the two models across all three shocks keeping the model parameters unchanged across models. In each panel, the stacked areas display the two components of the Solow residual introduced in Section 5.2: the blue area is the exogenous component $\log Z_t$, the orange area is the endogenous prefactor $\log \Phi_t$, and the top of the stack traces total measured log TFP. The left column shows the baseline (two-cutoff) model; the right column shows the no-default (one-cutoff) model. The y-axis scale is shared within each row so that magnitudes are directly comparable across the two models.

For the TFP shock (top row), both models generate a positive TFP response, with the endogenous component amplifying the exogenous shock. The amplification is stronger in the baseline model: strategic default provides an additional margin through which the upper cutoff $\overline{\overline{\omega}}_t$ responds to aggregate conditions, concentrating production among high-productivity firms and lifting the TFP prefactor further than the single-cutoff model can. For the preference and IST shocks (middle and bottom rows), Z_t is exactly zero—these shocks do not move the exogenous productivity process—so the entire TFP response is endogenous and confined to Φ_t. The contrast between the two models is striking here: the baseline model generates a smooth, monotonically decaying Φ_t response, while the no-default model generates a pronounced oscillatory pattern that alternates sign period by period.

8.1 Results for the No-Default Model

Two features of the no-default model interact to produce oscillatory behavior. The first feature is the role of the cutoff as a lagged state variable in production. As shown in equation (G.20) of Appendix G, aggregate output in period t depends on the cutoff from the *previous* period:

$$Y_t = E[\, \omega \mid \omega \geq \overline{\omega}_{t-1}] \cdot Z_t \, K_{t-1}^{\alpha} \, H_t^{1-\alpha}. \tag{44}$$

The mechanism is straightforward. Because capital is allocated equally across all producing firms, the aggregate capital–labor ratio is identical for every producer. The only source of heterogeneity in individual output is idiosyncratic productivity ω; summing across producers, aggregate output is therefore proportional to the conditional mean of ω among firms with $\omega \geq \overline{\omega}_{t-1}$. This is the *selection effect*: when the cutoff rises, the producing cohort is leaner and average productivity rises; when the cutoff falls, lower-productivity firms enter the cohort and drag down average productivity.

The second feature is the forward-looking determination of $\overline{\omega}_t$. The indifference condition (G.9) equates the expected discounted return from lending today with the expected discounted return from producing, so the cutoff is a function of current and expected future variables, not of past conditions. This creates a fundamental tension: $\overline{\omega}_t$ responds on impact to current shocks, but its effect on production materialises only one period later, through equation (44).

To fix ideas, consider a positive TFP shock; analogous dynamics set in for the preference and IST shocks, where the initiating impulse comes from the demand side rather than from a direct shift in Z_t. The feedback loop can be traced period by period.

- *Period t.* The positive TFP shock is persistent, so it raises the *expected* marginal product of capital in period $t+1$. Because the cutoff $\overline{\omega}_t$ is forward-looking—it equates the expected discounted return from producing in $t+1$ with the outside-option return—this higher expected return makes production more attractive. The indifference condition therefore shifts so that $\overline{\omega}_t$ *falls*, admitting additional lower-productivity producers into the active cohort.

- *Period* $t+1$. The lower $\overline{\omega}_t$ reduces $E[\omega \mid \omega \geq \overline{\omega}_t]$, so effective TFP is *lower* than it would have been at the steady-state cutoff. Output falls relative to what a higher cutoff would have delivered; consumption falls with it; and the bond Euler equation then requires R^B_{t+1} to *rise* to restore household indifference between consuming today and tomorrow. The higher risk-free rate tightens the lending threshold, pushing $\overline{\omega}_{t+1}$ *up*.

- *Period* $t+2$. The higher $\overline{\omega}_{t+1}$ raises $E[\omega \mid \omega \geq \overline{\omega}_{t+1}]$; effective TFP is now *higher*. Output recovers; consumption rises; R^B_{t+2} *falls*; and $\overline{\omega}_{t+2}$ *falls* again.

- *Subsequent periods.* The cycle repeats with alternating sign, decaying at a rate of approximately 0.69 per period.

The root cause of the oscillation is a sign mismatch between the cutoff's forward-looking response and the lagged selection effect it triggers. The cutoff $\overline{\omega}_t$ and the TFP factor $E[\omega \mid \omega \geq \overline{\omega}_t]$ always move in the same direction—a higher cutoff selects a more productive cohort. But the cutoff's forward-looking response to a positive shock is to *fall* on impact, and a lower cutoff then depresses effective TFP in the following period. This delayed negative feedback is what generates the negative autoregressive coefficient in the law of motion for $\overline{\omega}$ and the corresponding oscillatory dynamics in the right column of Figure 5.

8.2 Comparison of the No-default and Baseline Models

In the baseline model, aggregate output depends on the *upper* cutoff $\overline{\overline{\omega}}_{t-1}$, not on $\overline{\omega}_{t-1}$. As explained in Appendix G, the forward-looking condition that pins down the lower cutoff is unchanged and given by equation (31), but $\overline{\omega}_t$ does not enter the aggregate output equation directly in the baseline model. The two cutoffs are linked through the share of defaulting firms $\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t)-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}$, but they respond with a degree of independence to aggregate shocks.

Two additional stabilizing forces are present in the baseline model that are absent in the no-default model. First, in the baseline model, when aggregate conditions change, the mass of defaulting firms can expand or contract almost immediately, absorbing much of the shock before it has the opportunity to feed back through the production cutoff. Second, the aggregate equity return equation in the baseline model contains diversion-correction terms—involving the diversion scaling factor F_{t-1}, the aggregate diversion D_{t-1}, and the diversion function evaluated at the upper cutoff $\Theta(\overline{\overline{\omega}}_{t-1})$—that create a positive feedback channel from past defaulter activity to current returns. This positive channel offsets the negative selection channel described above, with the net result that the dominant eigenvalue for $\overline{\overline{\omega}}_t$ remains positive. The oscillatory mode is suppressed, and the IRFs in the left column of Figure 5 decay smoothly toward zero.

8.3 How Much TFP Variance is Endogenous?

The impulse responses in Figure 5 are informative about the dynamic structure of the endogenous TFP component, but they are computed shock by shock. To obtain an unconditional, model-wide assessment we simulate each model and ask what fraction of the variance of the Solow residual is attributable to the exogenous process Z_t alone.

The calculation proceeds as follows. We draw 50 simulated samples of 500 quarters each from the second-order perturbation solution of each model, using all three shocks simultaneously. All series are HP-filtered with smoothing parameter $\lambda = 1{,}600$ before moments are computed. Because production in period t uses the cutoff determined in period $t-1$ as shown in equation (G.20), the measured Solow residual decomposes as

$$\log \mathrm{TFP}_t = \log \Phi_t + \log Z_t, \tag{45}$$

where Φ_t is the endogenous TFP prefactor from equation (42), which already uses the lagged cutoff $\overline{\overline{\omega}}_{t-1}$. The variance of $\log \mathrm{TFP}_t$ therefore satisfies

$$\mathrm{Var}(\log \mathrm{TFP}_t) = \mathrm{Var}(\log Z_t) + \mathrm{Var}(\log \Phi_t) + 2\,\mathrm{Cov}(\log \Phi_t,\, \log Z_t). \tag{46}$$

The headline statistic is the ratio $\mathrm{Var}(\log Z_t)/\mathrm{Var}(\log \mathrm{TFP}_t)$, which equals one in a frictionless RBC economy where Φ_t is constant, and falls below one whenever the endogenous component adds variance to the Solow residual. The gap relative to the canonical RBC model,

$$1 - \frac{\text{Var}(\log Z_t)}{\text{Var}(\log \text{TFP}_t)}, \tag{47}$$

measures the share of Solow TFP variance that cannot be accounted for by the exogenous process alone. Although this gap combines $\text{Var}(\log \Phi_t)$ with its covariance with $\log Z_t$, both terms are properly attributed to the endogenous mechanism. Were the reallocation channel shut down, Φ_t would be constant and *both* $\text{Var}(\log \Phi_t)$ and $\text{Cov}(\log \Phi_t, \log Z_t)$ would be zero; the covariance is itself a product of the endogenous response of Φ_t to the shocks. The entire gap therefore measures the contribution of the endogenous channel, including the way it co-moves with the exogenous process.

Table 7 reports results averaged across the 50 simulations. In the baseline model, roughly 30 percent of the variance of the Solow TFP cannot be explained by exogenous productivity fluctuations alone; this 30 percent is accounted for by the endogenous mechanism. The no-default model also amplifies Solow TFP variance beyond what Z_t alone can generate, but to a lesser extent: the gap is roughly 20 percent, about three-quarters of the baseline figure. The difference between the two models reflects the additional default margin present in the baseline. When conditions change, the mass of defaulting firms adjusts rapidly, concentrating production in the most productive firms more aggressively than the single-cutoff mechanism can, and thereby generating a larger endogenous component of measured TFP.

9 Conclusion

We have developed a model in which informational frictions give rise to credit misallocation. In our model, as in Stiglitz and Weiss (1981), credit is rationed.

Each firm is endowed with two technologies: a production technology and a financial intermediation technology. The efficiency of the production technology is an idiosyncratic draw that is private information. Because households cannot observe this draw, they end up financing all firms equally. Depending on the realization of their production-technology draw, firms optimally choose which technology to deploy: Those with sufficiently low draws act as financial intermediaries and lend to firms with high draws. Borrowers are subject to a default decision.

We show that this tractable model can capture some key facets of the data. First and foremost, it is able to capture the average within-industry dispersion in productivity observed at the plant level in the United States. The model also reproduces the signs of a set of untargeted partial correlations: within-industry productivity dispersion narrows when GDP growth is above average, and business-loan delinquencies rise when growth is depressed and when real interest rates are high.

In our model, changes in productivity dispersion are linked to economic conditions, implying that a component of aggregate TFP evolves endogenously. We have shown that this endogenous component of productivity can induce notable TFP movements for realistically sized shocks.

We have further isolated the specific contribution of strategic default to the endogenous TFP component by contrasting the baseline model with a no-default variant in which the two-cutoff structure collapses to a single lending threshold. Unconditional simulations show that roughly 30 percent of the variance of the Solow residual is endogenously generated in the baseline model, compared with roughly 20 percent in the no-default variant. Strategic default therefore accounts for a meaningful share of the endogenous TFP component—acting not merely as an amplifier that concentrates production among high-productivity firms, but also as a stabiliser of the endogenous component.

We hope that extensions of our tractable model can serve as a building block in the exploration of how alternative policies affect productivity. Extensions could allow the study of fiscal and monetary policy choices. Going beyond our current analysis, we can speculate that expansionary fiscal policies could have additional desirable effects in our model as the associated increases in real interest rates would force some unproductive firms to quit. By contrast, monetary policy would face additional challenges in our setup. The typical New Keynesian rationale to lower policy rates in a downturn of aggregate demand could inefficiently keep low-productivity firms afloat. In sum, we see our work as opening promising avenues for further research.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., & Kerr, W. R. (2018). Innovation, Reallocation, and Growth. *American Economic Review*, *108*(11), 3450–3491.

Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, *60*(2), 323–351.

Akcigit, U., Celik, M. A., & Greenwood, J. (2016). Buy, Keep, or Sell: Economic Growth and the Market for Ideas. *Econometrica*, *84*(3), 943–984.

Arellano, C., Bai, Y., & Kehoe, P. J. (2019). Financial Frictions and Fluctuations in Volatility. *Journal of Political Economy*, *127*(5), 2049–2103.

Barlevy, G. (2002). The Sullying Effect of Recessions. *The Review of Economic Studies*, *69*(1), 65–96.

Bernanke, B., & Gertler, M. (1990). Financial fragility and economic performance. *The quarterly journal of economics*, *105*(1), 87–114.

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

Bernard, A. B., Eaton, J., Jensen, J. B., & Kortum, S. (2003). Plants and Productivity in International Trade. *American Economic Review*, *93*(4), 1268–1290.

Buera, F. J., & Moll, B. (2015). Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity. *American Economic Journal: Macroeconomics*, *7*(3), 1–42.

Caballero, R. J., & Hammour, M. L. (1994). The Cleansing Effect of Recessions. *American Economic Review*, *84*(5), 1350–1368.

Caballero, R. J., & Hammour, M. L. (1996). On the Timing and Efficiency of Creative Destruction. *The Quarterly Journal of Economics*, *111*(3), 805–852.

Caballero, R. J., & Hammour, M. L. (1998). The Macroeconomics of Specificity. *Journal of Political Economy*, *106*(4), 724–767.

Cao, D., Lorenzoni, G., & Walentin, K. (2019). Financial Frictions, Investment, and Tobin's q. *Journal of Monetary Economics*, *103*, 105–122.

Carlstrom, C. T., & Fuerst, T. S. (1997). Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis. *American Economic Review*, *87*(5), 893–910.

Christiano, L. J., Motto, R., & Rostagno, M. (2014). Risk Shocks. *American Economic Review*, *104*(1), 27–65.

Clementi, G. L., & Hopenhayn, H. A. (2006). A Theory of Financing Constraints and Firm Dynamics. *Quarterly Journal of Economics*, *121*(1), 229–265.

Comin, D., & Gertler, M. (2006). Medium-Term Business Cycles. *American Economic Review*, *96*(3), 523–551.

Cooley, T. F., & Quadrini, V. (2001). Financial Markets and Firm Dynamics. *American Economic Review*, *91*(5), 1286–1310.

Cunningham, C., Foster, L., Grim, C., Haltiwanger, J., Pabilonia, S. W., Stewart, J., & Wolf, Z. (2023). Dispersion in dispersion: Measuring establishment-level differences in productivity. *Review of Income and Wealth*, *69*(4), 999–1032.

Dang, J., Gillman, M., & Kejak, M. (2012). *Real Business Cycles with a Human Capital Investment Sector and Endogenous Growth: Persistence, Volatility and Labor Puzzles* (Working Paper). Cardiff Business School.

David, J. M., & Venkateswaran, V. (2019). The Sources of Capital Misallocation. *American Economic Review*, *109*(7), 2531–2567.

Dong, D., Liu, Z., & Wang, P. (2025). Turbulent business cycles. *Journal of Monetary Economics*, 103814.

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, *98*(1), 394–425.

Gilchrist, S., Sim, J. W., & Zakrajšek, E. (2014, April). *Uncertainty, Financial Frictions, and Investment Dynamics* (NBER Working Papers No. 20038). National Bureau of Economic Research, Inc.

Gomes, J. F., & Schmid, L. (2021). Equilibrium Asset Pricing with Leverage and Default. *Journal of Finance*, *76*(2), 977–1018.

Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L., & Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe. *Quarterly Journal of Economics*, *132*(4), 1915–1967.

Gornemann, N., Guerrón-Quintana, P. A., & Saffie, F. (2025). Real Exchange Rates and Endogenous Productivity. *American Economic Journal: Macroeconomics*, *17*(4), 204–261.

Guo, X., Ottonello, P., Winberry, T., & Whited, T. (2025). *Firm heterogeneity and adverse selection in external finance: Micro evidence and macro implications* (tech. rep.). National Bureau of Economic Research.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, *124*(4), 1403–1448.

Hsieh, C.-T., & Klenow, P. J. (2010). Development Accounting. *American Economic Journal: Macroeconomics*, *2*(1), 207–223.

Jo, I. H., Khan, A., Senga, T., & Thomas, J. K. (2021). *Firm debt and default over the pandemic and recovery* (tech. rep.). Mimeo, Ohio State University.

Jones, L. E., Manuelli, R. E., & Siu, H. E. (2005). Fluctuations in Convex Models of Endogenous Growth, II: Business Cycle Properties. *Review of Economic Dynamics*, *8*(4), 805–828.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica*, *50*(3), 649–670.

Khan, A., & Thomas, J. K. (2013). Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, *121*(6), 1055–1107.

Kim, J., Kim, S., Schaumburg, E., & Sims, C. A. (2008). Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control*, *32*, 3397–3414.

King, R. G., Plosser, C. I., & Rebelo, S. T. (1988). Production, growth and business cycles : I. the basic neoclassical model. *Journal of Monetary Economics*, *21*(2-3), 195–232.

King, R. G., & Rebelo, S. T. (1990). Public Policy and Economic Growth: Developing Neoclassical Implications. *Journal of Political Economy*, *98*(5), S126–S150.

Liu, Z., & Wang, P. (2014). Credit Constraints and Self-Fulfilling Business Cycles. *American Economic Journal: Macroeconomics*, *6*(1), 32–69.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, *22*(1), 3–42.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, *71*(6), 1695–1725.

Midrigan, V., & Xu, D. Y. (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, *104*(2), 422–458.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *American Economic Review*, *104*(10), 3186–3221.

Müller, U. K., & Watson, M. W. (2008). Testing Models of Low-Frequency Variability. *Econometrica*, *76*(5), 979–1016.

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

Pack, H. (1994). Endogenous Growth Theory: Intellectual Appeal and Empirical Shortcomings. *Journal of Economic Perspectives*, *8*(1), 55–72.

Paulson, A. L., Townsend, R. M., & Karaivanov, A. (2006). Distinguishing limited liability from moral hazard in a model of entrepreneurship. *Journal of political Economy*, *114*(1), 100–144.

Peters, M. (2020). Heterogeneous Markups, Growth, and Endogenous Misallocation. *Econometrica*, *88*(5), 2037–2073.

Rebelo, S. T. (1991). Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy*, *99*(3), 500–521.

Restuccia, D., & Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics*, *11*(4), 707–720.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, *98*(5), S71–S102.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, *71*(3), 393–410.

Tables and Figures

Table 1: Summary of firms' decisions and actions within each period

	Period t	Period $t+1$
1	Raise equity a_t	Produce, **outside option matures**
2	**Draw productivity level** $\omega \in [0, 1]$	Repay loans to other firms
3	**Lend or borrow in inter-firm market**	Pay households
4	**Some borrowing firms take the outside option and default**	
5	Purchase physical capital	

Note: The bolded entries denote actions or events specific to our model and not part of a standard RBC model.

Table 2: Model parameters

	Value	Description	
Conventional			
β	0.9925	Discount rate	
α	0.3	Capital share in production	
δ	0.01	Depreciation rate	
ν	2	Inverse Frisch elasticity of labor supply	
Estimated to match first moments with steady-state conditions			*Targets/Explanation*
η_1	1.543	First parameter of Beta distribution	jointly to match the average 1) 90-10 within-industry productivity dispersion, 2) spread between the return on assets and the outside-option return (prime rate minus one-year Treasury), 3) share of bank credit, and 4) delinquency rate on business loans
η_2	2.735	Second parameter of Beta distribution	
ξ	0.007	Haircut on the returns from the outside option	
θ	0.0003	Average fraction of funds that can be diverted	
Estimated by simulated methods of moments			*Targets/Explanation*
ϕ	0.239	Investment adjustment costs	estimated to hit the variances, correlations, and autocorrelations of 1) real output, 2) real consumption, 3) relative investment price, and 4) delinquency rate on business loans
ρ_z	0.407	Persistence of technology shock	
σ_z	0.010	s.d. of technology shock	
ρ_I	0.990	Persistence of investment technology shock	
σ_I	0.001	s.d. of investment technology shock	
ρ_ν	0.986	Persistence of consumption shock	
σ_ν	0.021	s.d. of consumption shock	
Specific			*Explanation*
ψ	0.5	Parameter in the function $\Theta_t(\omega) = \omega^\psi F_t$	Verifies $\Theta_t(\omega)$'s properties and the slope condition
ϑ	0.84236	Disutility of labor	Supports aggregate labor = 1 in the steady state

Table 3: SMM estimation: Data moments and model counterparts, 1987:Q1–2021:Q4

	Data	Data 5th perc.	Data 95th perc.	Model	Model 5th perc.	Model 95th perc.
Var(GDP)	1.70	0.77	2.63	2.29	1.60	3.08
Corr(GDP,Consumption)	0.91	0.25	1.57	0.84	0.76	0.90
Corr(GDP,Investment price)	0.31	0.14	0.48	0.70	0.62	0.78
Corr(GDP,Default rate)	-0.58	-0.78	-0.39	-0.59	-0.70	-0.46
Var(Consumption)	2.10	0.41	3.80	1.78	1.23	2.45
Corr(Consumption,Investment price)	0.22	0.10	0.35	0.46	0.32	0.59
Corr(Consumption,Default rate)	-0.42	-0.57	-0.27	-0.27	-0.43	-0.11
Var(Investment price)	0.56	0.40	0.72	0.62	0.47	0.78
Corr(Investment price,Default rate)	-0.48	-0.71	-0.26	-0.94	-0.96	-0.92
Var(Default rate)	0.34	0.24	0.45	0.34	0.19	0.52
Autocorr(GDP)	0.64	0.46	0.82	0.60	0.48	0.71
Autocorr(Consumption)	0.57	0.31	0.84	0.63	0.50	0.74
Autocorr(Investment price)	0.90	0.64	1.17	0.22	0.08	0.35
Autocorr(Default rate)	0.95	0.65	1.24	0.24	0.07	0.39

Note: The table reports the second moments targeted in the SMM procedure for the model calibration—variances, correlations, and autocorrelations at business-cycle frequencies. The model moments are computed from the average of 1,000 simulated samples of the same length as the observed data. Both observable and model-simulated data are HP-filtered setting the smoothing parameter to 1,600 as standard for quarterly data. The data 5th and 95th percentile columns show the bounds of a 90% confidence interval constructed under the assumption of asymptotic normality, using Newey–West HAC standard errors (Bartlett kernel with optimal bandwidth) for the sample moments. The model 5th and 95th percentile columns show the corresponding empirical percentiles from 1,000 simulated samples.

Table 4: Variance decomposition (percent)

	Output	Invest. Price	Consumption	Default Rate	Real Rate	90-10 TFP disp.
TFP	73.8	87.7	28.6	68.5	74.6	44.0
Consumption	25.1	11.5	69.4	29.0	22.9	52.9
ISP	1.1	0.9	2.0	2.5	2.6	3.2

Note: "TFP" refers to the shock to total factor productivity, Z_t. "Consumption" refers to the consumption shock, ν_{ct}. "IST" refers to the investment technology shock Z_{It}.

Table 5: Regressions of within-industry TFP dispersion (percent difference between the TFP for the 90^{th} percentile firm and that for the 10^{th} percentile firm) and delinquency rates, 1987–2021

	(1) 90-10 perc.	(2) 90-10 perc.	(3) Delinquency rate	(4) Delinquency rate
GDP per capita, growth	-2.86	-2.72	-0.22	-0.22
	[-4.16,-1.56]	[-3.87,-1.56]	[-0.31,-0.12]	[-0.31,-0.13]
Delinquency rate	-4.22	-4.38		
	[-5.56,-2.88]	[-5.57,-3.19]		
Lagged GDP per capita, growth		-2.31		
		[-3.56,-1.06]		
Lagged delinquency rate			0.84	0.78
			[0.74,0.94]	[0.68,0.88]
Real interest rate				0.14
				[0.04,0.24]

Note: 90% confidence intervals in brackets.

Table 6: Regressions of within-industry TFP dispersion (percent difference between the TFP for the 90^{th} percentile firm and that for the 10^{th} percentile firm) and delinquency rate, model-generated data

	(1) 90-10 perc.	(2) 90-10 perc.	(3) Delinquency rate	(4) Delinquency rate
GDP per capita, growth	-0.70	-0.77	-0.11	0.00
	[-1.20,-0.34]	[-1.33,-0.33]	[-0.19,-0.04]	[-0.03,0.03]
Delinquency rate	-4.28	-4.22		
	[-10.47,0.04]	[-10.59,0.09]		
Lagged GDP per capita, growth		-0.66		
		[-1.06,-0.25]		
Lagged delinquency rate			0.41	0.21
			[0.07,0.72]	[0.02,0.44]
Real interest rate				0.26
				[0.19,0.34]
N	34	33	34	34

90% range across replications in brackets

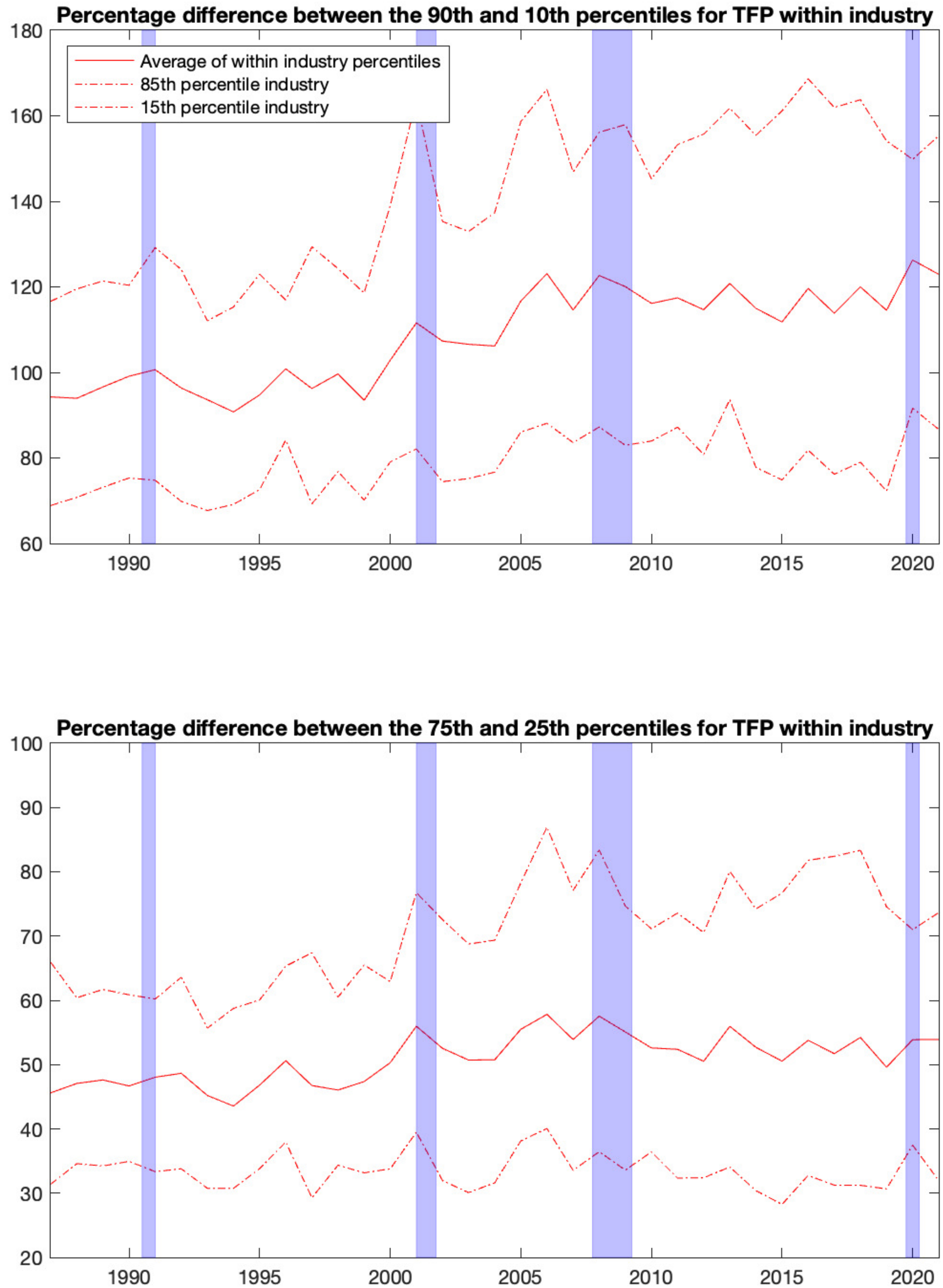
Note: The results are based on 1000 model-simulated samples of the same length as the observed data used for the counterpart regressions in Table 5. We use yearly means of quarterly model data to replicate the frequency of the observed data. 90% confidence intervals are reported in brackets. The confidence intervals for regressions on simulated data are based on the 5-95 percentiles from the 1000 data replications.

Table 7: Variance decomposition of the Solow residual

	$\text{Var}(\log Z_t)/\text{Var}(\log \text{TFP}_t)$	Gap vs. RBC (percent)
RBC benchmark	1.000	0.0
Two-cutoff model (baseline)	0.716	28.4
One-cutoff model (no strategic default)	0.788	21.2

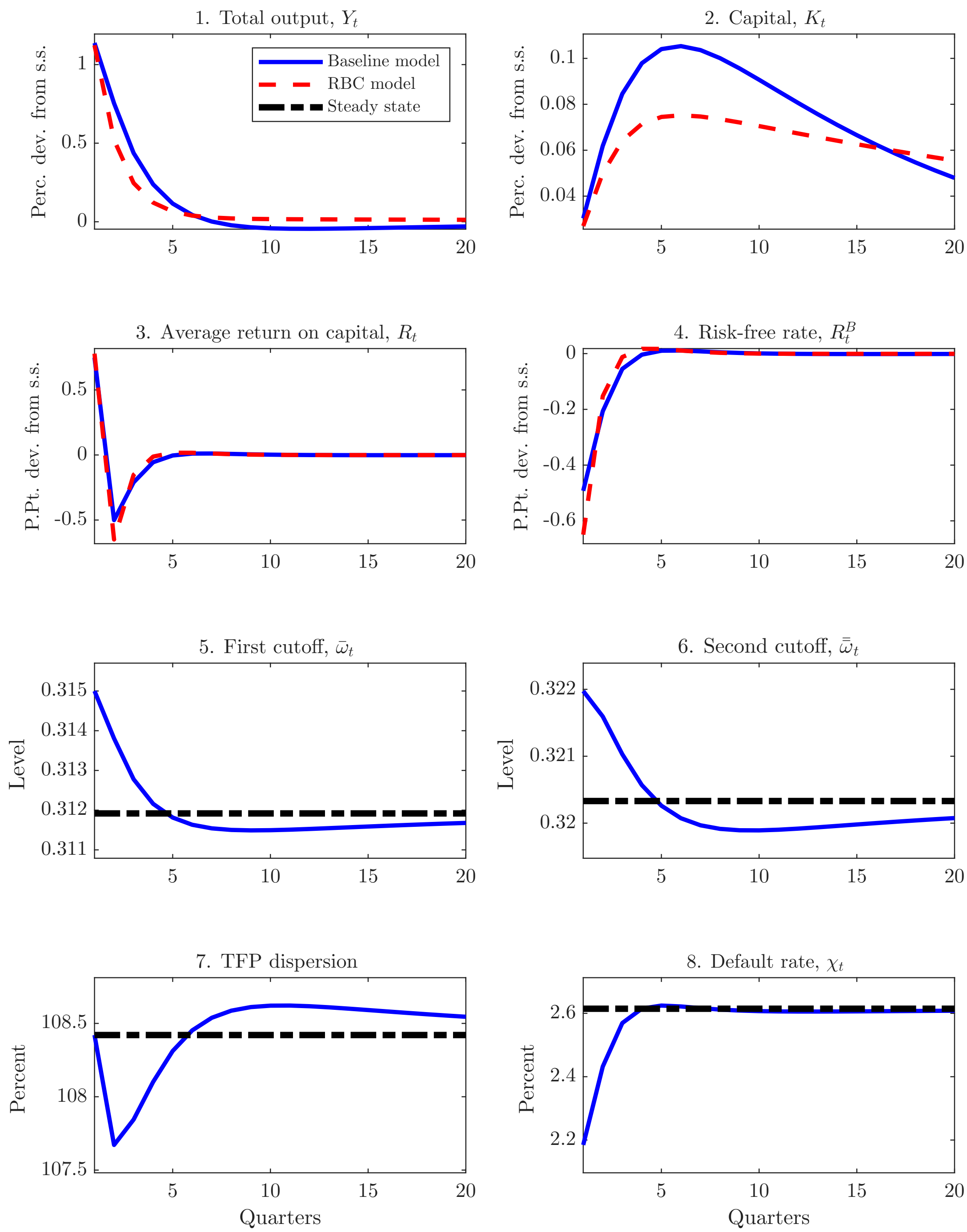
Note: HP-filtered simulations ($\lambda = 1{,}600$), 50 draws of 500 quarters each, all three shocks active. The ratio equals one when Φ_t is constant (RBC); the gap measures the endogenous share of Solow TFP variance.

Figure 1: TFP dispersion, 1987–2021



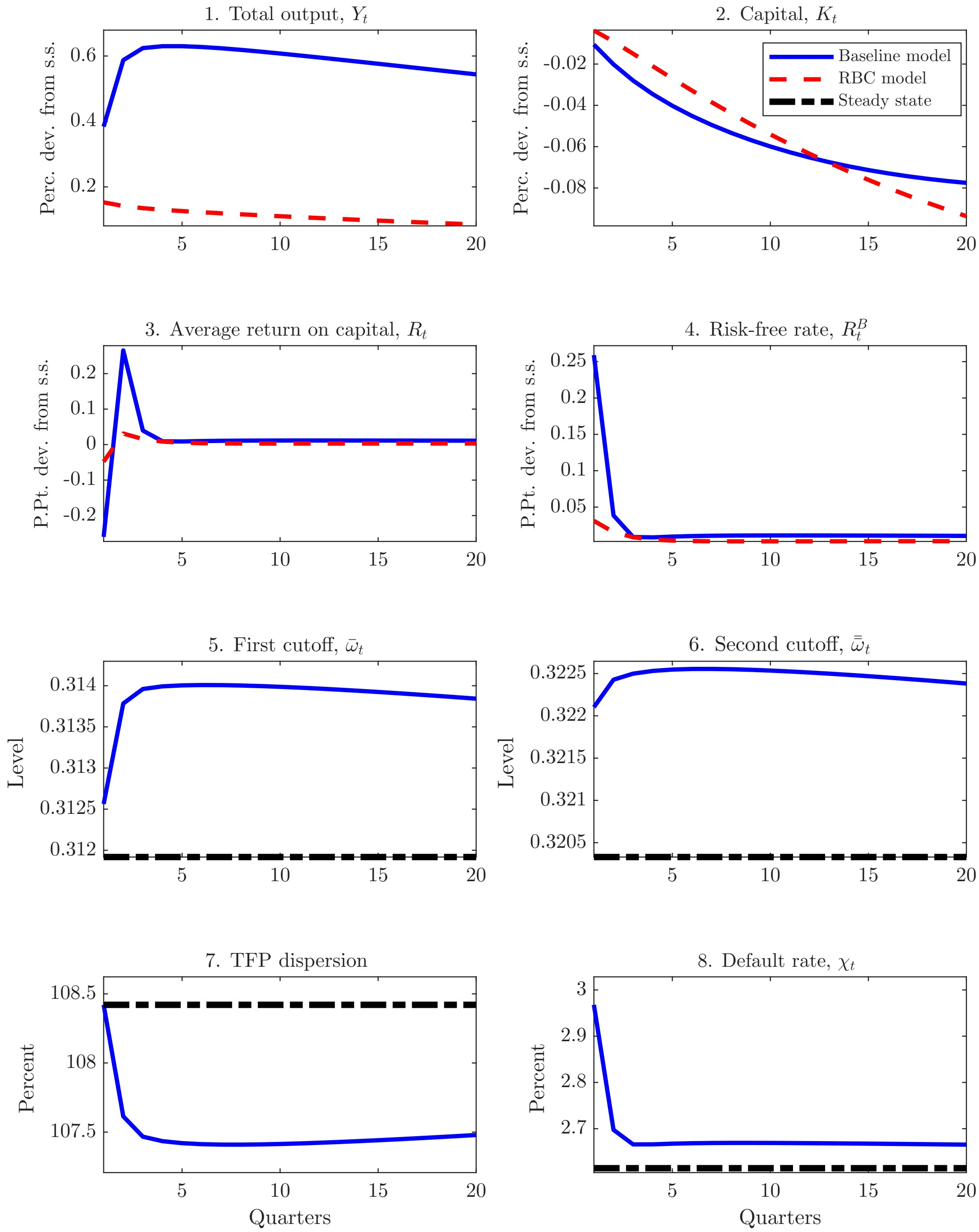
Note: Shaded areas denote recessions as dated by the NBER Business Cycle Dating Committee.
Source: Dispersion Statistics on Productivity, a dataset published jointly by the Bureau of Labor Statistics and the Census Bureau, and authors' calculations.

Figure 2: An expansionary one-standard-deviation TFP shock



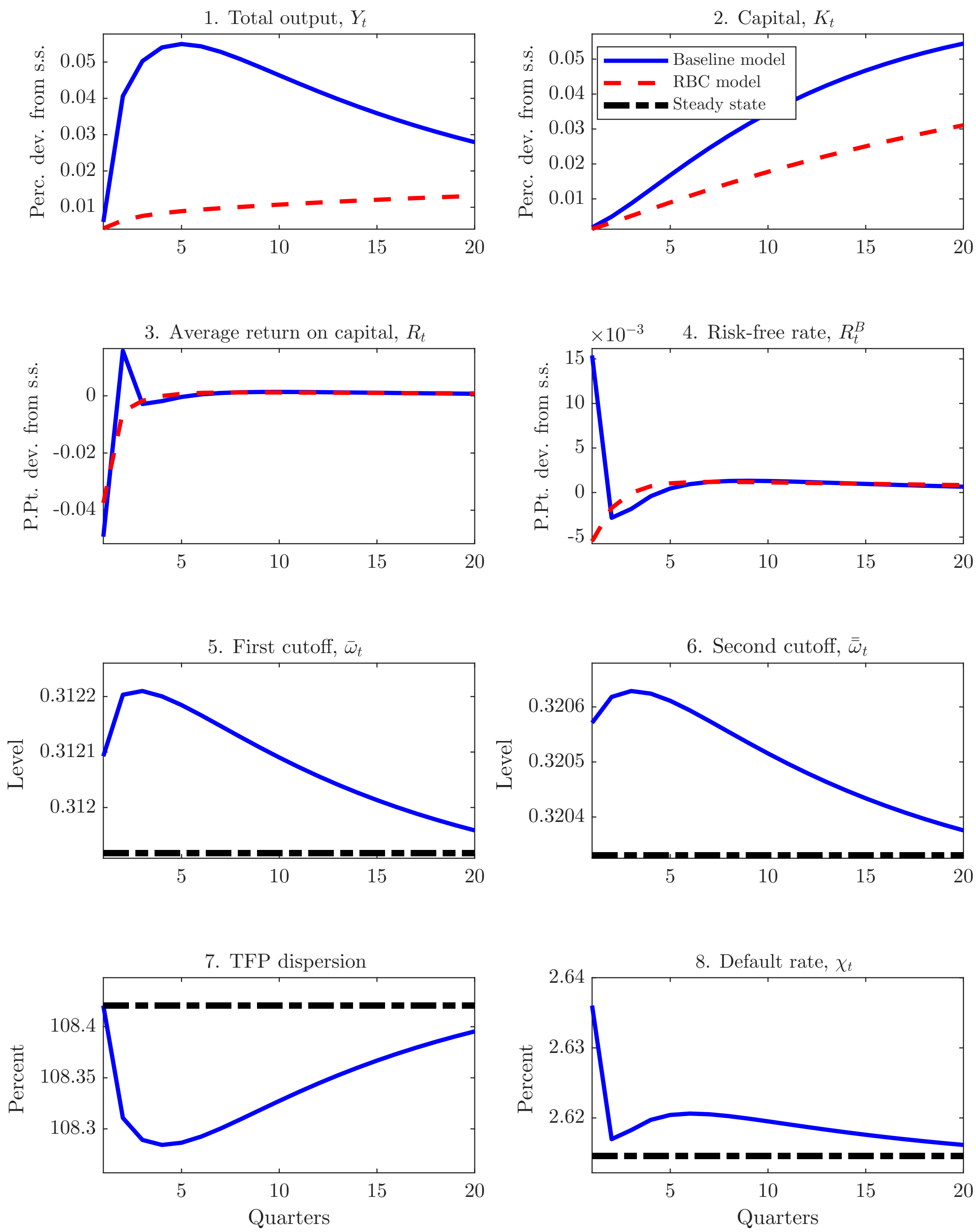
Note: "Omega bar" refers to $\overline{\omega}_t$ and "omega double bar" refers to $\overline{\overline{\omega}}_t$. The responses start from the stochastic steady state and are also shown in deviation from the stochastic steady state, where relevant.

Figure 3: An expansionary one-standard-deviation consumption shock



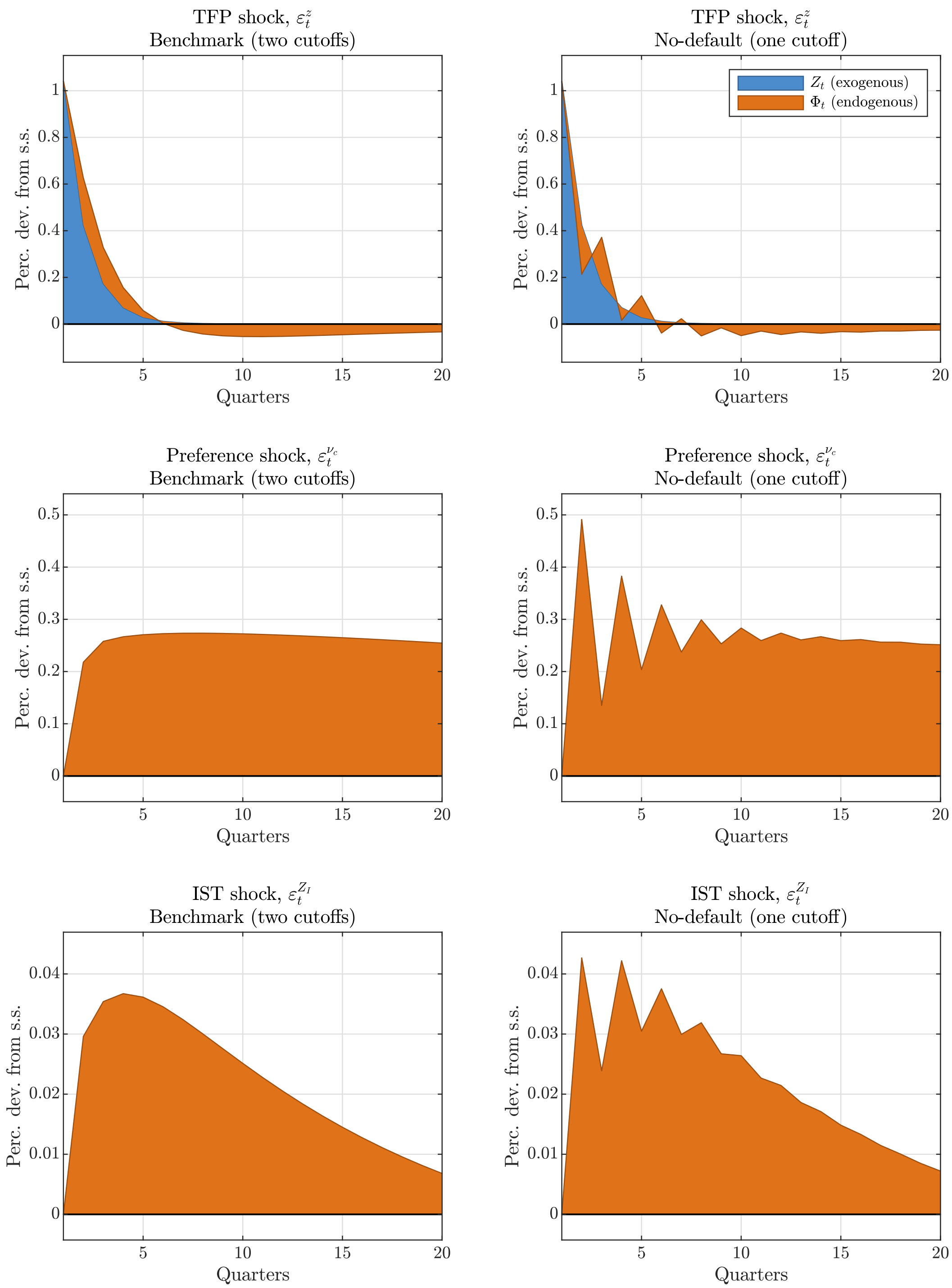
Note: "Omega bar" refers to $\overline{\omega}_t$ and "omega double bar" refers to $\overline{\overline{\omega}}_t$. The responses start from the stochastic steady state and are also shown in deviation from the stochastic steady state, where relevant.

Figure 4: An expansionary one-standard-deviation investment technology shock



Note: "Omega bar" refers to $\overline{\omega}_t$ and "omega double bar" refers to $\overline{\overline{\omega}}_t$. The responses start from the stochastic steady state and are also shown in deviation from the stochastic steady state, where relevant.

Figure 5: TFP decomposition: baseline vs. no-default model



Note: Each panel shows the impulse response of the Solow residual decomposed into its two components (Section 5.2). Because $\log \text{TFP}_t = \log \Phi_t + \log Z_t$, the decomposition is *additive in log-percent units*: the blue and orange areas stack to exactly total measured log TFP at each horizon. Blue area: exogenous component $\log Z_t$ (log % deviation from steady state). Orange area: endogenous component $\log \Phi_t$ (log % deviation from steady state). Rows correspond to the three shocks; columns correspond to the two models. The y-axis scale is identical within each row to facilitate direct comparison. Left column: baseline model with strategic default (two-cutoff model, $\overline{\omega}_t$ and $\overline{\overline{\omega}}_t$). Right column: no-default model (one-cutoff model, $\overline{\omega}_t$ only; see Section G of the online appendix).

Online Appendix

Cyclical Fluctuations, Financial Frictions, and Productivity Differences across Firms

Luca Guerrieri
Federal Reserve Board

Jinill Kim
Korea University
Bank of Korea

Arsenii Mishin
HSE University

This appendix is organized as follows.

Appendix A provides formal proofs of the three propositions that establish the existence and properties of the two firm-segment cutoff points and the inter-firm loan rate in equilibrium.

Appendix B provides a structural interpretation for our diversion function.

Appendix C derives the first-order conditions for the household and capital-producing-firm optimization problems; provides a formal definition of competitive equilibrium; derives key model aggregates—including the aggregate equity return and the aggregate resource constraint—by integrating firm-level outcomes across the three firm segments using the Beta distribution; and lists all equilibrium conditions in compact notation.

Appendix D solves for the non-stochastic steady state, describing an iterative algorithm that guesses the cutoff points and bond rate and verifies consistency across all equilibrium conditions.

Appendix E explains how to compute model-implied counterparts of the empirical TFP dispersion measures by rescaling the productivity distribution to the producing-firm segment.

Appendix F develops a canonical representative-firm RBC model that removes the informational friction by eliminating idiosyncratic productivity and decentralising the model via the standard household-owns-capital structure.

Appendix G develops a variant of the baseline model in which the outside option is unavailable, so that strategic default is eliminated and firms sort into only two segments (lenders and producers) rather than three; derives the single-cutoff equilibrium conditions, steady state, and connections to the baseline and canonical RBC model.

Appendix H compares alternative solution methods.

A Proofs of Propositions

We prove here the three propositions stated in Section 5.1.

A.1 Proof of Proposition 1

Equation (32) can be described as

$$\begin{aligned} & E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\rho_t\frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)}+(1-\theta)\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}\right)a_t\right]= \\ & E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)(a_t+\Theta_t(\omega)b_t)\right] \end{aligned} \tag{A.1}$$

when $\omega=\overline{\omega}_t$. The left-hand side of equation (A.1) does not depend on ω, while the right-hand side of equation (A.1) increases in ω due to our assumption that $\Theta_t(\omega)$ is increasing in ω. Therefore, firms with $\omega<\overline{\omega}_t$ will have no incentive to deviate to the outside option because they get higher expected profits from lending than from diverting funds.

It is left to show that firms with $\omega<\overline{\omega}_t$ will have no incentive to deviate to production. We need to establish that the expected profits from lending are higher than the expected profits from producing, i.e.,

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\rho_t\frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)}+(1-\theta)\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}\right)a_t\right]>E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\omega)a_t\right)\right], \tag{A.2}$$

for all firms with $\omega<\overline{\omega}_t$. Notice that these firms cannot borrow in the inter-firm market because of the screening technology.

Evaluating equation (13) at $\omega=\overline{\omega}_t$ and combining with the cutoff condition (31), we get

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega}_t)\right]=E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\right]. \tag{A.3}$$

Plug this result into equation (32):

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\rho_t\frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)}+(1-\theta)\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}\right)a_t\right]= E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega})\left(a_t+\Theta_t(\overline{\omega}_t)b_t\right)\right]. \quad (A.4)$$

Notice that

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega})a_t\right]>E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\omega)a_t\right)\right] \quad (A.5)$$

for all firms with $\omega<\overline{\omega}_t$ since $R_{t+1}(\omega)$ is increasing in ω. Combining this result with $\Theta_t(\overline{\omega}_t)b_t\geq 0$ and $\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega})\geq 0$, we get

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\rho_t\frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)}+(1-\theta)\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}\right)a_t\right]= E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega})\left(a_t+\Theta_t(\overline{\omega}_t)b_t\right)\right]\geq E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega})a_t+\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega})\Theta_t(\overline{\omega}_t)b_t\right]>E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\omega)a_t\right)\right], \quad (A.6)$$

for all firms with $\omega<\overline{\omega}_t$. Thus, equation (A.2) is verified. □

A.2 Proof of Proposition 2

Let us prove it by contradiction. Assume that $\overline{\overline{\omega}}_t^*<\overline{\omega}_t$. Therefore, by our definition $\overline{\overline{\omega}}_t=\max(\overline{\omega}_t,\overline{\overline{\omega}}_t^*)=\overline{\omega}_t$. Plugging this result into equation (32):

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\rho_t a_t\right]=E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\left(a_t+\Theta_t(\overline{\omega}_t)b_t\right)\right]. \quad (A.7)$$

Consider the right-hand-side of equation (33) and substitute for $E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\rho_t\right]$ from equation (A.7).

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\overline{\omega}}_t^*)(a_t+b_t)-\rho_t b_t\right)\right]= E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\overline{\omega}}_t^*)(a_t+b_t)-\left(R_t^B-\xi\right)\left(1+\Theta_t(\overline{\omega}_t)\frac{b_t}{a_t}\right)b_t\right)\right]. \quad (A.8)$$

Using that $R_{t+1}(\omega) > 0$ and strictly increasing in ω, together with equation (A.3) lead to the following inequality:

$$0 < E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\overline{\omega}}_t^*)\right] < E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}(\overline{\omega}_t)\right] = E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B - \xi\right)\right] \tag{A.9}$$

Plugging this inequality into equation (A.8) and collecting terms result in:

$$\begin{aligned}
&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\overline{\omega}}_t^*)(a_t+b_t) - \rho_t b_t\right)\right] = \\
&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\overline{\omega}}_t^*)(a_t+b_t) - \left(R_t^B - \xi\right)\left(1+\Theta_t(\overline{\omega}_t)\frac{b_t}{a_t}\right)b_t\right)\right] < \\
&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\left(R_t^B - \xi\right)(a_t+b_t) - \left(R_t^B - \xi\right)\left(1+\Theta_t(\overline{\omega}_t)\frac{b_t}{a_t}\right)b_t\right)\right] = \\
&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\left(R_t^B - \xi\right)\left(a_t - \Theta_t(\overline{\omega}_t)\frac{b_t^2}{a_t}\right)\right)\right] \leq E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\left(R_t^B - \xi\right)a_t\right)\right] \leq \\
&E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B - \xi\right)\left(a_t + \Theta_t(\overline{\overline{\omega}}_t^*)b_t\right)\right],
\end{aligned} \tag{A.10}$$

where we use that $\Theta_t(\overline{\omega}_t)\frac{b_t^2}{a_t} \geq 0$ and $\Theta_t(\overline{\overline{\omega}}_t^*)b_t \geq 0$. Therefore,

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B - \xi\right)\left(a_t + \Theta_t(\overline{\overline{\omega}}_t^*)b_t\right)\right] > E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\overline{\omega}}_t^*)(a_t+b_t) - \rho_t b_t\right)\right]. \tag{A.11}$$

However, inequality (A.11) contradicts equation (33). Therefore, $\overline{\overline{\omega}}_t^* \geq \overline{\omega}_t$ $\square$

A.3 Proof of Proposition 3

Note that equations (32) and (33) are constructed such that a marginal firm with productivity level $\overline{\omega}_t$ receives the same expected profits from lending and diverting funds, while a marginal firm with productivity level $\overline{\overline{\omega}}_t$ receives the same expected profits from diverting funds and producing. For this proposition and in the main text, we resolve the tie by assuming these marginal firms choose to lend and produce, respectively. Since the probability distribution of ω is continuous, the probability that $\omega = \overline{\omega}_t$ or $\omega = \overline{\overline{\omega}}_t$ is zero.

Given the results of Propositions 1 and 2, we are left to show that

1. No firm with idiosyncratic productivity $\omega > \overline{\overline{\omega}}_t$ has an incentive to deviate from bor-

rowing and producing.

2. No firm with idiosyncratic productivity $\overline{\omega}_t < \omega < \overline{\overline{\omega}}_t$ has an incentive to deviate from borrowing and defaulting.

We will show our proof in four smaller steps.

Step 1: Firms with $\omega > \overline{\omega}_t$ have higher expected profits from borrowing and defaulting than from lending.

Proof of step 1: Equation (32) can be described as

$$\begin{aligned} &E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\rho_t\frac{1-\mu(\overline{\overline{\omega}}_t)}{1-\mu(\overline{\omega}_t)}+(1-\theta)\frac{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}{1-\mu(\overline{\omega}_t)}\right)a_t\right]= \\ &E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\left(a_t+\Theta_t(\omega)b_t\right)\right]. \end{aligned} \tag{A.12}$$

when $\omega = \overline{\omega}_t$. The left-hand side of equation (A.12) does not depend on ω, while the right-hand side of equation (A.12) increases in ω due to our assumption that $\Theta_t(\omega)$ is increasing in ω. Therefore, firms with $\omega > \overline{\omega}_t$ get higher expected profits from diverting funds than from lending. □

Step 2: Firms with $\omega > \overline{\overline{\omega}}_t$ have higher expected profits from producing than from borrowing and defaulting.

Proof of step 2: Equation (32) can be described as

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\left(a_t+\Theta_t(\omega)b_t\right)\right]=E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\omega)(a_t+b_t)-\rho_t b_t\right)\right] \tag{A.13}$$

when $\omega = \overline{\overline{\omega}}_t$. Taking the first derivative of both sides of equation (A.13) with respect to ω, we need to show that

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\omega)b_t\right]<E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}'(\omega)(a_t+b_t)\right)\right] \tag{A.14}$$

for all $\omega > \overline{\overline{\omega}}_t$. According to this inequality, the expected profits from diverting funds grow less steeply than the expected profits from producing for all firms with the productivity level $\omega > \overline{\overline{\omega}}_t$. Together with the equalization of expected profits in equation (A.13) evaluated at

$\omega = \overline{\overline{\omega}}_t$, this inequality implies that the expected profits from diverting funds are lower than the expected profits from producing for all $\omega > \overline{\overline{\omega}}_t$.

Let us show that condition (35) is sufficient to ensure inequality (A.14) for all $\omega > \overline{\overline{\omega}}_t$. Finding the derivative of $R'_{t+1}(\omega)$ from equation (13) and plugging it into equation (A.14), we get

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\omega)b_t\right] < E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{1}{Q_t}\alpha\ Z_{t+1}\left(\frac{H_{t+1}}{K_t}\right)^{1-\alpha}\right)(a_t+b_t)\right]. \tag{A.15}$$

Since $\Theta_t(\omega)$ is concave (so $\Theta_t'(\omega)$ is non-increasing in ω) and $\overline{\overline{\omega}}_t \geq \overline{\omega}_t$, then

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\omega)b_t\right] \leq E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\overline{\omega}_t)b_t\right] \tag{A.16}$$

for all $\overline{\overline{\omega}}_t < \omega \leq 1$. Therefore, if

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\overline{\omega}_t)b_t\right] < E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{1}{Q_t}\alpha\ Z_{t+1}\left(\frac{H_{t+1}}{K_t}\right)^{1-\alpha}\right)(a_t+b_t)\right] \tag{A.17}$$

holds, then it implies

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\omega)b_t\right] < E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R'_{t+1}(\omega)(a_t+b_t)\right)\right] \tag{A.18}$$

for all $\omega > \overline{\overline{\omega}}_t$. It establishes the sufficiency of the global slope condition in equation (35). □

Step 3: Firms with $\omega > \overline{\overline{\omega}}_t$ have higher expected profits from producing than from lending.
Proof of step 3: It directly follows from the corollary of Proposition 2 and the two results shown in Steps 1 and 2. Since $\overline{\overline{\omega}}_t \geq \overline{\omega}_t$, the result of Step 1 implies that firms with $\omega > \overline{\overline{\omega}}_t$ have higher expected profits from diverting funds than from lending. Combining this implication with the result of Step 2 that establishes that firms with $\omega > \overline{\overline{\omega}}_t$ have higher expected profits from producing than from diverting funds, ensures that firms with $\omega > \overline{\overline{\omega}}_t$ have higher expected profits from producing than from lending.

To complete item 1 of the list above, combine Steps 2 and 3. □

Step 4: Firms with $\omega < \overline{\omega}_t$ have higher expected profits from borrowing and defaulting than from producing.

Proof of step 4: Since $\Theta_t(\omega)$ is concave and, for this part of the argument, we consider $\overline{\overline{\omega}}_t > \overline{\omega}_t$, we have

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\omega)b_t\right] \leq E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\overline{\omega}_t)b_t,\right] \tag{A.19}$$

for each $\omega < \overline{\overline{\omega}}_t$. When $\overline{\overline{\omega}}_t = \overline{\omega}_t$, the case is degenerate: there is zero mass of diverting firms. In this situation, it suffices to verify that no firm has an incentive to deviate from producing or from lending. This condition has already been established in the preceding steps of the proof of this Proposition and in Proposition 1. Combining equation (A.19) with the global slope condition in equation (35), we establish that

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B-\xi\right)\Theta_t'(\omega)b_t\right] < E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}'(\omega)(a_t+b_t)\right)\right] \tag{A.20}$$

for each $\omega < \overline{\overline{\omega}}_t$.

By definition of $\overline{\overline{\omega}}_t$, the expected payoff from diverting funds equals the expected payoff from producing at $\omega = \overline{\overline{\omega}}_t$. Equation (A.20) establishes that the slope of the production payoff strictly exceeds the slope of the diversion payoff for each ω. Therefore, for each $\omega < \overline{\overline{\omega}}_t$ the expected diversion payoff decreases more slowly than the expected production payoff, which implies that the expected payoff from diverting funds strictly exceeds the expected payoff from producing for all $\omega < \overline{\overline{\omega}}_t$.

To complete item 2 of the list above, we also need to combine the results of Steps 1 and 4 with the corollary of Proposition 2. Since $\overline{\overline{\omega}}_t \geq \overline{\omega}_t$, Steps 1 and 4 imply that the expected profits from borrowing and defaulting are higher than those from lending and from producing for all $\overline{\omega}_t < \omega < \overline{\overline{\omega}}_t$. □

B Structural Interpretation for the Diversion Technology

Our baseline specification adopts the parametric form $\Theta_t(\omega) = F_t \omega^{\psi}$ for the diversion technology. The objective of this section is to provide a decision-theoretic derivation that delivers exactly this functional form and pins down the qualitative dependence of F_t on the production cutoff $\overline{\overline{\omega}}_t$. We show that:

1. More productive firms can retain a larger fraction of the borrowed funds when they choose to default, justifying our assumption that $\Theta_t(\omega)$ is increasing in ω;

2. The diversion function $\Theta_t(\omega)$ is concave in ω;

3. The effectiveness of the diversion rate falls when more firms attempt to divert resources, justifying our normalization for F_t that varies endogenously and negatively with $\overline{\overline{\omega}}_t$.

The economy is populated by a continuum of firms and competitive lenders. Firms differ in productivity. After obtaining financing, firms choose between operating a productive project and diverting all available funds. We associate diversion with shifting resources into private assets or unobservable projects whose value depends on the firm's managerial ability (or as we call it organizational capital) that is linked to ω. Diversion leads to default and requires concealment services supplied by a separate sector. The supply of concealment services is fixed in the short run because labor is allocated to that sector before firm productivity is realized. Consequently, an increase in aggregate diversion activity raises the equilibrium price of concealment services and reduces the profitability of diversion.

The key mechanism relies on organizational capital. Firm productivity reflects both aggregate technology and firm-specific organizational capital. Organizational capital can be redeployed in diversion activities, and it can only be seized partially by creditors. As a result, more productive firms obtain larger private benefits from diversion. However, because diversion exhibits diminishing returns with respect to organizational capital, sufficiently productive firms prefer productive operation.

B.1 Environment

Consider a continuum of borrowing firms of mass 1. Firm i is indexed by productivity ω_i drawn from the cumulative distribution function $G(\omega)$ with the support $\left[\underline{\omega}, \underline{\underline{\omega}}\right]$.

Each firm has net worth a and borrows b. Total available resources equal

$$k = a + b. \tag{B.1}$$

As in the baseline model, lenders finance everyone identically and cannot condition contracts on productivity because firm-level production efficiency is not directly observed by investors and the claims of firms are not credible, so a and b are not firm-specific.

As in the baseline model, labor is committed before ω is realized, so a quantity h of labor is allocated equally to each borrowing firm.

The timing is:

1. Lenders extend credit.
2. Labor is allocated between production and concealment sectors.
3. Firm-specific productivity ω_i is realized.
4. Firms choose between production and diversion.
5. Output, repayment, and default occur.

B.2 Production

The production function of the firm i is given by

$$y_i = \omega_i z k^{\alpha} h^{1-\alpha}, \tag{B.2}$$

where z is the aggregate level of technology common to all firms and $0 < \alpha < 1$. The firm repays ρb, where ρ is the lending rate. The resulting payoff is

$$V_P(\omega_i) = \omega_i z k^{\alpha} h^{1-\alpha} - \rho b. \tag{B.3}$$

Define

$$A_P \equiv zk^{\alpha}h^{1-\alpha}. \tag{B.4}$$

Then

$$V_P(\omega_i) = A_P\omega_i - \rho b. \tag{B.5}$$

The value of productive operation is therefore linear in productivity as in the baseline model.

B.3 Concealment Services

Diversion requires concealment services. These services represent legal, financial, and organizational arrangements that facilitate hiding resources from creditors. We will measure concealment services as a fraction of the borrowed funds b.

A competitive concealment sector transforms aggregate labor employed in the diversion sector H_θ into concealment services Θ:

$$\Theta = A_\Theta H_\Theta^{\zeta}, \qquad 0 < \zeta < 1, \tag{B.6}$$

where A_Θ is the efficiency of the concealment technology.

Because labor is allocated before productivity is realized (so H_Θ is constant), we assume that lenders can control A_Θ to ensure that ex-post supply of concealment services is fixed (and corresponds to the constant θ in the baseline model):

$$\Theta_S = A_\Theta H_\Theta^{\zeta} = \theta. \tag{B.7}$$

Let q_Θ denote the market price of concealment services.

B.4 Diversion and Default

A diverting firm transfers all available resources away from production. Diversion automatically results in default. A diverting firm purchases concealment services $\Theta_i \equiv \Theta(\omega_i)$.

The private payoff from diversion is

$$V_D(\omega_i) = \omega_i^{\eta}(\Theta_i b)^{\gamma} a^{1-\gamma}, \tag{B.8}$$

where $0 < \gamma < 1$ and $0 < \eta < 1-\gamma$.[26] This payoff includes the private benefit from investing into unobserved projects minus the cost of concealment services.

The firm solves

$$\max_{\Theta_i} \left\{ \omega_i^{\eta}(\Theta_i b)^{\gamma} a^{1-\gamma} - q_{\Theta}\Theta_i b \right\}. \tag{B.9}$$

The first-order condition is

$$\gamma \omega_i^{\eta} \Theta_i^{\gamma-1} \left(\frac{a}{b}\right)^{1-\gamma} = q_{\Theta}. \tag{B.10}$$

Therefore, the optimal diversion

$$\Theta_i^* = \frac{a}{b}\left(\frac{\gamma\omega_i^{\eta}}{q_{\Theta}}\right)^{\frac{1}{1-\gamma}}. \tag{B.11}$$

From this equation, we can find that

$$\Theta'(\omega_i) \equiv \frac{\partial \Theta_i^*}{\partial \omega_i} > 0, \tag{B.12}$$

Substituting the optimal choice into the objective function yields

$$V_D(\omega_i) = \omega_i^{\frac{\eta}{1-\gamma}} q_{\Theta}^{-\frac{\gamma}{1-\gamma}} A_D, \tag{B.13}$$

where

$$A_D = (1-\gamma)\gamma^{\frac{\gamma}{1-\gamma}} a. \tag{B.14}$$

[26] Labor does not enter explicitly into this diversion payoff because concealment services are produced using labor, and the firm i purchases these services as a composite input Θ_i. Consequently, labor affects diversion indirectly through the equilibrium supply and price of concealment services.

The diversion payoff is increasing in productivity and decreasing in the equilibrium price of concealment services.

B.5 Occupational Choice

Borrowing firms choose the activity that yields the larger payoff:

$$V(\omega_i) = \max\left\{V_P(\omega_i), V_D(\omega_i)\right\}. \tag{B.15}$$

Because $0 < \eta < 1-\gamma$, we have that $\frac{\eta}{1-\gamma} < 1$. Therefore, the diversion payoff V_D is concave in the firm-specific productivity.

Since V_P is linear in ω_i and V_D is concave in ω_i, there exists a unique threshold ω^* satisfying

$$A_P\omega^* - \rho b = (\omega^*)^{\frac{\eta}{1-\gamma}} q_\Theta^{-\frac{\gamma}{1-\gamma}} A_D. \tag{B.16}$$

It guarantees that firms with $\omega_i < \omega^*$ choose diversion, whereas firms with $\omega_i > \omega^*$ choose production.[27] Note the assumption of $0 < \gamma < 1$ is required to ensure the global optimality of Θ_i^*, while $0 < \eta < 1-\gamma$ ensures the concavity of $V_D(\omega_i)$ which is needed to establish the firm sorting described.

B.6 Concealment-Market Equilibrium

Only diverting firms demand concealment services. Aggregate demand is

$$\Theta_D(q_\Theta) = \int_{\underline{\omega}}^{\omega^*} \Theta^*(\omega, q_\Theta)\, dG(\omega). \tag{B.17}$$

Plugging equation (B.11),

$$\Theta_D(q_\Theta) = \left[\gamma^{\frac{1}{1-\gamma}} \frac{a}{b}\right] q_\Theta^{-\frac{1}{1-\gamma}} \int_{\underline{\omega}}^{\omega^*} \omega^{\frac{\eta}{1-\gamma}}\, dG(\omega). \tag{B.18}$$

[27]The parameter values and the support of the productivity distribution are chosen such that $\underline{\omega} < \omega^* < \underline{\underline{\omega}}$.

Market clearing requires

$$\Theta_D(q_\Theta) = \Theta_S = \theta. \tag{B.19}$$

This condition determines the equilibrium price.

B.7 Lenders

Lenders do not observe firm productivity when extending credit. If a firm produces, lenders receive ρb. If a firm diverts, default occurs and lenders recover $1 - \theta$ fraction of b.

The lender revenues equal

$$\pi_L = \rho b(1 - G(\omega^*)) + G(\omega^*)b(1 - \theta). \tag{B.20}$$

B.8 Main Results

The main results of the model are summarized in 3 propositions.

Proposition 4. *More productive diverting firms retain a larger fraction of the diverted funds. Moreover, the diversion function is concave in firm-specific productivity, i.e.*

$$\Theta'(\omega_i) \equiv \frac{\partial \Theta_i^*}{\partial \omega_i} > 0, \qquad \Theta''(\omega_i) \equiv \frac{\partial^2 \Theta_i^*}{\partial \omega_i^2} < 0. \tag{B.21}$$

Proof. It follows directly from equation (B.11) given that $\frac{\eta}{1-\gamma} < 1$. □

Proposition 5. *There exists a unique threshold ω^* such that firms with productivity ω below the threshold divert funds and firms with productivity above the threshold produce.*

Proof. See Section B.5 □

Proposition 6. *The effectiveness of diversion decreases endogenously when more firms attempt to divert (when ω^* increases).*

Proof. An increase in the mass of diverting firms increases aggregate demand for concealment services. Since supply is fixed ex post, market clearing requires a higher equilibrium price q_Θ.

To formally show that q_Θ increases in equilibrium, notice that a higher ω^* increases the integral $\int_{\underline{\omega}}^{\omega^*} \omega^{\frac{\eta}{1-\gamma}} \, dG(\omega)$ in equation (B.18). Combining it with the market clearing condition in equation (B.19), it implies that q_Θ must increase since $-\frac{1}{1-\gamma} < 0$ for $0 < \gamma < 1$.

From equation (B.11), a rise in q_Θ decreases the optimal Θ_i^*. □

C Equilibrium Definition

This appendix is organized as follows. Sections C.1 and C.2 complete the derivations of the first-order conditions from the household and capital-producing-firm optimization problems, providing the steps omitted from the main text for brevity. Section C.3 presents the aggregation of firm-level decisions into aggregate variables and states a formal definition of competitive equilibrium. Sections C.4–C.9 derive the Beta-distribution closed forms for the aggregate equity return, aggregate output, average return on capital, aggregate equity-return market clearing, and the diversion scaling factor F_t. Finally, Section C.10 lists all equilibrium conditions in compact, Beta-distribution-specific notation.

C.1 Household Problem

The representative household maximizes

$$\max_{\left\{A_{t+\tau}, C_{t+\tau}, H_{t+\tau}, B^H_{t+\tau}\right\}_{\tau=0}^{\infty}} E_t \sum_{\tau=0}^{\infty} \beta^\tau \left[\ln(C_{t+\tau} - \nu_{ct+\tau}) - \frac{\vartheta}{1+\nu} (H_{t+\tau})^{1+\nu} \right],$$

subject to

$$C_{t+\tau} + A_{t+\tau} + B^H_{t+\tau} = R^A_{t+\tau} A_{t+\tau-1} + W_{t+\tau} H_{t+\tau} + R^B_{t+\tau-1} B^H_{t+\tau-1} + \Pi_{t+\tau} + T_{t+\tau} + \Xi_{t+\tau},$$

reproduced from equations (1) and (2). The first-order conditions for assets,

$$-\lambda_{ct} + \beta E_t \left\{ \lambda_{ct+1} R^A_{t+1} \right\} = 0, \tag{C.1}$$

consumption,

$$\frac{1}{C_t - \nu_{ct}} = \lambda_{ct}, \tag{C.2}$$

labor,

$$-\vartheta H_t^{\nu} + \lambda_{ct} W_t = 0, \tag{C.3}$$

and government bonds,

$$-\lambda_{ct} + \beta E_t \left\{ \lambda_{ct+1} R_t^B \right\} = 0, \tag{C.4}$$

follow from the stationarity conditions of the Lagrangian, where λ_{ct} is the multiplier on the budget constraint (2).

C.2 Capital-Producing Firms

Capital-producing firms solve

$$\max_{I_{t+i}^g} E_t \sum_{i=0}^{\infty} \beta^i \frac{\lambda_{ct+i}}{\lambda_{ct}} \left\{ Q_{t+i} Z_{It+i} \left[1 - \frac{\phi}{2} \left(\frac{I_{t+i}^g}{I_{t+i-1}^g} - 1 \right)^2 \right] I_{t+i}^g - I_{t+i}^g \right\},$$

reproduced from equation (24). The first-order condition implies

$$0 = E_t \left\{ \begin{array}{c} Q_t Z_{It} \left[-\phi \left(\frac{I_t^g}{I_{t-1}^g} - 1 \right) \frac{1}{I_{t-1}^g} \right] I_t^g + Q_t Z_{It} \left[1 - \frac{\phi}{2} \left(\frac{I_t^g}{I_{t-1}^g} - 1 \right)^2 \right] - 1 \\ + \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} Q_{t+1} Z_{It+1} \phi \left(\frac{I_{t+1}^g}{I_t^g} - 1 \right) \left(\frac{I_{t+1}^g}{I_t^g} \right)^2 \end{array} \right\} \tag{C.5}$$

C.3 Aggregation and Equilibrium

We proceed as follows: First, we link individual and aggregate variables, then we define a competitive equilibrium.

Since a mass $\mu(\overline{\omega}_t)$ of firms lend and the complement mass $1 - \mu(\overline{\omega}_t)$ of firms borrow, the inter-firm market clears when

$$\int_0^{\overline{\omega}_t} l_t(\omega) \mu'(\omega) \mathrm{d}\omega = \int_{\overline{\omega}_t}^1 b_t(\omega) \mu'(\omega) \mathrm{d}\omega, \tag{C.6}$$

which by defining

$$L_t = \int_0^{\overline{\omega}_t} l_t(\omega)\mu'(\omega)\mathrm{d}\omega, \tag{C.7}$$

$$B_t = \int_{\overline{\omega}_t}^1 b_t(\omega)\mu'(\omega)\mathrm{d}\omega \tag{C.8}$$

translates into:

$$L_t = B_t. \tag{C.9}$$

Since each type of firm borrows the same amount, we have $a_t(\omega) = a_t$ for all ω and $b_t(\omega) = b_t$ for borrowing firms ($\omega > \overline{\omega}_t$). Using the definition $A_t = \int_0^1 a_t(\omega)\mu'(\omega)\mathrm{d}\omega$ and equation (C.8), we can relate individual to aggregate variables as follows:

$$a_t = A_t, \tag{C.10}$$

$$b_t = \frac{B_t}{1-\mu(\overline{\omega}_t)}. \tag{C.11}$$

Next, consider the aggregation of budget constraints of firms in each segment. For lending firms, $l_t(\omega) = a_t(\omega)$ for each ω. Aggregating over the relevant mass of firms, i.e.,

$$\int_0^{\overline{\omega}_t} l_t(\omega)\mu'(\omega)\mathrm{d}\omega = \int_0^{\overline{\omega}_t} a_t(\omega)\mu'(\omega)\mathrm{d}\omega, \tag{C.12}$$

results into

$$L_t = \mu(\overline{\omega}_t)A_t. \tag{C.13}$$

Let $d_t(\omega) = a_t(\omega) + \Theta_t(\omega)b_t(\omega)$ define the amount of resources diverted and invested in the outside option by the firm of productivity ω. Aggregating over the relevant mass of firms, i.e.,

$$\int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \left(a_t(\omega) + \Theta_t(\omega)b_t(\omega)\right)\mu'(\omega)\mathrm{d}\omega = \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} d_t(\omega)\mu'(\omega)\mathrm{d}\omega, \tag{C.14}$$

and using the aggregation identity for firms' bond holdings,

$$D_t = \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} d_t(\omega)\mu'(\omega)\mathrm{d}\omega, \tag{C.15}$$

result into

$$D_t = \left(\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)\right) A_t + \frac{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)}{1 - \mu(\overline{\omega}_t)}\theta B_t. \tag{C.16}$$

For producing firms, $a_t(\omega) + b_t(\omega) = b_t^{tot}(\omega)$. Aggregating over the relevant mass of firms, i.e.,

$$\int_{\overline{\overline{\omega}}_t}^{1} \left(a_t(\omega) + b_t(\omega)\right) \mu'(\omega)\mathrm{d}\omega = \int_{\overline{\overline{\omega}}_t}^{1} b_t^{tot}(\omega)\mu'(\omega)\mathrm{d}\omega, \tag{C.17}$$

and defining

$$B_t^{tot} = \int_{\overline{\overline{\omega}}_t}^{1} b_t^{tot}(\omega)\mu'(\omega)\mathrm{d}\omega. \tag{C.18}$$

result into

$$B_t^{tot} = \left(1 - \mu(\overline{\overline{\omega}}_t)\right) A_t + \frac{1 - \mu(\overline{\overline{\omega}}_t)}{1 - \mu(\overline{\omega}_t)} B_t. \tag{C.19}$$

In equilibrium, all firms that produce raise the same amount of financing. Therefore $k_t(\omega) = \tilde{k}_t$ is uniform across firms. Because labor is also committed before ω is observed, the labor allocation (11) gives

$$h_{t+1}(\omega) = \tilde{h}_{t+1}, \tag{C.20}$$

uniform across producing firms. Individual output is therefore

$$y_{t+1}(\omega) = Z_{t+1}\,\omega\,\tilde{k}_t^{\alpha}\,\tilde{h}_{t+1}^{1-\alpha}, \tag{C.21}$$

linear in ω. Aggregate capital and labor are

$$K_t = \left(1 - \mu(\overline{\overline{\omega}}_t)\right)\tilde{k}_t, \tag{C.22}$$

$$H_{t+1} = \int_{\overline{\overline{\omega}}_t}^{1} h_{t+1}(\omega)\,\mu'(\omega)\,\mathrm{d}\omega = \left(1 - \mu(\overline{\overline{\omega}}_t)\right)\tilde{h}_{t+1}. \tag{C.23}$$

Integrating individual outputs and eliminating $\tilde{k}_t$ and $\tilde{h}_{t+1}$ using (C.22)–(C.23) yields

$$Y_{t+1} = Z_{t+1}\, K_t^{\alpha}\, H_{t+1}^{1-\alpha} \frac{\int_{\overline{\overline{\omega}}_t}^{1} \omega\, \mu'(\omega)\, \mathrm{d}\omega}{1-\mu(\overline{\overline{\omega}}_t)}. \tag{C.24}$$

Individual producing firms borrow to finance the purchase of capital, i.e., $b_t^{tot}(\omega) = Q_t k_t(\omega)$, so aggregating over this mass of firms results into the aggregate constraint

$$B_t^{tot} = Q_t K_t. \tag{C.25}$$

Let R_t define the average return on capital that producing firms receive, with the relevant density $\frac{\mu'(\omega)}{1-\mu(\overline{\overline{\omega}}_{t-1})}$:

$$R_t = \int_{\overline{\overline{\omega}}_{t-1}}^{1} R_t(\omega) \frac{\mu'(\omega)}{1-\mu(\overline{\overline{\omega}}_{t-1})} \mathrm{d}\omega. \tag{C.26}$$

Substituting $R_t(\omega) = (\kappa_t/Q_{t-1})\,\omega + (1-\delta)Q_t/Q_{t-1}$ from equation (13) and noting that $\kappa_t\, E[\omega \mid \omega \geq \overline{\overline{\omega}}_{t-1}] = \alpha Y_t/K_{t-1}$ (which follows from equation (C.24) applied at period t):

$$R_t = \frac{\alpha\, Y_t}{Q_{t-1}\, K_{t-1}} + \frac{(1-\delta)\, Q_t}{Q_{t-1}}. \tag{C.27}$$

The $E[\omega]$ factors cancel in the weighted average: the distribution-specific terms that appear in individual returns $R_t(\omega)$ drop out of the aggregate return, leaving a clean expression in observable aggregates. The aggregate equity return to households is defined as

$$R_t^A = \int_0^1 R_t^A(\omega)\mu'(\omega)\mathrm{d}\omega. \tag{C.28}$$

The following subsection derives that this equation can be used to get that

$$R_t^A A_{t-1} = B_{t-1}^{tot} R_t + (R_{t-1}^B - \xi) D_{t-1} + (1-\theta)\, \frac{\mu(\overline{\overline{\omega}}_{t-1}) - \mu(\overline{\omega}_{t-1})}{1-\mu(\overline{\omega}_{t-1})} B_{t-1}. \tag{C.29}$$

This equation says that the funds paid to the household are equal to the returns from lending to producing firms net of intermediation costs, plus the returns from diverting funds, and including the funds that were lent to diverting firms and recovered.

The goods market clears:

$$Y_t = C_t + I_t^g. \tag{C.30}$$

Now we are ready to define a competitive equilibrium. The equilibrium is an allocation

$$\left\{C_t, H_t, Y_t, K_t, B_t^{tot}, I_t^n, I_t^g, \Pi_t, T_t, B_t^G, B_t^H, D_t, \Xi_t, \overline{\omega}_t, \overline{\overline{\omega}}_t, Z_t, Z_{It}, \nu_{ct}, F_t, L_t, A_t, B_t\right\}_{t=0}^{\infty},$$

together with the sequence of prices $\left\{\lambda_{ct}, R_t^A, W_t, R_t^B, Q_t, R_t, \rho_t\right\}_{t=0}^{\infty}$ satisfying equations (C.1), (C.2), (C.3), (C.4), (3), (5), (22), (12), (21), (23), (C.5), (25), (26), (27), (28), (30), (31), (32), (33), (39), (C.9), (C.13), (C.16), (C.19), (C.24), (C.25), (C.27), (C.29), and (C.30), together with the slope conditions defined in Appendix A, given initial conditions K_0, I_0^g, B_0^{tot}, B_0^G, B_0^H, D_0, $\overline{\omega}_0$, $\overline{\overline{\omega}}_0$, A_0, Z_0, $Z_{I,0}$, ν_{c0}, R_0^B, Q_0, and the exogenous processes $\{\varepsilon_t^z, \varepsilon_{It}, \varepsilon_{\nu t}\}$.[28]

C.4 Aggregate Equity Return

By definition

$$\begin{aligned} R_t^A = &\int_0^{\overline{\omega}_{t-1}} R_t^A(\omega)\mu'(\omega)\mathrm{d}\omega + \int_{\overline{\omega}_{t-1}}^{\overline{\overline{\omega}}_{t-1}} R_t^A(\omega)\mu'(\omega)\mathrm{d}\omega + \int_{\overline{\overline{\omega}}_{t-1}}^{1} R_t^A(\omega)\mu'(\omega)\mathrm{d}\omega = \\ &\int_0^{\overline{\omega}_{t-1}} \left(\rho_{t-1}\frac{1-\mu(\overline{\overline{\omega}}_{t-1})}{1-\mu(\overline{\omega}_{t-1})} + \frac{\mu(\overline{\overline{\omega}}_{t-1})-\mu(\overline{\omega}_{t-1})}{1-\mu(\overline{\omega}_{t-1})}(1-\theta)\right)\mu'(\omega)\mathrm{d}\omega + \\ &\int_{\overline{\omega}_{t-1}}^{\overline{\overline{\omega}}_{t-1}} \frac{\left(R_{t-1}^B-\xi\right)\left(a_{t-1}(\omega)+\Theta_t(\omega)b_{t-1}(\omega)\right)}{a_{t-1}(\omega)}\mu'(\omega)\mathrm{d}\omega + \\ &\int_{\overline{\overline{\omega}}_{t-1}}^{1} \frac{R_t(\omega)\left(a_{t-1}(\omega)+b_{t-1}(\omega)\right)-\rho_{t-1}b_{t-1}(\omega)}{a_{t-1}(\omega)}\mu'(\omega)\mathrm{d}\omega, \end{aligned} \tag{C.31}$$

where we use equations (20), (17), and (15) to size the equity returns on each segment of firms.

[28] To express equations (32) and (33) in terms of the aggregate variables A_t, B_t, and F_t, we use equations (C.10), (C.11), and (38) to substitute for a_t, b_t, and $\Theta_t(\omega)$, respectively.

Simplifying

$$R_t^A = \rho_{t-1}\frac{1-\mu(\overline{\overline{\omega}}_{t-1})}{1-\mu(\overline{\omega}_{t-1})}\mu(\overline{\omega}_{t-1}) + \frac{\mu(\overline{\overline{\omega}}_{t-1})-\mu(\overline{\omega}_{t-1})}{1-\mu(\overline{\omega}_{t-1})}(1-\theta)\mu(\overline{\omega}_{t-1}) +$$

$$\int_{\overline{\omega}_{t-1}}^{\overline{\overline{\omega}}_{t-1}} \frac{(R_{t-1}^B-\xi)d_{t-1}(\omega)}{a_{t-1}}\mu'(\omega)\mathrm{d}\omega + \int_{\overline{\overline{\omega}}_{t-1}}^{1} \frac{R_t(\omega)(a_{t-1}+b_{t-1})}{a_{t-1}}\mu'(\omega)\mathrm{d}\omega - \int_{\overline{\overline{\omega}}_{t-1}}^{1} \frac{\rho_{t-1}b_{t-1}}{a_{t-1}}\mu'(\omega)\mathrm{d}\omega =$$

$$\rho_{t-1}\left(1-\mu(\overline{\overline{\omega}}_{t-1})\right)\frac{b_{t-1}}{a_{t-1}} + \left(\mu(\overline{\overline{\omega}}_{t-1})-\mu(\overline{\omega}_{t-1})\right)(1-\theta)\frac{b_{t-1}}{a_{t-1}} + \frac{(R_{t-1}^B-\xi)}{A_{t-1}}D_{t-1} +$$

$$\left(1-\mu(\overline{\overline{\omega}}_{t-1})\right)\frac{\left(A_{t-1}+\frac{B_{t-1}}{1-\mu(\overline{\omega}_{t-1})}\right)}{A_{t-1}}\int_{\overline{\overline{\omega}}_{t-1}}^{1} R_t(\omega)\mu'(\omega)\mathrm{d}\omega - \rho_{t-1}\left(1-\mu(\overline{\overline{\omega}}_{t-1})\right)\frac{b_{t-1}}{a_{t-1}} =$$

$$(1-\theta)\frac{\mu(\overline{\overline{\omega}}_{t-1})-\mu(\overline{\omega}_{t-1})}{1-\mu(\overline{\omega}_{t-1})}\frac{B_{t-1}}{A_{t-1}} + \frac{(R_{t-1}^B-\xi)}{A_{t-1}}D_{t-1} +$$

$$\left(1-\mu(\overline{\overline{\omega}}_{t-1})\right)\frac{\left(A_{t-1}+\frac{B_{t-1}}{1-\mu(\overline{\omega}_{t-1})}\right)}{A_{t-1}}\int_{\overline{\overline{\omega}}_{t-1}}^{1} R_t(\omega)\frac{\mu'(\omega)}{1-\mu(\overline{\overline{\omega}}_{t-1})}\mathrm{d}\omega =$$

$$(1-\theta)\frac{\mu(\overline{\overline{\omega}}_{t-1})-\mu(\overline{\omega}_{t-1})}{1-\mu(\overline{\omega}_{t-1})}\frac{B_{t-1}}{A_{t-1}} + \frac{(R_{t-1}^B-\xi)}{A_{t-1}}D_{t-1} + \frac{B_{t-1}^{tot}}{A_{t-1}}R_t, \quad \text{(C.32)}$$

where we use $a_{t-1}(\omega) = a_{t-1}$ and $b_{t-1}(\omega) = b_{t-1}$ for all ω and equations (C.10), (C.11), (C.14), and (C.19). Moreover, we use that

$$\frac{b_{t-1}}{a_{t-1}} = \frac{\mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})} = \frac{B_{t-1}}{A_{t-1}}\frac{1}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})} \quad \text{(C.33)}$$

by combining equations (C.10), (C.11), and (C.9).

We can re-write equation (C.32) as follows:

$$R_t^A A_{t-1} = B_{t-1}^{tot}R_t + (R_{t-1}^B-\xi)D_{t-1} + (1-\theta)\frac{\mu(\overline{\overline{\omega}}_{t-1})-\mu(\overline{\omega}_{t-1})}{1-\mu(\overline{\omega}_{t-1})}B_{t-1}. \quad \text{(C.34)}$$

C.5 Diversion Function: Limiting Case

Consider

$$I(\overline{\omega}_t, \overline{\overline{\omega}}_t) = \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \Theta_t(\omega)\frac{\mu'(\omega)}{\mu(\overline{\overline{\omega}}_t)-\mu(\overline{\omega}_t)}\mathrm{d}\omega. \quad \text{(C.35)}$$

Here we seek to find $\lim_{\overline{\overline{\omega}}_t \to \overline{\omega}_t} I(\overline{\omega}_t, \overline{\overline{\omega}}_t)$. We can write $I(\overline{\omega}_t, \overline{\overline{\omega}}_t)$ as a ratio, i.e.,

$$I(\overline{\omega}_t, \overline{\overline{\omega}}_t) = \frac{\int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \Theta_t(\omega)\mu'(\omega)\mathrm{d}\omega}{\int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \mu'(\omega)\mathrm{d}\omega}. \tag{C.36}$$

Both numerator and denominator go to zero as $\overline{\overline{\omega}}_t \to \overline{\omega}_t$. Applying L'Hopital's rule

$$\lim_{\overline{\overline{\omega}}_t \to \overline{\omega}_t} = \frac{\int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \Theta_t(\omega)\mu'(\omega)\mathrm{d}\omega}{\int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \mu'(\omega)\mathrm{d}\omega} = \lim_{\overline{\overline{\omega}}_t \to \overline{\omega}_t} \frac{\Theta_t(\overline{\overline{\omega}}_t)\mu'(\overline{\overline{\omega}}_t)}{\mu'(\overline{\overline{\omega}}_t)} = \Theta_t(\overline{\omega}_t), \tag{C.37}$$

as long as $\mu'(\overline{\omega}_t) \neq 0$ (a condition satisfied for our baseline calibration).

C.6 Beta Distribution: Aggregate Output

Plugging the Beta(η_1, η_2) density into equation (C.24), the key integral is

$$\int_{\overline{\overline{\omega}}_t}^{1} \omega \mu'_{\eta_1,\eta_2}(\omega)\,\mathrm{d}\omega = \int_{\overline{\overline{\omega}}_t}^{1} \frac{\omega^{\eta_1}(1-\omega)^{\eta_2-1}}{B(\eta_1,\eta_2)}\,\mathrm{d}\omega. \tag{C.38}$$

Multiplying and dividing by $B(\eta_1+1, \eta_2)$ identifies the integrand as a Beta(η_1+1, η_2) density kernel:

$$\begin{aligned}\int_{\overline{\overline{\omega}}_t}^{1} \omega \mu'_{\eta_1,\eta_2}(\omega)\,\mathrm{d}\omega &= \frac{B(\eta_1+1,\eta_2)}{B(\eta_1,\eta_2)} \int_{\overline{\overline{\omega}}_t}^{1} \frac{\omega^{\eta_1}(1-\omega)^{\eta_2-1}}{B(\eta_1+1,\eta_2)}\,\mathrm{d}\omega \\ &= c_\alpha\Big(1-\mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}}_t)\Big),\end{aligned} \tag{C.39}$$

where $c_\alpha = B(\eta_1+1, \eta_2)/B(\eta_1, \eta_2) = \eta_1/(\eta_1+\eta_2)$ as defined in (40). Therefore

$$E\Big[\omega \,\Big|\, \omega \geq \overline{\overline{\omega}}_t\Big] = c_\alpha \frac{1-\mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}}_t)}{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t)}, \tag{C.40}$$

and substituting into (C.24):

$$Y_{t+1} = c_\alpha \, \frac{1 - \mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}}_t)}{1 - \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t)} \cdot Z_{t+1} \, K_t^\alpha \, H_{t+1}^{1-\alpha}. \tag{C.41}$$

C.7 Beta Distribution: Average Return on Capital

The average return on capital follows directly from the general result (C.27), which no longer contains distribution-specific terms. For completeness, we verify it from the individual return. Substituting $R_t(\omega) = (\kappa_t/Q_{t-1})\,\omega + (1-\delta)Q_t/Q_{t-1}$ and using the Beta-distribution result for $E[\omega|\omega \geq \overline{\overline{\omega}}_{t-1}]$ from Appendix C.6:

$$\begin{aligned} R_t &= \frac{\kappa_t}{Q_{t-1}} \, E\Big[\omega \,\Big|\, \omega \geq \overline{\overline{\omega}}_{t-1}\Big] + \frac{(1-\delta)Q_t}{Q_{t-1}} \\ &= \frac{\kappa_t}{Q_{t-1}} \cdot c_\alpha \, \frac{1 - \mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}}_{t-1})}{1 - \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_{t-1})} + \frac{(1-\delta)Q_t}{Q_{t-1}}. \end{aligned} \tag{C.42}$$

Since $\kappa_t \, E[\omega \mid \omega \geq \overline{\overline{\omega}}_{t-1}] = \alpha Y_t/K_{t-1}$ (from equation (C.41) applied at period t), this simplifies to

$$R_t = \frac{\alpha \, Y_t}{Q_{t-1} \, K_{t-1}} + \frac{(1-\delta)\, Q_t}{Q_{t-1}}, \tag{C.43}$$

confirming (C.27). The distribution-specific ratio cancels because both Y_t and $\kappa_t \, E[\omega]$ contain the same Beta survival-function ratio.

C.8 Beta Distribution: Aggregate Equity Return

Plugging the functional form of the Beta distribution into equation (C.32), we get:

$$R_t^A = (1-\theta)\, \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_{t-1}) - \mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})} \, \frac{B_{t-1}}{A_{t-1}} + \frac{(R_{t-1}^B - \xi)}{A_{t-1}} D_{t-1} + \frac{B_{t-1}^{tot}}{A_{t-1}} R_t, \tag{C.44}$$

where R_t is given by equation (C.43).

C.9 Beta Distribution: Diversion Function

Expressing F_t from equation (39):

$$\theta = \frac{1}{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)} \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \omega^{\psi} F_t \frac{\omega^{\eta_1-1}(1-\omega)^{\eta_2-1}}{B(\eta_1,\eta_2)} \mathrm{d}\omega =$$

$$\frac{1}{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)} \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} F_t \frac{\omega^{\eta_1+\psi-1}(1-\omega)^{\eta_2-1}}{\Gamma(\eta_1)\Gamma(\eta_2)} \Gamma(\eta_1+\eta_2) \mathrm{d}\omega =$$

$$\frac{1}{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)} \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} F_t \frac{\omega^{\eta_1+\psi-1}(1-\omega)^{\eta_2-1}}{\Gamma(\eta_1+\psi)\Gamma(\eta_2)} \Gamma(\eta_1+\psi+\eta_2) \frac{\Gamma(\eta_1+\eta_2)\Gamma(\eta_1+\psi)}{\Gamma(\eta_1)\Gamma(\eta_1+\psi+\eta_2)} \mathrm{d}\omega =$$

$$\frac{F_t}{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)} \frac{\Gamma(\eta_1+\eta_2)\Gamma(\eta_1+\psi)}{\Gamma(\eta_1)\Gamma(\eta_1+\psi+\eta_2)} \left(\mu_{\eta_1+\psi,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1+\psi,\eta_2}(\overline{\omega}_t)\right) \quad \text{(C.45)}$$

Therefore,

$$F_t = \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)}{\mu_{\eta_1+\psi,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1+\psi,\eta_2}(\overline{\omega}_t)} \frac{\Gamma(\eta_1)\Gamma(\eta_1+\psi+\eta_2)}{\Gamma(\eta_1+\eta_2)\Gamma(\eta_1+\psi)} \theta \quad \text{(C.46)}$$

C.9.1 Limiting Case as $\overline{\overline{\omega}}_t \to \overline{\omega}_t$

In the limiting case as $\overline{\overline{\omega}}_t \to \overline{\omega}_t$, we can use the result of Appendix C.5

$$\lim_{\overline{\overline{\omega}}_t \to \overline{\omega}_t} \int_{\overline{\omega}_t}^{\overline{\overline{\omega}}_t} \Theta_t(\omega) \frac{\mu'(\omega)}{\mu(\overline{\overline{\omega}}_t) - \mu(\overline{\omega}_t)} \mathrm{d}\omega = \Theta_t(\overline{\omega}_t) = \overline{\omega}_t^{\psi} F_t. \quad \text{(C.47)}$$

Equating this result with θ, we find that, in the limiting case as $\overline{\overline{\omega}}_t \to \overline{\omega}_t$,

$$F_t = \frac{\theta}{\overline{\omega}_t^{\psi}}. \quad \text{(C.48)}$$

C.10 Equilibrium Conditions

This subsection lists the 29 equilibrium conditions that constitute a competitive equilibrium and identifies the derivation or earlier statement from which each is obtained.

Asset Euler equation. Equation (C.1) from Appendix C.1, rearranged to place λ_{ct} on the left-hand side:

$$\lambda_{ct} = \beta E_t \left\{ \lambda_{ct+1} R^A_{t+1} \right\}. \tag{C.49}$$

Consumption Euler equation. Equation (C.2) from Appendix C.1:

$$\beta E_t \left\{ \lambda_{ct+1} R^B_t \right\} = \frac{1}{C_t - \nu_{ct}}. \tag{C.50}$$

Labor supply. Compact form obtained by combining equations (C.3), (C.4), and (C.2) from Appendix C.1. Substituting $\beta E_t\{\lambda_{ct+1} R^B_t\} = \lambda_{ct}$ from (C.4) into (C.3) gives $-\vartheta H^\nu_t + \lambda_{ct} W_t = 0$. Applying $\lambda_{ct} = 1/(C_t - \nu_{ct})$ from (C.2) then yields:

$$W_t = \vartheta (C_t - \nu_{ct}) H^\nu_t. \tag{C.51}$$

Bond Euler equation. Equation (C.4) from Appendix C.1, rearranged to place λ_{ct} on the left-hand side:

$$\lambda_{ct} = \beta E_t \left\{ \lambda_{ct+1} R^B_t \right\}. \tag{C.52}$$

Aggregate output. Beta-distribution specialization of equation (C.24) from Appendix C.3, with time indices shifted by one period. The integral $\int_{\overline{\overline{\omega}}_{t-1}}^{1} \omega \, \mu'_{\eta_1,\eta_2}(\omega)\, \mathrm{d}\omega$ is evaluated in Appendix C.6 by rewriting $\omega \cdot \mu'_{\eta_1,\eta_2}(\omega)$ as a scalar multiple of the Beta$(\eta_1 + 1, \eta_2)$ density kernel, using $c_\alpha = B(\eta_1 + 1, \eta_2)/B(\eta_1, \eta_2) = \eta_1/(\eta_1 + \eta_2)$ from (40):

$$Y_t = c_\alpha \, \frac{1 - \mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}}_{t-1})}{1 - \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_{t-1})} \cdot Z_t K^\alpha_{t-1} H^{1-\alpha}_t. \tag{C.53}$$

Capital financing. Aggregate counterpart of the individual firm budget constraint (8) from Section 4.2.1. Derived in Appendix C.3 as equation (C.25) by integrating $b^{tot}_t(\omega) = Q_t k_t(\omega)$ over the mass of producing firms:

$$B^{tot}_t = Q_t K_t. \tag{C.54}$$

Labor demand. Aggregate form of the individual firm's optimal labor condition (11). Because $h_{t+1}(\omega) = \tilde{h}_{t+1}$ is uniform across producing firms, the aggregate wage equals the aggregate marginal product of labor:

$$W_t = (1-\alpha)\,\frac{Y_t}{H_t}. \tag{C.55}$$

Average return on capital. Equation (C.27) from Appendix C.3, verified for the Beta distribution in Appendix C.7. The distribution-specific terms cancel (see (C.27)), yielding the compact expression:

$$R_t = \frac{\alpha\, Y_t}{Q_{t-1}\, K_{t-1}} + \frac{(1-\delta)}{Q_{t-1}} Q_t. \tag{C.56}$$

Investment supply. Equation (21) from Section 4.3:

$$I_t^n = Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g. \tag{C.57}$$

Capital accumulation. Equation (23) from Section 4.3:

$$K_t = I_t^n + (1-\delta)K_{t-1}. \tag{C.58}$$

Capital-firm profits. Equation (25) from Section 4.3, the zero-profit condition for competitive capital-producing firms:

$$\Pi_t = Q_t Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g - I_t^g. \tag{C.59}$$

Government budget constraint. Equation (26) from Section 4.4:

$$T_t = B_t^G - R_{t-1}^B B_{t-1}^G. \tag{C.60}$$

Bond market identity. Equation (27) from Section 4.4:

$$B_t^G = B_t^H + D_t. \tag{C.61}$$

Household bond holdings. Equation (28) from Section 4.4. By assumption, government bonds are sold only to firms, so households hold none:

$$B_t^H = 0. \tag{C.62}$$

Haircut rebate to households. Equation (30) from Section 4.4:

$$\Xi_t = \xi D_{t-1}. \tag{C.63}$$

Lower cutoff condition. Equation (31) from Section 5.1. The expected return from producing at the marginal lending firm (productivity $\overline{\omega}_t$) equals the expected return from the outside option:

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{\kappa_{t+1}}{Q_t}\,\overline{\omega}_t + \frac{(1-\delta)}{Q_t}Q_{t+1}\right)\right] = E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B - \xi\right)\right]. \tag{C.64}$$

Aggregate lending. Equation (C.13) from Appendix C.3, with the generic cumulative distribution function μ replaced by its Beta-distribution counterpart μ_{η_1,η_2}:

$$L_t = \mu_{\eta_1,\eta_2}(\overline{\omega}_t)A_t. \tag{C.65}$$

Loan market clearing. Equation (C.9) from Appendix C.3:

$$B_t = L_t. \tag{C.66}$$

Aggregate diversion. Equation (C.16) from Appendix C.3, with μ replaced by μ_{η_1,η_2} throughout:

$$D_t = \left(\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)\right)A_t + \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}\theta B_t. \tag{C.67}$$

Aggregate total borrowing. Equation (C.19) from Appendix C.3, with μ replaced by μ_{η_1,η_2} throughout:

$$B_t^{tot} = \left(1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t)\right)A_t + \frac{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t)}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}B_t. \tag{C.68}$$

Equity return market clearing. Equation (C.29) from Appendix C.3, with μ replaced by μ_{η_1,η_2} throughout. The general form is derived in Appendix C.4 by aggregating the per-firm equity returns (20), (17), and (15) across the three firm segments; Appendix C.8 verifies the

Beta-specific version:

$$R_t^A A_{t-1} = B_{t-1}^{tot} R_t + (R_{t-1}^B - \xi) D_{t-1} + (1-\theta)\,\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_{t-1}) - \mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})} B_{t-1}. \tag{C.69}$$

Goods market clearing. Equation (C.30) from Appendix C.3:

$$Y_t = C_t + I_t^g. \tag{C.70}$$

Loan rate equation. Aggregate Beta-distribution form of equation (32) from Section 5.1. Individual-firm variables are replaced using $a_t = A_t$ from (C.10) and $b_t = B_t/(1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t))$ from (C.11); the diversion function $\Theta_t(\overline{\omega}_t)b_t$ is substituted using the functional form (38), giving $F_t\overline{\omega}_t^{\psi} B_t/(1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t))$, where F_t is defined in equation (C.77) below:

$$\left(\rho_t \frac{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t)}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)} + (1-\theta)\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}\right) A_t = \left(R_t^B - \xi\right)\left(A_t + F_t\overline{\omega}_t^{\psi}\frac{B_t}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}\right). \tag{C.71}$$

Upper cutoff condition. Aggregate Beta-distribution form of equation (33) from Section 5.1. The same substitutions as for the loan rate equation above are applied: a_t, b_t, and $\Theta_t(\overline{\overline{\omega}}_t)$ are replaced using (C.10), (C.11), and (38) respectively:

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_t^B - \xi\right)\left(A_t + F_t\overline{\overline{\omega}}_t^{\psi}\frac{B_t}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}\right)\right] = E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\left(\frac{1}{Q_t}\alpha\ Z_{t+1}\left(\frac{H_{t+1}}{K_t}\right)^{1-\alpha}\overline{\overline{\omega}}_t + \frac{(1-\delta)}{Q_t}Q_{t+1}\right)\left(A_t + \frac{B_t}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}\right) - \rho_t\frac{B_t}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)}\right)\right]. \tag{C.72}$$

Capital firm first-order condition (Tobin's q). Equation (C.5) from Appendix C.2, derived from the capital-producing firm's optimization problem (24) stated in Section 4.3:

$$0 = E_t\left\{\begin{array}{c} Q_t Z_{It}\left[-\phi\left(\frac{I_t^g}{I_{t-1}^g}-1\right)\frac{1}{I_{t-1}^g}\right] I_t^g + Q_t Z_{It}\left[1-\frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g}-1\right)^2\right] - 1 \\ +\beta\frac{\lambda_{ct+1}}{\lambda_{ct}} Q_{t+1} Z_{It+1}\phi\left(\frac{I_{t+1}^g}{I_t^g}-1\right)\left(\frac{I_{t+1}^g}{I_t^g}\right)^2 \end{array}\right\}. \tag{C.73}$$

Aggregate TFP process. Equation (5) from Section 4.2:

$$\log Z_t = \rho_z \log Z_{t-1} + \varepsilon_t^z. \tag{C.74}$$

Investment-specific technology process. Equation (22) from Section 4.3:

$$\ln(Z_{It}) = \rho_I \ln(Z_{It-1}) + \varepsilon_{It}. \tag{C.75}$$

Consumption demand shock process. Equation (3) from Section 4.1:

$$\nu_{ct} = \rho_\nu \nu_{ct-1} + \varepsilon_{\nu t}. \tag{C.76}$$

Diversion scaling factor (F_t). Beta-distribution closed form of the integral definition (39) from Section 5.1. The derivation in Appendix C.9 rewrites $\omega^\psi \mu'_{\eta_1,\eta_2}(\omega)$ as a scalar multiple of the Beta$(\eta_1 + \psi, \eta_2)$ density kernel, using $\Gamma(\eta_1 + \psi)/\Gamma(\eta_1)$ to convert the normalizing constants:

$$F_t = \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)}{\mu_{\eta_1+\psi,\eta_2}(\overline{\overline{\omega}}_t) - \mu_{\eta_1+\psi,\eta_2}(\overline{\omega}_t)} \frac{\Gamma(\eta_1)\Gamma(\eta_1 + \psi + \eta_2)}{\Gamma(\eta_1 + \eta_2)\Gamma(\eta_1 + \psi)}\theta \tag{C.77}$$

and ω follows the Beta distribution with parameters η_1 and η_2, i.e.,

$$\mu'_{\eta_1,\eta_2}(\omega) = \frac{\omega^{\eta_1-1}(1-\omega)^{\eta_2-1}}{B(\eta_1, \eta_2)},$$

where $B(\eta_1, \eta_2) = \frac{\Gamma(\eta_1)\Gamma(\eta_2)}{\Gamma(\eta_1+\eta_2)}$ and Γ is the Gamma function.

D Steady-State Conditions

We derived 29 equilibrium conditions for 29 endogenous variables:

$\lambda_{ct}, R_t^A, C_t, H_t, W_t, R_t^B, Y_t, \overline{\omega}_t, \overline{\overline{\omega}}_t, Z_t, Z_{It}, \nu_{ct}, K_t, B_t^{tot}, Q_t, R_t, I_t^n, I_t^g, \Pi_t, T_t, B_t^G, B_t^H, D_t, \Xi_t, L_t, A_t, B_t, \rho_t,$ and F_t.

Next, we consider a strategy for finding the non-stochastic steady state. The strategy will involve guessing the values of $\overline{\omega}$, $\overline{\overline{\omega}}$, and R and iterating to a fixed point.

Given our guesses, we can find

$$F = \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}) - \mu_{\eta_1,\eta_2}(\overline{\omega})}{\mu_{\eta_1+\psi,\eta_2}(\overline{\overline{\omega}}) - \mu_{\eta_1+\psi,\eta_2}(\overline{\omega})} \frac{\Gamma(\eta_1)\Gamma(\eta_1+\psi+\eta_2)}{\Gamma(\eta_1+\eta_2)\Gamma(\eta_1+\psi)}\theta \tag{D.1}$$

Continue by fixing

$$H = H^{ss} = 1. \tag{D.2}$$

We can support any choice of H by choosing ϑ appropriately. For example, from the intratemporal condition (C.51) evaluated in steady state, we can find that $\vartheta = \frac{W}{C\,H^{\nu}}$ supports $H = 1$ in the steady state.

From equation (C.52), we can see that in the steady state

$$R^B = \frac{1}{\beta}. \tag{D.3}$$

Similarly, from equation (C.49)

$$R^A = \frac{1}{\beta}. \tag{D.4}$$

From equation (C.74)

$$Z = 1. \tag{D.5}$$

From equation (C.75)

$$Z_I = 1. \tag{D.6}$$

From equation (C.76)

$$\nu_c = 0. \tag{D.7}$$

From equation (C.73)

$$Q = 1. \tag{D.8}$$

From equation (C.56) in the steady state ($Q = 1$, $Z = 1$) and using (C.53) to substitute $\alpha Y/K = \alpha\,\Phi\,(H/K)^{1-\alpha}$ (where Φ is the steady-state TFP prefactor (42) evaluated at $\overline{\overline{\omega}}$):

$$R = \alpha\,\Phi\left(\frac{H}{K}\right)^{1-\alpha} + 1 - \delta,$$

we can solve for K:

$$K = H\left[\frac{\alpha\,\Phi}{R-(1-\delta)}\right]^{\frac{1}{1-\alpha}}, \tag{D.9}$$

where $\Phi = c_\alpha\,(1-\mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}}))/(1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}))$ from (42). Combining equation (C.54) and our result (D.8), we get

$$B^{tot} = K. \tag{D.10}$$

From equation (C.53)

$$Y = c_\alpha\,\frac{1-\mu_{\eta_1+1,\eta_2}(\overline{\overline{\omega}})}{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})}\cdot K^\alpha H^{1-\alpha} = \Phi\,K^\alpha H^{1-\alpha}. \tag{D.11}$$

From equation (C.55)

$$W = (1-\alpha)\,\frac{Y}{H}. \tag{D.12}$$

From equation (C.58)

$$I^n = \delta K. \tag{D.13}$$

From equation (C.57)

$$I^g = I^n. \tag{D.14}$$

Therefore, from equation (C.70), we can find

$$C = Y - I^g. \tag{D.15}$$

Therefore, from equation (C.50), we can find

$$\lambda_c = \frac{1}{C}. \tag{D.16}$$

Combining equations (C.65) and (C.66) and plugging the result into equation (C.68),

$$B^{tot} = (1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}))\,A + \frac{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}\mu_{\eta_1,\eta_2}(\overline{\omega})A$$

we can find A

$$A = \frac{B^{tot}}{1 - \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}) + \frac{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}\mu_{\eta_1,\eta_2}(\overline{\omega})}. \tag{D.17}$$

Therefore, from the combination of equations (C.65) and (C.66), we can find

$$B = \mu_{\eta_1,\eta_2}(\overline{\omega})A. \tag{D.18}$$

From equation (C.66)

$$L = B. \tag{D.19}$$

From equation (C.67)

$$D = \left(\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}) - \mu_{\eta_1,\eta_2}(\overline{\omega})\right) A + \frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}) - \mu_{\eta_1,\eta_2}(\overline{\omega})}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega})}\theta B. \tag{D.20}$$

From equation (C.62)

$$B^H = 0. \tag{D.21}$$

Therefore, from equation (C.61)

$$B^G = D. \tag{D.22}$$

From equation (C.60)

$$T = B^G - R^B B^G, \tag{D.23}$$

From equation (C.63)

$$\Xi = \xi D. \tag{D.24}$$

From equation (C.59)

$$\Pi = 0. \tag{D.25}$$

From equation (C.71)

$$\left(\rho\frac{1 - \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega})} + (1-\theta)\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}) - \mu_{\eta_1,\eta_2}(\overline{\omega})}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega})}\right) A = \left(R^B - \xi\right)\left(A + \frac{F\overline{\omega}^{\psi} B}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega})}\right),$$

we can find

$$\rho = \frac{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})}\left(\left(R^B-\xi\right)\left(1+\frac{F\overline{\overline{\omega}}^{\psi}\mu_{\eta_1,\eta_2}(\overline{\omega})}{(1-\mu_{\eta_1,\eta_2}(\overline{\omega}))}\right)-(1-\theta)\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})-\mu_{\eta_1,\eta_2}(\overline{\omega})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}\right), \quad \text{(D.26)}$$

where we use equation (D.18) to substitute for $\frac{B}{A}$.

We have 3 equations to verify our 3 guesses

1. From equation (C.64), verify the guess for $\overline{\omega}$

$$\alpha\left(\frac{H}{K}\right)^{1-\alpha}\overline{\omega}+1-\delta = R^B-\xi,$$

which can be re-written as

$$0 = -1+\frac{R^B-\xi-(1-\delta)}{\alpha\left(\frac{H}{K}\right)^{1-\alpha}\overline{\omega}}. \quad \text{(D.27)}$$

2. From equation (C.72), verify the guess for $\overline{\overline{\omega}}$

$$\left(R^B-\xi\right)\left(A+F\overline{\overline{\omega}}^{\psi}\frac{B}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}\right) = \left(\alpha\left(\frac{H}{K}\right)^{1-\alpha}\overline{\overline{\omega}}+1-\delta\right)\left(A+\frac{B}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}\right)-\rho\frac{B}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})},$$

which can be re-written as

$$0 = -1+\left(\alpha\left(\frac{H}{K}\right)^{1-\alpha}\overline{\overline{\omega}}+1-\delta\right)\frac{1}{\rho}\left(\frac{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}{\mu_{\eta_1,\eta_2}(\overline{\omega})}+1\right)- \left(R^B-\xi\right)\frac{1}{\rho}\left(\frac{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}{\mu_{\eta_1,\eta_2}(\overline{\omega})}+F\overline{\overline{\omega}}^{\psi}\right), \quad \text{(D.28)}$$

where we use equation (D.18) to substitute for $\frac{A}{B}$

3. From equation (C.69), verify the guess for R

$$R^A A = B^{tot}R+(R^B-\xi)D+(1-\theta)\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})-\mu_{\eta_1,\eta_2}(\overline{\omega})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega})}B,$$

which can be re-written as

$$0 = -1 + \frac{B^{tot}R}{R^A A} + \frac{(R^B - \xi)D}{R^A A} + (1-\theta)\,\frac{\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}}) - \mu_{\eta_1,\eta_2}(\overline{\omega})}{(1-\mu_{\eta_1,\eta_2}(\overline{\omega}))R^A A}B. \tag{D.29}$$

E Data and Model Counterparts

We are interested in the segment for the producing firms. To calculate the 90^{th} percentile for that segment, we need to rescale the density so that it integrates to 1 over the segment. For the 10^{th} percentile, we solve for ω_{10} from

$$\int_{\overline{\overline{\omega}}}^{\omega_{10}} \frac{\mu'_{\eta_1,\eta_2}(\omega)}{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})} \mathrm{d}\omega = 0.10. \tag{E.30}$$

Working on the derivation:

$$\frac{1}{1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})}\left(\mu_{\eta_1,\eta_2}(\omega_{10}) - \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right) = 0.10.$$

Hence,

$$\omega_{10} = \mu^{-1}_{\eta_1,\eta_2}\left[0.10\left(1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right) + \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right], \tag{E.31}$$

where $\mu^{-1}_{\eta_1,\eta_2}$ is the quantile function of the Beta(η_1, η_2) distribution on $[0, 1]$.

Analogously for the 90^{th} percentile, the expression will be

$$\omega_{90} = \mu^{-1}_{\eta_1,\eta_2}\left[0.90\left(1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right) + \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right]. \tag{E.32}$$

Therefore, our target ratio is

$$\frac{\omega_{90}}{\omega_{10}} = \frac{\mu^{-1}_{\eta_1,\eta_2}\left[0.90\left(1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right) + \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right]}{\mu^{-1}_{\eta_1,\eta_2}\left[0.10\left(1-\mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right) + \mu_{\eta_1,\eta_2}(\overline{\overline{\omega}})\right]}. \tag{E.33}$$

F Canonical Representative-Firm RBC Model

In this appendix we show that equilibrium conditions of our baseline real model are equivalent to those of a canonical RBC model for the special case in which production is efficient, i.e. $\bar{\omega} = \bar{\bar{\omega}} = 1$ and $\xi = 0$.

We describe the canonical RBC model before tackling the equivalence proof.

F.1 The canonical RBC model

The model adopts the canonical RBC decentralization: the household directly owns the capital stock and rents it to the firm, rather than holding equity issued by firms as in the baseline. Capital-producing firms with investment adjustment costs are kept unchanged.

F.1.1 Firm's Problem

In each period t the representative goods-producing firm rents capital K_{t-1} from the household at the competitive rental rate r_t^K and hires labor H_t at the competitive wage W_t. It produces output

$$Y_t = Z_t K_{t-1}^{\alpha} H_t^{1-\alpha} \tag{F.1}$$

and maximizes static profit $\Pi_t^f = Y_t - r_t^K K_{t-1} - W_t H_t$. With constant returns to scale, equilibrium profits are zero and the first-order conditions give

$$W_t = (1-\alpha)\frac{Y_t}{H_t}, \tag{F.2}$$

$$r_t^K = \alpha \frac{Y_t}{K_{t-1}}. \tag{F.3}$$

F.1.2 Capital Producers' Problem

Competitive capital-producing firms solve exactly the same problem as in the baseline model (Section 4.3). They purchase I_t^g units of the consumption good, install

$$I_t^n = Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g \tag{F.4}$$

units of new capital, and sell it at the shadow price Q_t (Tobin's q), earning profits

$$\Pi_t = Q_t Z_{It} \left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g - I_t^g. \tag{F.5}$$

Profit maximization subject to the investment technology yields

$$0 = E_t \left\{ \begin{array}{c} Q_t Z_{It} \left[-\phi\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)\frac{1}{I_{t-1}^g}\right] I_t^g + Q_t Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] - 1 \\ +\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} Q_{t+1} Z_{It+1} \phi \left(\frac{I_{t+1}^g}{I_t^g} - 1\right)\left(\frac{I_{t+1}^g}{I_t^g}\right)^2 \end{array} \right\}. \tag{F.6}$$

F.1.3 Household's Problem

The representative household has the same KPR preferences as in the baseline model. It directly owns the capital stock K_{t-1}, rents it to the representative firm, receives the profits of capital-producing firms, and chooses consumption C_t, labor H_t, and the future capital stock K_t, purchasing the installed capital I_t^n from capital producers at price Q_t. The household solves

$$\max_{\{C_{t+\tau}, H_{t+\tau}, I_{t+\tau}^n, K_{t+\tau}\}_{\tau=0}^{\infty}} E_0 \sum_{\tau=0}^{\infty} \beta^\tau \left[\ln(C_{t+\tau} - \nu_{c,t+\tau}) - \frac{\vartheta}{1+\nu}(H_{t+\tau})^{1+\nu}\right], \tag{F.7}$$

subject to the flow budget constraint

$$C_t + Q_t I_t^n = W_t H_t + r_t^K K_{t-1} + \Pi_t \tag{F.8}$$

and the capital accumulation equation

$$K_t = I_t^n + (1-\delta) K_{t-1}, \tag{F.9}$$

where the household purchases the installed capital I_t^n produced through the capital producers' technology (F.4) at the price Q_t, and Π_t denotes the profits of capital-producing firms rebated to the household.

Let λ_{ct} be the Lagrange multiplier on (F.8) and Q_t the multiplier (in units of the consumption good) on (F.9). With KPR log utility in consumption, the first-order conditions are:

$$\lambda_{ct} = \frac{1}{C_t - \nu_{ct}}, \tag{F.10}$$

$$W_t = \vartheta(C_t - \nu_{ct})H_t^{\nu}, \tag{F.11}$$

$$\lambda_{ct} = \beta E_t\{\lambda_{ct+1}R_{t+1}\}\,, \tag{F.12}$$

where the gross return on capital is

$$R_{t+1} \equiv \frac{r^K_{t+1} + (1-\delta)Q_{t+1}}{Q_t} = \frac{\alpha Z_{t+1}(H_{t+1}/K_t)^{1-\alpha} + (1-\delta)Q_{t+1}}{Q_t}. \tag{F.13}$$

Gross investment I_t^g is chosen by the capital producers, whose optimality condition (F.6) determines the price Q_t of installed capital.

F.1.4 Equilibrium Conditions

The equilibrium consists of sequences for $\{Y_t, C_t, H_t, W_t, K_t, I_t^n, I_t^g, R_t, Q_t, \Pi_t, \lambda_{ct}, Z_t, Z_{It}\}$ satisfying the following conditions:

$$\lambda_{ct} = \beta E_t\{\lambda_{ct+1} R_{t+1}\}\,, \tag{F.14}$$

$$\lambda_{ct} = \frac{1}{C_t - \nu_{ct}}, \tag{F.15}$$

$$W_t = \vartheta (C_t - \nu_{ct}) H_t^{\nu}, \tag{F.16}$$

$$Y_t = Z_t K_{t-1}^{\alpha} H_t^{1-\alpha}, \tag{F.17}$$

$$W_t = (1-\alpha)\frac{Y_t}{H_t}, \tag{F.18}$$

$$R_t = \frac{1}{Q_{t-1}} \alpha\, Z_t \left(\frac{H_t}{K_{t-1}}\right)^{1-\alpha} + \frac{(1-\delta)}{Q_{t-1}} Q_t, \tag{F.19}$$

$$I_t^n = Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g, \tag{F.20}$$

$$K_t = I_t^n + (1-\delta) K_{t-1}, \tag{F.21}$$

$$\Pi_t = Q_t Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g - I_t^g, \tag{F.22}$$

$$0 = E_t \left\{ \begin{array}{c} Q_t Z_{It}\left[-\phi\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)\frac{1}{I_{t-1}^g}\right] I_t^g + Q_t Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] - 1 \\ +\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} Q_{t+1} Z_{It+1} \phi \left(\frac{I_{t+1}^g}{I_t^g} - 1\right)\left(\frac{I_{t+1}^g}{I_t^g}\right)^2 \end{array} \right\}, \tag{F.23}$$

$$Y_t = C_t + I_t^g, \tag{F.24}$$

$$\log Z_t = \rho_z \log Z_{t-1} + \varepsilon_t^z, \tag{F.25}$$

$$\ln(Z_{It}) = \rho_I \ln(Z_{It-1}) + \varepsilon_{It}, \tag{F.26}$$

$$\nu_{ct} = \rho_\nu \nu_{ct-1} + \varepsilon_t^{\nu}. \tag{F.27}$$

Equations (F.14)–(F.16) are the household's first-order conditions. Under KPR preferences, equation (F.15) gives $\lambda_{ct} = 1/(C_t - \nu_{ct})$ and equation (F.16) gives $W_t = \vartheta(C_t - \nu_{ct})H_t^{\nu}$, so a wealth effect on labor is present. Equation (F.27) is the consumption preference shock process. Equations (F.17)–(F.19) describe goods production and factor returns. Equations (F.20)–(F.22) govern investment and capital accumulation. Equation (F.23) is the capital producers' optimality condition (Tobin's q equation). Equations (F.24)–(F.25) are the goods market clearing condition and the TFP process, respectively. Equation (F.26) is the investment-efficiency shock process.

The household budget constraint (F.8) is not listed among the equilibrium conditions because it is redundant: by Walras' law, goods-market clearing (F.24) together with the factor-income identity $W_t H_t + r_t^K K_{t-1} = Y_t$ (which follows from CRS via equations (F.17)–(F.18)) implies $C_t + I_t^g = W_t H_t + r_t^K K_{t-1} + \Pi_t = W_t H_t + r_t^K K_{t-1}$, so the budget constraint forces $\Pi_t = 0$. This holds exactly only in the non-stochastic steady state, where $Q_t = 1$ and equation (F.22) confirms $\Pi = 0$; away from steady state, the household budget is simply dropped from the system as the redundant equation by Walras' law.

F.2 Equivalence with the Baseline Real Model

We show that the equilibrium conditions of the RBC model (Section F.1.4) are equivalent to those of the baseline real model (Appendix C.10) when $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$ for all t, in the sense that each RBC condition (F.14)–(F.27) coincides with the corresponding baseline condition once the financial variables absent from the RBC model are eliminated by substitution. We derive a key limit first and then proceed in three steps.

A key limit. The TFP prefactor $\Phi(\bar{\bar{\omega}})$ defined in equation (42) satisfies $\Phi(\bar{\bar{\omega}}) \to 1$ as $\bar{\bar{\omega}} \to 1$. To see this, note that for any shape parameter $\eta > 0$ the Beta c.d.f. satisfies

$$1 - \mu_{\eta,\eta_2}(\omega) \;\sim\; \frac{(1-\omega)^{\eta_2}}{B(\eta,\eta_2)\,\eta_2} \qquad \text{as } \omega \to 1,$$

so both survival functions in Φ share the same $(1-\omega)^{\eta_2}$ tail and their ratio converges to the ratio of Beta normalisation constants:

$$\lim_{\bar{\bar{\omega}}\to 1} c_\alpha \frac{1-\mu_{\eta_1+1,\eta_2}(\bar{\bar{\omega}})}{1-\mu_{\eta_1,\eta_2}(\bar{\bar{\omega}})} \;=\; c_\alpha \cdot \frac{B(\eta_1,\eta_2)}{B(\eta_1+1,\eta_2)} \;=\; c_\alpha \cdot \frac{1}{c_\alpha} \;=\; 1, \tag{F.28}$$

and therefore $\Phi(\bar{\bar{\omega}}) \to 1$.

Step 1: Financial variables degenerate. At $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$ the CDF satisfies $\mu_{\eta_1,\eta_2}(1) = 1$, so the measures of the three firm segments collapse:

- Equation (C.62): $B_t^H = 0$.
- Equation (C.67): $D_t = \big(\mu(1) - \mu(1)\big)A_t + \frac{\mu(1)-\mu(1)}{1-\mu(1)}\theta B_t = 0$.
- Equations (C.61)–(C.60): $B_t^G = 0,\ T_t = 0$.
- Equation (C.63): $\Xi_t = \xi D_{t-1} = 0$.
- Equations (C.65)–(C.66): $L_t = B_t = \mu(1)A_t = A_t$.

For B_t^{tot}, equation (C.68) reads $B_t^{tot} = (1 - \mu(\bar{\bar{\omega}}_t))A_t + \frac{1-\mu(\bar{\bar{\omega}}_t)}{1-\mu(\bar{\omega}_t)}B_t$. Since $\bar{\omega}_t = \bar{\bar{\omega}}_t$, the second ratio equals unity for every $\bar{\omega}_t < 1$; taking the limit as $\bar{\omega}_t \to 1$ (and using $B_t = A_t$) gives

$$B_t^{tot} = A_t. \tag{F.29}$$

Combined with equation (C.54) this yields $A_t = Q_t K_t$.

Substituting $D_{t-1} = 0$, $B_{t-1}^{tot} = A_{t-1}$, and $\mu(\bar{\bar{\omega}}_{t-1}) - \mu(\bar{\omega}_{t-1}) = 0$ into equation (C.69):

$$R_t^A A_{t-1} = A_{t-1} R_t \implies R_t^A = R_t. \tag{F.30}$$

Step 2: Real conditions. Table A.1 lists each RBC equilibrium condition alongside the Appendix B condition(s) from which it follows at $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$.

Step 3: Remaining baseline conditions are vacuous or determine financial variables only.

- *Government block* ((C.60)–(C.63)): shown in Step 1 to give $T_t = \Xi_t = B_t^G = B_t^H = 0$, trivially consistent.
- *Aggregation* ((C.65)–(C.68) and (C.69)): used in Step 1 to derive $B_t^{tot} = A_t = Q_t K_t$ and $R_t^A = R_t$, both already absorbed into conditions (F.14) and (F.19).
- *Bond Euler equation* (C.52): $\lambda_{ct} = \beta E_t\{\lambda_{ct+1} R_t^B\}$ pins the financial variable R_t^B given the real allocation (which determines λ_{ct} through (F.15)).
- *Lower cutoff* (C.64): at $\bar{\omega}_t = 1$, the left-hand side equals $E_t[\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} R_{t+1}] = 1$ (by (F.14)), so the condition becomes $\lambda_{ct} = \beta E_t\{\lambda_{ct+1}(R_t^B - \xi)\}$. This must hold simultaneously with the bond Euler (C.52), which gives $\lambda_{ct} = \beta E_t\{\lambda_{ct+1} R_t^B\}$. Subtracting yields $\xi\, \beta E_t\{\lambda_{ct+1}\} = 0$; since $\lambda_{ct+1} > 0$, this requires $\xi = 0$. The full RBC equivalence therefore requires both $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$ and $\xi = 0$. Under these conditions the lower cutoff

Table A.1: Correspondence between the RBC model's equilibrium conditions (Section F.1.4) and the baseline model's conditions (Appendix C.10) at $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$.

RBC eq.	Condition	Baseline source
(F.14)	$\lambda_{ct} = \beta E_t\{\lambda_{ct+1} R_{t+1}\}$	Substitute $R_t^A = R_t$ from (F.30) into (C.49).
(F.15)	$\lambda_{ct} = 1/(C_t - \nu_{ct})$	Equations (C.50)–(C.52): both give $\lambda_{ct} = \beta E_t\{\lambda_{ct+1} R_t^B\} = 1/(C_t - \nu_{ct})$.
(F.16)	$W_t = \vartheta(C_t - \nu_{ct})H_t^{\nu}$	Identical to (C.51).
(F.17)	$Y_t = Z_t K_{t-1}^{\alpha} H_t^{1-\alpha}$	Equation (C.53) with the limit (F.28) at $\bar{\bar{\omega}}_t = 1$.
(F.18)	$W_t = (1-\alpha)Y_t/H_t$	Identical to (C.55).
(F.19)	$R_t = \frac{1}{Q_{t-1}}\alpha Z_t\left(\frac{H_t}{K_{t-1}}\right)^{1-\alpha} + (1-\delta)\frac{Q_t}{Q_{t-1}}$	Equation (C.56); substitute (C.53) with $\Phi_t = 1$ (from (F.28)) to expand Y_t.
(F.20)	$I_t^n = Z_{It}[1 - \frac{\phi}{2}(\cdot)^2]I_t^g$	Identical to (C.57).
(F.21)	$K_t = I_t^n + (1-\delta)K_{t-1}$	Identical to (C.58).
(F.22)	$\Pi_t = Q_t Z_{It}[1 - \frac{\phi}{2}(\cdot)^2]I_t^g - I_t^g$	Identical to (C.59).
(F.23)	Tobin's q equation	Identical to (C.73).
(F.24)	$Y_t = C_t + I_t^g$	Identical to (C.70).
(F.25)	$\log Z_t = \rho_z \log Z_{t-1} + \varepsilon_t^z$	Identical to (C.74).
(F.26)	$\ln(Z_{It}) = \rho_I \ln(Z_{It-1}) + \varepsilon_{It}$	Identical to (C.75).
(F.27)	$\nu_{ct} = \rho_\nu \nu_{ct-1} + \varepsilon_t^{\nu}$	Identical to (C.76).

Note: "Identical" means the equation is structurally unchanged between the two models. Equations (C.60)–(C.63) (government block), (C.64) (lower cutoff), (C.67)–(C.68) (aggregation), (C.69) (equity return), (C.71)–(C.72) (cutoffs and loan rate), (C.76)–(C.77) (preference shock, diversion shock, and diversion scaling) are used only to determine financial variables absent from the RBC model or become vacuously satisfied at $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$; see Step 3 below.

degenerates to an identity and imposes no further restriction on real quantities.

- *Upper cutoff and loan rate* ((C.71)–(C.72)): at $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$ these conditions degenerate to equalities between ρ_t and $R_t^B - \xi$, jointly pinning the financial variables R_t^B and ρ_t without restricting real quantities.

Hence every condition determining a real quantity in the RBC model corresponds to a baseline condition, and the baseline conditions that do not appear in the RBC model either become trivially satisfied or serve only to determine financial variables. The equivalence holds when $\bar{\omega}_t = \bar{\bar{\omega}}_t = 1$ and $\xi = 0$; the haircut parameter ξ must vanish because the bond Euler (C.52) and the lower cutoff (C.64) are jointly consistent at $\bar{\omega}_t = 1$ only in that case. □

F.3 Steady-State Conditions

We solve for the non-stochastic steady state sequentially. For every variable that also appears in the baseline model we verify that the steady-state value coincides with the expression in Appendix D evaluated at $\bar{\omega} = \bar{\bar{\omega}} = 1$; the equivalence of equilibrium conditions established in Section F.2 implies that any verification of a dynamic condition in steady state automatically checks the corresponding Appendix C condition.

Sequential derivation. From equation (F.25),

$$Z = 1. \tag{F.31}$$

From equation (F.26),

$$Z_I = 1. \tag{F.32}$$

From equation (F.23) with $Z_I = 1$,

$$Q = 1. \tag{F.33}$$

Matches Appendix D (D.5)–(D.8).

From equation (F.14),

$$R = \frac{1}{\beta}. \tag{F.34}$$

Matches Appendix D (D.4) (the baseline gives $R^A = 1/\beta$; here $R_t^A = R_t$ by (F.30)).

Fix $H = H^{ss} = 1$ and choose $\vartheta = W/(C\,H^{\nu})$ to support this normalisation. From equation (F.19) in steady state ($Z = Q = 1$),

$$R = \alpha\left(\frac{H}{K}\right)^{1-\alpha} + (1-\delta), \tag{F.35}$$

so

$$K = H\left[\frac{R-(1-\delta)}{\alpha}\right]^{1/(\alpha-1)}. \tag{F.36}$$

Matches Appendix D (D.9): the prefactor Φ in (D.9) equals 1 at $\bar{\bar{\omega}} = 1$ by (F.28).

From equation (F.17),

$$Y = K^{\alpha}H^{1-\alpha}. \tag{F.37}$$

Matches Appendix D (D.11) at $\bar{\bar{\omega}} = 1$ by (F.28).

From equation (F.18),

$$W = (1-\alpha)\frac{Y}{H}. \tag{F.38}$$

Matches Appendix D (D.12).

From equations (F.20)–(F.21) in steady state,

$$I^n = \delta K, \qquad I^g = I^n. \tag{F.39}$$

Matches Appendix D (D.13)–(D.14).

From equation (F.24),

$$C = Y - I^g. \tag{F.40}$$

Matches Appendix D (D.15).

From equation (F.15),

$$\lambda_c = \frac{1}{C}. \tag{F.41}$$

Matches Appendix D (D.16).

From equation (F.22) with $Q = 1$ and $I^g = I^n$,

$$\Pi = 0. \tag{F.42}$$

Matches Appendix D (D.25).

G Model Without Strategic Default

This appendix develops a variant of the baseline model in which the outside option is unavailable. Without an outside option, no firm has an incentive to divert borrowed funds, so strategic default is impossible. The two-cutoff structure of the baseline model—which is needed to demarcate lenders, strategic defaulters, and producers—collapses to a single cutoff that separates lenders from firms that borrow and produce. The resulting model is structurally simpler than the baseline and provides a useful intermediate case between the full model and the frictionless canonical RBC model of Appendix F: financial intermediation and capital misallocation are still present, but the default margin is shut down.

The appendix proceeds as follows. Section G.1 describes which elements of the baseline model change and which are preserved. Section G.2 restates the firm problems. Section G.3 derives the single cutoff condition and the equilibrium loan rate. Section G.5 lists all equilibrium conditions. Section G.6 derives the non-stochastic steady state.

G.1 Changes Relative to the Baseline Model

The following building blocks of the baseline model are *unchanged*: the household problem (Section 4.1 and Appendix C.1); the capital-producing firm problem (Section 4.3 and Appendix C.2); the aggregate technology shock process (equation (5)); the investment-specific technology process (equation (22)); and the consumption preference shock process (equation (3)).

The following elements are *removed or simplified*:

No outside option. Firms cannot divert borrowed funds into the government bond. Parameters ξ (the haircut on the outside-option return) and θ (the average diversion share) play no role. The firm-specific diversion function $\Theta_t(\omega)$, its scaling factor F_t, and the aggregate diversion D_t are all undefined. Because no funds are ever diverted, the haircut rebate to households is $\Xi_t = 0$ for all t.

Two firm segments instead of three. Since strategic default is impossible, the inter-firm intermediation market sorts firms into only two groups depending on the realization of idiosyncratic productivity ω:

1. Firms with $\omega < \overline{\omega}_t$ that lend their equity in the inter-firm market.

2. Firms with $\omega \geq \overline{\omega}_t$ that borrow and produce.

The defaulter segment $\overline{\omega}_t \leq \omega \leq \overline{\overline{\omega}}_t$ of the baseline model is absent. The upper cutoff $\overline{\overline{\omega}}_t$ does not arise.

Simplified government sector. Because no firm uses the government bond as an outside option, aggregate firm holdings of government bonds are $D_t = 0$. Retaining the assumption from the baseline that households do not hold government bonds, $B_t^H = 0$, the government budget constraint (equation (27)) implies $B_t^G = 0$ and hence $T_t = 0$ for all t. The government sector drops out of the model entirely. The risk-free rate R_t^B nonetheless remains well defined as the shadow rate implied by the household's bond Euler equation.

Loan rate. Because all inter-firm loans are repaid with certainty, the inter-firm loan market is free of default risk. To preserve the financial friction from the baseline model, we retain $\xi > 0$ as a parameter that separates lenders and borrowers via a single cutoff $\omega = \overline{\omega}_t$, defined by condition (31) as in the baseline. In the present specification the default margin is switched off, so the second cutoff is not operational; this design allows us to isolate the role of the default margin. In the version of the model without an outside option, $\xi > 0$ functions solely as an intermediation-cost parameter that governs the intermediation spread in the inter-firm lending market, rather than also acting as a haircut on the outside option as in the baseline specification. Competition in the inter-firm market drives

$$\rho_t = R_t^B - \xi. \tag{G.1}$$

Keeping ξ at its baseline-calibrated value facilitates direct comparison across models.

The household budget constraint simplifies to

$$C_t + A_t = R_t^A A_{t-1} + W_t H_t + \Pi_t, \tag{G.2}$$

dropping the terms B_t^H, $R_{t-1}^B B_{t-1}^H$, T_t, and Ξ_t that appear in equation (2) of the baseline.

G.2 Firm Problems

G.2.1 Firms Choosing to Produce ($\omega \geq \overline{\omega}_t$)

The problem of a producing firm is identical to that in Section 4.2.1. Each producing firm finances its capital purchase by combining household equity $a_t(\omega) = a_t$ with an inter-firm loan $b_t(\omega) = b_t$, so that total financing is $b_t^{tot}(\omega) = a_t + b_t = Q_t k_t(\omega)$. The first-order conditions for labor and capital are equations (12) and (13) of the baseline model, reproduced here for convenience:

$$W_t = (1-\alpha)\frac{Y_t}{H_t}, \tag{G.3}$$

$$R_{t+1}(\omega) = \frac{\kappa_{t+1}}{Q_t}\,\omega + \frac{(1-\delta)}{Q_t} Q_{t+1}, \tag{G.4}$$

where $\kappa_{t+1} \equiv \alpha Z_{t+1}(H_{t+1}/K_t)^{1-\alpha}$ is the common factor defined in equation (13) of Section 4.2.1. The return paid to households by a producing firm is

$$R_{t+1}^A(\omega) = \frac{R_{t+1}(\omega)(a_t + b_t) - \rho_t\, b_t}{a_t}, \tag{G.5}$$

identical to equation (15) of the baseline model.

G.2.2 Firms Choosing to Lend ($\omega < \overline{\omega}_t$)

The problem of a lending firm simplifies substantially relative to Section 4.2.3. Because all borrowers repay at rate ρ_t, the lending firm faces no default loss. Its problem is

$$\max_{l_t(\omega)|\omega} E_t\left[\beta \frac{\lambda_{ct+1}}{\lambda_{ct}}\Big(\rho_t\, l_t(\omega) - R_{t+1}^A(\omega)\, a_t(\omega)\Big)\right], \tag{G.6}$$

subject to $l_t(\omega) \leq a_t(\omega)$. Since revenues are increasing in $l_t(\omega)$, the constraint holds with equality: $l_t(\omega) = a_t(\omega)$. The zero-profit condition that holds under every state of nature is

$$R^A_{t+1}(\omega) = \rho_t \qquad \text{for all } \omega < \overline{\omega}_t. \tag{G.7}$$

Compare this to the baseline equation (20): the two terms involving the fraction of defaulting borrowers vanish, leaving a simple pass-through of the loan rate to households.

G.3 The Single Cutoff Point and the Loan Rate

With only lenders and producers, a single cutoff $\overline{\omega}_t$ characterizes the equilibrium sorting of firms. A firm with productivity exactly at $\overline{\omega}_t$ is indifferent between lending—earning ρ_t on its equity—and borrowing to produce—earning $R_{t+1}(\overline{\omega}_t)$ on total funds net of loan repayments. Formally,

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\,\rho_t\right] = E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R_{t+1}(\overline{\omega}_t)\frac{a_t+b_t}{a_t} - \rho_t\frac{b_t}{a_t}\right)\right].$$

Rearranging and collecting terms in ρ_t:

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\,\rho_t\right] = E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\,R_{t+1}(\overline{\omega}_t)\right]. \tag{G.8}$$

Substituting $\rho_t = R^B_t - \xi$ from equation (G.1) and expanding $R_{t+1}(\overline{\omega}_t)$ using equation (13), the *single cutoff condition* is

$$\boxed{E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(\frac{\kappa_{t+1}}{Q_t}\,\overline{\omega}_t + \frac{(1-\delta)}{Q_t}Q_{t+1}\right)\right] = E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left(R^B_t - \xi\right)\right].} \tag{G.9}$$

This condition is *identical* to the baseline lower-cutoff condition (31): ξ enters in exactly the same way in both models, which preserves a clean mapping between the two calibrations.

The following proposition establishes that equation (G.9) together with (G.1) is sufficient to support the two-segment equilibrium.

Proposition 7. *Suppose $R_{t+1}(\omega)$ is strictly increasing in ω and $\rho_t = R^B_t - \xi$ with $\xi \geq 0$.*

Assuming an interior solution exists (i.e., there is a $\overline{\omega}_t \in (0,1)$ *satisfying equation* (G.9)*), then equation* (G.9) *defines a unique cutoff* $\overline{\omega}_t$ *such that*

1. *firms with* $\omega < \overline{\omega}_t$ *strictly prefer to lend;*

2. *firms with* $\omega \geq \overline{\omega}_t$ *weakly prefer to borrow and produce.*

Proof. Equation (G.8) equates the expected discounted return from lending, ρ_t (independent of ω), with the expected discounted return from producing at $\overline{\omega}_t$. Since $R_{t+1}(\omega)$ is strictly increasing in ω, the return from producing is strictly higher than the return from lending for all $\omega > \overline{\omega}_t$ and strictly lower for all $\omega < \overline{\omega}_t$. Because there is no outside option, no firm defaults strategically; the only alternatives are lending and producing, and the sorting follows directly from the monotonicity of $R_{t+1}(\omega)$. □

Remark. The baseline model requires two conditions (equations (31), (32)) and a verification that no firm deviates to the outside option (Proposition 3). Here a single condition (G.9) suffices because the outside option is unavailable and the only relevant comparison is lending versus producing.

G.4 Aggregation and Equilibrium Definition

With two firm segments instead of three, the aggregation of individual decisions into aggregate variables simplifies relative to Appendix C.3. The steps below follow the same procedure; we highlight the equations that change.

Inter-firm loan market clearing. A mass $\mu_{\eta_1,\eta_2}(\overline{\omega}_t)$ of firms lend and a mass $1-\mu_{\eta_1,\eta_2}(\overline{\omega}_t)$ borrow. The aggregate supply and demand of inter-firm loans are

$$L_t = \int_0^{\overline{\omega}_t} l_t(\omega)\, \mu'(\omega)\, \mathrm{d}\omega = \mu_{\eta_1,\eta_2}(\overline{\omega}_t)\, A_t, \tag{G.10}$$

$$B_t = \int_{\overline{\omega}_t}^{1} b_t(\omega)\, \mu'(\omega)\, \mathrm{d}\omega, \tag{G.11}$$

and market clearing requires $L_t = B_t$. Individual borrowing is therefore $b_t = B_t/\Big(1 - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)\Big)$.

Aggregate total borrowing. Every producing firm uses total funds $b_t^{tot}(\omega) = a_t + b_t$. Integrating over the producing segment:

$$B_t^{tot} = \Big(1 - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)\Big)A_t + B_t. \tag{G.12}$$

Compare this with the baseline equation (C.19): when $\overline{\overline{\omega}}_t = \overline{\omega}_t$ the two expressions coincide, confirming that the no-default model is the special case in which the defaulter segment has zero mass.

Aggregate equity return. Using equations (G.7) and (G.5) and integrating over both segments:

$$\begin{aligned} R_t^A A_{t-1} &= \int_0^{\overline{\omega}_{t-1}} \rho_{t-1}\,\mu'(\omega)\,\mathrm{d}\omega\, A_{t-1} + \int_{\overline{\omega}_{t-1}}^{1} \Big[R_t(\omega)(A_{t-1}+b_{t-1}) - \rho_{t-1}b_{t-1}\Big]\mu'(\omega)\,\mathrm{d}\omega \\ &= \rho_{t-1}L_{t-1} + R_t B_{t-1}^{tot} - \rho_{t-1}B_{t-1} \\ &= R_t\, B_{t-1}^{tot}, \end{aligned} \tag{G.13}$$

where the last equality uses loan market clearing $L_{t-1} = B_{t-1}$. This is the simplification of the baseline equation (C.29) obtained by setting $D_{t-1} = 0$ and $\theta = 0$: without diversion losses, all capital returns flow through to households intact.

Aggregate output and average return on capital. With $\omega_i \geq \overline{\omega}_t$ for every producing firm, output and capital returns aggregate to the same Beta-distribution expressions as in Appendix C, replacing the baseline $\overline{\overline{\omega}}_t$ with $\overline{\omega}_t$ throughout:

$$Y_t = c_\alpha \frac{1 - \mu_{\eta_1+1,\eta_2}(\overline{\omega}_{t-1})}{1 - \mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})} \cdot Z_t K_{t-1}^{\alpha} H_t^{1-\alpha}, \tag{G.14}$$

$$R_t = \frac{\alpha\, Y_t}{Q_{t-1}\, K_{t-1}} + \frac{(1-\delta)}{Q_{t-1}}\, Q_t. \tag{G.15}$$

A competitive equilibrium of the no-default model is an allocation

$$\Big\{C_t,\, H_t,\, Y_t,\, K_t,\, B_t^{tot},\, I_t^n,\, I_t^g,\, \Pi_t,\, \overline{\omega}_t,\, Z_t,\, Z_{It},\, \nu_{ct},\, L_t,\, A_t,\, B_t\Big\}_{t=0}^{\infty}$$

together with a sequence of prices $\{\lambda_{ct},\, R_t^A,\, W_t,\, R_t^B,\, Q_t,\, R_t,\, \rho_t\}_{t=0}^{\infty}$ satisfying the 22 equations listed in the next subsection, given initial conditions K_0, I_0^g, B_0^{tot}, $\overline{\omega}_0$, A_0, B_0, Z_0, $Z_{I,0}$, ν_{c0}, R_0^B, Q_0, and exogenous processes $\{\varepsilon_t^z,\, \varepsilon_{It},\, \varepsilon_{\nu t}\}$.

G.5 Equilibrium Conditions

This subsection lists the 22 equilibrium conditions of the no-default model, identifying the source of each. Equations that are *identical* to their baseline counterparts in Appendix C.10 are noted as such; the remaining equations are new or simplified.

Asset Euler equation. Identical to baseline equation (C.49):

$$\lambda_{ct} = \beta\, E_t\Big\{\lambda_{ct+1}\, R_{t+1}^A\Big\}. \tag{G.16}$$

Consumption Euler equation. Identical to baseline equation (C.50):

$$\beta\, E_t\Big\{\lambda_{ct+1}\, R_t^B\Big\} = \frac{1}{C_t - \nu_{ct}}. \tag{G.17}$$

Labor supply. Identical to baseline equation (C.51):

$$W_t = \vartheta\,(C_t - \nu_{ct})\, H_t^{\nu}. \tag{G.18}$$

Bond Euler equation. Identical to baseline equation (C.52):

$$\lambda_{ct} = \beta\, E_t\Big\{\lambda_{ct+1}\, R_t^B\Big\}. \tag{G.19}$$

Aggregate output. Beta-distribution form of equation (G.14). The single cutoff $\overline{\omega}_{t-1}$ replaces the baseline $\overline{\overline{\omega}}_{t-1}$:

$$Y_t = c_\alpha \frac{1-\mu_{\eta_1+1,\eta_2}(\overline{\omega}_{t-1})}{1-\mu_{\eta_1,\eta_2}(\overline{\omega}_{t-1})}\cdot Z_t\, K_{t-1}^{\alpha}\, H_t^{1-\alpha}. \tag{G.20}$$

Capital financing. Identical to baseline equation (C.54):

$$B_t^{tot} = Q_t\, K_t. \tag{G.21}$$

Labor demand. Identical to baseline equation (C.55):

$$W_t = (1-\alpha)\,\frac{Y_t}{H_t}. \tag{G.22}$$

Average return on capital. Equation (G.15): the $E[\omega]$ factors cancel between output and average return, giving the same simplified form as baseline equation (C.56):

$$R_t = \frac{\alpha\, Y_t}{Q_{t-1}\, K_{t-1}} + \frac{(1-\delta)}{Q_{t-1}}\, Q_t. \tag{G.23}$$

Investment supply. Identical to baseline equation (C.57):

$$I_t^n = Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g. \tag{G.24}$$

Capital accumulation. Identical to baseline equation (C.58):

$$K_t = I_t^n + (1-\delta)\, K_{t-1}. \tag{G.25}$$

Capital-firm profits. Identical to baseline equation (C.59):

$$\Pi_t = Q_t\, Z_{It}\left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g - I_t^g. \tag{G.26}$$

Capital-firm first-order condition (Tobin's q). Identical to baseline equation (C.73):

$$0 = E_t\left\{\begin{array}{c} Q_t Z_{It}\left[-\phi\left(\dfrac{I_t^g}{I_{t-1}^g} - 1\right)\dfrac{1}{I_{t-1}^g}\right] I_t^g + Q_t Z_{It}\left[1 - \dfrac{\phi}{2}\left(\dfrac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] - 1 \\ +\beta\dfrac{\lambda_{ct+1}}{\lambda_{ct}} Q_{t+1} Z_{It+1}\,\phi\left(\dfrac{I_{t+1}^g}{I_t^g} - 1\right)\left(\dfrac{I_{t+1}^g}{I_t^g}\right)^2 \end{array}\right\}. \tag{G.27}$$

Single cutoff condition. Equation (G.9) from Section G.3, reproduced here. The expected discounted return from producing at the marginal lending firm equals the inter-firm lending rate $\rho_t = R_t^B - \xi$. This replaces the two baseline conditions (C.64) and (C.71)–(C.72):

$$E_t\left[\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left(\frac{\kappa_{t+1}}{Q_t} \overline{\omega}_t + \frac{(1-\delta)\, Q_{t+1}}{Q_t} \right)\right] = E_t\left[\beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left(R_t^B - \xi\right)\right]. \tag{G.28}$$

Aggregate lending. Equation (G.10), with $\overline{\omega}_t$ in place of the baseline $\overline{\omega}_t$ (same label, but here it is the *only* cutoff):

$$L_t = \mu_{\eta_1,\eta_2}(\overline{\omega}_t)\, A_t. \tag{G.29}$$

Loan market clearing. Identical to baseline equation (C.66):

$$B_t = L_t. \tag{G.30}$$

Aggregate total borrowing. Equation (G.12), the two-segment simplification of the baseline equation (C.68):

$$B_t^{tot} = \Big(1 - \mu_{\eta_1,\eta_2}(\overline{\omega}_t)\Big)\, A_t + B_t. \tag{G.31}$$

Equity return market clearing. Equation (G.13), the simplification of the baseline equation (C.69) with $D_{t-1} = 0$ and $\theta = 0$. All capital returns flow to households without any diversion losses:

$$R_t^A\, A_{t-1} = R_t\, B_{t-1}^{tot}. \tag{G.32}$$

Goods market clearing. Identical to baseline equation (C.70):

$$Y_t = C_t + I_t^g. \tag{G.33}$$

Loan rate. Equation (G.1):

$$\rho_t = R_t^B - \xi. \tag{G.34}$$

Aggregate TFP process. Identical to baseline equation (C.74):

$$\log Z_t = \rho_z\, \log Z_{t-1} + \varepsilon_t^z. \tag{G.35}$$

Investment-specific technology process. Identical to baseline equation (C.75):

$$\ln(Z_{It}) = \rho_I \ln(Z_{It-1}) + \varepsilon_{It}. \tag{G.36}$$

Consumption demand shock process. Identical to baseline equation (C.76):

$$\nu_{ct} = \rho_\nu \nu_{ct-1} + \varepsilon_{\nu t}. \tag{G.37}$$

Table A.2 summarizes the 22 equations and their relationship to the baseline model.

G.6 Steady-State Conditions

We solve for the non-stochastic steady state sequentially. The strategy mirrors Appendix F.3 for the canonical RBC model and Appendix D for the baseline model; it is considerably simpler than for the baseline because the government sector is absent and the loan rate is pinned directly by the risk-free rate.

The 22 endogenous variables in steady state are

$$\lambda_c,\ R^A,\ C,\ H,\ W,\ R^B,\ Y,\ \overline{\omega},\ Z,\ Z_I,\ \nu_c,\ K,\ B^{tot},\ Q,\ R,\ I^n,\ I^g,\ \Pi,\ L,\ A,\ B,\ \rho.$$

The solution uses one free initial guess: the single cutoff $\overline{\omega}$.

Step 1: Exogenous variables and prices that do not depend on $\overline{\omega}$.

From equations (G.35), (G.36), (G.37):

$$Z = 1, \qquad Z_I = 1, \qquad \nu_c = 0. \tag{G.38}$$

From equation (G.27) with $Z_I = 1$ and no investment growth:

$$Q = 1. \tag{G.39}$$

Table A.2: Equilibrium conditions of the no-default model and their relation to the baseline model (Appendix C.10)

#	Condition	No-default eq.	Relation to baseline model
1	Asset Euler	(G.16)	Identical to (C.49)
2	Consumption Euler	(G.17)	Identical to (C.50)
3	Labor supply	(G.18)	Identical to (C.51)
4	Bond Euler	(G.19)	Identical to (C.52)
5	Aggregate output	(G.20)	Baseline (C.53) with $\overline{\overline{\omega}}_{t-1} \to \overline{\omega}_{t-1}$
6	Capital financing	(G.21)	Identical to (C.54)
7	Labor demand	(G.22)	Identical to (C.55)
8	Average return on capital	(G.23)	Baseline (C.56) with $\overline{\overline{\omega}}_t \to \overline{\omega}_t$
9	Investment supply	(G.24)	Identical to (C.57)
10	Capital accumulation	(G.25)	Identical to (C.58)
11	Capital-firm profits	(G.26)	Identical to (C.59)
12	Tobin's q	(G.27)	Identical to (C.73)
13	Single cutoff	(G.28)	Replaces (C.64), (C.71), (C.72); identical to (C.64)
14	Aggregate lending	(G.29)	Baseline (C.65), same form
15	Loan market clearing	(G.30)	Identical to (C.66)
16	Aggregate total borrowing	(G.31)	Baseline (C.68) with $\overline{\overline{\omega}}_t \to \overline{\omega}_t$
17	Equity return MC	(G.32)	Baseline (C.69) with $D_{t-1} = 0$, $\theta = 0$
18	Goods market clearing	(G.33)	Identical to (C.70)
19	Loan rate	(G.34)	Simplified from (C.71); same ξ, no default term
20	TFP process	(G.35)	Identical to (C.74)
21	IST process	(G.36)	Identical to (C.75)
22	Consumption shock	(G.37)	Identical to (C.76)

Note: The baseline model has 29 endogenous variables and 29 equations (Appendix C.10). Removing the outside option eliminates seven variables ($\overline{\overline{\omega}}_t$, F_t, D_t, Ξ_t, T_t, B_t^G, B_t^H) and replaces three cutoff/loan-rate conditions by two simpler ones (the single cutoff (G.28) and the no-default loan rate (G.34)), yielding 22 equations in 22 unknowns.

From the bond Euler equation (G.19) in steady state:

$$R^B = \frac{1}{\beta}. \tag{G.40}$$

From equation (G.34):

$$\rho = R^B - \xi = \frac{1}{\beta} - \xi. \tag{G.41}$$

Step 2: Real allocation given $\overline{\omega}$.

Because no default occurs, the equity return market clearing condition (G.32) in steady state gives $R^A A = RB^{tot}$. Since the asset Euler equation (G.16) implies $R^A = 1/\beta$ and the financial aggregation (shown in Step 3 below) implies $A = B^{tot}$, we obtain

$$R = R^A = \frac{1}{\beta}. \tag{G.42}$$

This result also obtains in the baseline model (Appendix D, equation (D.4)) and the canonical RBC model (Appendix F.3, equation (F.34)).

With $R = 1/\beta$, $Q = 1$, and $Z = 1$, the simplified equation (G.23) in steady state gives

$$\frac{1}{\beta} = \frac{\alpha Y}{K} + (1-\delta) = \alpha\, \Phi^{nd}\left(\frac{H}{K}\right)^{1-\alpha} + (1-\delta),$$

where $\Phi^{nd} = \Phi^{nd}(\overline{\omega})$ is the no-default TFP prefactor (equation (G.14) with $Z = 1$, $H = 1$). Setting $H = H^{ss} = 1$ (choosing $\vartheta = W/(C\, H^{\nu})$ to support this normalization) and solving for K:

$$K = \left[\frac{\alpha\, \Phi^{nd}}{1/\beta - (1-\delta)}\right]^{1/(1-\alpha)}. \tag{G.43}$$

Compare with Appendix D equation (D.9): the formula is identical in structure, with $\Phi^{nd}(\overline{\omega})$ in place of $\Phi(\overline{\overline{\omega}})$.

The remaining real variables follow directly:

$$B^{tot} = Q\,K = K, \tag{G.44}$$

$$Y = \Phi^{nd}\,K^{\alpha} H^{1-\alpha}, \tag{G.45}$$

$$W = (1-\alpha)\,\frac{Y}{H}, \tag{G.46}$$

$$I^n = \delta\,K, \tag{G.47}$$

$$I^g = I^n = \delta\,K, \tag{G.48}$$

$$C = Y - I^g, \tag{G.49}$$

$$\lambda_c = \frac{1}{C}, \tag{G.50}$$

$$\Pi = 0. \tag{G.51}$$

Step 3: Financial variables given $\overline{\omega}$.

From the aggregate total borrowing equation (G.31) and the lending and market-clearing equations (G.29)–(G.30):

$$B^{tot} = \Big(1 - \mu_{\eta_1,\eta_2}(\overline{\omega})\Big)A + B,$$
$$B = L = \mu_{\eta_1,\eta_2}(\overline{\omega})\,A.$$

Substituting the second into the first and using (G.44):

$$K = B^{tot} = \Big(1 - \mu_{\eta_1,\eta_2}(\overline{\omega})\Big)A + \mu_{\eta_1,\eta_2}(\overline{\omega})\,A = A, \tag{G.52}$$

confirming that household equity equals total capital financing, and hence $R^A = R$ as used in Step 2.

Aggregate lending and borrowing are then

$$B = \mu_{\eta_1,\eta_2}(\overline{\omega})\,A = \mu_{\eta_1,\eta_2}(\overline{\omega})\,K, \tag{G.53}$$

$$L = B. \tag{G.54}$$

Step 4: Verifying the cutoff $\overline{\omega}$.

The steady-state cutoff condition follows from equation (G.28) with $Q = 1$, $Z = 1$, and the stochastic discount factor evaluated in steady state. Writing $R_{t+1}(\overline{\omega})$ from equation (13) at $Z = Q = 1$:

$$\kappa\,\overline{\omega} + (1-\delta) = \rho = R^B - \xi = \frac{1}{\beta} - \xi, \tag{G.55}$$

where $\kappa \equiv \alpha(H/K)^{1-\alpha}$ is the steady-state value of κ_{t+1} (see equation (13)). This condition states that the return from producing at the marginal lender equals the inter-firm lending rate ρ. Together with equation (G.43), this forms a scalar fixed-point problem in $\overline{\omega}$ alone, solvable by one-dimensional root finding.

Iterative algorithm.

1. Guess $\overline{\omega} \in (0, 1)$.

2. Compute K from equation (G.43) and Y from equation (G.45).

3. Check whether equation (G.55) holds. Specifically, evaluate the residual

$$\mathcal{R}(\overline{\omega}) \;\equiv\; \kappa(\overline{\omega})\,\overline{\omega} + (1-\delta) - \left(\frac{1}{\beta} - \xi\right), \tag{G.56}$$

   where $\kappa(\overline{\omega}) = \alpha(H/K(\overline{\omega}))^{1-\alpha}$, and update $\overline{\omega}$ until $\mathcal{R}(\overline{\omega}) = 0$.

4. Recover all remaining variables from Steps 2 and 3.

Unlike for the baseline model (Appendix D), which requires guessing *three* values ($\overline{\omega}$, $\overline{\overline{\omega}}$, R) and iterating on a three-dimensional fixed point, the no-default steady state requires guessing only $\overline{\omega}$, since $R = 1/\beta$ is determined analytically.

Comparison with the canonical RBC model. The frictionless canonical RBC model of Appendix F is recovered in the joint limit $\xi \to 0$ and $\overline{\omega} \to 1$. As $\xi \to 0$, the fixed-point condition (G.55) forces $\overline{\omega} \to 1$: the two equations become consistent only when the cutoff approaches unity, so that production is concentrated in the most productive firms (those with ω arbitrarily close to 1). In that limit the TFP prefactor $\Phi^{nd}(\overline{\omega}) \to 1$ by the limit established in equation (F.28) of Appendix F, and equations (G.43)–(G.51) coincide exactly with the

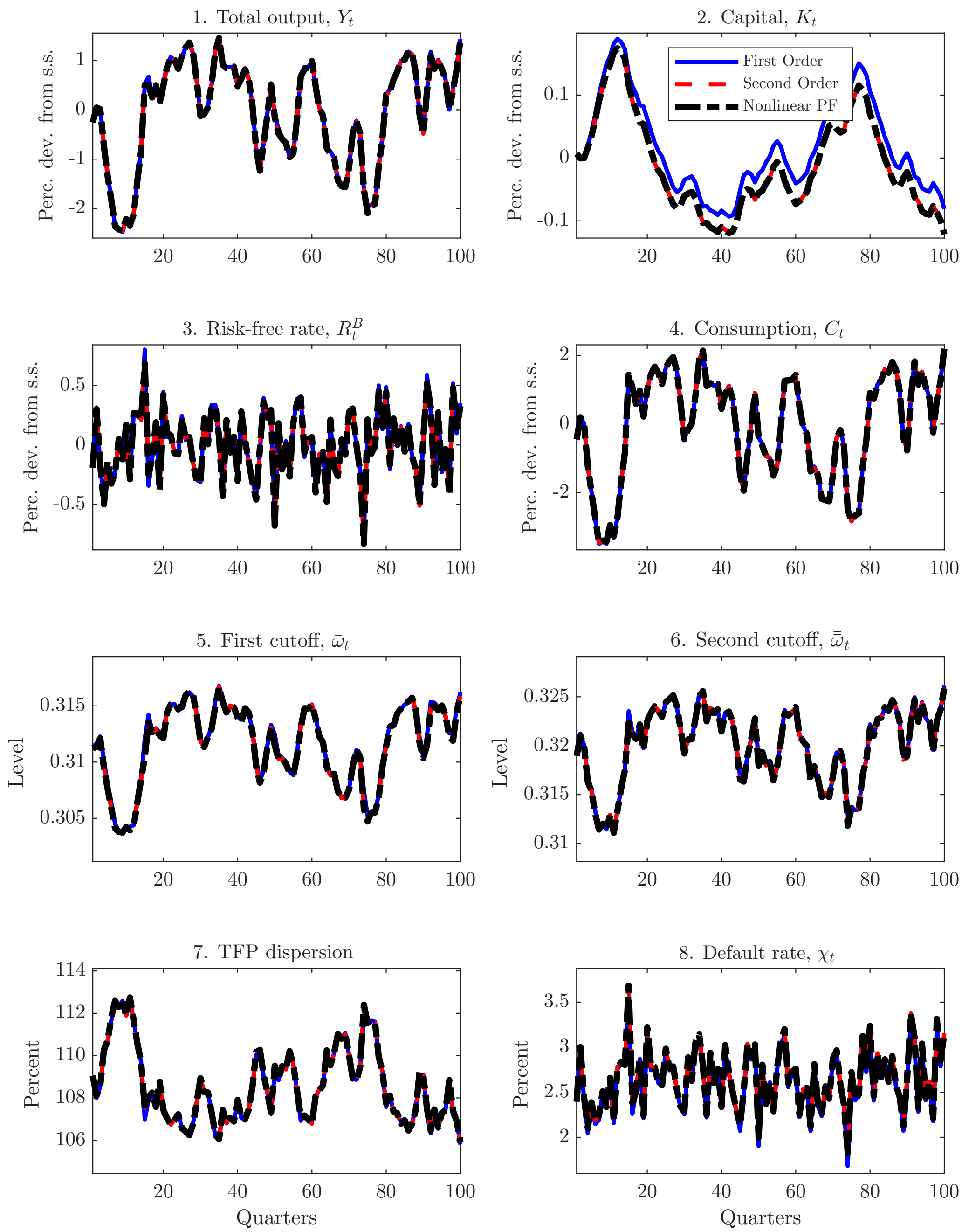
RBC steady-state equations (F.36)–(F.42). For $\xi > 0$, the no-default model provides an intermediate case between the baseline model (which adds strategic default on top of the same ξ) and the frictionless RBC model.

H Comparing Alternative Solution Methods

In the main body of the paper, all model calculations are based on an exact second-order perturbation solution. We show here a comparison between alternative solution methods, a first-order solution and a solution based on the fully nonlinear perfect-foresight shooting algorithm based on Fair and Taylor (1983) as implemented in Adjemian et al. (2026). We set the precision of the nonlinear solution close to machine precision. We have generalized the implementation of Adjemian et al. (2026) by taking the first guess of the solution paths from the first-order perturbation solution.

In a random sample, such as the ones that we simulate for our SMM estimation, the accumulation of deviations from the linearization point can render the first-order solution inaccurate. To compare the alternative solution methods, we draw a random sample including all the shocks with parameters set as described in Section 6. We find it reassuring that in Figure A.1, the second-order perturbation solution is closely consistent with the fully-nonlinear solution. Conditional on a realized shock sequence, the perfect-foresight path captures the full nonlinearity of the model's transition but abstracts from the stochastic risk terms that the second-order solution prices; the comparison therefore validates the second-order solution against nonlinearity in the deterministic transition rather than against all risk-adjustment terms. We prefer the second-order perturbation solution based on computational speed. By contrast, a noticeable gap opens up between the first-order perturbation solution and the perfect foresight solution. Note that, in the figure, the path from the first-order solution and from the perfect-foresight solution are shown in deviation from the non-stochastic steady state. For the second-order solution, the paths are shown in deviation from the stochastic fixed point.

Figure A.1: Comparing the evolution of a random sample across solutions methods



Note: "Omega bar" refers to $\overline{\omega}_t$ and "omega double bar" refers to $\overline{\overline{\omega}}_t$. For the first-order perturbation solution and the nonlinear perfect-foresight solution, the paths start from the non-stochastic steady state and are also in deviation from that point (where relevant). For the second-order solution the paths start from the stochastic steady state and are also in deviation from that point (where relevant).

References

Adjemian, S., Juillard, M., Karamé, F., Mutschler, W., Pfeifer, J., Ratto, M., Rion, N., & Villemot, S. (2026). *Dynare: Reference manual, version 7* (Dynare Working Papers No. 87). CEPREMAP.

Fair, R. C., & Taylor, J. B. (1983). Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models. *Econometrica*, *51*(4), 1169–1185.