Default Risk and Aggregate Fluctuations in an Economy with Production Heterogeneity

Aubhik Khan
Ohio State University

Tatsuro Senga
Queen Mary, University of London

Julia K. Thomas
Ohio State University

December 2021

ABSTRACT

We develop an economy with a time-varying measure of firms featuring persistent differences in productivity, capital and financial position. Firms fund their investments using retained earnings and non-contingent debt. Lenders offer firms individual loan schedules accounting for default risk and recovery rates, so unit borrowing costs rise with debt and fall with collateral value. As a result, firms with inadequate accumulated wealth to eliminate default risk have investment influenced by their financial positions. Larger firms with more collateral invest more than smaller firms with less collateral on average, although they draw from the same productivity distribution, implying insufficient capital allocated to small firms. That in turn discourages entry, reducing the number of production locations, since young firms tend to be small and heavily reliant on external finance. These inefficiencies together reduce aggregate productivity, capital, employment and GDP.

We consider business cycles driven by shocks to aggregate productivity and by financial shocks affecting firms' cash positions and borrowing terms. Our nonlinear equilibrium loan rate schedules drive countercyclical default risk and exit, alongside procyclical entry. As a negative TFP shock raises default risks, its effects are amplified by a slight rise in incumbent-firm capital misallocation and a modest loss of firms. A negative financial shock intensifies both channels and drives a recession similar to the U.S. 2007 recession in several respects. Its disparate effects on cash-poor firms concentrate investment declines among young, small firms, causing persistently high incumbent-firm misallocation. Its negative consequences for firm values also boost default and exit while suppressing entry, driving large, persistent losses in the number of firms. Measured TFP falls over several periods, as do employment, investment and GDP, and the ultimate declines in investment and employment are large relative to that in TFP. The recovery that follows is gradual given slow repair to aggregate capital, its distribution across firms and the stock of firms. The underlying distribution of firms matters for this. The steepness of the downturn and gradualism of recovery hinge on the relative strengths of incumbent misallocation changes versus damage along the extensive margin, which we trace to specific aspects of the average age-size distribution of firms.

Khan: mail@aubhik-khan.net, Senga: tatsuro.senga@gmail.com, Thomas: mail@juliathomas.net. We thank Fernando Alvarez, Hugo Hopenhayn, Urban Jermann, David Levine, Vincenzo Quadrini, Edward Prescott, Victor Rios-Rull, Immo Schott and Mathieu Taschereau-Dumouchel for helpful comments and suggestions. We also thank the National Science Foundation for research support under grant SES-1357725.

1 Introduction

Following the crisis in financial markets accompanying the 2007 recession in the U.S. and abroad, researchers have worked to better understand the extent to which the fall in real economic activity came in response to shocks originating in financial markets. Beyond the sharp contraction in aggregate lending and steep declines in aggregate investment and employment over this recession, disaggregated data reveal an unusual disparity in the impact on firms. Large declines in firm entry rates alongside losses disproportionately concentrated in small and young firms each indicate the importance of a shock affecting firms' ability to borrow.[1] At the same time, the corporate sector as a whole held unusually large levels of cash at the start of the recession.[2] These observations suggest that a quantitative analysis of the real effect of a financial shock should include a nontrivial distribution of firms that vary in their capital, debt, and retained earnings.

We develop a quantitative general equilibrium business cycle model accommodating the elements noted above while, at the same time, maintaining consistency with aggregate data. We introduce financial shocks worsening firms' cash positions and the underlying value of collateral securing loans to assess the relevance of such shocks over the most recent recession. More specifically, we explore the extent to which such a shock may have prompted the unusual reduction in lending volumes over the 2007 recession, the large declines in real macroeconomic variables relative to measured productivity, the disproportionate contractions among small firms, young firms and firms more reliant on external finance, as well as the unusual reduction in the number of operating firms.

Our environment is unique in featuring a distribution of firms over capital, debt and persistent idiosyncratic productivity, endogenous entry and exit, and shocks influencing default risk and the loan rate schedules offered to borrowers. It is further distinguished by qualitative differences in the business cycles it predicts following a financial shock versus a real shock. It is yet more rare in the sensitivity of its financially-led recession and recovery episodes to changes in the average underlying age-size distribution of firms, with no such sensitivity during productivity-shock driven episodes.

Firms in our model may default on their non-contingent debt. If a firm defaults, its creditors

---

[1] Using quarterly BED data maintained by the Bureau of Labor Statistics, Khan and Thomas (2013) find firms with fewer than 100 employees, as a whole, contracted employment twice as much as large firms and accounted for 46 percent of total employment losses, far beyond their 38 percent average employment share. The average relative size of a firm aged 0 through 5 years in the Business Dynamics Statistics database fell 4.2 percentage points over 2007 - 2015. Elsewhere, Almeida et al. (2009) find that firms that had to refinance a significant fraction of their debt in the year following August 2007 experienced a one-third fall in their investment relative to otherwise similar firms.

[2] Bates et al (2009) show cash held among nonfinancial firms rose through 2005; see also Khan and Thomas (2013).

recover only a fraction of its capital. Competitive lenders choose loan rate schedules for each loan delivering them an expected return equal to the real return on risk-free debt, so the interest rate charged on any given loan rises with the probability of default and falls in the collateral recoverable under default. Firms' probabilities of default vary with their expected future productivity, capital stock and debt; thus, so do their costs of borrowing.

By assuming that a defaulting firm loses its assets and must immediately exit, we can derive several results allowing us to characterize borrowing and lending in general equilibrium. First, absent firm-level capital adjustment costs, we show that idiosyncratic productivity and cash-on-hand are sufficient to describe a firm's state. Second, we show that firm values are continuous and weakly increasing in cash. Third, we establish the existence of cash thresholds with which we can fully summarize firms' decision rules regarding default and thereby isolate equilibrium loan schedules. Specifically, given the current aggregate state, we derive threshold cash levels varying with idiosyncratic productivity, with each representing the minimum cash-on-hand such that a firm realizing a given productivity will repay its loan and continue operating. We exploit these results to numerically characterize borrowing and lending in our model where the aggregate state includes the distribution of a time-varying number of firms over idiosyncratic productivity, capital and debt.

In general, firms with higher expected future earnings or capital have a lower risk of default and so can borrow more, and at lower cost. The risk of default rises with the debt taken on; beyond some critical level of debt, that risk becomes certainty, and there is no interest rate at which such a loan will be conveyed. Thus, each firm's ability to borrow is endogenously limited, and these borrowing limits vary across firms as functions of their individual state. We characterize firm decision rules conditional on membership in one of three groups, and then identify threshold cash levels determining membership in these groups. Firms having stockpiled sufficient cash to ensure they can invest efficiently in every possible future date and state without ever again incurring default risk are henceforth *impervious* to financial shocks; these firms invest efficiently, borrow at the risk-free rate and follow debt-dividend policies ensuring they can always do so. Were it not for exogenous death risk, such firms would live forever and ultimately take over all production. Outside this group, firms with adequate cash to invest efficiently in the current period with no immediate default risk take on the necessary loans to do so at the risk-free interest rate; these firms pay no dividends in efforts to reach the absorbing impervious state. The remaining group of firms face risk-premia to invest efficiently, so their investments are affected by their cash; they also face immediate default risk and so may exit the economy, eliminating useful production locations.

We calibrate our model using firm-level employment and financial data, data on default rates and debt recovery rates, real aggregate quantity data, data from the flow of funds, and evidence on the frequency and duration of financial crises. We show that borrowing and lending with non-contingent debt leads to a long-run misallocation of resources relative to a setting without financial frictions. Larger firms with more collateral invest more than smaller firms with less collateral on average, although they draw from the same productivity distribution. The implication of this is that insufficient capital is allocated to small firms. That in turn discourages entry, reducing the number of production locations, since young firms tend to be small and heavily reliant on external finance. This misallocation, alongside endogenous exit in the event of default, reduces average aggregate total factor productivity, alongside GDP and its inputs.

Examining business cycle moments from our model driven by both real and financial shocks, we find that they resemble those from a version of the model without financial shocks, which are in turn similar to those from a frictionless representative agent model. The latter is because aggregate TFP shocks affect our firms in a largely even way, yielding little change in the extent of misallocation. The former is due to the infrequent nature of financial shocks, not because their consequences resemble a TFP-driven recession.[3] When our economy experiences exogenous shocks to aggregate TFP, non-contingent loans drive countercyclical default and so are reinforced by modest endogenous TFP effects. Nonetheless, irrespective of our model's micro-level calibration, its aggregate quantity responses following a shock to TFP closely resemble those from a representative agent model such as Hansen (1985). Thus, a productivity shock selected to yield the observed decline in measured TFP over the 2007 recession falls far short of explaining the observed declines in GDP, investment and employment. It also implies a reduction in lending an order of magnitude too low, and it does not yield disproportionate employment declines among small and young firms.

When we examine the response to a financial shock in our model, the greatest declines in output, employment and investment do not occur at the onset. This distinction relative to the response following a TFP shock stems from a resulting allocative disturbance that compounds through time. Unlike an across-the-board productivity shock, a shock affecting firms' cash positions and borrowing terms has disparate implications throughout the distribution of firms, concentrating immediate losses and lasting damage among firms most reliant on credit for their growth and survival.

When hit by a financial shock, our model economy experiences a large, slowly unfolding recession ultimately followed by slow recovery. At the shock's impact, a larger than usual set of young,

---

[3] Large financial shocks are rare in the postwar U.S. (Reinhart and Rogoff (2009) and Bianchi and Mendoza (2012)).

cash-poor firms with relatively high productivities that should be growing rapidly find themselves undertaking suboptimal investments, with these investments further below optimal than in normal times. As a result, beginning with the first date thereafter, the allocation of capital, and thus production, across incumbent firms is unusually distorted. This problem worsens with time as new generations of firms are affected and each generation requires more time to reach efficient mature size. Furthermore, because the shock causes a sharp fall in the values of cash-poor firms, and entrants tend to be cash-poor, there are fewer new firms than usual entering production in each period while exit rates rise, particularly among young firms made increasingly vulnerable by each date's inefficiently low investment and production. Given decreasing returns-to-scale production, the resulting destruction to the stock of firms delivers a second form of capital misallocation in the form of lost production locations. Misallocation rises over time as these forces compound, so endogenous productivity and real activity gradually unravel, with the consequent reductions in the returns to saving driving large declines in investment, and hence employment, GDP and consumption. Thereafter, the recovery episode is far slower than what follows an aggregate TFP shock. Even by comparison to the Khan and Thomas (2013) setting where an aggregate shock tightens exogenous collateral constraints, our model economy exhibits significantly more gradualism in recovery due to its second, extensive-margin, channel of misallocation. When financial conditions revert to normal, it takes many periods to repair the effects of our two misallocation channels on the distribution and stock of firms. So long as the measure of producers has not recovered, the return on capital remains low, slowing the rise in aggregate investment and thus the recovery in aggregate capital, with that in turn hindering recovery in the stock of firms and its distribution.

It turns out that the microeconomic details of our model's calibration that shape the usual firm age-size distribution can have first-order effects on its aggregate predictions following a financial shock. We present results from two cases of our model to make this point. When we emphasize a large set of moments from the U.S. firm age-size distribution in its calibration, our model has a significantly greater elasticity of response in GDP, more gradual downturn, and brisker recovery relative to its predictions when the calibration heavily weights moments from the unconditional size distribution at the expense of a protracted cohort growth phase. A far larger financial shock is required to generate the same percent GDP decline in the latter case as in the former. With the average firm maturing rapidly in normal times, the incumbent misallocation channel rises only weakly, and the recession is largely driven by a sharp rise in extensive margin misallocation. That implies an immediate and large loss of firms that hastens the speed of the downturn and ultimately causes enormous damage to the stock of firms and thus very slow economic recovery. By contrast, in the version of our model more evenly weighting age-size and size distribution targets, the two misallocation channels have similar importance in driving the recession, so the downturn is more gradual and the ultimate damage to the stock of firms more empirically plausible.

We find that a financial shock to our baseline model economy can do well in explaining several unique features of the U.S. 2007 recession episode. First, it generates unusually steep declines in GDP, investment and employment, while simultaneously predicting a comparatively modest decline in measured TFP. Second, these declines are more concentrated among small firms than large firms and more concentrated among young firms than old firms. Third, the recession builds through time, with aggregate investment and employment steadily worsening over several years. Fourth, the recovery episode is unusually gradual because of a pronounced reduction in the number of firms driven by large increases in exit and reductions in entry over the downturn. Given decreasing returns to scale at individual firms, the number of available production units is itself a valuable stock that affects measured aggregate total factor productivity. Until this stock recovers, the return on capital remains low. This slows the return in investment in comparison to the recovery following a real shock, causing weaker, more gradual, recoveries in production and employment.

2 Related work

Recent studies have begun exploring how shocks to financial markets affect aggregate fluctuations. An early example is Jermann and Quadrini (2012), which examines a representative firm model wherein investment is financed with debt and equity. Given limited enforceability of debt contracts, the firm faces endogenous limits on its debt issuance. That friction is not trivially evaded using changes in equity, because there are convex costs associated with such adjustments. In this setting, financial shocks are shocks affecting the fraction of the firm's value recoverable under default, and thus the severity of borrowing limits. Because the firm's enforceability constraint always binds, a shock tightening this constraint causes large reductions in real economic activity. Beyond our emphasis on heterogeneity and equilibrium default, a notable difference in our setting is that financial frictions do not dampen the response of the aggregate economy to non-financial shocks.

Gomes, Jermann and Schmid (2014) examine a model with equilibrium default and nominally denominated console debt. They find that, so long as the expected maturity of debt exceeds one period, a temporary fall in inflation increases default rates and also generates a debt overhang channel affecting future investment and production decisions. As in our model, the outside value of a defaulting firm is zero. However, default in the Gomes, Jermann and Schmid economy triggers costly restructuring whereby some fraction of a defaulting firm's capital is destroyed, but implies no change in the number of firms. We instead lump-sum return to households any capital not recovered from a defaulting firm by the financial intermediary, and associate default with exit. This aspect of our model combines with endogenous entry decisions to deliver a time-varying measure of producers. Our model is also distinguished by persistent firm-level heterogeneity across periods. Because the distribution of firms enters into our aggregate state vector, there are persistent real effects from a purely financial shock despite the fact that debt is one-period lived.

Khan and Thomas (2013) study financial frictions in the form of collateralized borrowing. There, a credit shock is an unanticipated change in the fraction of their collateral firms can borrow against. When hit by such a shock, that model delivers a large, persistent recession with some similar features to the 2007 U.S. recession. Our current work departs from this study in three main ways. First, while capital serves as collateral, our borrowing terms depend on firm-level and aggregate state variables and arise from endogenous forward-looking lending schedules. Second, given the complexity associated with solving for equilibrium loan schedules, we omit the partial capital irreversibility considered there. Third, our model generates endogenous movements in entry and exit rates, and these series exhibit cyclical properties consistent with the U.S. data (see Campbell (1998)). We find this is important in generating more gradual recoveries in employment and investment following a shock affecting financial markets, and thus greater consistency with the post 2009Q2 U.S. experience.

Buera and Moll (2015) also study a heterogenous agent model with borrowing subject to collateral constraints. In their setting, entrepreneurs have constant returns production, face i.i.d. productivity shocks and observe shock realizations a period in advance. As a result, the distribution in their model does not evolve gradually over time as in ours, and they find shocks to collateral constraints are isomorphic to shocks to aggregate total factor productivity. Shourideh and Zetlin-Jones (2016) also study financial shocks in a model where heterogeneous firms face collateral constraints. Their model features two types of intermediate goods producers, publicly owned and privately owned. They argue that financial shocks are a promising source of aggregate fluctuations when there are strong linkages through intermediate goods trades across firms.

The financial frictions we study stem from non-contingent loans that introduce default into the model. This type of loan contract was first characterized by Eaton and Gersovitz's (1981) study of international lending. Aguiar and Gopinath (2006) and Arellano (2008) undertake quantitative analyses of sovereign debt. Chatterjee et al. (2008) study unsecured lending to households, and

Nakajima and Rios-Rull (2005) extend the framework to include real aggregate shocks.

Our emphasis on productivity dispersion, non-contingent debt and equilibrium default is shared by Arellano, Bai and Kehoe (2012), who explore the extent to which aggregate fluctuations are explained by movements in the labor wedge driven by uncertainty shocks.[4] Our study differs from theirs in that we explore aggregate responses to financial shocks, our employment levels are not predetermined, and we include capital investment. We share in common the simplifying assumption that defaulting firms must exit the economy.

The model we develop is consistent with the observation that the aggregate U.S. nonfinancial business sector can fully finance investment from cash flows. At the same time, changes in financial conditions have aggregate implications, given firms' differing reliance on external finance. On average, new firms begin with relatively small capital stocks. They grow gradually, maintaining their leverage in a narrow range. Conditional on survival, they ultimately achieve a capital level consistent with their expected productivity and begin reducing their debt, eventually changing its sign to build financial savings. In short, our firms have a natural maturing phase and tend to eventually outgrow default risk. Thus, the incidence of a credit shock differs, and we can explore the extent to which small firms are disproportionately affected. Following such shocks, shifts in the distribution of capital drive movements in aggregate total factor productivity through misallocation. In this respect, our study is also related to Buera and Shin (2013), who show that collateral constraints can protract the transition path to economic development if capital is initially misallocated.

3 Model

Our model economy has three types of agents: households, firms, and a perfectly competitive representative financial intermediary. Only firms are heterogeneous. They face persistent differences in their individual total factor productivities. Furthermore, their only source of external finance is non-contingent one-period debt provided by the financial intermediary at loan rates determined by their individual characteristics. These two aspects of the model combine to yield substantial heterogeneity in production.

---

[4] Gomes and Schmid (2014) develop a model with endogenous default where firms vary with respect to their leverage and study the implication for credit spreads. Credit spreads are also a focus of Gertler and Kiyotaki (2010), who study a model where such spreads are driven by agency problems arising with financial intermediaries.

3.1 Production, credit and capital adjustment

We assume a large number of firms, each able to produce a homogenous output using predetermined capital stock k and labor n, via an increasing and concave production function; $y = z\varepsilon F(k,n)$, where $F(k,n) = k^{\alpha}n^{\nu}$, with $\alpha > 0$, $\nu > 0$ and $\alpha + \nu < 1$. Here, z represents exogenous stochastic total factor productivity common across firms, while ε is a firm-specific counterpart. We assume z is a Markov chain, $z \in \mathbf{Z} \equiv \{z_1, \ldots, z_{N_z}\}$, where $\Pr\left(z' = z_g \mid z = z_f\right) \equiv \pi^z_{fg} \geq 0$, and $\sum_{g=1}^{N_z} \pi^z_{fg} = 1$ for each $f = 1, \ldots, N_z$. The idiosyncratic component of firm total factor productivity $\varepsilon \in \mathbf{E} \equiv \{\varepsilon_1, \ldots, \varepsilon_{N_\varepsilon}\}$, where $\Pr\left(\varepsilon' = \varepsilon_j \mid \varepsilon = \varepsilon_i\right) \equiv \pi^{\varepsilon}_{ij} \geq 0$, and $\sum_{j=1}^{N_\varepsilon} \pi^{\varepsilon}_{ij} = 1$ for each $i = 1, \ldots, N_\varepsilon$.

In our model, firms' financial positions and their cost of borrowing are affected by credit shocks. These are determined by changes in ζ, where $\zeta \in \left\{\zeta_1, \ldots, \zeta_{N_\zeta}\right\}$ with $\Pr\left\{\zeta' = \zeta_k \mid \zeta = \zeta_h\right\} \equiv \pi^{\zeta}_{hk} \geq 0$ and $\sum_{k=1}^{N_\zeta} \pi^{\zeta}_{hk} = 1$ for each $h = 1, \ldots, N_\zeta$. Let $s = (z, \zeta)$ be the joint stochastic process for the exogenous aggregate state with transition matrix π^s derived from the Markov Chains $\{\pi^z\}$ and $\left\{\pi^{\zeta}\right\}$. The bivariate process s has a support with $N_s = N_z N_\zeta$ values.

At the opening of each period, a firm is identified by its predetermined stock of capital, $k \in \mathbf{K} \subset \mathbf{R}_+$, the level of debt it took on in the previous period, $b \in \mathbf{B} \subset \mathbf{R}$, and its current idiosyncratic productivity level, ε. We summarize the distribution of firms over (k, b, ε) using the probability measure μ defined on the Borel algebra generated by the open subsets of the product space, $\mathbf{K} \times \mathbf{B} \times \mathbf{E}$.

The aggregate state of the economy is fully summarized by (s, μ), and the distribution of firms evolves over time according to a mapping, Γ, from the current aggregate state; $\mu' = \Gamma(s, \mu)$. The evolution of the firm distribution is determined in part by the actions of continuing firms and in part by entry and exit, as will be made clear below.

No microeconomic frictions impede capital reallocation in our model, so any firm's individual state in a period can be effectively summarized by its cash on hand, x. Given real wage $\omega(s, \mu)$ and capital depreciation rate δ, the cash on hand of a type (k, b, ε) firm that operates is:

$$x(k, b, \varepsilon; s, \mu) = y(k, \varepsilon; s, \mu) - \omega(s, \mu)\, n(k, \varepsilon; s, \mu) + (1 - \delta)k - b - [\xi_0 + \chi_\zeta(s)\, \xi_1(\varepsilon)],$$

where y and n represent the firm's chosen output and employment. The fixed cost of operation, $\xi_0 + \chi(s)\, \xi_1(\varepsilon)$, must be paid for the firm to operate in this or any future period. A firm can avoid it only by permanently exiting the economy prior to production. If it exits, the firm avoids both its debt repayment and operating cost, but it forfeits current flow profits and its capital stock, and achieves a value of 0. The fixed cost has two components, a real resource cost and a financial cost. Both must be paid whenever the firm wishes to continue operation. The real cost is state-invariant, while the financial cost $\chi_\zeta(s)\,\xi_1(\varepsilon)$, which varies with firm's idiosyncratic productivity, is positive only when the indicator function $\chi_\zeta(s) = 1$. This is the case whenever ζ in $s = (z, \zeta)$ is associated with a credit shock; otherwise $\chi_\zeta(s) = 0$.

Because our interest is in understanding how imperfect credit markets shape the decisions taken by firms, we require that firms' dividends always be non-negative to prevent them using equity to circumvent frictions in debt markets. Similarly, we prevent all firms growing so large that none ever faces a borrowing limit or a cost of borrowing exceeding the current risk-free interest rate. To do this, we impose some state-independent exit in the model. In particular, we assume each firm faces a fixed probability, $\pi_d \in (0, 1)$, that it will be forced to exit the economy after production in any given period. To maintain the number of firms in our economy at 1 on average, we also have exogenous birth of potential firms. At the start of any period, $\overline{\mu}^0$ potential firms are born with (ε, k, b) drawn from a distribution we will describe below. Each potential entrant becomes a new firm if it pays operating cost, produces and repays its start-up loan.[5]

Given the aggregate state (s, μ) and its start-of-period individual state (k, b, ε), each firm takes a series of actions to maximize the expected discounted value of its dividends. First, it chooses whether to exit or remain in operation. To remain, the firm must be prepared to pay the fixed operating cost $\xi_0 + \chi_\zeta(s)\,\xi_1(\varepsilon)$ and repay its existing debt b. If a firm defaults on its debt, it must immediately exit the economy, forfeiting all its remaining revenues and capital. If a firm fails to pay its operating cost, it also must immediately exit and so will at the same time default on its debt. Second, conditional on operating, the firm chooses its current level of employment and production, pays its wage bill, and repays its existing debt. After current production, wage payments and debt repayment, but prior to investment, each operating firm learns whether it will be permitted to continue into the next period.[6] If the firm is forced to leave the economy by an exogenous exit shock, it takes on no new debt and sells its remaining capital, thus achieving value x, which is paid to its shareholders as it exits. A continuing firm, by contrast, chooses its investment, current dividends, and the level of debt with which it will enter the next period.

Before turning to continuing firms' end of period decisions, we first examine choices among all firms operating in the current period. Each such firm chooses its employment to solve: $\pi(k, \varepsilon; s, \mu) = \max_n [z\varepsilon k^\alpha n^\nu - \omega(s, \mu)\, n]$ where z is given by its value in $s = (z, \zeta)$. The firm's optimal labor and

---

[5] Potential firms either become entrants or lump-sum return their capital to households and leave the economy.

[6] We have adopted this timing to ensure that no default arises from the exogenous exit shock in our model.

production are independent of its existing debt, b, and given by:

$$n\left(k,\varepsilon;s,\mu\right) = \left(\frac{\nu z\varepsilon k^{\alpha}}{\omega\left(s,\mu\right)}\right)^{\frac{1}{1-\nu}} \tag{1}$$

$$y\left(k,\varepsilon;s,\mu\right) = \left(\varepsilon z\right)^{\frac{1}{1-\nu}}\left(\frac{\nu}{\omega\left(s,\mu\right)}\right)^{\frac{\nu}{1-\nu}} k^{\frac{\alpha}{1-\nu}}, \tag{2}$$

which in turn imply its flow profits net of labor costs,

$$\pi\left(k,\varepsilon;s,\mu\right) = \left(1-\nu\right) y\left(k,\varepsilon;s,\mu\right). \tag{3}$$

These values are common to type (k,ε) firms that choose to operate given their state, (k,b,ε).

At the end of the period, conditional on it producing, repaying its debt, and escaping the exit shock, a firm determines its future capital, k', future debt, b', alongside current dividends, D. Given investment, i, the firm's capital stock for the start of next period is given by:

$$k' = \left(1-\delta\right)k + i, \tag{4}$$

where $\delta \in (0,1)$ is the rate of capital depreciation, and primes indicate one-period-ahead values. For each unit of debt it incurs for the next period, the firm receives $q(\cdot)$ units of output for use toward investment or current dividends. Thus a loan of $q(\cdot)b'$ implies debt of b' to be repaid in the next period, and the continuing firm's current dividends are $D(\cdot) = x - k' + q(\cdot)b'$, where x is its cash on hand including current profits and the value of nondepreciated capital, after loan repayment and the payment of the fixed operating cost:

$$x = \pi\left(k,\varepsilon;s,\mu\right) + \left(1-\delta\right)k - b - \xi_0 - \chi_{\zeta}\left(s\right)\xi_1\left(\varepsilon\right). \tag{5}$$

In contrast to models with exogenous collateral constraints, our default risk implies that the loan discount factor faced by a borrowing firm, $q(\cdot)$, depends on that firm's chosen debt and capital and its current productivity. Given a level of debt, a firm's capital choice for next period affects the distribution of its earnings and thus the probability it will repay.

3.2 Cash on hand and firm values

Firms selecting the same k' and b' will not all have the same income next period because there is uncertainty in the firm-specific component of total factor productivity, ε'. Among firms with common (k',b'), those realizing high productivities may repay their loans, while those realizing low ones may default. If a firm defaults, the financial intermediary recovers a fraction, $\lambda(\zeta)$, of the firm's nondepreciated capital; this notation allows the possibility that the recovery fraction may depend on current financial conditions. We assume the remainder of any such firm's capital is lump-sum rebated to households, so that default implies no direct loss of resources. Because a defaulting firm forfeits all its assets, only those firms that repay their debts will pay operating costs to produce.

When a continuing firm with current idiosyncratic productivity ε chooses to take on a debt b', alongside a future capital stock k', that firm receives $q\left(k', b', \varepsilon; s, \mu\right) b'$ units of output in the current period. As noted above, the loan discount factor, $q\left(k', b', \varepsilon; s, \mu\right)$, is determined as a function of the firm's repayment probability. Competitive lending equates the financial intermediary's expected return on each of its loans to the risk-free real interest rate. Letting $\pi^s_{lm} d_m\left(s_l, \mu\right)$ be the price of an Arrow security that pays off if $s' = s_m$, the risk-free real rate is $\frac{1}{q_0(s_l,\mu)} - 1$, where:

$$q_0\left(s_l, \mu\right) = \sum_{m=1}^{N_s} \pi^s_{lm} d_m\left(s_l, \mu\right). \tag{6}$$

Each firm's loan discount factor is bounded above by the risk-free factor, $q_0\left(s, \mu\right)$, and below by 0. Given a chosen $\left(k', b'\right)$, and given $\mu' = \Gamma\left(s, \mu\right)$, the firm will face $q\left(k', b', \varepsilon; s, \mu\right) < q_0\left(s, \mu\right)$ so long as there is some possible realization of $\left(\varepsilon', s'\right)$ next period under which it will default on its b'. Among firms selecting a common $\left(k', b'\right)$, those realizing higher ε' next period will be less likely to default, as will be clear below. Thus, given persistence in the firm productivity process, $q\left(k', b', \varepsilon; s, \mu\right)$ rises (weakly) in ε. For the same reasons, the firm's q rises in k' and falls in b'.

Recall the definition of an operating firm's cash on hand, x, from (5) above. In considering the lending schedule each firm faces, it is useful to note that a firm's individual levels of k and b do not separately determine any of its choices beyond their effect in x. To see this, note that a firm's resource constraint (determined by the non-negativity constraint on dividends) may be written as simply: $x - k' + q\left(k', b', \varepsilon; s, \mu\right) b' \geq 0$. This means that the firm's feasible capital and debt combinations are given by the set $\Phi\left(x, \varepsilon; s, \mu\right)$, where:

$$\Phi\left(x, \varepsilon; s, \mu\right) = \left\{\left(k', b'\right) \in \mathbf{K} \times \mathbf{B} \mid \; x - k' + q\left(k', b', \varepsilon; s, \mu\right) b' \geq 0\right\}. \tag{7}$$

Since x fully captures previous decisions influencing its current choice set, the firm's value is a function only of x and ε, and does not depend separately upon k and b. This important result allows us to reduce the firm-level state vector, and thus the dimension of the value and default functions that characterize competitive equilibrium.

Let $V^0\left(x, \varepsilon_i; s_l, \mu\right)$ represent the beginning of period value of a firm just before its default decision, and let $V^1\left(x, \varepsilon_i; s_l, \mu\right)$ represent its value conditional on repaying its debt and operating. If $\Phi\left(x, \varepsilon_i; s_l, \mu\right) \neq \{\emptyset\}$, the firm can cover its operating cost and repay its debt while paying a non-negative current dividend. In that case, the firm operates in the current period, achieving a non-negative value $V^1(x, \varepsilon_i; s_l, \mu)$. Otherwise, it defaults on its debt and immediately exits the economy with zero value.

$$V^0(x, \varepsilon_i; s_l, \mu) = \max\Big\{V^1(x, \varepsilon_i; s_l, \mu), 0\Big\}. \tag{8}$$

Given the constraint set in (7), it is straightforward to show that a firm's value is increasing in its cash on hand, x, and in its productivity, ε.

The firm's value of operating, V^1, must account for the possibility of receiving the exogenous exit shock after current production and thus being unable to continue into the next period. Recall that, with probability π_d, the firm is forced to exit at the end of the period. In that case, it simply pays out its cash on hand as dividend as it exits. Otherwise, it moves to the next period with continuation value V^2 determined below.

$$V^1(x, \varepsilon_i; s_l, \mu) = \pi_d x + (1 - \pi_d) V^2(x, \varepsilon_i; s_l, \mu) \tag{9}$$

Given the current aggregate state, and given $\mu' = \Gamma(s_l, \mu)$, firms continuing to the next period solve the following problem.

$$\begin{aligned}
V^2(x, \varepsilon_i; s_l, \mu) &= \max_{k', b'}\Big[x - k' + q(k', b', \varepsilon_i; s_l, \mu) b' \qquad (10)\\
&\quad + \sum_{m=1}^{N_s} \pi^s_{lm} d_m(s_l, \mu) \sum_{j=1}^{N_\varepsilon} \pi^\varepsilon_{ij} V^0(x'_{jm}, \varepsilon_j; s_m, \mu')\Big],\\
\text{subject to} &: \\
(k', b') &\in \Phi(x, \varepsilon_i; s_l, \mu) \qquad (11)\\
x'_{jm} &= \pi(k', \varepsilon_j; s_m, \mu') + (1 - \delta)k' - b' - [\xi_0 + \chi_\zeta(s_m)\xi_1(\varepsilon_j)], \qquad (12)
\end{aligned}$$

where $V^0(\cdot)$ is defined in (8), $\Phi(\cdot)$ is given by (7), and $\pi(\cdot)$ is from (3).

3.3 Loan rates

We now turn to the determination of the loan discount factors, $q(\cdot)$. Let $\chi(x', \varepsilon'; s', \mu')$ be an indicator for a firm entering next period with cash on hand x' and productivity ε', given aggregate state (s', μ'), with this indicator taking on the value 1 if the firm chooses to repay its debt, and 0 otherwise. Threshold values of cash on hand, $x^d(\varepsilon'; s', \mu')$, solve $V^1(x^d, \varepsilon'; s', \mu') = 0$ and separate those firms of a given productivity level for which $\chi(\cdot) = 1$ (those with $x \geq x^d$) from those for which $\chi(\cdot) = 0$ (those that default).

Recall that the financial intermediary providing loans to firms is perfectly competitive. Thus, the interest rate it offers on any loan is determined by a zero expected profit condition. Taking into account the fact that the intermediary recovers no more than $\lambda(\zeta)$ fraction of a firm's remaining capital in the event of default (or b' if that is smaller), and recalling the determination of x'_{jm} from (12), we arrive at the following implicit solution for the loan discount factor.

$$q\left(k', b', \varepsilon_i; s_l, \mu\right) b' = \sum_{m=1}^{N_s} \pi^s_{lm} d_m\left(s_l, \mu\right) \sum_{j=1}^{N_\varepsilon} \pi^\varepsilon_{ij} \Big[\chi\left(x'_{jm}, \varepsilon_j; s_m, \mu'\right) b' + [1 - \chi\left(x'_{jm}, \varepsilon_j; s_m, \mu'\right)] \min\{b', \lambda(\zeta)\left(1-\delta\right) k'\}\Big]. \tag{13}$$

Note that the loan price determined by (13) gives the risk-neutral lender the same per-unit expected return as that associated with risk-free real discount factor, $q_0\left(s_l, \mu\right)$. If a loan involves no probability of default, then $\chi\left(x'_{jm}, \varepsilon_j; s_m, \mu'\right) = 1$ for every $\left(\varepsilon_j, s_m\right)$ with $\pi^\varepsilon_{ij} > 0$ and $\pi^s_{lm}(s_l) > 0$. In that case, $q\left(k', b', \varepsilon_i; s_l, \mu\right) b' = \sum_{m=1}^{N_s} \pi^s_{lm} d_m\left(s_l, \mu\right) b'$, so $q\left(k', b', \varepsilon_i; s_l, \mu\right) = q_0\left(s_l, \mu\right)$.

3.4 Households

We close the model with a unit measure of identical households Household wealth is held as one-period shares in firms identified using the measure $\widetilde{\lambda}$, and in one-period noncontingent bonds, ϕ.[7] Given the (dividend inclusive) prices received for their current shares, $m_0\left(x, \varepsilon; s, \mu\right)$, the risk-free bond price $q_0\left(s, \mu\right)^{-1}$, and the real wage, $w\left(s, \mu\right)$, households determine their current consumption, c, hours worked, n^h, new bond holdings ϕ', and the numbers of new shares, $\widetilde{\lambda}'\left(x', \varepsilon'\right)$, to purchase at ex-dividend prices $m_1\left(x', \varepsilon'; s, \mu\right)$.

The lifetime expected utility maximization problem of the representative household is:

$$V^h\left(\widetilde{\lambda}, \phi; s_l, \mu\right) = \max_{c, n^h, \phi', \widetilde{\lambda}'} \left[U\left(c, 1-n^h\right) + \beta \sum_{m=1}^{N_s} \pi^s_{lm} V^h\left(\widetilde{\lambda}', \phi'; s_m, \mu'\right)\right] \tag{14}$$

subject to

$$c + q_0(s,\mu)\phi' + \int m_1\left(x', \varepsilon'; s_l, \mu\right) \widetilde{\lambda}'\left(d\left[x' \times \varepsilon'\right]\right) \leq \Big[\omega\left(s_l, \mu\right) n^h + \phi + \int m_0\left(x, \varepsilon; s_l, \mu\right) \widetilde{\lambda}\left(d\left[x \times \varepsilon\right]\right)\Big]$$

and given $\mu' = \Gamma(s, \mu)$. Let $C^h\left(\widetilde{\lambda}, \phi; s, \mu\right)$ and $N^h\left(\widetilde{\lambda}, \phi; s, \mu\right)$ be the household decision rules for consumption and hours worked. Let $\Phi^h(\widetilde{\lambda}, \phi; s, \mu)$ be the household decision rule for bonds, and let

---

[7] Households also have access to a complete set of state-contingent claims. As there is no household heterogeneity, these assets are in zero net supply in equilibrium; thus we do not explicitly model them here.

$\Lambda^h\left(x', \varepsilon', \widetilde{\lambda}, \phi; s, \mu\right)$ be the quantity of shares purchased in firms that will begin next period with cash on hand x' and productivity ε'.

4 Computing equilibrium

In *recursive competitive equilibrium*, each firm solves the problem described by (8) - (12), households solve the problem described in (14), loans are priced according to (13), the markets for labor, output and firm shares clear, and the resulting individual decision rules for firms and households are consistent with the aggregate law of motion, Γ. Using $C(s,\mu)$ and $N(s,\mu)$ to describe the market-clearing values of household consumption and hours worked, it is straightforward to show that market-clearing requires that (a) the real wage equal the household marginal rate of substitution between leisure and consumption, (b) the risk-free bond price, q_0^{-1}, equal the expected gross real interest rate, and (c) firms' state-contingent discount factors be consistent with the household marginal rate of substitution between consumption across states.

$$w\left(s,\mu\right) = D_2U\Big(C(s,\mu), 1-N(s,\mu)\Big)/D_1U\Big(C(s,\mu), 1-N(s,\mu)\Big)$$

$$q_0\left(s,\mu\right) = \beta\sum_{m=1}^{N_s}\pi^s_{lm}D_1U\Big(C(s_m,\mu'), 1-N(s_m,\mu')\Big)/D_1U\Big(C(s,\mu), 1-N(s,\mu)\Big)$$

$$d_m\left(s,\mu\right) = \beta D_1U\Big(C(s_m,\mu'), 1-N(s_m,\mu')\Big)/D_1U\Big(C(s,\mu), 1-N(s,\mu)\Big).$$

We compute equilibrium in our economy by combining the firm-level optimization problem with the equilibrium implications of household utility maximization listed above, effectively subsuming households' decisions into the problems faced by firms. Without loss of generality, we assign $p(s,\mu)$ as an output price at which firms value current dividends and payments and correspondingly assume that firms discount their future values by the household subjective discount factor. Given this alternative means of expressing firms' discounting, the following three conditions ensure all markets clear in our economy.

$$p\left(s,\mu\right) = D_1U\Big(C(s,\mu), 1-N(s,\mu)\Big) \tag{15}$$

$$\omega\left(s,\mu\right) = D_2U\Big(C(s,\mu), 1-N(s,\mu)\Big)/p\left(s,\mu\right) \tag{16}$$

$$q_0\left(s,\mu\right) = \beta\sum_{m=1}^{N_s}\pi^s_{lm}p\left(s_m,\Gamma\left(s,\mu\right)\right)/p\left(s,\mu\right) \tag{17}$$

We reformulate (8) - (12) here to obtain an equivalent, more convenient, representation of the firm problem with each firm's value measured in units of marginal utility, rather than output.

$$v^0\left(x,\varepsilon_i;s_l,\mu\right)=\max\Big\{v^1\left(x,\varepsilon_i;s_l,\mu\right),0\Big\}. \tag{18}$$

$$v^1\left(x,\varepsilon_i;s_l,\mu\right)=\pi_d xp\left(s_l,\mu\right)+\left(1-\pi_d\right)v^2\left(x,\varepsilon_i;s_l,\mu\right) \tag{19}$$

$$\begin{aligned} v^2\left(x,\varepsilon_i;s_l,\mu\right) &= \max_{k',b'}\Big([x-k'+q\left(k',b',\varepsilon_i;s_l,\mu\right)b']p\left(s_l,\mu\right) \\ &\quad +\beta\sum_{m=1}^{N_s}\pi^s_{lm}\sum_{j=1}^{N_\varepsilon}\pi^\varepsilon_{ij}v^0\left(x'_{jm},\varepsilon_j;s_m,\mu'\right)\Big),\text{ subject to (11) - (12).}\end{aligned} \tag{20}$$

The problem listed in equations (18) - (20) forms the basis for solving equilibrium allocations in our economy, so long as the prices p,ω and q_0 taken as given by firms satisfy the restrictions in (15) - (17), and the loan price schedules offered satisfy (13).

As noted above, a firm of type (k,b,ε) hires labor and produces only if $\Phi\left(x,\varepsilon_i;s_l,\mu\right)\neq\{\emptyset\}$ and $v^1\left(x,\varepsilon;s,\mu\right)\geq 0$, where $x=\pi\left(k,\varepsilon;s,\mu\right)+(1-\delta)k-b-[\xi_0+\chi_\zeta\left(s\right)\xi_1\left(\varepsilon\right)]$. In that case, its decision rules for labor and output are given by (1) - (2), and its flow profits are given by (3). The more challenging objects we must determine are D, k' and b' for firms continuing into the next period. These decisions are dynamic, inter-related functions of firm productivity, ε, and cash on hand, x.

To solve for the forward-looking decisions of continuing firms, we use a partitioning analogous to that in Khan and Thomas (2013), here extended for the fact that we study noncontingent debt with default and exit. In particular, we assign firms across three distinct categories reflecting the extent to which their investment activities can be affected by financial frictions and identify their decision rules accordingly. Firms termed (financially) *impervious* are those that have permanently outgrown the implications of financial frictions. Firms that are (financially) *exposed type 1* can undertake efficient investment in the current period while borrowing at the risk-free interest rate, but face the possibility of paying a risk premium in future. *Exposed type 2* firms cannot finance efficient investment this period without incurring default risk and thus a risk premium.

4.1 Decisions among impervious firms

An impervious firm has accumulated sufficient cash on hand such that, in every possible future state, it will be able to finance its efficient level of investment at the risk-free interest rate. Because any such firm has effectively outgrown financial frictions, its marginal valuation on retained earnings equals the household marginal valuation of consumption, p, so it is indifferent between financial savings and dividends. Viewed another way, the firm's value function is linear in its debt or financial savings, so b' does not affect its k' decision. Any such firm not forced by the exit shock to leave at the end of the period adopts the efficient capital stock $k^*(\varepsilon_i; s_l, \mu)$ solving (21), achieving value $w^2(\cdot)$, and, given its indifference to financing arrangements, it is content to adopt the debt policy $B^w(\cdot)$ we isolate below to maintain that indifference permanently.

$$\begin{aligned} w^2(x, \varepsilon_i; s_l, \mu) = \max_{k'} \Big[& [x - k' + q_0(s_l, \mu) B^w(\varepsilon_i; s_l, \mu)] p(s_l, \mu) \\ & + \beta \sum_{m=1}^{N_s} \pi^s_{lm} \sum_{j=1}^{N_\varepsilon} \pi^\varepsilon_{ij} w^0\left(x'_{jm}, \varepsilon_j; s_m, \mu'\right) \Big], \end{aligned} \tag{21}$$

where x'_{jm} is given by (12) and $w^0(x, \varepsilon; s, \mu) = p(s, \mu)\pi_d x + (1 - \pi_d) w^2(x, \varepsilon; s, \mu)$.

We assign an impervious firm a *minimum savings* debt policy solving (22) - (23) to just ensure that it always maintains sufficient wealth to implement its optimal investments with no default risk under all possible future paths of (ε, s).[8] Let $\widetilde{B}\Big(k', \varepsilon_j; s_m, \Gamma(s, \mu)\Big)$ define the maximum debt level at which a firm entering next period with k' and realizing (ε_j, s_m) will be impervious. This requires that the firm can adopt $k^*(\varepsilon_j; s_m, \mu')$ and $B^w(\varepsilon_j; s_m, \mu')$ next period while maintaining $D \geq 0$, and that it will choose to remain in the economy. The impervious firm debt policy, $B^w(\varepsilon_i; s_l, \mu)$, is the minimum $\widetilde{B}_{jm}$; i.e., it is the maximum debt with which the firm can exit this period and be certain to be impervious next period, given that it adopts $k^*(\varepsilon_i; s_l, \mu)$.

$$B^w(\varepsilon_i; s_l, \mu) = \min_{\{\varepsilon_j | \pi^\varepsilon_{ij} > 0 \text{ and } s_m | \pi^s_{lm} > 0\}} \widetilde{B}\Big(k^*(\varepsilon_i; s_l, \mu), \varepsilon_j; s_m, \Gamma(s_l, \mu)\Big), \tag{22}$$

$$\begin{aligned} \text{where } \widetilde{B}(k, \varepsilon_i; s_l, \mu) \equiv{} & \pi(k, \varepsilon_i; s_l, \mu) + (1 - \delta) k - \xi_0 - \chi_\zeta(s_l)\, \xi_1(\varepsilon_i) \\ & + \min\Big\{ \Big[-k^*(\varepsilon_i; s_l, \mu) + q_0(s_l, \mu) B^w(\varepsilon_i; s_l, \mu) \Big], -x^d(\varepsilon_i; s_l, \mu) \Big\}, \end{aligned} \tag{23}$$

and $x^d(\varepsilon_i; s_l, \mu)$ solves $v^1\left(x^d, \varepsilon_i; s_l, \mu\right) = 0$. Given their decision rules for capital and debt, we retrieve impervious firms' dividend payments as:

$$D^w(x, \varepsilon_i; s_l, \mu) = x - k^*(\varepsilon_i; s_l, \mu) + q_0(s_l, \mu) B^w(\varepsilon_i; s_l, \mu). \tag{24}$$

4.2 Decisions among exposed firms

We now consider the decisions made by a firm that has not yet been identified as impervious. We begin by evaluating whether the firm has crossed the relevant cash threshold to become impervious. This is verified from equation 24 using the impervious firm decision rules (21) - (23). If

---

[8] We adopt this policy rather than an alternative minimizing currrent dividends so as to bound the financial savings of long-lived firms.

$D^w(x, \varepsilon; s, \mu) \geq 0$, the firm is impervious and those decision rules apply. If not, it is still exposed in that financial considerations may continue to influence its investment decisions now or in future, so its choice of capital and debt remain intertwined.

Exposed firms of type 1 can invest to the efficient capital $k^*(\varepsilon_i; s_l, \mu)$ from (21) in the current period while ensuring that all possible resulting x'_{jm} for next period will imply zero probability of default; in other words, $\chi\left(x'_{jm}, \varepsilon_j; s_m, \mu'\right) = 1$ for all j, m such that $\pi^{\varepsilon}_{ij} > 0$ and $\pi^{s}_{lm}(s_l) > 0$, where x'_{jm} is from (12). Such firms optimally adopt $k^*(\varepsilon_i; s_l, \mu)$ and borrow at the risk-free rate. Because they are sure to remain in the economy throughout the next period, but have not permanently outgrown financial frictions, their shadow value of internal funds exceeds the household valuation on dividends. Thus, they set $D = 0$ and $b' = q_0(s_l, \mu)^{-1}[k^*(\varepsilon_i; s_l, \mu) - x]$. This means that, to determine whether an exposed firm is type 1, we need only determine whether the following inequality is satisfied for all possible j, m combinations.

$$\begin{aligned}&\pi\left(k^*(\varepsilon_i; s_l, \mu), \varepsilon_j; s_m, \Gamma(s_l, \mu)\right) + (1-\delta)k^*(\varepsilon_i; s_l, \mu)\\ &\qquad -\xi_0 - \chi(s_m)\,\xi_1(\varepsilon_j) - q_0(s_l, \mu)^{-1}[k^*(\varepsilon_i; s_l, \mu) - x] \geq x^d(\varepsilon_j; s_m, \Gamma(s_l, \mu))\end{aligned}$$

For exposed firms that do not satisfy the type 1 check just above, we know of no convenient way to separate the loan implied by a given k' choice to distinctly identify the corresponding debt level, b'. Unlike type 1 firms, a type 2 exposed firm may find $D = 0$ suboptimal; it may in fact achieve higher expected discounted value by paying dividends in the current period if it faces sufficiently high probability that it will be forced to default and exit the economy with zero value at the start of the next period. Thus, we must isolate the k', b' and D choices of any such firm by directly solving the problem listed in (18) - (20), which we do using a two dimensional grid on k' and b'.

Through the presence of type 2 exposed firms, the financial frictions in our economy generate two types of misallocation reducing aggregate TFP. First, these firms are led to adopt inefficiently small capital stocks either because they cannot borrow to k^* or because they are unwilling to suffer the implied risk premium. Second, with their low cash on hand and poor financing terms, these firms may default on their loans and exit the economy. Given decreasing returns to scale, the loss of a production unit further distorts the allocation of aggregate capital away from the efficient one.

5 Calibration

We explore the firm-level and aggregate implications of imperfect credit markets across a series of numerical exercises below. We assume that the representative household's period utility is the result of indivisible labor (Rogerson (1988)): $u(c, L) = \log c + \varphi L$. The firm-level production function is Cobb-Douglas: $z\varepsilon F(k, n) = z\varepsilon k^{\alpha} n^{\nu}$.

We set the length of a period to be one year and select parameters governing preferences and aspects of technology common across firms as follows.[9] First, we set the household discount factor, β, to imply an average real interest rate of 4 percent, consistent with findings by Gomme, Ravikumar and Rupert (2008). Next, the production parameter ν is set to yield an average labor share of income at 0.60 (Cooley and Prescott (1995)). The depreciation rate, δ, is taken to imply an average investment-to-capital ratio of roughly 0.069, matching the average value for the private capital stock between 1954 and 2002 in the U.S. Fixed Asset Tables, controlling for growth. Given this value, we determine capital's share, α, so that our model matches the average private capital-to-output ratio over the same period, at 2.3, and we set the parameter governing the preference for leisure, φ, to imply an average of one-third of available time is spent in market work.

Exact aggregation obtains in an otherwise identical reference model without financial frictions. We use that model to estimate an exogenous stochastic process for aggregate productivity. We begin by assuming the shock follows a mean zero AR(1) process in logs: $\log z' = \rho_z \log z + \eta'_z$ with $\eta'_z \sim N\left(0, \sigma^2_{\eta_z}\right)$. Next, we estimate the values of ρ_z and σ_{η_z} from Solow residuals measured using data on real U.S. GDP and private capital, together with the total employment hours series constructed by Cociuba, Prescott and Ueberfeldt (2012) from CPS household survey data, over the years 1959-2012. Then, we discretize this process as a 3-state Markov Chain; $N_z = 3$.

As noted above, our exogenous aggregate state also includes credit shocks. These shocks, ζ, follow a 2-state Markov chain with realizations $\{\zeta_o, \zeta_l\}$ and the transition matrix below. We associate ζ_o with ordinary credit conditions and ζ_l with a credit shock (low credit conditions). We assume $\zeta = \zeta_o$ in calibrating our steady state and assume no financial component of operating costs there, $\xi_1(\varepsilon_i) = 0$. In times of ordinary credit, the capital recovery rate λ_o is selected to imply a 43.4 percent debt recovery rate consistent with Moody's debt recovery rates over 1982-2010, stripping away years associated with the S&L crisis and the recent financial crisis.

$$\Pi^{\zeta} = \begin{bmatrix} p_o & 1 - p_o \\ 1 - p_l & p_l \end{bmatrix}$$

---

[9] Our annual calibration allows us to be consistent with establishment-level firm size data.

The probability of continuing in ordinary borrowing conditions is p_o, $\Pr\{\zeta' = \zeta_o | \zeta = \zeta_o\}$, while $1 - p_l$ is the probability of escape from crisis conditions, $\Pr\{\zeta' = \zeta_o | \zeta = \zeta_l\}$. We select the parameters of the Π^ζ matrix using evidence on banking crises from Reinhart and Rogoff (2009). Their definition of a banking crisis includes episodes where bank runs lead to the closure or public takeover of financial institutions as well as those without bank runs where the closure, merging, takeover or government bailout of one important financial institution is followed by similar outcomes for others. They document 13 crises in the U.S. since 1800 and the share of years spent in crises at 13 percent, which together imply an average crisis duration of 2.09 years. Given our use of postwar targets to calibrate the remaining parameters of our model, the more appropriate statistics for our purposes are those from the period 1945-2008, wherein the U.S. has had two banking crises (the 1989 savings and loan crisis and the 2007 subprime lending crisis).[10] Unfortunately, it is not possible to determine the average length of a U.S. crisis from this sample period, without knowing the ending date of the most recent crisis. Given this difficulty, alongside Reinhart and Rogoff's argument that the incidence and number of crises is similar across the extensive set of countries they consider, we focus instead on their data for advanced economies. The average number of banking crises across advanced economies over 1945 - 2008 was 1.4, while the share of years spent in crisis was 7 percent. Combining these observations, we set $p_o = 0.9765$ and $1 - p_l = 0.3125$ so that the average duration of a credit crisis in our model is 3.2 years, and the economy spends 7 percent of time in the crisis state.

We will examine two cases of our model in the results to follow, with these cases differing only in the calibration weights we ascribe to reproducing the maturing phase evident in the firm age-size distribution versus the skewness of the unconditional size distribution. Whereas our baseline case will weight both sets of moments roughly equally, our alternate case will focus almost exclusively on reproducing the thin right tail of the size distribution. We will use these two cases below to show that, over recessions led by a financial shock sufficient to drive the roughly 5.6 percent GDP decline over the 2007-9 U.S. recession, the steepness of the downturn and gradualism of recovery are critically affected by which misallocation channel (intensive- or extensive-margin) is dominant in hastening recession, and that this in turn depends on the shape of the average firm age-size distribution. This average distribution in fact determines the aggregate elasticity of response to credit conditions, so we specify what constitutes the financial shock differently across our two cases.

A credit shock ($\zeta = \zeta_l$) in our baseline model reduces the default recovery rate to $\lambda_l = 0$ and

---

[10] These observations are consistent with findings by Bianchi and Mendoza (2012); they document a frequency of financial crisis at 3 percent, consistent with three financial crisis in the U.S. over the past hundred years. Mendoza (2010) estimates a crisis frequency of 3.6 percent across emerging economies since 1980.

triggers financial operating costs, $\xi_1(\varepsilon_i) = 0.035\pi^*(k^*(\varepsilon_i), \varepsilon_i)$, where $\pi^*(k^*(\varepsilon_i), \varepsilon_i)$ is the steady state profit of a firm operating with the efficient level of capital for ε_i. This generates an endogenous decline in measured TFP in our model matching the exogenous TFP shock we separately consider below toward comparing quantity responses to real versus financial shocks. We allow ξ_1 to vary in ε_i in an effort to even the incidence of the balance sheet shock across firms of differing sizes. Note that $\xi_1(\varepsilon_i)$ is termed a balance sheet shock because it applies only when $\zeta = \zeta_l$, and it is a purely financial shock redistributing cash from firms to households. To be clear, in contrast to the real operating cost (ξ_0) paid by firms in every period, $\xi_1(\varepsilon_i)$ does not enter the aggregate resource constraint. Thus, recalling that defaulting firms' capital not recovered by lenders is always lump-sum rebated to households, a credit shock has no direct consequence for real resources. Because a fall in recovery rates has minimal effect in our alternate case emphasizing the firm size distribution, we leave $\lambda_l = \lambda_o$ there and consider the balance sheet shock alone. However, for reasons that will be clear below, the scale factor is raised from 0.035 in this case to 0.42.

The distribution of idiosyncratic productivity for potential new firms is consistent with the invariant distribution of ε. These firms have common initial debt, b_0, but draw an initial level of capital from a Pareto distribution with lower bound k_0 and curvature parameter κ_0. We adopt the common b_0 choice to reduce the dimension of the distribution of entrants and have it vary over ε and k. The level of debt is set to reproduce the average level of indebtedness of entrants reported in the data discussed below. Given the distribution of productivity across entrants, the lower bound and curvature parameter of the distribution of their capital influences the relative employment shares among young firms and the relative size of new firms.

A firm's productivity process depends on its permanent type drawn at birth, normal versus special, a distinction we include to reproduce the empirical firm size distribution. Normal firms are $\omega_n = 97.5$ percent of potential entrants in our model emphasizing the size distribution and 100 percent in our baseline model. These firms draw productivities from a log-normal distribution with persistence ρ_ε and standard deviation σ_ε. We discretize ε using 15 values, and then we shift each of the resulting 15 support points up by one index and add a zero value as the lowest of $N_\varepsilon = 16$ points. We assume the 0 draw occurs with probability π_0^ε, independent of the prior idiosyncratic productivity draw, and we scale remaining transition probabilities by $1 - \pi_0^\varepsilon$. We assume that firms realizing $\varepsilon = 0$ face the same transition probabilities as do new firms drawing what is now $\varepsilon = \varepsilon_8$. Because young firms rely heavily on flow income to repay their debts, adding risk of a zero productivity period allows us to reproduce a realistic life-cycle wherein young firms do not rapidly overcome borrowing constraints and adopt efficient levels of capital.[11] To reproduce the highly skewed firm size distribution in the alternate case of our model, we add a measure 0.025 of special firms with productivities drawn from a distinct stochastic process from that above. Such firms have one of two idiosyncratic productivity levels, either a high value $\overline{\varepsilon}_s$ unattainable for normal firms or the highest normal-firm draw (ε_{16}). The probability of the latter is π_0^{ε} every period for each firm.

TABLE 1. Parameter Values

(a) Common

β	φ	α	ν	δ	ρ_ε	ρ_z	σ_{η_z}	p_o	p_l
0.96	2.15	0.265	0.60	0.067	0.757	0.909	0.015	0.977	0.687

(b) Baseline case emphasizing life-cycle aging

ξ_0	λ_o	π_d	$\overline{\mu}_0$	k_0	κ_0	b_0	σ_ε	π_0^{ε}	ω_n	$\overline{\varepsilon}_s$
0.009	0.37	0.08	0.20	0.023	3.00	0.04	0.076	0.100	0.000	n/r

(c) Alternate case emphasizing skewed size distribution

ξ_0	λ_o	π_d	$\overline{\mu}_0$	k_0	κ_0	b_0	σ_ε	π_0^{ε}	ω_s	$\overline{\varepsilon}_s$
0.024	0.46	0.05	0.09	0.063	2.14	0.05	0.050	0.025	0.025	1.8

Having chosen the first 5 parameters in Table 1 to match long-run aggregate moments, we adopt Chad Syverson's estimated persistence for normal firms' productivities. Given other parameter values, we set the measure of potential new firms each period, $\overline{\mu}_0$, so that our long-run average number of firms in production is 1. Our remaining parameters are chosen jointly to reproduce the average debt recovery rate, the aggregate debt-to-asset ratio and a series of moments from the Business Dynamics Statistics database (BDS) reflecting the average U.S. firm age-size distribution.

We select k_0 (lower bound for capital in new firms), κ_0 (curvature parameter for new firm capital distribution), b_0 (a common level of debt for potential entrants), $\overline{\varepsilon}_s$ and ω_s (top productivity draw for special firms and their proportion), ξ_0 (real operating cost), π_d (exogenous departure rate among producing firms), λ_o (fraction of capital stock recovered under default in times of ordinary credit), σ_ε^2 (normal firms' idiosyncratic productivity variance) and π_0^{ε} (probability of unusual ε draw) to reproduce the following empirical targets drawn from the BDS except where otherwise noted. Our targets are: (1) the average debt-to-asset ratio of nonfarm nonfinancial businesses (0.372; Flow of Funds 1954-2006), (2) average debt recovery rates under default in non-crisis years (0.434; Moodys

---

[11] There are alternatives to ensure a similar growth phase for firms; one is to add a random component to firm operating costs that is proportional to capital.

Analytics), (3) unconditional average firm exit rate (10 percent), (4) average debt-to-asset ratio among entrants (40.0 percent; Kauffman Firm Survey), (5) average employment (size) among entrants, relative to incumbent firms (28.5 percent), (6) population and employment share of firms with fewer than 20 employees (88.4 and 20.3 percent, respectively), (7) population and employment share of firms with 20-499 employees (11.3 and 32.2 percent, respectively), (8) employment share of firms aged 0 - 5 years (17.0 percent), (9) survival rates through age 5 (45.4 percent), and (10-12) firm exit rates in years 1, 2 and 3, respectively (21.3, 15.9 and 13.4 percent).

6 Results

Hindrances to borrowing arise endogenously in our model and are backward- *and forward*-looking. Among any group of firms with the same cash, irrespective of their current capitals, a firm with higher current productivity can borrow more at a better rate than one with lower productivity. Compared to settings with exogenous borrowing limits or collateral constraints, allowing productivity to matter in loan arrangements tends to reduce capital misallocation across firms. This creates a tension in efforts to reproduce the slow maturing phase evident in the U.S. firm age-size distribution (see Table 2) alongside the extreme skewness of the unconditional size distribution (Table 3) simultaneously in a model of this class where heterogeneity stems from mean-reverting productivity differences. We will show below that our model's aggregate responses to financial shocks are influenced significantly by which of these goals is emphasized more heavily in calibration. The more closely the model fits the age distribution, the greater the rise in incumbent misallocation following a credit shock. By contrast, under our alternative calibration heavily weighting the size distribution, changes along the extensive margin play the dominant role.

TABLE 2. U.S. Employment and Population Shares by Age 1988-2006

age	0	1	2	3	4	5	6-10	11-15	16+
pop.	10.66	8.21	6.97	6.11	5.46	4.89	18.84	11.87	26.99
emp.	2.94	2.87	2.65	2.55	2.43	2.34	10.54	9.20	64.47
rel. emp.	28.02	36.39	39.46	42.68	45.52	48.60			

Whereas slow growth comes very naturally in a model with exogenous collateral constraints, it is difficult to reproduce the moments of Table 2 in our model with forward-looking loan rate determination, plus selection effects in firms' exit decisions and absent real frictions. Our baseline calibration does a respectable job in matching these only with the inclusion of the 0 idiosyncratic shock realization raising default risk for young firms that have not yet accumulated much cash. However, this calibration fails to reproduce the most striking feature of the unconditional size distribution in Table 3, the long thin right tail making up a huge fraction of aggregate employment.

TABLE 3. U.S. Employment and Population Shares by Size 1979-2006

employees	1-19	20-499	500+	1-9	10-49	50-249	250+
emp.	20.3	32.2	47.6	12.5	18.5	16.0	53.0
pop.	88.4	11.3	0.4	76.1	19.7	3.5	0.7

By contrast, our alternative calibration succeeds in reproducing extreme skewness in the size distribution by introducing special firms drawing productivities far above ordinary firms'. However, because special firms imply little default risk for lenders, such firms are able to reach their mature size very rapidly. This makes special firms huge relative to other firms in their cohort, thereby destroying our model's fit to the aging profile in Table 2. In this case of our model, we can either fit the conditional exit rates at each age or the employment shares, but not both. This has an unfortunate implication for the alternate model's response to a credit shock below, making the extensive margin misallocation channel too important relative to the usual incumbent firm channel in driving the 5.6 percent GDP decline we seek, and thus sharply over-stating reductions in the number of firms relative to the 5.5 percent lost over the Great Recession. Our baseline model has a more even mix of both channels at work in its credit shock response, because it always has at least 4-5 cohorts struggling towards maturity, so we will place greater emphasis on this case.

We will show below that shocks to financial conditions cause large reductions in real economic activity in both specifications of our model. This is in part because, in endogenizing firms' borrowing limits, we have introduced procyclical entry and countercyclical exit, which themselves amplify the effects of shocks in our model. We begin this section with a summary of our model's steady state. Next, we present moments drawn from a long simulation with real and financial shocks. Thereafter, we compare our results to the most recent U.S. recession.

6.1 Steady state

The aggregate implications of financial frictions in our model are considerable even in steady state. Relative to an otherwise identical economy with perfect credit markets, the allocation of capital is distorted by the fact that firms' choices of future capital depend upon their cash on hand. Absent financial frictions, each firm with current productivity ε_i would simply select k' to equate the unit purchase price of investment to its expected discounted marginal return to capital, $q_0 \sum \pi^{\varepsilon}_{ij} [\pi_k (k', \varepsilon_j) + 1 - \delta]$. Thus, k' would depend only upon ε_i, and the coefficient of variation of firms' capital choices, conditional on ε, would be 0. Alongside this intensive margin misallocation among continuing firms, realized default generates extensive margin misallocation in the form of lost firms and their capital sitting idle while that exit takes place.

TABLE 4. Percent Population Shares [cut-off cash] by Type: Median Productivity

firm type	*decisions*	*baseline case*	*size-oriented case*
exposed 2 (default)	immediate exit	0.04	3.15
exposed 2 (continue)	(k',b') affected by x	37.10 [0.00]	22.90
exposed 1	$k^*(\varepsilon)$ and $b' = \frac{k^*(\varepsilon)-x}{q_0}$	55.40 [0.19]	67.60
impervious	$k^*(\varepsilon)$ and $b' = B^w(\varepsilon)$	7.50 [4.44]	6.30
	$cv\Big(k'(x, \varepsilon_M) > 0\Big)$	1.73	0.14

Table 4 summarizes the distribution of cash-on-hand for firms entering the period drawing the median positive productivity in each case of our model. The top two rows are cash-poor, exposed type-2 firms that have not yet reached their efficient capital stocks. These firms' investments depend on their cash, and they have an immediate risk of default. The small group in row 1 that would have cash on hand below 0 if they operated are type 2 firms that default and immediately exit.[12] The group in the next row are type 2 firms that produce and enter the coming period with capital levels influenced by their cash. These firms are responsible for the non-zero coefficient of variation listed in the table's final row, a convenient statistic with which to gauge the strength of the traditional incumbent firm misallocation channel in our model. Notice that this steady-state measure of intensive-margin misallocation is 1.73 under our baseline calibration emphasizing the slow growth phase of young firms, whereas it is only 0.14 in the alternative case that largely ignores slow-growth in favor of the thin right tail in the firm size distribution.

---

[12] Threshold cash levels required to prevent default are negative at higher productivity realizations.

Aggregating over all values of ε, type 2 continuing firms are one-third of all producers in the baseline case of our model, and make up 10 percent of total production. Such firms are 20 percent of producers in our alternate, size-oriented case and responsible for only 3.7 percent of production. There, the presence of extreme right-tail firms implies almost no maturing phase for the mean firm and so far weaker traditional misallocation. In both instances, however, the misallocation associated with type 2 exposed firms has large effects. Steady-state GDP is 26 percent lower in the baseline model compared to its frictionless counterpart if we make no adjustment to prevent a natural rise in the average number of firms absent default; even when we reduce the number of potential entrants in the frictionless model to imply the same measure 1 firms, a 4.6 percent gap remains. Interestingly, this producer-controlled measure of GDP loss is larger, at 6.8 percent, in the alternate model despite its weaker incumbent-firm misallocation; this stems from that model's higher default rate and the indirect losses arising from idle capital in the event of endogenous exit.

6.2 Business cycles

We next consider business cycles in our model. We solve for stochastic equilibrium using the method of Krusell and Smith (1998), approximating the endogenous component of the aggregate state using the first moment of the distribution of capital. Table 5 reports the results of a 5000 period simulation of our full baseline model, where business cycles are driven by shocks to total factor productivity and by financial shocks.

TABLE 5. Baseline Business Cycles with TFP Shocks and Credit Shocks

$x =$	GDP	C	I	N	K	r
$mean(x)$	0.512	0.424	0.088	0.337	1.182	0.042
σ_x/σ_{GDP}	(2.099)	0.493	4.019	0.618	0.540	0.421
$corr(x, GDP)$	1.000	0.872	0.949	0.921	0.106	0.679

Results of 5000 period simulation. Columns: GDP, consumption, investment, hours worked, capital and risk-free real interest rate. Rows 2 and 3 report second moments for HP-filtered series using weight 100; $corr(exit, GDP) = -0.306$; $corr(entry, GDP) = 0.152$.

Given our section 5 calibration, credit shocks are rare, and business cycles are largely driven by TFP shocks. As such, the response of our economy with a rich distribution of firms broadly resembles a typical equilibrium business cycle model. GDP is more variable than consumption, investment is more variable than output, and both series are strongly procyclical, as is hours worked. Unlike a typical business cycle model, however, this setting has an endogenous number of producers. As in U.S. firm-level data, our baseline model's exit rate is countercyclical and its entry rate is procyclical.

Each feature noted in Table 5 also holds in the size-oriented version of our model (not shown), save one. While the level of entry is procyclical in that alternate case, the entry *rate* is essentially acyclical with a -0.037 contemporaneous correlation with GDP. There, countercyclical changes in exit generate such large movements in the denominator that greater variation in the number of entrants is needed to reproduce the weak procylicality of entry rates seen in the data.

TABLE 6. Baseline Business Cycles with TFP Shocks Only

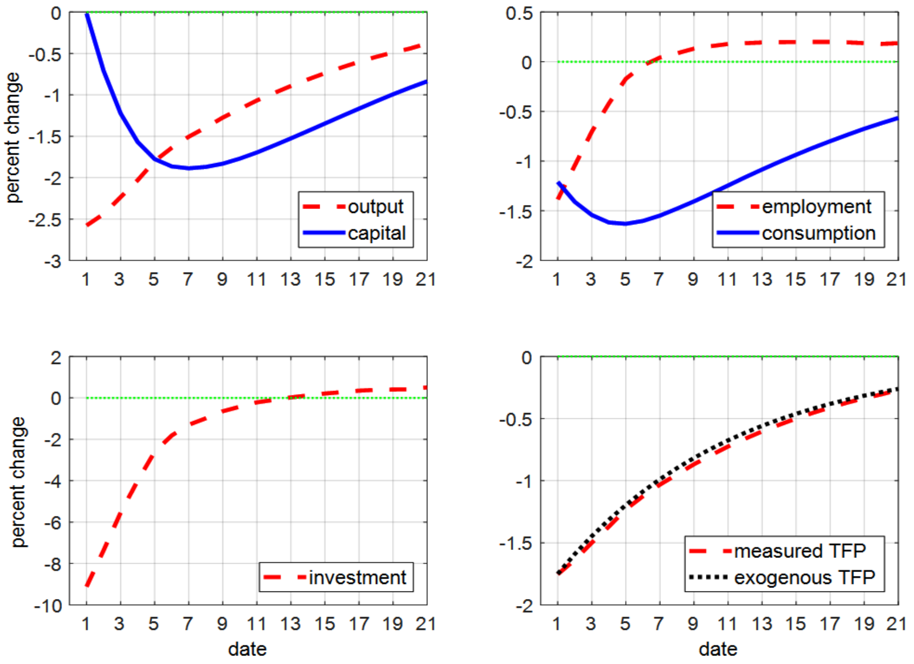
$x =$	GDP	C	I	N	K	r
$mean(x)$	0.514	0.425	0.088	0.337	1.188	0.042
σ_x/σ_{GDP}	(1.969)	0.515	3.624	0.549	0.501	0.458
$corr(x, GDP)$	1.000	0.936	0.968	0.944	0.090	0.650

Results of 5000 period simulation; $corr(exit, Y) = -0.367$; $corr(entry, Y) = 0.296$.

Table 6 examines a version of our baseline model with only aggregate productivity shocks. Comparing this to Table 5, we see that the presence of credit shocks marginally raises the cyclical volatility of GDP a bit and reduces its correlations with consumption, employment and investment somewhat. Otherwise, business cycle moments across the two tables are similar. We will confirm using impulse responses below that this apparent invariance to financial shocks is entirely because such shocks occur in only 7 percent of all dates and not because fluctuations driven by such shocks are near-isomorphic to those driven by TFP shocks. Those results are foreshadowed to an extent by the larger relative volatilities in employment and investment in Table 5 versus Table 6. Each of these observations applies equally well for the analogues to Tables 5 - 6 from our alternate calibration, including the ordering of volatility and correlation changes, so we omit those tables.

Figure 1 shows our baseline model's impulse responses following a persistent 1.75 percent shock to the exogenous component of total factor productivity. In most respects, the economy's response is similar to that of a representative firm equilibrium business cycle model without financial frictions. The largest response is in investment, and the GDP response exceeds that of consumption. Hours worked, consumption and investment are procyclical. Responses in output, hours and investment are essentially monotone, and consumption exhibits the customary hump. The lower right panel of the figure shows a small gap between the exogenous change in TFP and measured TFP, which widens slightly over time. This happens due to a small, gradual fall in the number of producers (not shown here) and, in turn, causes a slight flattening in the GDP recovery. Responses to the TFP shock in the alternate version of our model (not shown) share these features.

FIGURE 1. TFP-Shock Recession in the Baseline Model



With the persistent decline in productivity, the risk-free real interest rate falls, implying a rise in the risk-free discount rate on loans. However, many firms cannot borrow at that rate, or cannot borrow what they wish. While not shown here, our analysis in section 3.3 implies that the negative TFP shock yields a general worsening of firms' credit situations. Lower productivity raises default probabilities; that reduces firms' ability to borrow at the rates they otherwise would, raising the fraction of type 2 exposed firms and further rationing such firms' investment activities. With this said, the 2 percent gradual decline in the overall level of debt in the economy this generates is similar to that from a model with exogenous collateral constraints (Khan and Thomas (2013)).

A new feature of our environment is its allowance for endogenous changes in the number of producers, and thus a second channel affecting capital allocation. We noted above that the fall in aggregate productivity worsens credit conditions for firms. Given worsened loan rate schedules in light of reduced productivity, alongside lower productivity on its own, firm values conditional on operating fall, driving increased exit. At the impact of the shock, the exit rate rises roughly 0.06 percentage points, reducing the number of producers by 0.2 percent. Thereafter, exit rates continue rising for several periods as firms vulnerable to default continue experiencing a slight worsening in their borrowing terms that renders them more so. Because the same forces prompting these changes also reduce the number of entering firms relative to normal, the overall measure of producers steadily declines over about 7 periods and ultimately reaches 0.7 percent below its starting value. Given decreasing returns to scale at the firm ($\alpha + \nu = 0.865$), these changes deliver a separate source of misallocation beyond that associated with inefficient capital choices among financially exposed continuing firms. That extra source of propagation, changes in the measure of producers, drives the small wedge between exogenous and measured TFP seen in Figure 1.

We next explore aggregate responses in our economy to a financial shock effectively reducing firms' cash-on-hand and raising their costs of borrowing. This shock is an unanticipated change in credit conditions from ζ_o to ζ_l and lasts for 4 periods. While ζ is at its low value in dates 1 through 4, all firms experience additional financial costs, $\xi_1(\varepsilon)$, of operation that amount to balance sheet shocks. Because loan rate schedules are influenced by x (section 3.3), this directly increases the credit frictions facing firms. However, recall that the shock has no direct real effect on the economy, as all defaulting firms' capital not recouped by the intermediary is lump-sum returned to households, and the additional costs reducing firms' balance sheets are purely financial.

FIGURE 2. Credit-Shock Recession in the Baseline Model

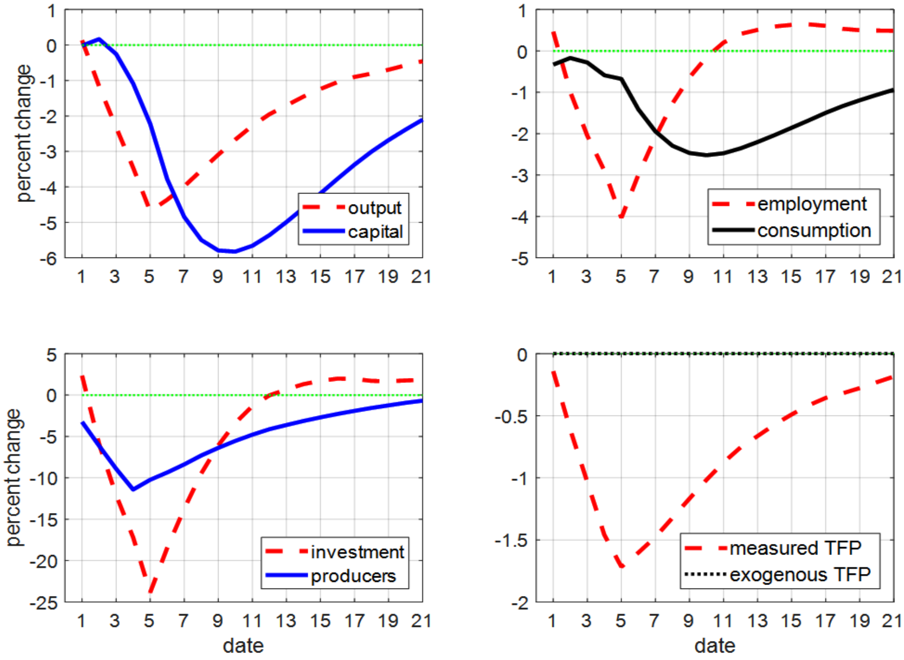


Figure 2 shows the financial shock recession in the baseline case of our model emphasizing the empirical age and size distribution in its calibration. In this case, the financial shock lowers the capital recovery rate in the event of default from 0.37 to 0, and it turns on the financial operating costs $\xi_1(\varepsilon)$ described in section 5, with a 3.5 percent scale factor. Anticipating a rise in default next period due to firms' worsened cash positions, alongside greater loss from each default, the financial intermediary demands higher interest rates for loans involving a nonzero probability of default. This persistently worsens the situations of type 2 financially exposed firms. Moreover, it causes an unusually large number of type 1 firms to become type 2, so their investment decisions are no longer purely a function of productivity. Reductions in the numbers of producers are a second important contributing factor; as noted above, our economy generates procyclical entry and countercyclical exit in response to the financial shock.

The top right and lower left panels in Figure 2 show a small initial rise in employment and investment as households respond to the loss in future income they expect the two forms of increased misallocation will cause. Thereafter, GDP, employment and investment steadily fall over several periods. The driving force behind this is an anticipated decline in endogenous productivity in the lower right panel, implying low returns to investment over several periods. Part of that endogenous TFP declines comes because type 2 firms' capital choices shift further from the frictionless capitals associated with their productivities, and because there are more such firms than usual. However, this increased intensive-margin misallocation among incumbents is amplified by changes in the numbers of firms (production locations) shown in the lower left panel.

The worsening of credit conditions simultaneously raises default thresholds and pushes an unusually large number of firms' cash positions down into regions involving default risk. While not shown here, this drives unusually high exit rates, and unusually low entry rates, for so long as the credit situation does not improve. At the impact of the shock, when capital stocks are already in place, the aggregate exit rate rises about 0.7 percentage points, driving roughly a 3.4 percent fall in the measure of producers in the lower left panel of Figure 2. As the same time, the entry rate falls by almost 2.5 percentage points. Over the next few periods, the entry rate regains some ground, to around 1.75 percentage points below normal, as the number of firms in its denominator shrinks, whereas accumulated damage to type 2 firms' production levels drives a slight further increase in the exit rate. On balance, the number of producers steadily falls over dates 2 - 4, with an ultimate loss of about 11 percent of the economy's producers by the time credit conditions are restored. Given the usual per-period supply of potential entrants and the fact that only a fraction of them ever chose to enter production, it takes many periods for the number of firms to recover from such damage. Because the stock of firms is itself a valuable input for aggregate production, these losses slow the aggregate productivity recovery, and so hold down returns to savings after credit conditions are restored in date 5. That, in turn, delivers sluggish recoveries in investment and employment, slowing the recovery in aggregate capital and, hence, in GDP and consumption.

FIGURE 3. Credit-Shock Recession in the Alternate Model

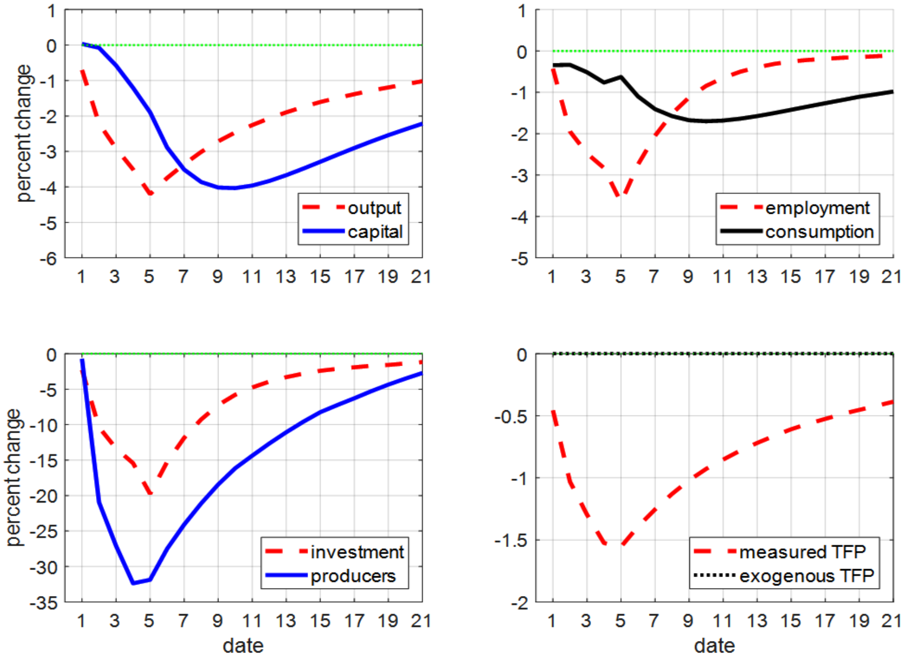


Figure 3 presents responses to the financial shock in the alternate case of our model where the calibration exercise emphasizes moments from the unconditional size distribution at the expense of those from the age distribution. The same mechanisms are present here as in the discussion immediately above. However, recall that this case emphasizing extreme skewness in the employment distribution is achieved through the presence of special firms whose productivity process virtually ensures they will reach the far right tail of the size distribution. That fact implies that, even at early ages when they have very little cash, such firms can access virtually risk-free loans to race toward their optimal size. Since that size is extreme relative to other firms, special firms dominate employment in every cohort, implying that the mean firm matures to full size almost immediately after birth. Thus, as a proportion of investment, there is little incumbent firm misallocation in steady state, and lowering the capital recovery rate in the event of default from its usual 0.47 even to zero does little to increase that misallocation. This makes this case of our model extremely reliant on the extensive-margin channel for its financial-shock recession and implies a weak aggregate elasticity of response. Whereas the scale factor on balance sheet shocks in the baseline case of our model was

3.5, that scale factor must be 42 to generate the same ultimate GDP decline here. While it may be difficult to dismiss the plausibility of this size balance sheet shock, the decline it implies for total producers is an order of magnitude larger that the 5.5 percent counterpart in the data.

6.3 The recent recession

Table 7 compares the peak-to-trough changes from our baseline model in response to the TFP and credit shocks described above with observations from the 2007 U.S. recession, and Table 8 presents the same comparison for our alternate model calibration. In each case, we set the size of the TFP shock to match the endogenous TFP fall caused by the credit shock.[13] We may confine our discussion to the baseline case in Table 7, noting that Table 8 shares very similar features.

TABLE 7. Peak-to-trough Declines: U.S. 2007 Recession and Baseline Model

	trough	GDP	I	N	C	TFP	Debt
data	2009$Q2$	5.59	18.98	6.03	4.08	2.18	25.94
credit shock*	5	4.69	23.85	4.03	0.68 (2.53)	1.72	9.47
tfp shock	1	2.58	9.07	1.39	1.22 (1.64)	1.75	2.04

*credit shock: 4-pds of $\lambda = 0$ (versus 0.37) and 3.5%-scaled financial operating costs

TABLE 8. Peak-to-trough Declines: Alternate Model Calibration

	trough	GDP	I	N	C	TFP	Debt
credit shock*	5	4.20	19.75	3.66	0.63 (1.70)	1.56	10.69
tfp shock	1	2.26	7.20	1.19	1.22 (1.38)	1.55	2.83

*credit shock: 4-pds of 42%-scaled financial operating costs with λ at normal value

The data row in Table 7 reports seasonally adjusted, HP-filtered real quarterly series and measures declines between 2007Q4 and 2009Q2. The one exception to this is the debt entry, where we report the ultimate drop in the stock of real commercial and industrial loans, which came later. The credit shock row reports the declines in our model in response to the credit shock described above. Declines in real series are reported as of the GDP trough date, period 5. In contrast to other real series, the

---

[13] The measured TFP drop is marginally higher because we chose the exogenous TFP shock to match the 1.72 percent (1.56 percent) endogenous decline over the credit shock recession without adjusting for the slight rise in misallocation the productivity shock causes by raising default and thus reducing the measure of producers.

declines in consumption continue for several further periods, as GDP recovers gradually while the rebuilding of the capital stock begins. Consumption ultimately reaches 2.53 percent below average around date 10. As in the data, debt in the model falls more slowly than GDP, and we report the ultimate decline at its trough, which occurs in date 8.

The losses in GDP, investment, employment and debt are disproportionate in the credit shock row relative to the endogenous fall in measured TFP. This unusual aspect of our model's response to a credit shock response resembles that over the 2007 recession. In our model, it appears to stem from the strong disincentives for investment in both physical capital and firms when misallocation not only worsens, but is anticipated to worsen further. The non-monotone path of measured TFP (despite the monotone path of ζ) itself happens as an increasing number of young cohorts are affected by tightened credit conditions, and vulnerable incumbents are forced to exit.

By contrast, our model's GDP, investment and employment responses to the TFP shock are monotone. Thus, the GDP trough occurs at date 1 in the TFP shock row, and we report date 1 declines for all series except debt. Debt falls negligibly at date 1, so here again we report the ultimate drop in that series, which happens 5 periods after the impact of the TFP shock.

The credit and TFP shocks are equivalent in their effects on measured productivity; both model rows capture roughly 80 percent of the observed decline. However, there are large differences in other columns. The credit shock drives roughly 84 percent of the observed reduction in GDP, 67 percent of the observed decline in hours worked, 125 percent of the empirical drop in investment, and about 37 percent of the ultimate decline in debt. By contrast, confronting the same model with the same-sized exogenous TFP shock delivers only 46 percent of the reduction in GDP seen over the 2007 recession. The fall in investment there is also about half that in the data, there is no significant reduction in debt, and the fall in employment is less than one-fourth the observed decline.

7 Concluding remarks

We have developed a model where the aggregate state includes a distribution of firms over idiosyncratic productivity, capital and debt. Firms are risky borrowers, and equilibrium loan rate schedules reflect each borrower's default risk. We associate default with exit, and find that our model economy's response to a persistent shock to exogenous total factor productivity is similar to a representative firm business cycle model without financial frictions. In contrast, while our economy is consistent with the average aggregate debt-to-asset ratio in the U.S., and only a small fraction of production takes place in firms with investment hindered by financial frictions ordinarily, a credit shock affecting firms' balance sheets and debt recovery rates in the event of default generates a large and lasting recession with some unique features.

Our model's credit shock recession unravels through two sources of capital misallocation that grow over several periods. First, the misallocation of capital across continuing firms gradual worsens as several subsequent cohorts of young firms with little cash but high expected future productivity encounter worsened borrowing terms. Second, there are large endogenous reductions in firm entry and increases in exit following the credit shock that sharply reduce the number of firms over the downturn. Because the number of production locations is itself a valuable input affecting the productivity of the aggregate stock of capital, this second misallocation effect compounds the first in moving the distribution of production further from the efficient one and reducing measured productivity. Furthermore, because this substantial damage to the stock of firms takes time to repair, our model delivers a far more gradual rebuilding of the aggregate capital stock, and thus more gradual recoveries in employment and investment than would otherwise occur.

Despite marked differences in response to real versus financial shocks, our setting delivers business cycle moments similar to a snapshot of postwar U.S. business cycles, because credit shocks in our model, as in the data, are infrequent. While aggregate responses following a productivity shock are impervious to the details of micro-level calibration, we have seen that the average age-size distribution in the model affects its responses to a credit shock in important ways. The explanation for this rather unique aspect of our model lies in the fact that, whereas an aggregate TFP shock's consequences are for the most part level across firms, a financial shock has disparately large effects on young, small firms most reliant on external finance and most vulnerable to default and exit.

References

[1] Aguiar, M. and Gopinath, G. (2006) ‘Defaultable debt, interest rates and the current account,’ *Journal of International Economics* 69(1), 64-83.

[2] Almeida, H., M. Campello, B. Laranjeira, and S. Weisbenner (2011) ‘Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis,’ *Critical Finance Review* 1, 3-58.

[3] Arellano, C. (2008) “Default Risk and Income Fluctuations in Emerging Economies,” *American Economic Review* 98(3), 690-712.

[4] Arellano, C., Y. Bai and P. Kehoe (2012) “Financial Markets and Fluctuations in Uncertainty,” *Federal Reserve Bank of Minneapolis Staff Report* 466.

[5] Bates, T. W., K. M. Kahle and R. M. Stulz. (2009) “Why Do U.S. Firms Hold So Much More Cash than They Used To?” *Journal of Finance* LXIV(5): 1985-20.

[6] Bianchi, J. and E. G. Mendoza (2012) “Overborrowing, Financial Crisis and Macro-prudential Policy,” University of Maryland working paper.

[7] Buera, F. and Moll, B. (2015) “Aggregate Implications of a Credit Crunch,” *American Economic Journal: Macroeconomics* 7, 1-42.

[8] Buera, F. and Shin, Y. (2013) “Financial Frictions and the Persistence of History: A Quantitative Exploration,” *Journal of Political Economy* 121, 221-272.

[9] Campbell, J. (1998) “Entry, Exit, Embodied Technology, and Business Cycles,” *Review of Economic Dynamics* 1: 371-408.

[10] Chatterjee, S., D. Corbae, M. Nakajima and J. V. Ríos-Rull (2007) “A Quantitative Theory of Unsecured Consumer Credit with Risk of Default,” *Econometrica* 75, 1525-89.

[11] Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno. 2014. "Risk Shocks." *American Economic Review*, 104(1): 27-65.

[12] Cociuba, S. E., Prescott, E. C. and A. Ueberfeldt (2012) “U.S. Hours and Productivity Behavior Using CPS Hours Worked Data: 1947-III to 2011” University of Western Ontario working paper.

[13] Cooley, T. F., Prescott, E. C., 1995. “Economic Growth and Business Cycles.” in *Frontiers of Business Cycle Research*, Cooley, T. F. Editor, Princeton University Press.

[14] Cooper, R. and J. C. Haltiwanger (2006) “On the Nature of Capital Adjustment Costs,” *Review of Economic Studies* 73, 611-33.

[15] David, Joel, Hugo A. Hopenhayn and Venky Venkateswaran, 2013. "Information, Misallocation and Aggregate Productivity" Working paper.

[16] Dunne, T., M. Roberts, and L. Samuelson (1989): “The Growth and Failure of US Manufacturing Plants,” *Quarterly Journal of Economics* 94, 671–98.

[17] Eaton, J. and M. Gersovitz (1981) “Debt with Potential Repudiation: A Theoretical and Empirical Analysis" *Review of Economic Studies* 48, 289-309.

[18] Gertler, M. and N. Kiyotaki (2010) “Financial Intermediation and Credit Policy in Business Cycle Analysis” prepared for *The Handbook of Monetary Economics.*

[19] Gilchrist, Simon, Jae W. Sim, and Egon Zakrajsek. 2010. “Uncertainty, Financial Frictions, and Aggregate Investment.” Working paper, Boston University and Federal Reserve Board.

[20] Gomes, J. F., U. Jermann and L. Schmid (2014) “Sticky Leverage” *American Economic Review* forthcoming.

[21] Gomes, J. F. and L. Schmid (2014) “Equilibrium Credit Spreads and the Macroeconomy” Wharton School working paper.

[22] Gomme, P., Ravikumar, B. and P. Rupert (2008), “The Return to Capital and the Business Cycle,” *Review of Economic Dynamics* 14(2): 262-78.

[23] Hsieh, Chang-Tai and Peter J. Klenow, 2009. “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics* 124(4): 1403-1448.

[24] Jermann, U. and V. Quadrini (2012) “Macroeconomic Effects of Financial Shocks” *American Economic Review* 102(2), 238-71.

[25] Khan, A. and J. K. Thomas (2013) “Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity,” *Journal of Political Economy* 121(6), 1055-1107.

[26] Krusell, Per and Anthony A. Smith (1998) ‘Income and Wealth Heterogeneity in the Macroeconomy’ *Journal of Political Economy*, 106(5), 867 - 96.

[27] Nakajima, M. and J. V. Ríos-Rull (2005) 'Default and Aggregate Fluctuations in Storage Economies,' *Frontiers in Applied General Equilibrium Modeling*, Cambridge University Press, chapter 7, 127-150.

[28] Reinhart, C. M. and K. S. Rogoff (2009) *This Time Is Different: Eight Centuries of Financial Folly*, Princeton University Press.

[29] Restuccia, Diego and Richard Rogerson, 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics* 11(4): 707-720.

[30] Rogerson, R. (1988) "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics* 21, 3-16.

[31] Shourideh, A. and A. Zetlin-Jones (2016) "External Financing and the Role of Financial Frictions over the Business Cycle: Measurement and Theory," working paper.