

Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity

Aubhik Khan

The Ohio State University

Julia K. Thomas

The Ohio State University and NBER

June 2011

ABSTRACT

We study the cyclical implications of credit market imperfections in a dynamic, stochastic general equilibrium model wherein firms face persistent shocks to both aggregate and individual productivity. In our model economy, optimal capital reallocation is distorted by two frictions. First, collateralized borrowing constraints limit the investment undertaken by small firms with relatively high productivities. Second, specificity in firm-level capital implies partial investment irreversibilities that lead firms to pursue (S,s) decision rules. This second friction compounds the first in implying that a subset of firms carry a share of the aggregate capital stock disproportionate to their productivity, thereby reducing endogenous aggregate total factor productivity.

In the presence of persistent heterogeneity in capital, debt and total factor productivity, the effects of a financial shock are amplified and propagated through large and long-lived disruptions to the distribution of capital that, in turn, imply large and persistent reductions in aggregate total factor productivity. This paper measures the strength of these effects in a calibrated DSGE setting. We find that an unanticipated tightening in borrowing conditions can, on its own, generate a large recession that is far more persistent than the financial shock itself. This recession, and the subsequent recovery, is distinguished both quantitatively and qualitatively from that driven by exogenous shocks to total factor productivity.

Keywords: Financial frictions, irreversibilities, (S,s) policies, business cycles

Aubhik Khan: mail@aubhik-khan.net, Julia Thomas: mail@juliathomas.net. This paper is a revision of an earlier work, ‘Collateral constraints, capital specificity and the distribution of production: the role of real and financial frictions in aggregate fluctuations.’ We thank Gian Luca Clementi, Simon Gilchrist, Nicolas Petrosky-Nadeau, Adriano Rampini, and Vincenzo Quadrini for helpful comments and suggestions. We also thank seminar participants at Austin, The Bank of England, Cambridge, Edinburgh, The Federal Reserve Banks of Chicago, Cleveland and San Francisco, Maryland, Oxford, Rochester, St. Andrews, UCLA, UQAM, Washington University in St. Louis, Yale and conference participants at the 2010 SED meeting, the 2010 Carnegie-Mellon Conference on Advances in Macro-Finance, the 2010 Wisconsin Conference on Money, Banking and Asset Markets, and the 2011 Philadelphia Workshop on Macroeconomics.

1 Introduction

Can a large shock to an economy's financial sector produce a large and lasting recession? Can it amplify and propagate the effects of a real shock sufficiently to transform recession into depression? Over the past few years, events in the real and financial sectors of the U.S. and other large, developed economies have been difficult to disentangle. If these conditions have reawakened interest in business cycle research, they have also raised concerns about our existing macroeconomic models' ability to address such topics.

In this paper, we develop a quantitative, dynamic, stochastic general equilibrium model to explore how real and financial shocks interact in determining the size and frequency of aggregate fluctuations. In our model, firms experience persistent shocks to both aggregate and individual productivity, while credit market frictions interact with real frictions to yield persistent disruptions to the efficient allocation of capital across them, and thus persistent reductions in endogenous aggregate productivity. Calibrating our model to aggregate and firm-level data, we use it as a laboratory in which to obtain answers to the questions raised above.

Considering the matter from the perspective of a representative agent model, one might expect that the reductions in aggregate capital implied by a temporary tightening in credit markets could not yield sizeable or long-lived real aggregate effects, since investment is a small fraction of GDP. However, disaggregated data reveals that there is substantial heterogeneity across firms in their individual productivity levels, and that there are real frictions limiting the reallocation of capital across them.¹ Indeed, these elements are essential to understanding microeconomic investment patterns. In light of the first fact, a reduction in credit may sharply reduce aggregate total factor productivity by distorting the allocation of production away from the efficient one, placing too little capital in a subset of firms with relatively high productivities.² To the extent that real frictions slow the reversal of such an allocative disruption, the second fact compounds the first, propagating shocks to the provision of credit.

As mentioned above, capital reallocation is distorted by two frictions in our model, one finan-

¹For direct evidence of large and increasing heterogeneity in firm-level productivity, see Comin and Philippon (2005) and the empirical studies cited therein. Elsewhere, Cooper and Haltiwanger (2006) find it is impossible to reproduce microeconomic investment patterns without both large idiosyncratic shocks and adjustment costs limiting capital reallocation.

²Restuccia and Rogerson (2008) show that this endogenous TFP effect is an important component in explaining cross-country per-capita GDP differences.

cial and one real. First, collateralized borrowing constraints limit the investment undertaken by small firms. Second, specificity in capital implies partial investment irreversibilities that lead firms to pursue (S,s) rules with respect to their capital adjustments. The second friction further tilts the distribution of production towards larger firms, further reducing endogenous aggregate total factor productivity. This added element of realism in our setting relative to existing DSGE financial frictions models may be quite important to the transmission and propagation of a financial shock, as we discuss below.

The inclusion of the real friction hindering capital reallocation makes our model difficult to solve and analyze, in that it implies a distribution of firms distinguished by three individual state variables: productivity, debt and capital. While it would be more convenient to omit the partial irreversibility of investment and track firms entirely by their productivity and net worth, we see its presence as essential to the proper calibration of the model along the dimensions that matter most. Capital reallocation is at the heart of what is new in our model's mechanics. As such, it is important that we ensure reasonable agreement there between our model economy and microeconomic evidence from the actual economy it is intended to emulate.

The extent to which changes in the availability of credit disrupt real economic activity depends upon the amount of capital reallocation that would otherwise occur. Alternatively, the response to a credit shock hinges on the initial shape of the cross-sectional distribution of firms, as well as its elasticity with respect to the shock. Both aspects of our model are crucially affected by the volatility and persistence of firm-level productivity shocks. As such, these parameters must be chosen to ensure that the reallocation it predicts in a typical data is plausible relative to what we see in the data. We measure our model's fit in this respect using observations on establishment investment rates drawn from U.S. data. Certainly, the idiosyncratic productivity parameters can be selected so that the model reproduces the mean and standard deviation of establishment-level investment rates. However, unless we include some real friction inhibiting reallocation alongside our calibrated financial friction, no plausible idiosyncratic productivity process can overcome our model's counterfactual tendency for negative serial correlation in firms' investment rates. Thus, we introduce a partial irreversibility in investment and calibrate this jointly with the firm-level productivity process to ensure the microeconomic predictions of our model regarding capital reallocation are sensible before turning to its predictions regarding macroeconomic time series.

When calibrated to the microeconomic evidence on capital reallocation, our model gives rise to

a rich distribution. Within this distribution, a subset of firms have investment curtailed by their current ability to borrow, while a second subset have sufficient resources as to have permanently outgrown the implications of collateral constraints. Most firms fall into a third class, one where borrowing constraints do not currently bind, but the prospect that they may bind in future affects current decisions. In the aggregate, revenues net of labor costs exceed total investment. Thus, our model is consistent with evidence from the U.S. Flow of Funds that the average firm can has sufficient funds to internally finance investment. Nonetheless, because individual firms have differing credit needs and access, a change in financial conditions can affect the level of real economic activity.

Our primary question in this study is whether a temporary crisis in financial markets can generate a large and persistent drop in aggregate productivity by disrupting the distribution of capital further from that implied by firms' relative productivities, thereby further distorting the distribution of production. We are to our knowledge the first to explore this endogenous TFP channel in a quantitative DSGE setting where real frictions slow the reallocation of capital across firms, and where that reallocation is essential in determining the marginal product of the aggregate stock. In keeping with previous results in the literature, we find that aggregate responses to real shocks are largely unaffected by the presence of financial frictions. However, changes in the distribution of capital can have large and long-lived effects in our model economy. Because there are substantial and protracted changes in this distribution when individual firms encounter unexpectedly tight collateralized borrowing limits, and because these changes in turn imply persistent reductions in aggregate productivity, we find that an unanticipated disruption to the availability of credit can, on its own, generate a large and protracted recession.

We also find that the response to a credit shock is qualitatively different from that following a real shock, both at its impact and in the recovery episode. Unlike the response to a productivity shock, the greatest declines in output, employment and investment do not occur at the onset of a credit crisis, and consumption does not fall immediately. Moreover, once credit conditions return to normal, our model predicts the subsequent recovery will be slow, and it will be led by employment and business fixed investment, rather than household consumption spending.

Given the widely-held view that crises in financial markets had a large role in the most recent U.S. recession, we are led to compare the changes in our model following a credit shock to those observed in the data after 2007Q4. Ohanian (2009) has argued that, in comparison to other

episodes over the postwar period, this recession is particularly difficult to explain using a standard equilibrium business cycle model driven by exogenous shocks to total factor productivity. When driven by such shocks, our own model behaves similarly to that standard representative firm model. As such, when we introduce TFP shocks sufficient to generate the observed decline in GDP, we find that the resulting declines in employment, investment, and lending are far weaker than those in the data. Moreover, the changes in TFP required to generate the empirical decline in total production are sharply overstated relative to the observed changes in measured TFP.

When we instead consider a temporary shock affecting individual firms' access to credit in our model, we find the aggregate changes it predicts resemble those from the 2007 U.S. recession in several respects. It delivers a gradual deterioration of GDP, an initial rise in consumption, and an unusually steep decline in investment. It also succeeds in capturing the magnitudes of decline in GDP and investment, and it generates equilibrium reductions in total lending consistent with several measures from the aggregate data. Further, our credit shock leads to shifts in the distribution of production across firms that produce endogenous changes in total factor productivity consistent with those measured from the aggregate data. Tighter borrowing limits worsen the allocation of capital and increase the dispersion in the returns to investment. This is consistent with the finding of Eisfeldt and Rampini (2006) that the benefits to capital reallocation rise in recessions.

We argue that our analytical framework captures real and financial frictions that are vital in explaining actual microeconomic reallocation and that thus have the potential to shape macroeconomic outcomes. As such, our findings here are suggestive that changes in firms' access to credit are important in understanding the recent US recession. With this said, we would not suggest that our model *explains* the recent U.S. recession. Our recession is generated by an unanticipated one-time tightening of collateral constraints that persists for several periods, while the recovery that follows is initiated by the instantaneous return of lending conditions to normal. This simple exercise cannot, in itself, explain the U.S. recovery that began in the second half of 2009. Most notably, while its GDP recovery is gradual, even more so than in the data, the model cannot simultaneously account for the growth in consumption alongside much weaker growth in employment and investment.

The solution method we use to derive competitive equilibrium in our model may be of independent interest. As noted above, firms differ in their total factor productivity, capital and debt,

and the partial irreversibility of investment implies that firms' net worth is insufficient to describe the endogenous component of their individual state. Beyond this, our collateral constraints and our investment irreversibility together lead to nonlinear firm-level decision rules that preclude a solution based on linear approximation.

To solve for dynamic stochastic general equilibrium, with firms' decision rules evolving in response to firm-specific shocks, their existing capital and debt, and the aggregate state of the economy, we use nonlinear methods, alongside aggregate state space approximation. Beyond these elements, our solution method relies on identifying a set of firms that have accumulated sufficient real and/or financial assets such that collateralized borrowing limits will never again affect their choice of capital. We characterize the behavior of these firms, including both their physical investment in capital and their savings in financial assets. This provides the starting point to solve for the decision rules of all other firms in the economy.

The remainder of the paper is organized as follows. Section 2 briefly summarizes the literature most closely related to our work. Next, in section 3, we present our model economy. Section 4 provides some analysis useful in developing a numerical algorithm capable of its solution. In section 5, we describe our calibration to moments drawn from postwar U.S. aggregate and firm-level data and also explain our numerical approach. Section 6 presents results. There, we begin by exploring the mechanics of our model in its deterministic steady state. Next we consider business cycles in our model driven by shocks to the exogenous component of total factor productivity, contrasting the results to those obtained from a reference model without financial frictions as well as a frictionless benchmark with neither real nor financial frictions. Thereafter, we explore the effects of a credit shock. Finally, section 7 concludes.

2 Related literature

Until recently, there has been little quantitative research examining the channels through which changes in the availability of credit influence macroeconomic series like business investment, employment and production in well-articulated dynamic stochastic general equilibrium settings. There is a large related literature exploring how financial frictions influence the aggregate response to non-financial shocks. Leading this literature, Kiyotaki and Moore (1997) develop a model of credit cycles and show that collateral constraints can have a large role in amplifying and

propagating shocks to the value of collateral.³ Our own work follows in the spirit of Kiyotaki and Moore in that the financial frictions we explore are collateralized borrowing constraints. We adopt this approach in part because collateral appears to have an important role in loan contracts and in part for computational tractability in our heterogeneous firm DSGE setting.

While we assume that firms face collateral constraints, there are well-known alternative approaches. Cooley, Marimon and Quadrini (2004) study constrained-optimal dynamic contracts under limited enforceability.⁴ Elsewhere, a large literature examines agency costs as the source of financial frictions.⁵ However, these papers do not consider financial shocks as such. Moreover, they abstract from potentially important heterogeneity across firms under which the allocation of capital, and thus credit, becomes relevant.

Over the past few years, several studies have begun exploring how financial shocks affect aggregate fluctuations. A leading example is Jermann and Quadrini (2010), which examines a representative firm model wherein investment is financed using both debt and equity, while costs of adjusting dividends prevent the avoidance of financial frictions. These frictions stem from limited enforceability of intra-temporal debt contracts, which gives rise to endogenous borrowing limits. Specifically, the firm retains its working capital under default, but the lender is able to recover a fraction of the firm's future value. Shocks to the fraction that the lender can confiscate alter the severity of borrowing limits. Measuring these credit shocks, Jermann and Quadrini find that they have been an important source of business cycles.⁶ In contrast to the Jermann and Quadrini model, the financial frictions in our setting do not significantly dampen the response of the aggregate economy to non-financial shocks. We also introduce a real friction in the form of capital specificity.⁷ Because this hinders the reallocation of capital across firms, it leads to

³Cordoba and Ripoll (2004) and Kocherlakota (2000) argue that these effects are quantitatively minor in calibrated versions of the model.

⁴Firm-level dynamics in our model have some similarities to those in models with constrained optimal dynamic contracts. For example, because our borrowing and investment decisions are inter-related, young firms grow as their ability to borrow rises and mean growth rates fall with age and size. Albuquerque and Hopenhayn (2004) derive these regularities in a model with limited enforceability, while Clementi and Hopenhayn (2006) derive them under private information. However, in contrast to these papers, our firm-level debt is not contingent.

⁵See Bernanke and Gertler (1989), Carlstrom and Fuerst (1997), and Bernanke, Gertler and Gilchrist (1999).

⁶Jermann and Quadrini (2009) adapt this model to address the evolving variability of real and financial variables in the past 25 years. In a related setting, Christiano, Motto and Rostagno (2009) study a New Keynesian model with lending subject to agency costs; they too find that financial shocks are an important source of economic fluctuations.

⁷See Veracierto (2002) for a DSGE analysis of how these frictions affect aggregate responses to productivity

a more gradual evolution of our distribution of firms that, in itself, both dampens and protracts the real effects of credit shocks.

Our emphasis on firm-level productivity dispersion is shared by Arellano, Bai and Kehoe (2010), who examine the role of uncertainty shocks in a model with non-contingent debt and equilibrium default. Gomes and Schmid (2009) also develop a model with endogenous default, where firms vary with respect to their leverage, and study the implication for credit spreads.⁸ In contrast to these papers, we study firm-level capital reallocation and the aggregate response to credit shocks. We find that credit shocks can generate recessions through reductions in aggregate TFP that, in turn, have sharp implications for investment and employment. In emphasizing the endogenous TFP channel, our study is also related to Buera and Shin (2007). They examine the effect of collateral constraints on economic development and show that these frictions can protract the transition to the balanced growth path if capital is initially misallocated.

3 Model

In our model economy, firms face both partial capital fixity and collateralized borrowing limits, which together compound the effects of persistent differences in their total factor productivities to yield substantial heterogeneity in production. We begin our description of the economy with an initial look at the optimization problem facing each firm, then follow with a brief discussion of households and equilibrium. Next, in section 4, we will use a simple implication of equilibrium alongside some immediate observations about firms' optimal allocation of profits across dividends and retained earnings to characterize the capital adjustment decisions of our firms. This analysis will show how we derive a convenient, computationally tractable algorithm to solve for equilibrium allocations in our model, despite its three-dimensional heterogeneity in production.

3.1 Production, credit and capital adjustment

We assume a large number of firms, each producing a homogenous output using predetermined capital stock k and labor n , via an increasing and concave production function, $y = z\varepsilon F(k, n)$.

shocks. Caggese (2007) considers both irreversible capital and collateral constraints; our study is distinguished from his by general equilibrium analysis, partial reversibility in investment, and frictionless within-period borrowing.

⁸Gilchrist, Sim and Zakrajšek (2011) study credit spreads under uncertainty shocks in a model with default. Credit spreads are also emphasized by Gertler and Kiyotaki (2010); they study a model where such spreads are driven by agency problems arising with financial intermediaries.

The variable z represents exogenous stochastic total factor productivity common across firms, while ε is a firm-specific counterpart. For convenience, we assume that ε is a Markov chain, $\varepsilon \in \mathbf{E} \equiv \{\varepsilon_1, \dots, \varepsilon_{N_\varepsilon}\}$, where $\Pr(\varepsilon' = \varepsilon_j \mid \varepsilon = \varepsilon_i) \equiv \pi_{ij} \geq 0$, and $\sum_{j=1}^{N_\varepsilon} \pi_{ij} = 1$ for each $i = 1, \dots, N_\varepsilon$. Similarly, $z \in \{z_1, \dots, z_{N_z}\}$, where $\Pr(z' = z_m \mid z = z_l) \equiv \pi_{lm}^z \geq 0$, and $\sum_{m=1}^{N_z} \pi_{lm}^z = 1$ for each $l = 1, \dots, N_z$.

Because our interest is in understanding how financial constraints interact with the specificity of capital in shaping the investment decisions taken by firms in our economy, we must prevent firms growing so large that none will ever again experience a binding borrowing limit. To ensure this does not occur, we impose exit and entry in the model. In particular, we assume that each firm faces a fixed probability, $\pi_d \in (0, 1)$, that it will be forced to exit the economy following production in any given period. Within a period, prior to investment, firms learn whether they will survive to produce in the next period. Exiting firms are replaced by an equal number of new firms whose initial state will be described below.

At the beginning of each period, a firm is defined by its predetermined stock of capital, $k \in \mathbf{K} \subset \mathbf{R}_+$, by the level of one-period debt it incurred in the previous period, $b \in \mathbf{B} \subset \mathbf{R}$, and by its current idiosyncratic productivity level, $\varepsilon \in \{\varepsilon_1, \dots, \varepsilon_{N_\varepsilon}\}$. Immediately thereafter, the firm learns whether it will survive to produce in the next period.⁹ Given this individual state, and having observed the current aggregate state, the firm then takes a series of actions to maximize the expected discounted value of dividends returned to its shareholders, the households in our economy. First, it chooses its current level of employment, undertakes production, and pays its wage bill. Thereafter, it repays its existing debt and, conditional on survival, it chooses its investment, i , current dividends, and the level of debt with which it will enter into the next period, b' . For each unit of debt it incurs for the next period, a firm receives q units of output that it can use toward paying current dividends or investing in its future capital. The relative price q^{-1} reflecting the interest rate at which firms can borrow and lend is a function of the economy's aggregate state, as is the wage rate ω paid to workers. For expositional convenience, we suppress the arguments of these equilibrium price functions until we have described the model further.

In contrast to the typical setting with firm-level capital adjustment frictions, and unlike a

⁹We have adopted this timing to ensure there is no equilibrium default in our model, so that all firms borrow at a common real interest rate. Because the only firms borrowing are those that will produce in the next period, and the debt they take on is limited by a collateral constraint, firms are always able to repay their debt in the quantitative exercises to follow.

typical environment with financial frictions, real and financial frictions are allowed to interact in our model economy. Our firms' borrowing and investment decisions are inter-related, because each firm faces a collateralized borrowing constraint inside of any period. This constraint takes the form: $b' \leq \theta_b \theta_k k$. Two external forces together determine what fraction of its capital stock a firm can borrow against - the degree of specificity in capital and enforceability of financial arrangements. Here, we simply impose both, deferring the question of their foundations for a future study. In particular, we assume that $\theta_k \in [0, 1]$ is a parameter determining what fraction of a firm's capital stock survives when it is uninstalled and moved to another firm, and $\theta_b \in \mathbf{R}_+$ is the fraction of that collateral firms can borrow against.¹⁰ A financial shock in our model is represented by an unanticipated change in the collateral term, θ_b .

If a firm undertakes any nonnegative level of investment, then its capital stock at the start of the next period is determined by a familiar accumulation equation,

$$k' = (1 - \delta)k + i \quad \text{for } i \geq 0,$$

where $\delta \in (0, 1)$ is the rate of capital depreciation, and primes indicate one-period-ahead values. Because there is some degree of specificity in capital, the same equation does not apply when the firm undertakes negative investment. In this case, the effective relative price of investment is θ_k rather than 1, so the accumulation equation is instead:

$$\theta_k k' = \theta_k (1 - \delta)k + i \quad \text{for } i < 0.$$

In the analysis section to follow, we will show how the asymmetry that firms face in the cost of capital adjustment naturally gives rise to two-sided (S, s) investment decision rules. Firms have nonzero investment only when their capital falls outside a range of inactivity.¹¹ In contrast to a nonconvexity in the capital adjustment technology, this type of adjustment friction implies not only investment inaction among firms within their (S, s) adjustment bands, but also some inertia among firms outside of their (S, s) bands. Because there are no increasing returns in the adjustment technology, and there is instead a linear penalty for negative adjustments, a firm finding itself with an unacceptably high capital stock (given its current productivity) will reduce

¹⁰Throughout our numerical exercises in section 6, we assume that the degree of capital irreversibility, $1 - \theta_k$, is a fixed technological parameter. In ordinary times when aggregate fluctuations arise from changes in productivity alone, θ_b is also a fixed parameter. However, we allow for an unanticipated change in θ_b when we consider the aggregate implications of a credit shock in section 6.3.

¹¹The problem of costly investment reversibility was originally solved by Abel and Eberly (1996).

its stock only to the upper bound of its (S, s) inactivity range. Similarly, a firm with too little capital recognizes that it will incur a linear penalty should it later need to shed capital, so it invests only to the lower bound of its inactivity range.

It should be clear from the discussion above that, alongside its current productivity, a firm's capital adjustment may also be influenced by its ability to borrow (now and in the future). This is in turn affected by the capital (collateral) it currently holds. Note also that the firm's current investment decision may influence the level of debt it carries into the next period. These observations imply that we must monitor the distinguishing features of firms along three dimensions: their capital, k , their debt, b , and their idiosyncratic productivity, ε . Thus, in contrast to models with loan market frictions, but without irreversible investment, a firm's net worth is an insufficient description of its state; capital and debt are distinct state variables.

We summarize the distribution of firms over (k, b, ε) using the probability measure μ defined on the Borel algebra, \mathcal{S} , generated by the open subsets of the product space, $\mathbf{S} = \mathbf{K} \times \mathbf{B} \times \mathbf{E}$. The aggregate state of the economy is then described by (z, μ) , and the distribution of firms evolves over time according to a mapping, Γ , from the current aggregate state; $\mu' = \Gamma(z, \mu)$. The evolution of the firm distribution is determined in part by the actions of continuing firms and in part by entry and exit. As already mentioned, fraction π_d of firms exit the economy after production in each period. These firms invest negatively to shed their remaining capital, returning the proceeds to households, and are replaced by the same number of new firms. Each new firm has zero debt and productivity $\varepsilon_0 \in \mathbf{E}$ drawn from an initial distribution $H(\varepsilon_0)$, and each enters with an initial capital stock $k_0 \in \mathbf{K}$.¹²

We now turn to the problem solved by each firm in our economy. Let $v_0(k, b, \varepsilon_i; z_l, \mu)$ represent the expected discounted value of a firm that enters the period with (k, b) and firm-specific productivity ε_i , when the aggregate state of the economy is (z_l, μ) , just before it learns whether it will survive into the next period. We state the firm's dynamic optimization problem using a functional equation defined by (1) - (4) below.

$$v_0(k, b, \varepsilon_i; z_l, \mu) = \pi_d \max_n [z_l \varepsilon_i F(k, n) - \omega(z_l, \mu) n + \theta_k (1 - \delta) k - b] \quad (1)$$

$$+ (1 - \pi_d) v(k, b, \varepsilon_i; z_l, \mu)$$

After the start of the period, the firm knows which line of (1) will prevail. If it is not continuing

¹²We select k_0 below so that each entrant's capital is χ fraction of the typical stock held across all firms in the long-run of our economy.

beyond the period, the firm simply chooses labor to maximize its current dividend payment to shareholders. Because it will carry no capital or debt into the future, an exiting firm's dividends are its output, less wage payments and debt repayment, together with the remaining capital it can successfully uninstall at the end of the period. The problem conditional on continuation is more involved, because a continuing firm must choose its current labor and dividends alongside its future capital and debt. For expositional convenience, given the partial irreversibility in investment, we begin to describe this problem by defining the firm's value as the result of a binary choice between upward versus downward capital adjustment in (2), then proceed to identify the value associated with each option in (3) and (4).

$$v(k, b, \varepsilon_i; z_l, \mu) = \max \left\{ v^u(k, b, \varepsilon_i; z_l, \mu), v^d(k, b, \varepsilon_i; z_l, \mu) \right\} \quad (2)$$

Assume that $d_m(z_l, \mu)$ is the discount factor applied by firms to their next-period expected value if aggregate productivity at that time is z_m and the current aggregate state is (z_l, μ) . Taking as given the evolution of ε and z according to the transition probabilities specified above, and given the evolution of the firm distribution, $\mu' = \Gamma(z, \mu)$, the firm solves the following two optimization problems to determine its values conditional on (weakly) positive and negative capital adjustment. In each case, the firm selects its current employment and production, alongside the debt and capital with which it will enter into next period and its current dividends, D , to maximize its expected discounted dividends. As above, dividends are determined by the firm's budget constraint as the residual of its current production and borrowing after its wage bill and debt repayment have been covered, net of its investment expenditures.

Conditional on an upward capital adjustment, the firm solves the following problem constrained by (i) the fact that investment must be non-negative, (ii) a borrowing limit determined by its collateral, and (iii)-(iv) the requirements that dividends be non-negative and satisfy the firm's budget constraint.

$$v^u(k, b, \varepsilon_i; z_l, \mu) = \max_{n, k', b', D} \left[D + \sum_{m=1}^{N_z} \pi_{lm}^z d_m(z_l, \mu) \sum_{j=1}^{N_\varepsilon} \pi_{ij} v_0(k', b', \varepsilon_j; z_m, \mu') \right] \quad (3)$$

$$\text{subject to: } k' \geq (1 - \delta)k, \quad b' \leq \theta_b \theta_k k,$$

$$0 \leq D \leq z_l \varepsilon_i F(k, n) - \omega(z_l, \mu)n + q(z_l, \mu)b' - b - [k' - (1 - \delta)k],$$

$$\text{and } \mu' = \Gamma(z, \mu)$$

The downward adjustment problem differs from that above only in that investment must be

non-positive and, thus, its relative price is θ_k .

$$v^d(k, b, \varepsilon_i; z_l, \mu) = \max_{n, k', b', D} \left[D + \sum_{m=1}^{N_z} \pi_{lm}^z d_m(z_l, \mu) \sum_{j=1}^{N_\varepsilon} \pi_{ij} v_0(k', b', \varepsilon_j; z_m, \mu') \right] \quad (4)$$

subject to: $k' \leq (1 - \delta)k$, $b' \leq \theta_b \theta_k k$,

$$0 \leq D \leq z_l \varepsilon_i F(k, n) - \omega(z_l, \mu) n + q(z_l, \mu) b' - b - \theta_k [k' - (1 - \delta)k],$$

$$\text{and } \mu' = \Gamma(z, \mu)$$

Notice that there is no friction associated with the firm's employment choice, since the firm pays its current wage bill after production takes place, and its capital choice for next period also has no implications for current production. Thus, irrespective of their current debt or their continuation into the next period, all firms sharing in common the same (k, ε) combination select the same employment, which we will denote by $N(k, \varepsilon; z, \mu)$, and hence have common production, $y(k, \varepsilon; z, \mu)$.¹³ The same cannot be said for the intertemporal decisions of continuing firms, given the presence of both borrowing limits and irreversibilities. Let $K(k, b, \varepsilon; z, \mu)$ and $B(k, b, \varepsilon; z, \mu)$ represent the choices of next-period capital and debt, respectively, made by firms sharing in common a complete individual type (k, b, ε) . We will characterize these decision rules below in section 4.

3.2 Households

The economy is populated by a unit measure of identical households. Household wealth is held as one-period shares in firms, which we identify using the measure λ .¹⁴ Given the prices they receive for their current shares, $\rho_0(k, b, \varepsilon; z, \mu)$, and the real wage they receive for their labor effort, $\omega(z, \mu)$, households determine their current consumption, c , hours worked, n^h , as well as the numbers of new shares, $\lambda'(k', b', \varepsilon')$, to purchase at prices $\rho_1(k', b', \varepsilon'; z, \mu)$. The lifetime

¹³Here forward, except where necessary for clarity, we suppress the indices for current aggregate and firm productivity, l and i , respectively.

¹⁴Households also have access to a complete set of state-contingent claims. However, as there is no heterogeneity across households, these assets are in zero net supply in equilibrium. Thus, for sake of brevity, we do not explicitly model them here.

expected utility maximization problem of the representative household is listed below.

$$V^h(\lambda; z_l, \mu) = \max_{c, n^h, \lambda'} \left[U(c, 1 - n^h) + \beta \sum_{m=1}^{N_z} \pi_{lm}^z V^h(\lambda'; z_m, \mu') \right] \quad (5)$$

subject to

$$c + \int_{\mathbf{S}} \rho_1(k', b', \varepsilon'; z_l, \mu) \lambda'(d[k' \times b' \times \varepsilon']) \leq \omega(z_l, \mu) n^h + \int_{\mathbf{S}} \rho_0(k, b, \varepsilon; z_l, \mu) \lambda(d[\varepsilon \times k])$$

$$\text{and } \mu' = \Gamma(z, \mu)$$

Let $C^h(\lambda; z, \mu)$ describe the household decision rule for current consumption, and let $N^h(\lambda; z, \mu)$ be the rule determining the allocation of current available time to working. Finally, let $\Lambda^h(k', b', \varepsilon', \lambda; z, \mu)$ be the quantity of shares purchased in firms that will begin the next period with k' units of capital, b' units of debt, and idiosyncratic productivity ε' .

3.3 Recursive equilibrium

A *recursive competitive equilibrium* is a set of functions,

$$\left(\omega, q, (d_m)_{m=1}^{N_z}, \rho_0, \rho_1, v_0, N, K, B, D, V^h, C^h, N^h, \Lambda^h \right),$$

that solve firm and household problems and clear the markets for assets, labor and output, as described by the following conditions.

- (i) v_0 solves (1) - (4), N is the associated policy function for exiting firms, and (N, K, B, D) are the associated policy functions for continuing firms
- (ii) V^h solves (5), and (C^h, N^h, Λ^h) are the associated policy functions for households
- (iii) $\Lambda^h(k', b', \varepsilon_j, \mu; z, \mu) = \mu'(k', b', \varepsilon_j; z, \mu)$, for each $(k', b', \varepsilon_j) \in \mathbf{S}$
- (iv) $N^h(\mu; z, \mu) = \int_{\mathbf{S}} [N(k, \varepsilon; z, \mu)] \mu(d[k \times b \times \varepsilon])$
- (v) $C^h(\mu; z, \mu) = \int_{\mathbf{S}} \left[z \varepsilon F(k, N(\varepsilon, k; z, \mu)) - (1 - \pi_d) \mathcal{J} \left(K(k, b, \varepsilon; z, \mu) - (1 - \delta)k \right) \left(K(k, b, \varepsilon; z, \mu) - (1 - \delta)k \right) + \pi_d [\theta_k (1 - \delta)k - k_0] \right] \mu(d[k \times b \times \varepsilon])$, where $\mathcal{J}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ \theta_k & \text{if } x < 0 \end{cases}$

$$(vi) \mu'(A, \varepsilon_j) = (1 - \pi_d) \int_{\{(k, b, \varepsilon_i) \mid (K(k, b, \varepsilon_i; z, \mu), B(k, b, \varepsilon_i; z, \mu)) \in A\}} \pi_{ij} \mu(d[k \times b \times \varepsilon_i]) + \pi_d \chi(k_0) H(\varepsilon_j),$$

for all $(A, \varepsilon_j) \in \mathcal{S}$, defines Γ , where $\chi(k_0) = \{1 \text{ if } (k_0, 0) \in A; 0 \text{ otherwise}\}$

Let C and N describe the market-clearing values of household consumption and hours worked satisfying conditions (iv) and (v) above, and denote next period's equilibrium consumption and hours worked when $z' = z_m$ as C'_m and N'_m , respectively. 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, $\omega(z, \mu) = D_2 U(C, 1 - N) / D_1 U(C, 1 - N)$, that (b) the bond price, q^{-1} , equal the expected gross real interest rate, $q(z, \mu) = \beta \sum_{m=1}^{N_z} \pi_{lm}^z D_1 U(C'_m, 1 - N'_m) / D_1 U(C, 1 - N)$, and that (c) firms' state-contingent discount factors agree with the household marginal rate of substitution between consumption across states $d_m(z, \mu) = \beta D_1 U(C'_m, 1 - N'_m) / D_1 U(C, 1 - N)$. We compute equilibrium by solving the firm-level optimization problem with these implications of household utility maximization imposed, thereby effectively subsuming households' decisions into the problems faced by firms.

Without loss of generality, we assign $p(z, \mu)$ 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.

$$p(z, \mu) = D_1 U(C, 1 - N) \tag{6}$$

$$\omega(z, \mu) = D_2 U(C, 1 - N) / p(z, \mu) \tag{7}$$

$$q(z, \mu) = \beta \sum_{m=1}^{N_z} \pi_{lm}^z p(z_m, \mu') / p(z, \mu) \tag{8}$$

Our reformulation of (1) - (4) below yields an equivalent description of the firm-level problem where each firm's value is measured in units of marginal utility, rather than output, with no change in the resulting decision rules. Suppressing the arguments of the price functions, exploiting the fact that the choice of n is independent of the k' and b' choices, and using the indicator function $\mathcal{J}(x) = \{1 \text{ if } x \geq 0; \theta_k \text{ if } x < 0\}$ to distinguish the relative price of nonnegative versus negative investment, we have:

$$V_0(k, b, \varepsilon_i; z_l, \mu) = \pi_d \max_n p(z_l, \mu) \left[z_l \varepsilon_i F(k, n) - \omega(z_l, \mu) n + \theta_k (1 - \delta) k - b \right] + (1 - \pi_d) V(k, b, \varepsilon_i; z_l, \mu), \tag{9}$$

$$\text{where } V(k, b, \varepsilon_i; z_l, \mu) = \max_{n, k', b', D} \left[p(z_l, \mu)D + \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} V_0(k', b', \varepsilon_j; z_m, \mu') \right] \quad (10)$$

subject to

$$0 \leq D \leq z_l \varepsilon_i F(k, n) - \omega(z_l, \mu)n + q(z_l, \mu)b' - b - \mathcal{J}(k' - (1 - \delta)k) [k' - (1 - \delta)k] \quad (11)$$

$$\text{and } b' \leq \theta_b \theta_k k. \quad (12)$$

4 Analysis

The problem listed in equations (9) - (12) forms the basis for solving equilibrium allocations in our economy, so long as the prices p, ω and q taken as given by our firms satisfy the restrictions in (6) - (8) above.¹⁵ From here, we begin to characterize the decision rules arising from this problem. Each firm chooses its labor $n = N(k, \varepsilon; z, \mu)$ to solve $z\varepsilon D_2 F(k, n) = \omega$, which immediately returns its current production, $y(k, \varepsilon) = z\varepsilon F(k, N(k, \varepsilon; z, \mu))$. Let $\pi(k, b, \varepsilon)$ represent the earnings of a firm of type (k, b, ε) net of labor costs and debt.

$$\pi(k, b, \varepsilon, z, \mu) \equiv z\varepsilon F(k, N(k, \varepsilon; z, \mu)) - \omega N(k, \varepsilon; z, \mu) - b \quad (13)$$

The challenging objects to determine are D, k' and b' for continuing firms. Turning to these, we will use a simple observation about the implications of borrowing constraints for the value a firm places on retained earnings versus dividends. If the firm places non-zero probability weight on encountering a future state in which its borrowing constraint will bind, the shadow value of retained earnings (which includes the discounted sequence of multipliers on future borrowing constraints) will necessarily exceed the shadow value of current dividends, p .¹⁶ This means that it will set $D = 0$. In this case, the binding budget constraint from equation 11 establishes that the firm's choice of k' directly implies the level of debt with which it will enter into the next period. We refer to any such firm as a *constrained firm*. To be clear, a constrained firm need not currently face a binding borrowing constraint; our definition includes any firm that can now or in future encounter a binding constraint. We will return to the problem solved by a constrained firm below. It is useful to first characterize the decisions of a firm whose capital choices are never affected by borrowing limits.

¹⁵Here, and in many instances below, we suppress the z, μ arguments of price functions, decision rules and firm-level state vectors to reduce notation.

¹⁶This is easily proved using a sequence approach with explicit multipliers on each constraint; see Caggese (2007).

4.1 Decisions among unconstrained firms

Consider a firm that has accumulated sufficient wealth (via $k > 0$ or $b < 0$) such that collateral constraints will never again affect its investment activities. In this case, the sequence of multipliers on all possible future borrowing constraints are zero, and the firm is indifferent between allocating earnings to savings versus paying dividends. We refer to any such firm as *unconstrained*. Importantly, as it is indifferent between savings and paying dividends, an unconstrained firm's marginal value of retained earnings is equal to that of households.

Let W_0 represent the beginning-of-period expected value of an unconstrained firm and W its value if it will continue beyond the current period. These functions are analogous to those defined for any firm in (1).

$$W_0(k, b, \varepsilon_i; z_l, \mu) = \pi_d p \left[\pi(k, b, \varepsilon_i) + \theta_k (1 - \delta) k \right] + (1 - \pi_d) W(k, b, \varepsilon_i; z_l, \mu). \quad (14)$$

As in (2), a continuing unconstrained firm has a binary choice involving capital adjustment. Let $W^u(k, b, \varepsilon_i; z, \mu)$ represent its value it chooses to undertake an upward capital adjustment, and $W^d(k, b, \varepsilon_i; z_l, \mu)$ its value conditional on a downward capital adjustment.

$$W(k, b, \varepsilon_i; z_l, \mu) = \max\{W^u(k, b, \varepsilon_i; z_l, \mu), W^d(k, b, \varepsilon_i; z_l, \mu)\} \quad (15)$$

An unconstrained firm must never again experience a binding borrowing constraint (in any conceivable future state). We assign any such indifferent firm a savings policy just ensuring that, under all possible future paths of (ε, z) , it will have sufficient wealth to implement its optimal investment plan while borrowing more than is permitted by (12). Below, we will define this *minimum savings policy*. While an unconstrained firm's minimum savings policy is affected by its capital choice, its capital choice is independent of its savings or debt, b . Thus, before we derive the savings policy for unconstrained firms, we characterize their capital adjustment.

By construction, an unconstrained firm has the same marginal valuation of savings as a household. It then follows from equation 13 that, if such a firm enters a period with any non-zero debt or savings, b , its value is affected only through the change in current earnings. As current earnings are valued by p , we can express the value of a continuing unconstrained firm of type (k, b, ε) as $w(k, \varepsilon) - pb$, where $w(k, \varepsilon) \equiv W(k, 0, \varepsilon)$. The firm's beginning-of-period expected value inherits the same property; $W_0(k, b, \varepsilon_i; z_l, \mu) = w_0(k, \varepsilon) - pb$, where $w_0(k, \varepsilon) \equiv W_0(k, 0, \varepsilon)$. Given these

observations, we have:

$$\begin{aligned}
W^u(k, b, \varepsilon_i; z_l, \mu) &= p\pi(k, b, \varepsilon_i) + p(1 - \delta)k \\
&+ \max_{k' \geq (1-\delta)k} \left[-pk' + \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} w_0(k', \varepsilon_j; z_m, \mu') \right]
\end{aligned} \tag{16}$$

$$\begin{aligned}
W^d(k, b, \varepsilon_i; z_l, \mu) &= p\pi(k, b, \varepsilon_i) + p\theta_k(1 - \delta)k \\
&+ \max_{k' \leq (1-\delta)k} \left[-p\theta_k k' + \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} w_0(k', \varepsilon_j; z_m, \mu') \right],
\end{aligned} \tag{17}$$

where (13) defines $\pi(k, b, \varepsilon)$, and $\mu' = \Gamma(z, \mu)$. In the above, W^u and W^d are both strictly increasing in k . This in turn implies that W and W_0 are increasing functions of the unconstrained firm's capital, as are the w and w_0 functions defined above.

We may characterize the capital decision rule for an unconstrained firm by reference to two target capital stocks, the upward and downward adjustment targets that would solve the problems in (16) and (17), respectively, were there no sign restrictions on investment. Define the upward target, k_u^* , as the capital a firm would choose given a unit relative price of investment, and define the downward target, k_d^* , as the capital a firm would choose given a relative price at θ_k .

$$k_u^*(\varepsilon_i) = \arg \max_{k'} \left[-pk' + \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} w_0(k', \varepsilon_j; z_m, \Gamma(z_l, \mu)) \right] \tag{18}$$

$$k_d^*(\varepsilon_i) = \arg \max_{k'} \left[-p\theta_k k' + \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} w_0(k', \varepsilon_j; z_m, \Gamma(z_l, \mu)) \right] \tag{19}$$

Notice that each target is independent of current capital and depends only on the aggregate state and the firm's current ε . As such, all unconstrained firms that share in common the same current productivity ε have the same upward and downward target capitals. Note also that, because $\theta_k < 1$ (and because the value function w_0 is strictly increasing in k), the upward adjustment target necessarily lies below the downward target: $k_u^* < k_d^*$.

We are now in a convenient position to retrieve the unconstrained firm's capital decision rule. Given a constant price associated with raising (lowering) its capital stock, and because w_0 is increasing in k , the firm selects a future capital as close to the upward (downward) target as its constraint set allows. Thus, the firm's decision rules conditional on upward adjustment and downward adjustment are as follow.

$$k_u(\varepsilon) = \max \{(1 - \delta)k, k_u^*(\varepsilon)\} \quad \text{and} \quad k_d(\varepsilon) = \min \{(1 - \delta)k, k_d^*(\varepsilon)\}$$

Given these conditional adjustment rules, we know that an unconstrained firm of type (k, b, ε) selects one of three future capital levels, $k' \in \{k_u^*(\varepsilon), k_d^*(\varepsilon), (1 - \delta)k\}$. Which one it selects depends only on where its current capital lies in relation to its two targets. Recalling that $k_u^*(\varepsilon) < k_d^*(\varepsilon)$, it is straightforward to obtain the following decision rule for an unconstrained firm.

$$K^w(k, \varepsilon; z, \mu) = \begin{cases} k_u^*(\varepsilon; z, \mu) & \text{if } k < \frac{k_u^*(\varepsilon; z, \mu)}{1 - \delta} \\ (1 - \delta)k & \text{if } k \in \left[\frac{k_u^*(\varepsilon; z, \mu)}{1 - \delta}, \frac{k_d^*(\varepsilon; z, \mu)}{1 - \delta} \right] \\ k_d^*(\varepsilon; z, \mu) & \text{if } k > \frac{k_d^*(\varepsilon; z, \mu)}{1 - \delta} \end{cases} \quad (20)$$

Unconstrained firms maintain sufficient assets to prevent borrowing limits affecting their future investment. In this sense, they accumulate precautionary savings. Given the decision rule for capital, we isolate a minimum level of financial savings that insures that an unconstrained firm of type (ε, k) will never be affected by borrowing constraints across all possible future $(\varepsilon'; z', \mu')$. Any such firm that maintains a level of debt not exceeding the threshold defined by the minimum savings policy will be indifferent to paying additional revenues in the form of dividends, or accumulating further savings. This, in turn, implies that the firm is willing to follow the minimum savings policy.¹⁷

Let $\tilde{B}\left(K^w(k, \varepsilon; z, \mu), \varepsilon_j; z_m, \Gamma(z, \mu)\right)$ define the maximum debt level at which a firm entering next period with capital $K^w(k, \varepsilon; z, \mu)$ and (ε_j, z_m) will remain unconstrained. The following pair of equations recursively defines the minimum savings policy, $B^w(k, \varepsilon; z, \mu)$.

$$B^w(k, \varepsilon; z, \mu) = \min_{\{\varepsilon_j | \pi_{ij} > 0 \text{ and } z_m | \pi_{im}^z > 0\}} \tilde{B}\left(K^w(k, \varepsilon), \varepsilon_j; z_m, \Gamma(z, \mu)\right), \quad (21)$$

$$\begin{aligned} \tilde{B}(k, \varepsilon; z, \mu) = z\varepsilon F(k, N(k, \varepsilon)) - \omega N(k, \varepsilon) + q \min \left\{ B^w(k, \varepsilon; z, \mu), \theta_b \theta_k k \right\} \\ - \mathcal{J}\left(K^w(k, \varepsilon) - (1 - \delta)k\right) \left[K^w(k, \varepsilon) - (1 - \delta)k \right] \end{aligned} \quad (22)$$

In equation 21, $B^w(k, \varepsilon; z, \mu)$ is derived as the maximum level of debt with which the firm can exit this period and be certain to remain unconstrained next period, given that it adopts the unconstrained capital decision rule. Next, (22) defines the beginning of period maximum debt level under which a firm can adopt the unconstrained capital rule and debt not exceeding that

¹⁷ In effect, our minimum savings policy maximizes the dividends paid by unconstrained firms each period. We adopt this policy rather than the natural alternative minimizing unconstrained firms' period-by-period dividends so as to bound the $b < 0$ levels carried by firms that can with positive probability outlive any set finite age we consider.

identified by the minimum savings policy without paying negative dividends, and hence satisfy the definition of an unconstrained firm.

Notice that \tilde{B} is increasing in the firm's current earnings, since these may be used to cover outstanding debt. The minimum operator imposes the borrowing constraint; if the firm does not have sufficient collateral to borrow to B^w , it can still be unconstrained if it has entered this period with sufficient savings to finance its investment.

Given the decision rule for capital and the minimum savings policy, we can now retrieve unconstrained firms' dividend payments.

$$D^w(k, b, \varepsilon, z, \mu) = \pi(k, b, \varepsilon, z, \mu) - \mathcal{J}\left(K^w(k, \varepsilon) - (1 - \delta)k\right)[K^w(k, \varepsilon) - (1 - \delta)k] \quad (23)$$

$$+ q \min\left\{B^w(k, \varepsilon, z, \mu), \theta_b \theta_k k\right\}$$

The firm's value, listed above in (15), may be expressed as

$$W(k, b, \varepsilon; z, \mu) = pD^w(k, b, \varepsilon, z, \mu) \quad (24)$$

$$+ \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} W_0\left(K^w(k, \varepsilon; z, \mu), B^w(k, \varepsilon; z, \mu), \varepsilon_j; z_m, \mu'\right),$$

where $W_0(k, b, \varepsilon, z, \mu)$ is given by (14), and $\mu' = \Gamma(z, \mu)$.

4.2 Decisions among constrained firms

We now consider the decisions made by a firm that has, until now, been constrained. We begin by evaluating whether or not the firm has crossed the relevant wealth threshold to become unconstrained. If it has, the decision rules described above apply. If it has not, the collateralized borrowing constraint will continue to influence its investment decisions, and its choice of capital and debt will remain intertwined.

To ascertain whether a firm of type (k, b, ε) has become unconstrained, we need only consider whether it is feasible for the firm to adopt the capital rule $K^w(k, \varepsilon)$ and a level of debt not exceeding that implied by the rule $B^w(k, \varepsilon)$, while maintaining non-negative dividends in the current period. If the firm of type (k, b, ε) is able to adopt the decision rules in (20) and (21) without violating the non-negativity of dividends, then it achieves the value given by (24), and it exits the period indistinguishable from any other unconstrained firm that entered the period

with (k, ε) .

$$\begin{aligned} V(k, b, \varepsilon; z, \mu) &= W(k, b, \varepsilon; z, \mu) \text{ iff } D^w(k, b, \varepsilon; z, \mu) \geq 0 \\ &= V^c(k, b, \varepsilon; z, \mu) \text{ otherwise} \end{aligned} \quad (25)$$

Any constrained firm that can adopt the decision rules of an unconstrained firm will always choose to do so, since $V \leq W$. However, when the inequality in the top line of (25) is not satisfied, the firm remains constrained, with value $V^c(k, b, \varepsilon; z, \mu)$.

We approach a continuing constrained firm's problem as follows. First, given its (k, ε) , we isolate a cutoff debt level under which non-negative investment is a feasible option. The lowest level of k' associated with non-negative investment is $(1 - \delta)k$. If this choice is not affordable given the firm's borrowing constraint in (12), it cannot undertake even a trivial upward capital adjustment. Using (11), it follows that, among any group of firms sharing a common (k, ε) , only those with $b \leq q\theta_b\theta_k k + z\varepsilon F(k, N(k, \varepsilon)) - \omega N(k, \varepsilon)$ can consider an upward capital adjustment. Firms with higher levels of debt must choose a downward capital adjustment and repay debt by selling capital.

We identify the maximum capital stocks permitted by the borrowing constraint under upward and downward capital adjustment.

$$\begin{aligned} \bar{k}_u(k, b, \varepsilon) &\equiv (1 - \delta)k + \left[q\theta_b\theta_k k + \pi(k, b, \varepsilon) \right] \\ \bar{k}_d(k, b, \varepsilon) &\equiv (1 - \delta)k + \frac{1}{\theta_k} \left[q\theta_b\theta_k k + \pi(k, b, \varepsilon) \right] \end{aligned}$$

Next we determine the associated choice sets for upward and downward capital adjustment.

$$\begin{aligned} \Lambda^u(k, b, \varepsilon) &= [(1 - \delta)k, \bar{k}_u(k, b, \varepsilon)] \\ \Lambda^d(k, b, \varepsilon) &= [0, \max\{0, \min\{(1 - \delta)k, \bar{k}_d(k, b, \varepsilon)\}\}] \end{aligned}$$

Using these results and recalling V_0 in equation (9), we may express a continuing constrained firm's value as follows.

$$V^c(k, b, \varepsilon; z, \mu) = \max\{V^u(k, b, \varepsilon; z, \mu), V^d(k, b, \varepsilon; z, \mu)\}, \quad (26)$$

$$\begin{aligned} V^u(k, b, \varepsilon; z, \mu) &= \max_{k' \in \Lambda^u(k, b, \varepsilon)} \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} V_0(k', b'_u(k'), \varepsilon_j; z_m, \mu') \\ \text{subject to } b'_u(k') &= \frac{1}{q} \left(-\pi(k, b, \varepsilon) + [k' - (1 - \delta)k] \right) \\ \text{and } \mu' &= \Gamma(z, \mu) \end{aligned} \quad (27)$$

$$\begin{aligned}
V^d(k, b, \varepsilon; z, \mu) &= \max_{k' \in \Lambda^d(k, b, \varepsilon)} \beta \sum_{m=1}^{N_z} \sum_{j=1}^{N_\varepsilon} \pi_{lm}^z \pi_{ij} V_0(k', b'_d(k'), \varepsilon_j; z_m, \mu') \\
\text{with } b'_d(k') &= \frac{1}{q} \left(-\pi(k, b, \varepsilon) + \theta_k [k' - (1 - \delta)k] \right) \\
\text{and } \mu' &= \Gamma(z, \mu)
\end{aligned} \tag{28}$$

Denoting the capital stocks that solve the conditional adjustment problems in (27) and (28) above by $\widehat{k}^u(k, b, \varepsilon)$ and $\widehat{k}^d(k, b, \varepsilon)$, respectively, we obtain the following decision rules for capital and debt.

$$K^c(k, b, \varepsilon; z, \mu) = \begin{cases} \widehat{k}^u(k, b, \varepsilon) & \text{if } V^c(k, b, \varepsilon; z, \mu) = V^u(k, b, \varepsilon; z, \mu) \\ \widehat{k}^d(k, b, \varepsilon) & \text{if } V^c(k, b, \varepsilon; z, \mu) = V^d(k, b, \varepsilon; z, \mu) \end{cases} \tag{29}$$

$$\begin{aligned}
B^c(k, b, \varepsilon; z, \mu) &= \frac{1}{q} [\mathcal{J}(K^c(k, b, \varepsilon; z, \mu) - (1 - \delta)k) [K^c(k, b, \varepsilon; z, \mu) \\
&\quad - (1 - \delta)k] - \pi(k, b, \varepsilon; z, \mu)] \tag{30}
\end{aligned}$$

4.3 Numerical Method

The numerical algorithm we use to solve our model builds on that described in Khan and Thomas (2003, 2008) using the analysis above. However, the discrete choices and three-dimensional heterogeneity arising here from the presence of investment irreversibility and collateralized borrowing, alongside the firm-level productivity shocks, necessitate a nonlinear solution method that is more involved than that used in these papers.

We compute equilibrium by solving the problems of unconstrained and constrained firms in a setting where prices are consistent with market-clearing. Because the distribution in the model's aggregate state is a high-dimensional object, we approximate it with the first-moment of the distribution of capital, applying the algorithm of Krusell and Smith (1998). More specifically, we assume that agents perceive (z, m) as the economy's aggregate state (rather than (z, μ)), where m represents the unconditional mean of the distribution of capital across firms.

We replace Γ with a forecasting rule $m' = \Gamma_m^l(z_i, m)$, $i = 1, \dots, N(z)$. The solution method iterates over forecasting rules, $l = 1, 2, \dots$, until they converge. In each iteration, given the l^{th} set of forecasting rules, we solve the unconstrained and constrained firm value functions, thereby obtaining their decision rules. In this first step of the iteration, firms forecast the prices, p , q , and w as functions of m and z . Next, we run a T period simulation, determining equilibrium prices and aggregate quantities at each date of the simulation given the true aggregate state, the

forecasting rule $\Gamma_m^l(z_i, m)$, and firms' decision rules. Thereafter, we use the simulation data to update the forecasting rules using ordinary least squares regressions.

In the first step of each iteration, where value functions are solved, firms assume that prices are given by log-linear functions of z_i and m , $p^l(z_i, m)$ and $q^l(z_i, m)$.¹⁸ Given these forecasting rules for prices and the forecasting rule for the future state, Γ_m^l , we solve for unconstrained firms' value w using (14) - (17) and $W(k, b, \varepsilon_i; z_l, m) = w(k, \varepsilon) - pb$. In doing so, we obtain the decision rules K^w , B^w , and D^w from (20) - (23), which serve as inputs for solving the constrained firm value function. Next, we solve for the constrained firm value function using (25) - (28) with (9), with W_0 serving as the initial guess for V_0 . In each case, we solve for firm values at a set of points from the firm-level state vector; then we use non-linear multivariate piecewise polynomial spline interpolation to approximate the firm value function consistent with these data. When solving the constrained firm problem, we find it useful to solve for a value function $\widehat{V}(k, \frac{b}{k}, \varepsilon_i; z, m)$ rather than the primitive $V(k, b, \varepsilon_i; z, m)$; this lets us restrict the points where values are solved to avoid irrelevant areas of insolvency where firm debt is so high relative to productivity and capital that it cannot be repaid.

In the second step of each iteration, we simulate the model for 5041 periods. The exact length of our simulation is chosen to best fit our parallel computing environment.¹⁹ We draw a fixed vector of aggregate shock levels, $\{z_t\}_{t=1}^T$, $T = 5041$, and assume an initial distribution of firms that is equal to that obtained in the steady state of the model. This distribution over (k, b, ε) is stored using a fine grid of values. As in the value function solution step, however, firms' decisions are not restricted to these grid points. Rather, when determining the future state, their choices are allocated across the nearest pair of grid points using a weighting scheme that, in expectation, yields the selected value.

At each date of the simulation, the distribution μ_t implied by the previous date's equilibrium decisions is used to compute m_t . Based on this, alongside z_t , firms use W , \widehat{V} and Γ_m^l to forecast their future values associated with any choice of k' and b' . This lets them select their optimal production, debt, and investment in response to any given set of prices (p, q, w) . We solve for the

¹⁸The utility function we assume below implies that the real wage is determined by p , so a separate forecasting rule for it is unnecessary.

¹⁹We solve the model using MPI with 71 computational cores in a Beowulf Cluster. Parallel methods are required, despite our use of the Krusell Smith algorithm, because constrained firms' decisions depend on their productivity, debt and capital, implying a computationally intensive numerical algorithm.

equilibrium (p_t, q_t, w_t) so that the market-clearing conditions (iv) - (v) in section 3.3 are satisfied and asset markets clear. This give us μ_{t+1} , the distribution of firms at the start of the next period. After we have completed the simulation, we derive $\{m_t\}_{t=1}^T$ from $\{\mu_t\}_{t=1}^T$ and estimate new forecasting rules. The resulting rules for our baseline calibrated model, where $\theta_b = 1.35$, are listed in Table 5. Despite the rich distribution of firms in our economy, we find that agents are quite successful in forecasting the aggregate state with no more information on the current distribution than the mean capital stock. The results presented in section 6.2 below will show that this is because changes in z have little effect on the shape of the distribution.

5 Calibration

The data on establishment-level investment dynamics are reported annually. As the mechanics of the reallocation of capital across firms are at the core of our model, we reproduce salient empirical regularities from this data. Accordingly, we set the length of a period to one year.

In the section to follow, we will consider how the mechanics of our model with real and financial frictions compare to those in two relevant reference models - one where there are no borrowing limits and one where there are neither financial nor real frictions ($\theta_k = 1$). These two reference models will help us to isolate how much the *interaction* between credit constraints and micro-level capital rigidities influences our economy's aggregate dynamics. Aside from the values of θ_b and θ_k , all three models share a common parameter set that is selected in our full model to best match moments drawn from postwar U.S. aggregate and firm-level data. However, as the average capital-to-output ratio and hours worked vary little across the three models, the results to follow are unaffected by our decision to maintain a fixed parameter set.

Across our model economies, 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: $z\varepsilon F(k, n) = z\varepsilon k^\alpha n^\nu$. The initial capital stock of each entering firm is a fixed χ fraction of the typical stock held across all firms in the long-run of our full economy; that is, $k_0 = \chi \int k \tilde{\mu}(d[k \times b \times \varepsilon])$, where $\tilde{\mu}$ represents the steady-state distribution therein.

5.1 Aggregate data

We determine the values of β , ν , δ , α , φ and θ_b using moments from the aggregate data as follows. First, we set the household discount factor, β , to imply an average real interest rate

of 4 percent, consistent with recent 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, which corresponds to 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.

We calibrate our model to reproduce an aggregate measure of the indebtedness of firms in the U.S. economy. Specifically, we set the parameter determining our collateral constraint, θ_b , to imply an average debt-to-assets ratio at 0.366, which matches that of nonfarm nonfinancial businesses over 1952-05 in the Flow of Funds. The extent to which the resulting financial frictions affect firm-level and aggregate outcomes depends on the productivity process individual firms face, as well as the extent of investment irreversibility. We will determine these aspects of the model using firm-level data below.

Exact aggregation obtains in the reference model without real or financial frictions; in particular, it has an aggregate production function. We use this reference model to estimate an exogenous stochastic process for aggregate productivity. We begin by assuming a continuous shock following a mean zero AR(1) process in logs: $\log z' = \rho_z \log z + \eta'_z$ with $\eta'_z \sim N(0, \sigma_{\eta_z}^2)$. Next, we estimate the values of ρ_z and σ_{η_z} from Solow residuals measured using NIPA data on US real GDP and private capital, together with the total employment hours series constructed by Prescott, Ueberfeldt, and Cociuba (2005) from CPS household survey data, over the years 1959-2002, and we discretize the resulting productivity process using a grid with 3 shock realizations ($N_z = 3$) to obtain (z_l) and (π_{lm}^z) . We apply this exogenous shock process across all three models; as we will see below, there are no changes in the endogenous component of aggregate total factor productivity in our full model with both real and financial frictions when aggregate fluctuations are driven by exogenous shocks to aggregate TFP.

5.2 Firm-level data

The costly reversibility of investment and the dispersion of firm-level total factor productivity are calibrated to reproduce microeconomic evidence on establishment-level investment dynamics.

We begin by assuming that firm-specific productivity follows an AR(1) log-normal process, $\log \varepsilon' = \rho_\varepsilon \log \varepsilon + \eta'$, with $\eta' \sim N(0, \sigma_\eta^2)$. Next we choose θ_k , ρ_ε and σ_η jointly to reproduce three aspects of establishment-level investment data documented by Cooper and Haltiwanger (2006) based on a 17-year sample drawn from the Longitudinal Research Database. These targets are (i) the average mean investment rate (i/k) across establishments: 0.122, (ii) the average standard deviation of investment rates: 0.337, and (iii) the average serial correlation of investment rates: 0.058.²⁰

While our model has life-cycle aspects affecting firms' investments, the Cooper and Haltiwanger (2006) dataset includes only large manufacturing establishments that remain in operation throughout their sample period. Thus, in undertaking this part of our calibration, we must select an appropriate model-generated sample for comparability with their sample. This we do by simulating a large number of firms for 30 years, retaining only those firms that survive throughout, and then restricting the dates over which investment rates are measured to eliminate life-cycle effects. This restricts attention to firms whose investment decisions are unaffected by their borrowing limits. In implementing this algorithm, we discretize firms' log-normal productivity process using 7 values ($N_\varepsilon = 7$) to obtain $\{\varepsilon_i\}_{i=1}^{N_\varepsilon}$ and $(\pi_{ij})_{i,j=1}^{N_\varepsilon}$.

The idiosyncratic shock process we calibrate has a persistence of 0.653 and a standard deviation of innovations of 0.135. As firms in the model sample are unaffected by borrowing constraints, their investments would respond immediately to changes in their total factor productivities in the absence of costs of uninstalling capital. This implies a negative autocorrelation in investment rates, since capital is determined by lagged investment. The costly reversibility of capital is then essential in reproducing the investment moments reported above, and we set $\theta_k = 0.95$ to reproduce the serial correlation of investment rates in the data.

If we eliminate the real friction, firm-level capital reallocation dramatically increases. Maintaining our idiosyncratic shock process but setting $\theta_k = 1$, the mean and standard deviation of firm-level investment rates rises to 0.31 and 0.85, respectively, while the serial correlation falls to -0.16 . Alternatively, if we reset $(\rho_\varepsilon, \sigma_\eta)$ to match the mean and standard deviation of $\frac{i}{k}$ in the data, the serial correlation falls to -0.2 . When we impose a cost of uninstalling capital, firms become unresponsive to moderate changes in ε . This reduces the variability of their investment, and increases its persistence.

²⁰While not a target in the calibration, our model also closely matches a fourth moment drawn from the Cooper and Haltiwanger study, the fraction of establishment-year observations wherein a positive investment spike ($i/k > 0.20$) occurs: 0.186.

Finally, we choose the exit rate, π_d , and the fraction of the steady-state aggregate capital stock held by each entering firm, χ . We set π_d at 0.10, so that 10 percent of firms enter and exit the economy each year. Next, we set $\chi = 0.10$ so that, in an average date, each entering firm begins with an initial capital that is one-tenth the size of the aggregate stock. If we had assumed constant returns to scale in production, this would imply an employment size of entering firms averaging one-tenth the size of the typical firm in our economy, matching the Davis and Haltiwanger (1992) data. In our model economy, where returns to scale is 0.87, the relative employment size of a new firm is 21 percent. In this sense, our choice of χ_0 is conservative.²¹

The table below lists the parameter set obtained from our calibration.

β	ν	δ	α	φ	ρ_z	σ_{η_z}	θ_b	π_d	χ	θ_k	ρ_ε	σ_η
0.96	0.60	0.065	0.27	2.15	0.852	0.014	1.35	0.10	0.10	0.95	0.653	0.135

Note that these parameters imply only a 5 percent loss incurred in uninstalling capital, and a moderate degree of financial frictions, with firms able to take on debt up to 135 percent of the value of their tangible assets. Also note that firm-level shocks are far more volatile and less persistent than aggregate shocks. Given these aspects of the calibration, our model gives rise to a stationary distribution of firms over (k, b, ε) wherein roughly 86 percent of firms are constrained when one applies the definition from section 4 above. By contrast, the fraction of firms facing a currently binding borrowing limit is 27 percent.

6 Results

6.1 Steady state

We begin by considering the implications of borrowing limits and irreversibilities for the typical decisions made in our economy. Figure 1 overviews the stationary distribution of firms in the baseline case of our full model, presenting three slices of the full distribution. In the top panel, we see the distribution of firms over capital and debt-to-capital levels at the lowest firm-level productivity, while the middle and bottom present the counterparts at the median and highest levels of productivity.

²¹To match the Davis and Haltiwanger relative employment size of an entrant in our setting, firms would have to enter with only 3.3 percent the capital of the typical firm. In that case, firms would take far longer to mature, amplifying the effects of financial frictions as well as the effect of a credit shock increasing those frictions.

Each panel of Figure 1 contains what are essentially two separate distributions. The first, in the foreground, has a distinctly curved shape that reflects an inverse relation between firms' capital stocks and their savings rates. This corresponds to older, wealthier firms that are unconstrained and following the minimum savings policy described in section 4. Such firms have higher capital and higher savings relative to constrained firms, which are distributed near the back of each panel. As would be expected, the mean capital among constrained firms rises with firm-level productivity, and the same is true for unconstrained firms.

The 10 percent of firms newly entering the economy each period are scattered across each ε level according to the ergodic productivity distribution. These firms enter with zero debt and low initial capital (roughly 0.14), and are found in a large spike near the left edge of each panel in Figure 1. After its first date in production, each new firm begins to take on debt in effort to build up its capital. In the absence of the collateralized borrowing limits, young firms would immediately take on a large, temporary debt that would allow them to jump to the capital stock selected by unconstrained firms with the same current productivity level. Here, however, firms with little collateral have a relatively limited ability to borrow, so their capital accumulation is necessarily gradual. As a result, ripples of these entering firms slowly move into higher ranges of k and b/k as they age. In the figure, the youngest firms are found along the back edge of each panel; as they mature, they steadily raise their capital while maintaining a roughly constant borrowing rate typically below the maximum permitted. Those firms that survive long enough eventually reach a level of capital such that they can adopt the unconstrained capital choices consistent with their current productivity while beginning to reduce their debt. Those surviving longer still will, at some point, attain a level of capital and savings such that their investment decisions become impervious to borrowing limits. At this point, they join the distribution of unconstrained firms.

The life-cycle aspects of our model described above may be seen from Figure 2. There we display the average capital and debt choices within a cohort of (initially) 25,000 firms as they age. Notice that the typical firm raises its capital and debt over its first six periods of life. Thereafter, starting in period 7, it begins to reduce its debt and finances the remaining rise in its capital stock fully out of earnings. By age 16, the typical firm has become a net saver, and thereafter joins the distribution of permanently unconstrained firms.

We next consider how the predicted relation between firm size and leverage ratios in our model

compares with the data. Empirical evidence presented by Fama and French (2002) and Rajan and Zingales (1995) indicates that leverage ratios rise with size. Reexamining the finding using a panel of nonfinancial firms in the Compustat, Rampini and Viswanathan (2011) argue that the positive relation between size (book value of assets) and leverage disappears when one accounts for an often overlooked form of leverage heavily used by small firms, the rental of capital.

Figure 2 does not imply that our model generates a strong negative relation between size and leverage. Recall that this figure is not drawn from a balanced panel of firms. Given the constant exit rate, there are fewer firms in the right half of the figure than there are on the left where leverage is roughly constant. For this reason, when we draw a large random sample of firms from our stationary distribution, we find that the sample correlation between size and leverage is -0.22 , far less negative than Figure 2 would suggest. Note also that, because our model has no theory of firm ownership, all leverage in our model is associated with investment loans. By contrast, the data reflects not only borrowing for investment activities, but other activities such as restructuring and mergers and acquisitions. Indeed, Ivashina and Scharfstein (2009) report that two-thirds of syndicated loans in 2007 were associated with these other activities.

Extensions of our model allowing for different borrowing constraints across firms could eliminate the small negative correlation between size and leverage, and even make it positive. For example, consider a version of the model with two types of firms distinguished by different levels of θ_b^i , $i = h, l$, with $\theta_b^h > \theta_b^l$. If θ_b^h was sufficiently large, firms of type h reaching the unconstrained financial status and hence adopting the minimum savings policy described in section 4 would hold substantial debt. This would imply leverage ratios rising with size among type h firms. Since our existing model has a correlation between size and leverage near 0, the fraction type h firms need not be large to deliver an overall positive correlation in this example. Considering the response to a credit shock yielding large reductions in each θ_b^i in this setting, we expect that the real effects following a credit shock would be amplified relative to the results we present below, since large firms would, in that case, have high levels of debt at the onset of the shock.

While relatively simple in its microeconomic elements, our existing model is consistent with various aspects of firm-level behavior observed in the data. For example, our unconditional stationary firm size distribution is right-skewed, firm employment growth is negatively correlated with age (Dunne, Roberts and Samuelson (1989)), and larger and older firms pay more dividends

(Fama and French (2001)).²² Moreover, we remind the reader that the model is calibrated deliberately to reproduce the aspects of firm behavior most crucial in affecting the core misallocation mechanism therein. Specifically, our model matches the mean, variability and autocorrelation of microeconomic investment rates.

Figure 3 is the no-financial frictions counterpart to Figure 2, depicting the average capital among the same cohort of firms in a version of our model where the collateral constraint is removed, so that debt becomes irrelevant to investment. As in the previous figure, the cohort enters the economy with low initial capital. However, in this case, young firms can immediately reach their unconstrained capital targets for the start of the next period. Thus, we see a much larger initial rise in capital between dates 1 and 2 relative to Figure 2. Notice, however, that the elimination of financial frictions does not entirely eliminate life-cycle aspects from our model. Firms still face a real friction that causes bands of inaction in investment. Thus, as a firm transits from one ε to a lower one nearby, it will at times choose not to lower its capital stock, given the forfeit of 5 percent of any capital uninstalled. Likewise, when a firm's relative productivity rises, it is slow to respond fully to that rise given the partial irreversibility in investment. As a result, we see the average capital stock of the cohort gradually continuing to rise from age 2 to age 7. Nonetheless, this rise is quite modest relative to that between age 1 and 2; after taking into account the implications of irreversibility, all but the newest firms operate at a scale appropriate to their productivity. The quantitative impact of the more efficient allocation of production this implies is that steady state output rises by 4 percent relative to our full economy, with measured TFP rising roughly 1 percent.

Returning to our full model economy with both frictions in place, Figure 4 illustrates the pure effects of the irreversibility in cases where it does not interact with the financial friction in our economy. It summarizes the capital choices made by unconstrained firms entering the period with various levels of capital (measured on the x-axis) and debt (measured on the y-axis), conditional on a current productivity draw. The top panel depicts firms entering with the lowest productivity value, the middle panel shows those with the median value, and the bottom panel shows those the highest productivity. The z-axis in each panel reports an indicator variable that takes on a value of 1 for unconstrained firms that invest positively to the upward target capital consistent with their current productivity, a value of 2 for those investing negatively to the relevant downward target,

²²Figures are available on request.

and a value of 5 for those that remain inactive with respect to their capital, setting investment to zero. (The right, rear areas with a zero indicator value are combinations of (k, b) where firms are not unconstrained.)

The region of (k, b) where firms invest to their upward target capital expands into higher current capital levels as one looks from the top panel downward, since rises in current productivity predict higher marginal product of capital schedules next period. To the left of these regions are the areas with zero investment induced by the irreversibility in capital. While the loss associated with uninstalling capital in our economy is only 5 percent, it nonetheless makes some firms reluctant to shed capital. Those with higher current productivities are more so, given the persistence in ε alongside depreciation. As such, the inactivity region expands to higher capital levels as productivity rises, while the region associated with downward investment shrinks, finally disappearing from view by the bottom panel. On balance, of the 14 percent of firms in our stationary distribution that are (permanently) unconstrained, 36 percent adjust to the upward targets consistent with their productivities, 50 percent are inactive with respect to their capital, while 14 percent undertake negative investment.

Figure 5 is analogous to Figure 4. Again conditional on currently productivity, it illustrates the capital decisions taken by firms, this time considering those that are affected by both the real friction in our economy and the financial one. Such firms represent 86 percent of the population in our model's steady state, and are located in regions of the (k, b) space to the right and back where capital is low and/or debt is high. (Foreground areas where the indicator is zero are combinations of (k, b) where firms are unconstrained.)

Because constrained firms' capital choices are largely determined by life-cycle factors stemming from the collateral constraints, they tend to avoid negative investment. Those investing positively to the maximum capital permitted by their ability to borrow are reflected by a value of 3 on the z -axis. These are firms with higher current productivity, comparatively low capital, and comparatively high debt. They make up 32 percent of constrained firms in our model's steady state and are the only firms facing a currently binding borrowing limit. Looking just left and in front of that region, firms with slightly higher capital (or slightly lower debt) adjust to their upward capital targets. This region, reflected by a value of 1 on the z -axis, expands into higher values of capital as ε rises, since the target itself rises. In the stationary distribution, roughly 29 percent of constrained firms are of this type. Finally, looking further left in each panel, we have

firms selecting inaction with respect to investment due to the irreversibility (with a z-axis value of 5), and thereafter those whose capital is sufficiently high relative to their productivity that they disinvest (with a z-value of 2). These categories represent 32 and 7 percent of constrained firms in our model's stationary distribution.

A comparison of Figure 5 to the preceding figure summarizing the capital decision rules of permanently financially unconstrained firms suggests that, for a large fraction of our economy's firms, financial considerations interfere with the optimal firm-level investment responses to information about the future marginal product of capital conveyed current productivity draws. Even in ordinary times, this generates a misallocation of production across firms. One indication of this is the fact that the average capital stock among unconstrained firms in our model's stationary distribution is 2.14, while the average stock of constrained firms is 1.19. As such, the typical old firm in our economy has far more capital than the typical young firm, despite the fact that each group draws from the same productivity distribution.

Old firms in our economy do not carry excess capital; the inefficiency lies in the fact that young, small firms carry too little. This is clarified by Figure 6, which again examines a cohort of (initially) 25,000 firms, this time focusing on the expected discounted return to investment for a cohort as it ages over time. Absent real or financial frictions, firms would always select investment to equate this return to the unit purchase price of investment goods. In that case, the mean investment return across firms in the top panel of our figure would be constant at 1, and the coefficient of variation in this return in the bottom panel would always be 0. By contrast, in our model, the mean expected discounted return to investment is 1.12 for a cohort as it ends its first year of production, and the coefficient of variation in this return is 4.0. Thereafter, over each subsequent year of life, we see ever less dispersion in the return to investment across surviving members of the cohort, and the mean expected return falling towards 1.

After a certain age, surviving members of the cohort have sufficient assets that financial frictions no longer affect their investment decisions. Thus, we see the mean expected investment return for the cohort ultimately reach a value around 1. Even then, there remains some variation in the expected return, and its mean is not precisely 1, given the 5 percent irreversibility of capital in our economy. Recall the capital decision rules of unconstrained firms summarized in Figure 4.

6.2 Business Cycles

We begin to examine business cycle results by first considering the effect each friction in our economy has on its typical business cycle. Table 1 presents some commonly reported business cycle statistics derived from an HP-filtered 5041 period simulation of our model economy under the assumption that aggregate productivity shocks are the only source of aggregate fluctuations, Table 2 presents the corresponding moments when we eliminate financial frictions, and Table 3 is the same economy with neither collateral constraints nor capital specificity. As expected, each friction acts to reduce the average levels of output, capital, and consumption over our simulation. Most notably, average output rises by roughly 4.1 percent when financial frictions are stripped away, then another 2.3 percent when the irreversibility is also eliminated.

Moving to consider second moments, there are some small differences across the three tables. Output volatility rises between our full economy and the counterpart model without limits to borrowing, and it rises again between that model and the one with no frictions. Despite this, as each friction is lifted, the representative household grows more effective in smoothing its consumption. As the contemporaneous correlation between consumption and production is slightly weakened from one table to the next, consumption's standard deviation (raw and relative) falls. Elsewhere, the volatility of hours worked rises steadily, and the hours series is marginally more correlated with output as each friction is eliminated. The same monotone pattern does not follow for investment expenditures, however. There, the relative standard deviation falls from 3.83 percent to 3.77 percent as the financial friction is stripped away, allowing the inertia associated with irreversibility more prominence, while it rises to 4.04 percent when the irreversibility is eliminated.²³

While we have mentioned some minor differences in the business cycle moments across Tables 1 through 3, two points are surely more important. The first is that the business cycle moments drawn from our full model in Table 1 are similar to those of a typical real business cycle model without its complications. Output volatility is roughly 2 percent, consumption is about half as volatile as output, and investment roughly four times as volatile as output. We also see the customary strong positive contemporaneous correlations with output in consumption, investment, and total hours worked. While the usual difficulties of excessive investment volatility and weak hours volatility are a bit more pronounced here relative to some representative firm real business

²³This evidence that firm-level capital irreversibilities dampen changes in aggregate investment is consistent with earlier findings by Bertola and Caballero (1994).

cycle models, these distinctions come from our differing returns to scale in production rather than either friction we mean to study; the same features are present in Table 3 with both removed.

This brings us to our second point. Despite the differences noted above, the second moments across all three tables are quite similar on the whole. Comparing Table 1 to Table 2, in particular, it appears that the typical business cycle in our economy is relatively impervious to some ordinary, ongoing degree of financial frictions. This observation is reinforced by Figure 7, which presents our full model economy's impulse responses following a persistent negative shock to the exogenous component of total factor productivity. As may be seen from the close match between the exogenous and measured TFP series in the top panel, a persistent real shock has inconsequential implications for the endogenous component of aggregate productivity. Thus, examining output, consumption, employment and investment, we see impulse responses closely resembling those of a counterpart economy without real or financial frictions, such as that summarized in Table 3. Just as in a frictionless business cycle model, there are immediate declines in all four series. Further, aside from the customary U-shaped consumption response, we see the largest declines at the impact of the shock, with each series thereafter monotonically reverting toward its long-run level. Note also that the largest percent drop in hours worked is half that of GDP, while investment's greatest drop is roughly 3.5 times that in GDP.

When driven solely by exogenous changes to total factor productivity, we have seen that our model economy closely resembles a standard frictionless business cycle economy. This, in turn, implies a failure to capture some key aspects of the latest U.S. recession, dated by the NBER as 2007Q4 - 2009Q2. Figure 8 reports the recent movements in GDP, consumption, investment, employment hours and measured TFP, plotting each series' percent deviations relative to their 2007Q4 levels.²⁴ While the figure includes data over 2009Q3 - 2010Q4, we defer discussion of the recovery period for now and focus here on the economic downturn within the NBER recession dates.

Figure 8 reveals that the initial response in GDP was negligible, while real personal consumption expenditure actually rose by roughly 1 percent and stayed high until 2008Q4. Moreover, the immediate declines in investment were modest relative to what came later. While total private

²⁴Our hours series extends through 2010Q4 the Cociuba et. al (2009) series representing total civilian and military hours worked per noninstitutional population aged 16 to 64. The results presented here change in no notable way if we omit the military component from this series. Our measured TFP series is a direct Solow Residual calculation based on the productive shares of labor and capital to which our model is calibrated.

investment fell immediately, this was initially driven by housing. Non-residential investment did not begin to fall until 2008Q3, at which point it began to drop off sharply relative to the more gradual declines in GDP and consumption. Measured TFP falls until the first quarter of 2009, where it is 1.6 percent below its initial level. Elsewhere, the greatest declines in all other series in this figure do not occur until 2009Q2 or later. As of 2009Q2, GDP was roughly 4 percent below its initial level, total hours had declined 1.7 times as much as GDP, while the investment drop was five to six times that in GDP.²⁵

On balance, the recent U.S. economic downturn presents several challenges for any equilibrium business cycle model driven by shocks to technology. First, it shows an initial rise in consumption where the model would predict a clear decline. Next, the ultimate losses in employment and investment by the end of the recession are far greater than in the model. The magnitudes of change in these series relative to output are also several times larger than predicted.

6.3 A Credit Crisis

Clearly, the challenges regarding the latest U.S. recession apply not only to a standard business cycle model but also to our model, so long as its aggregate fluctuations arise solely from changes in exogenous productivity. However, our focus is on understanding what happens when the extent of financial frictions suddenly and unexpectedly grows more severe than is normal. Given its assumption of a representative firm, the standard business cycle model is unaffected by such events. We now explore their effect in our model where firms have different access to, and need for, credit.

Before continuing, we discuss the evidence for an exogenous shock to the availability of credit that, in our model, corresponds to an unanticipated decline in θ_b . As is now well understood, it is hard to find conclusive evidence that the U.S. has experienced an exogenous reduction in lending to businesses. In an early exploration of this issue, Chari et al. (2008) argued there was little evidence that the financial crisis had affected lending to nonfinancial firms. Examining the Flow of Funds, they found that the stock of commercial and industrial loans across regulated banks had actually risen as of the third quarter of 2008.

Chari et al. (2008) also argue that, in the aggregate, business fixed investment is less than

²⁵Ohanian (2010) presents evidence that the magnitudes of these declines in investment and hours relative to GDP are dramatic in comparison with those seen in previous postwar U.S. recessions, as well as other G7 countries' 2007-9 recessions.

firms' revenues net of labor costs; that is, the mean firm can self-finance investment. It is worth noting that our calibrated model is consistent with this observation in the aggregate data. Nonetheless, borrowing limits bind for some smaller firms, and this leads to insufficient capital in those firms, reducing both aggregate TFP and GDP.

Ivashina and Scharfstein (2010) re-examine the issue of lending over the financial crisis. They study Reuters DealScan data on syndicated loans, which captures new lending to large corporations. While these loans originate with banks, the pool of lenders, which includes non-bank financial institutions, is larger. Moreover, the data on syndicated lending covers new loans, as opposed to the stock of outstanding debt that is reported in the Flow of Funds. Ivashina and Scharfstein find strong evidence of a reduction in lending. Between 2007 and 2008, total syndicated lending fell 54 percent, while loans used to fund investment in equipment and structures fell 48 percent.

Koepke and Thomson (2011) examine loans from FDIC-insured commercial banks and savings institutions. They find that total lending declined between 2008 and 2009 and since then has not recovered to its pre-recession level. Importantly, commercial and industrial loans, which are an important source of borrowing for small and medium-sized businesses, fell 18.7 percent between 2008Q4 and 2009Q4. These loans continued to fall through 2010Q2, when they were 19.1 percent below their level at the start of the recession, in 2007Q4.

While these papers provide evidence that there has been a reduction in lending to nonfinancial firms, they do not establish whether this represents an exogenous reduction in credit or instead an equilibrium response to reductions in business fixed investment. However, both Almeida et al. (2009) and Duchin et al. (2010) provide support for a credit shock interpretation. Controlling for firm characteristics using a matching estimator, Almeida et al. study the investment behavior of firms that, given their existing loan maturity structures, needed to refinance a substantial fraction of their long-term debt over the year following August 2007, the onset of the financial crisis. They find that investment spending among such firms fell by one-third. By contrast, other firms with similar characteristics, but without a large refinancing in the period following the start of the financial crisis, showed no investment reduction. Since the fraction of long-term debt maturing after August 2007 is likely exogenous to the financial crisis, this suggests that there was an exogenous reduction in the supply of credit. Duchin et al. (2010) compare the behavior of firms that were carrying more cash prior to the onset of the crisis with that of firms carrying

less cash. Using a difference-in-difference approach, they find that firms with less liquid assets before the financial crisis exhibited a larger reduction in investment.

Given this evidence that the recent recession involved an exogenous reduction in lending, we now examine a credit shock in our model economy. Figure 9 depicts our economy's response to a financial crisis, absent any technology shock. More specifically, it is the response to a 55 percentage point drop in the value of firms' collateral, as generated by a reduction in θ_b , which we will see below implies an eventual 23 percent reduction in debt. This reduction is not implausibly large relative to the actual declines in various measures of lending since the onset of the 2007 credit crisis, as discussed above. While it is slightly larger than the reduction in the commercial and industrial loans of regulated banks, it is substantially smaller than the fall in syndicated loans used to finance investment.

From the first date of the credit shock onwards, we assume that households and firms expect an eventual return to normal financial conditions. Each period, they place 40 percent probability on a full financial recovery in the subsequent period. Thus, when the shock occurs in period 1, they expect it will persist for 2.5 years.²⁶ However, we begin by focusing on the downturn following a credit shock. Thus we omit the actual recovery from this figure.

Although the distribution of capital is predetermined when the financial shock hits in year 1, the top left panel of Figure 9 reveals that aggregate production immediately falls by about 1.5 percent (relative to its simulated mean in normal financial times). This is, of course, a direct consequence of the 2.4 percent fall in the labor input (top right panel), which is, in turn, a reaction to the reduced expected return to investment (bottom right panel). With the sudden reduction in credit, there is a drop in the fraction of firms that are financially unconstrained and a sharp rise in the fraction of firms facing currently binding borrowing limits. Underlying these changes, young firms are now far more hindered in their investment activities relative to the pre-shock economy, and thus will take considerably longer to outgrow financial frictions and begin producing at a scale consistent with their productivities. If these financial conditions persisted, the resulting stationary distribution would have 51 percent of firms constrained in their current upward capital adjustments and 1 percent of firms forced to undertake some negative investment to repay outstanding debt. In ordinary financial times, these percentages are 27 and 0, respectively. In the meantime, firms that are unconstrained after the financial shock must increase their precautionary

²⁶This represents an average of 2 and 3 years it took GDP and consumption to recover, respectively, took to their 2007Q4 level.

savings to remain so. As a result, aggregate savings among firms that are not borrowers rises roughly 30 percent at the date of the shock; thereafter, while the stock of debt declines, this series slowly falls toward its initial level, reaching about 10 percent above normal by date 4. (Figure available on request.)

Unlike the response that would follow a negative productivity shock, consumption does not immediately fall when the financial shock hits our economy. Anticipating a more distorted distribution of production over coming years, and thus unusually low endogenous total factor productivity (in the lower right panel), the representative household in our economy expects a lowered return to saving. This leads to a 0.5 percent rise in consumption at the impact of the shock, and also a rise in leisure. This effect of reduced future TFP is compounded by the fact that the initial aggregate capital stock is roughly 9 percent above that consistent with the tighter borrowing conditions, which further encourages consumption and leisure.²⁷ The fall in investment (at lower left) does not support consumption for long, however; consumption falls to its pre-shock level by year 3, then steadily declines for roughly 8 more years before it levels off. Elsewhere, labor falls at the impact of the shock as described above. Thereafter, given the severe misallocation of capital at the start of date 2, alongside reductions in the total capital stock, the marginal product of labor drops, yielding further large reductions in employment. By year 3, employment is 3.9 percent below its pre-shock level, and it does not rise back to the level consistent with the new financial setting until around period 15. This long adjustment period is a reflection of the time that it takes for the capital distribution to settle, as may be inferred from the measured TFP response in the lower right panel.

Figure 10 illustrates how the unanticipated tightening of credit distorts the allocation of production in our economy, giving rise to the initial aggregate changes in the previous figure. The top panel plots the distribution of firms over capital and productivity that is in place in date 1 when the shock occurs. The lower panel shows the same distribution at the start of date 2, once the tightened borrowing conditions have begun affecting firm-level investment activities. Comparing these two panels, we see a noticeable rise in dispersion within just one period of the shock. The mass of firms with capital stocks between 1.5 and 3 in the top panel is considerably

²⁷Fernald and Matoba (2009) argue that utilization-adjusted total factor productivity rose over the recent U.S. recession. The counterpart to their series in our model, the exogenous component of TFP, is held constant over our credit shock exercise. A positive shock to this series could be introduced, provided the credit shock was adjusted so that measured TFP still fell as in the data.

reduced, with much of that mass pushed into lower regions of capital by date 2. In other words, the shock creates fewer medium-sized firms and more small firms. At the same time, we see a few firms, the very largest, growing larger.²⁸

Given a heightened misallocation of production coming in date 2, the largest firms that are otherwise unaffected by credit concerns respond to the reduced real interest rate by expanding in size. Such firms adopt the efficient capital levels dictated by their (S,s) investment policies, given the productivities and real interest rate they face. The increased inefficiency that reduces TFP arises because small firms, now facing more severe collateral requirements, see the gap between their expected discounted return to capital and the real rate widen.

Figure 11 shows that these problems do not abate in nearby dates, in that dispersion continues to rise. The top panel shows that the mean of the (ex-post) marginal product of capital moves further above its pre-crisis value (0.134) during the first 8 periods over which the overall stock of capital is being reduced. The lower panel plots the coefficient of variation in firms' marginal products. This measure of dispersion rises over the first several dates of the financial crisis, and then begins to return toward its pre-crisis level (0.537). That initial level is never fully recovered in this figure, however, because financial conditions do not improve.

Critically, the increase in the mean marginal product of capital coincides with a fall in the ex-post real interest rate. The marginal product of capital across large, unconstrained firms falls, with the contrasting rise in the economywide average driven entirely by tighter borrowing limits for other firms. Capital falls in these firms over time, so their marginal product of capital rises. With the increased dispersion in the returns to capital, the coefficient of variation is seen to rise. These results in our model following a credit shock are reminiscent of empirical results in Eisefeldt and Rampini (2006), who show that the benefits of capital reallocation rise in recessions while the level of reallocation falls. In our model, both forces operate.

On balance, we take the following observations from Figures 9 - 11. A tightening of collateral constraints alone, a purely financial shock, drives large and persistent real effects in our model economy. It does so because it increases the dispersion in firm-level capital further away from

²⁸Using BEDS data from the Bureau of Labor Statistics, Gilchrist (2011) compares average employment growth for firms with more than 500 employees to average employment growth rate of firms with less than 500 employees and finds that, after 2007, small firms reduced employment more rapidly than large firms. More generally, our increased dispersion in production is consistent with evidence from Bloom et al (2009) that various measures of firm-level dispersion rise during recessions.

the efficient level consistent with productivity dispersion (and capital specificity) and it allocates insufficient capital to an expanded group of small firms, putting downward pressure on the interest rate. In the example we have shown here, the misallocation of production arising from tight financial conditions is compounded by the reductions in aggregate capital, productivity, and labor that it causes. As a result, there are protracted adjustments in aggregate quantities lasting a decade or more, and GDP is ultimately reduced by 3.6 percent, while aggregate consumption is reduced by 1.3 percent.

We next consider what implications the prolonged financial crisis from above can have if its onset is followed by a 1 standard deviation negative technology shock. As seen in the lower right panel of Figure 12, the exogenous component of TFP falls one year after the financial shock hits, and thereafter gradually reverts to its mean. Were credit markets functioning as normal when this TFP shock appeared, output would fall 3.8 percent, labor would fall 2 percent, and the half-life of the output response would be roughly 5 years, as seen in Figure 7. While the overall declines in quantities here are larger and protracted, they are no more so than would be expected once we account for the effects of the credit shock in Figure 8.

Until now, we have considered the implications of a *persistent* financial crisis, in that borrowing conditions do not recover throughout the exercises depicted in Figures 9 - 12. We next consider the recovery. We assume the same financial shock studied in Figure 9 persists for 4 periods; thereafter, beginning in date 5, we allow a complete recovery of financial conditions, returning the value of collateral to normal. Over the first four periods, agents have the same expectations regarding financial recovery as we assumed above. Thereafter, beginning in date 5, they understand that borrowing conditions have permanently recovered. We omit the negative TFP shock from this exercise for expositional simplicity, as we have seen above that its implications do not add unexpected or noteworthy features to the impulse responses.

In Table 4, we compare the peak-to-trough behavior of our model with the 2007 US recession. Driven solely by a credit shock, our model generates the same decline in GDP as in the data. Furthermore, the 23 percent fall in investment is almost as large as the 27 percent in the data. Elsewhere, the model explains roughly half of the actual reduction in employment and 43 percent of the fall in consumption. As shown below, the overall fall in consumption in our model is larger, but it occurs after the trough in GDP.

Consider these results in comparison with the responses to a technology shock, whether in

our model or in a standard representative firm model. On its own, a technology shock cannot reconcile the empirical declines in GDP, measured TFP and investment. A one-standard deviation exogenous shock to TFP generates almost as large a change in GDP, -3.85 percent, and about half the fall in employment. However the required change in TFP, -2.67 percent, is far larger than the largest change measured in the data, which was -1.65 percent in 2009Q1.²⁹ Furthermore, the drop in investment is only 14 percent. In contrast, following a credit shock, our model reproduces the changes in output and investment observed in the data, without implying a counterfactually large decline in measured TFP. Furthermore, this exercise implies a reduction in debt of 22.8 percent, which is comparable to the 19 percent fall in commercial and industrial loans by regulated banks through the end of 2009, and far less than the 48 drop in syndicated lending between 2007 and 2008. In contrast, the TFP-shock driven recession generates less than a 4 percent fall in debt.

Figure 13 shows the full credit shock driven recession in our model, including the recovery. Three aspects of the responses there are worthy of note. First, so long as GDP or consumption is adopted as our measure, the effects of a large credit shock are not rapidly reversed. Although loan markets are operating normally in year 5, GDP remains 3.1 percent below trend in that date. Moreover, GDP recovers twice as slowly as in the data, with a half-life of 1.5 years. Consumption takes longer to return to its average with a half life of 4 periods. The slow recovery of output and consumption after real and financial frictions have been restored to their ordinary levels arises in part from the fact that the distribution of capital does not immediately settle back to its pre-shock state. As a result, aggregate productivity remains below normal until year 8, as seen in the third panel of the figure. This compounds the fact that the aggregate capital stock is more than 5 percent below its usual level by the start of the recovery.

Second, consumption does not begin to recover in date 5. Given a high demand for investment goods, and output's failure to rebound rapidly, households actually allow their consumption to fall for two more periods and thereafter raise it only very slowly. Third, during this episode, it is the labor input that drives the recovery. Anticipating the subsequent rise in endogenous productivity, and thus a raised return to savings, households abruptly raise their hours worked from 3.6 percent to only 1.1 percent below normal within date 5. In the next date, the allocation of capital across firms has begun to move back toward the long-run distribution, and the resulting improvement in productivity directly encourages a further large rise in the labor input. At this

²⁹Because TFP led GDP in the 2007 recession, the change reported in Table 4 reflecting the 2007Q4 to 2009Q2 drop in TFP was less, at 0.60 percent.

point, it overshoots its average level by just over 1 percent. Thereafter, it remains high for many periods while the capital stock is being rebuilt.

One and a half years after the start of the recovery in 2009Q2, there is little net change in total hours worked in the data. In the model, the recovery is stronger. In part because it fell by less, employment regains two-thirds of its pre-crisis level within a year. In 2010Q4, private investment is still more than 20 percent below its pre-recessionary level. Here again, the model predicts a more rapid recovery, with investment less than ten percent below its steady state level after one year. Nonetheless, these responses are more gradual than would be seen following an equivalent shock to TFP.³⁰

The rapid growth in investment and employment, relative to the data, is in part a result our assumption that lending conditions, as represented by the value of θ_b , are completely restored in one period. We have not examined gradual changes in θ_b in the exercises presented above largely because data restricting θ_b period-by-period is unavailable. Moreover, our assumption of a constant probability of complete reversal in θ_b has allowed us to solve for impulse responses under uncertainty. We have explored the economy's response following a gradual recovery in θ_b under the assumption of perfect foresight.³¹ The recovery in investment and employment is somewhat slower there, but that in output and consumption is also slowed.

7 Concluding remarks

We have developed a dynamic stochastic general equilibrium model with persistent, firm-level shocks to total factor productivity, costly investment reversibility and collateralized borrowing constraints. We have calibrated the model to be consistent with U.S. evidence on establishment-level investment dynamics, as well as overall measures of borrowing by non-financial firms. Our resulting economy is characterized by a nontrivial distribution of firms over productivity, debt and capital that shapes aggregate output and total factor productivity.

³⁰Consider an exogenous one standard deviation reduction in aggregate total factor productivity. Absent any credit shock, this TFP shock would result in a similar fall in GDP to that seen in figure 10. If this shock was eliminated at date 5, with the exogenous component of TFP immediately reset to its long-run level, then GDP would complete 94 percent of its recovery instantaneously. By contrast, the corresponding instantaneous rise in GDP in figure 10, following the elimination of the credit shock, is only 27 percent. Furthermore, the rapid increases in employment and investment seen in figure 10 would be instantaneous with the elimination of a technology shock.

³¹The results of these exercises are available upon request.

Firms respond endogenously to the frictions they face and, over time, build sufficient precautionary savings so as to ensure that borrowing limits will not affect their investment. Only a small subset of the firms in our economy have investment activities curtailed by their current ability to borrow. Nonetheless, borrowing and investment decisions are inter-related for most firms, and the risk associated with future collateral constraints affects their current decisions with respect to both.

Our model predicts that, in a modern developed economy such as the U.S., the typical business cycle may be relatively unaffected by financial frictions. Nonetheless, absent any real shock to the economy, we find that a credit crisis on its own can generate a recession that is not only large, but persistent. Because tight borrowing conditions deliver a long-lived disruption to the distribution of capital, and thus to endogenous aggregate productivity, their aftermath is a gradual recovery in output.

The recession generated in our model by a credit shock is qualitatively different from that following a negative shock to aggregate productivity, and it more closely resembles the 2007 U.S. recession in several respects. The decline in GDP is gradual. Consumption initially rises. The response in investment is unusually severe relative to the fall in GDP. The decline in measured TFP that accompanies these movements in output, consumption and investment is similar to that in the data.

While capturing several aspects of the recent U.S. recession, the credit shock we have considered here does not deliver the unusually slow recovery in investment and employment over the 18 months of data since the trough of the recession in 2009Q2. It is possible that tight credit has affected not only business fixed investment, but also firms' ability to finance working capital used to pay wages, and that lending conditions did not fully recover by the end of 2010 (Koepeke and Thomson (2011)). However, no such explanation is likely to reconcile the observed changes in investment and employment with the growth in consumption and GDP. A more likely suggestion proposed by Ohanian (2010) is that time-varying distortions in the labor market have been important in shaping employment over this recession. Given the complexity of our current model, micro-founding such additional frictions is beyond the scope of this paper.

References

- [1] Albuquerque, R. and H. A. Hopenhayn (2004) “Optimal Lending Contracts and Firm Dynamics,” *Review of Economic Studies*, 71(2), 285-315.
- [2] Almeida, H., M. Campello, B. Laranjeira and S. Weisbenner (2009) “Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis,” NBER working paper 14990.
- [3] Abel, A. B. and J. C. Eberly (1996) “Optimal Investment with Costly Reversibility,” *Review of Economic Studies* 63(4), 581-93.
- [4] Arellano, C., Y. Bai and P. Kehoe (2010) “Financial Markets and Fluctuations in Uncertainty,” *Federal Reserve Bank of Minneapolis Staff Report*.
- [5] Bernanke, B. and M. Gertler (1989) “Agency Costs, Net Worth, and Business Fluctuations,” *American Economic Review* 79, 14-31.
- [6] Bernanke, B., M. Gertler and S. Gilchrist (1999) “The financial accelerator in a quantitative business cycle framework,” Chapter 21, pages 1341-93, in J. B. Taylor and M. Woodford (eds.) *Handbook of Macroeconomics*, Volume 1, Part 3, Elsevier.
- [7] Bertola, G. and R. J. Caballero (1994) “Irreversibility and Aggregate Investment,” *Review of Economic Studies* 61, 223-246.
- [8] Bloom, N., M. Floetotto and N. Jaimovich (2009) “Really Uncertain Business Cycles” Stanford University working paper.
- [9] Buera, F. and Shin, Y. (2007) “Financial Frictions and the Persistence of History: A Quantitative Exploration,” Northwestern University working paper.
- [10] Caballero, R. J. (1999) “Aggregate Investment,” Chapter 12 in M. Woodford and J. Taylor (eds.) *Handbook of Macroeconomics*, vol. IB. Amsterdam: North Holland.
- [11] Caballero, R. J and E. M. R. A. Engel (1999) “Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S, s) Approach,” *Econometrica* 67, 783-826.
- [12] Caggese, A. (2007) “Financing Constraints, Irreversibility, and Investment Dynamics” *Journal of Monetary Economics* 54, 2102–30.

- [13] Carlstrom, C. T. and T. S. Fuerst (1997) “Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis,” *American Economic Review* 87, 893-910.
- [14] Chari, V. V., L. Christiano and P. J. Kehoe (2008) “Facts and Myths about the Financial Crisis of 2008,” Federal Reserve Bank of Minneapolis Working Paper 666.
- [15] Christiano, L., R. Motto and M. Rostagno (2010) “Financial Factors in Economic Fluctuations” European Central Bank Working Paper no. 1192.
- [16] Clementi, G. L. and H. A. Hopenhayn (2006) “A Theory of Financing Constraints and Firm Dynamics” *Quarterly Journal of Economics* 121(1), 229-265.
- [17] Comin, D. and T. Philippon (2005) “The Rise in Firm-Level Volatility: Causes and Consequences,” NBER Working Paper No. 11388.
- [18] Cooley, T. F., R. Marimon and V. Quadrini (2004) “Aggregate Consequences of Limited Contract Enforceability” *Journal of Political Economy*, 2004, vol. 112, no. 4, 817-47.
- [19] 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.
- [20] Cooper, R. W. and J. C. Haltiwanger (2006) “On the Nature of Capital Adjustment Costs,” *Review of Economic Studies* 73, 611-633.
- [21] Cordoba, J. C. and M. Ripoll (2004) “Credit Cycles Redux” *International Economic Review* 45, 1011-46.
- [22] Davis, S. J. and J.C. Haltiwanger (1992) “Gross Job Creation, Gross Job Destruction, and Employment Reallocation,” *Quarterly Journal of Economics* 107(3), 19-63.
- [23] Duchin, R., O. Ozbas and B. A. Sensoy (2010) “Costly external finance, corporate investment, and the subprime mortgage credit crisis” *Journal of Financial Economics* 97, pages 418 -435.
- [24] Fernald, J. and K. Matoba (2009) “Growth Accounting, Potential Output, and the Current Recession” *FRBSF Economic Letter* 2009-26.
- [25] Gertler, M. and N. Kiyotaki (2010) “Financial Intermediation and Credit Policy in Business Cycle Analysis” prepared for *The Handbook of Monetary Economics*.

- [26] Gilchrist, S. (2011) “Small versus Large Firm Employment Growth over the Business Cycle” Boston University working paper.
- [27] Gilchrist, S., J. W. Sim and E. Zakrajšek (2011) “Uncertainty, Financial Frictions, and Investment Dynamics” Boston University working paper.
- [28] Gomes, J. F. and L. Schmid (2009) “Equilibrium Credit Spreads and the Macroeconomy” Wharton School working paper.
- [29] Gomme, P., Ravikumar, B. and P. Rupert (2008), “The Return to Capital and the Business Cycle,” Concordia University Working Paper 08-002.
- [30] Hansen, G. D. (1985) “Indivisible Labor and the Business Cycle,” *Journal of Monetary Economics* 16, 309-327.
- [31] Ivashina, V. and D. S. Scharfstein. “Bank Lending During the Financial Crisis of 2008.” *Journal of Financial Economics* 97(September 2010): 319-338.
- [32] Jermann, U. and V. Quadrini (2010) “Macroeconomic Effects of Financial Shocks” forthcoming *American Economic Review* .
- [33] Jermann, U. and V. Quadrini (2009) “Financial Innovations and Macroeconomic Volatility” Wharton School working paper.
- [34] Johnson, S. A. (1989) *Spline approximation in discrete dynamic programming with application to stochastic multi-reservoir systems* Unpublished dissertation (Cornell, Ithaca, NY).
- [35] Khan, A. and J. K. Thomas (2003) “Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?” *Journal of Monetary Economics* 50, 331-360.
- [36] Khan, A. and J. K. Thomas (2008) “Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics” *Econometrica* 76, 395-436.
- [37] Kiyotaki, N. and J. Moore (1997) “Credit Cycles” *Journal of Political Economy* 105, 211-248.
- [38] Kocherlakota, N.R. (2000) “Creating Business Cycles Through Credit Constraints” *Federal Reserve Bank of Minneapolis Quarterly Review* 24, 2-10.
- [39] Koepke, M. and J. B. Thomson (2011) “Bank Lending” Economic Trends Federal Reserve Bank of Cleveland.

- [40] Krusell, P. and A. A. Smith Jr. (1997) “Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns,” *Macroeconomic Dynamics* 1, 387-422.
- [41] Modigliani, F. and M. Miller (1958). “The Cost of Capital, Corporation Finance, and the Theory of Investment,” *American Economic Review* 48, 261–297.
- [42] Ohanian, L. E. (2010) “The Economic Crisis from a Neoclassical Perspective,” *Journal of Economic Perspectives* 24 (4), 45-66.
- [43] Prescott, E. C. (1986) “Theory Ahead of Business Cycle Measurement,” *Federal Reserve Bank of Minneapolis Quarterly Review* 10, 9-22.
- [44] Restuccia D. and R. Rogerson (2008) “Policy distortions and aggregate productivity with heterogeneous establishments” *Review of Economic Dynamics* 11, 707-720.
- [45] Rogerson, R. (1988) “Indivisible Labor, Lotteries and Equilibrium,” *Journal of Monetary Economics* 21, 3-16.
- [46] Veracierto, M. L. (2002) “Firm-Level Irreversible Investment and Equilibrium Business Cycles,” *American Economic Review* 92, 181-197.

FIGURE 1. Steady state distribution in the full model

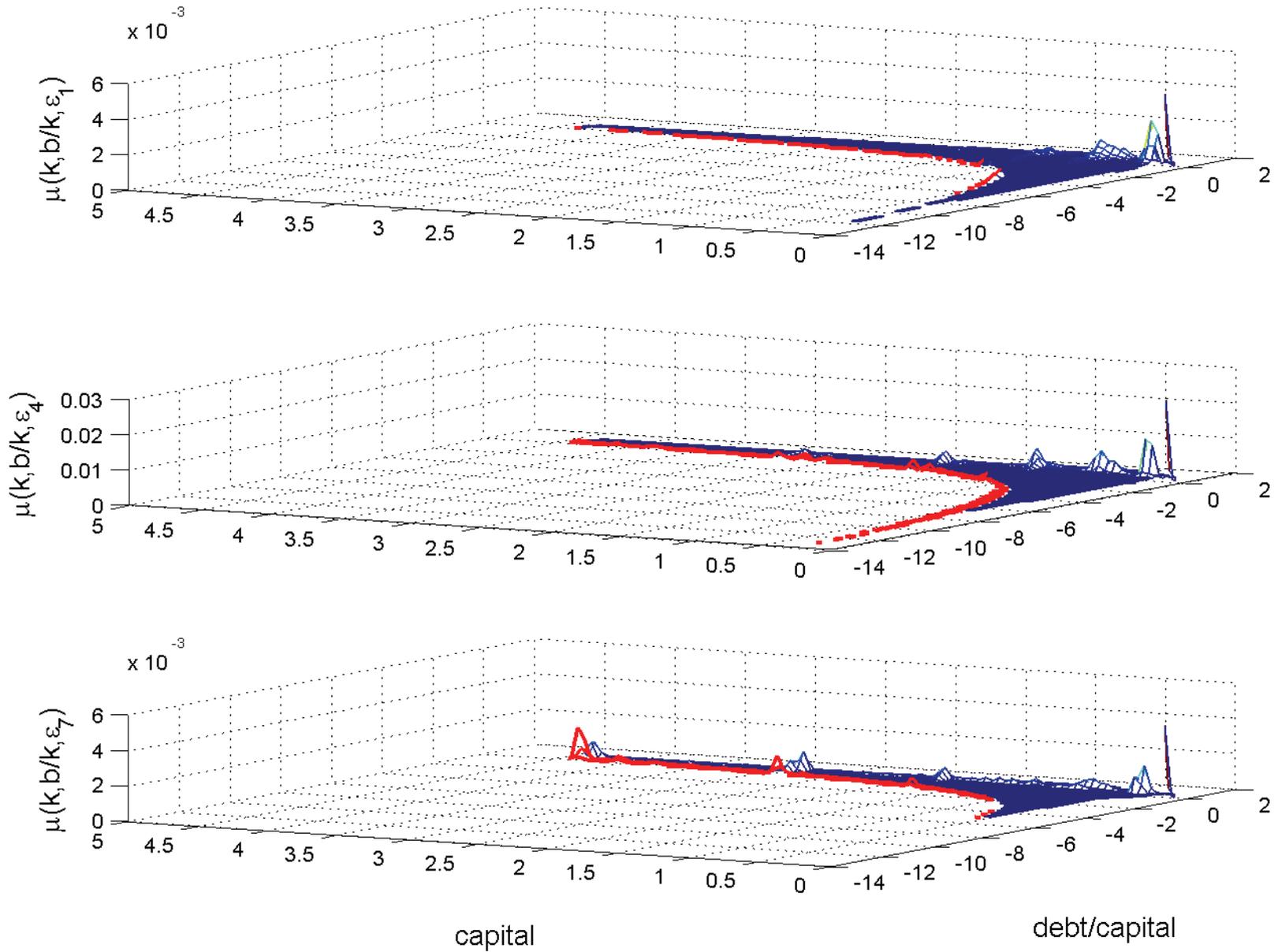


FIGURE 2. Cohort in steady state

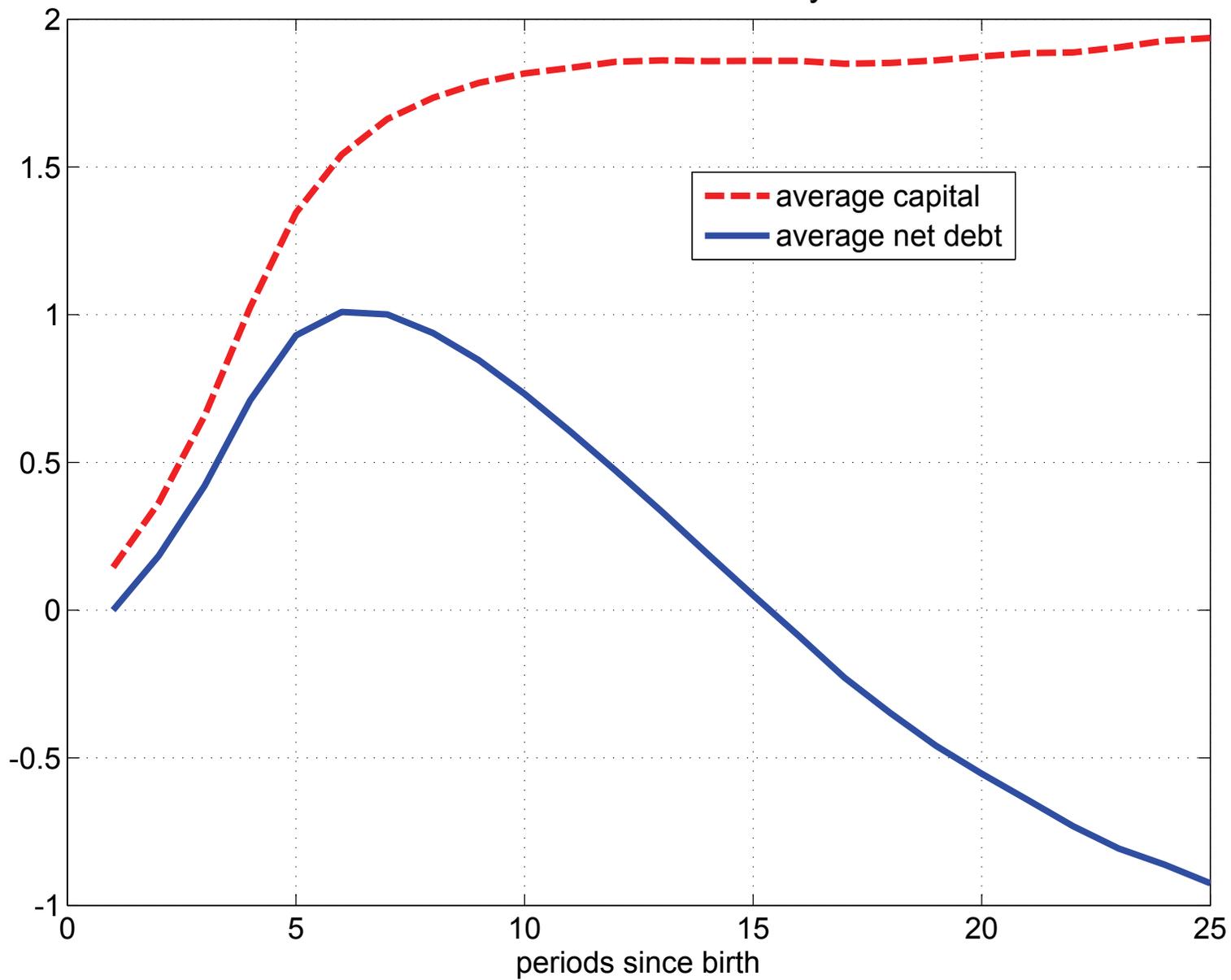


FIGURE 3. Cohort in no-financial-frictions steady state

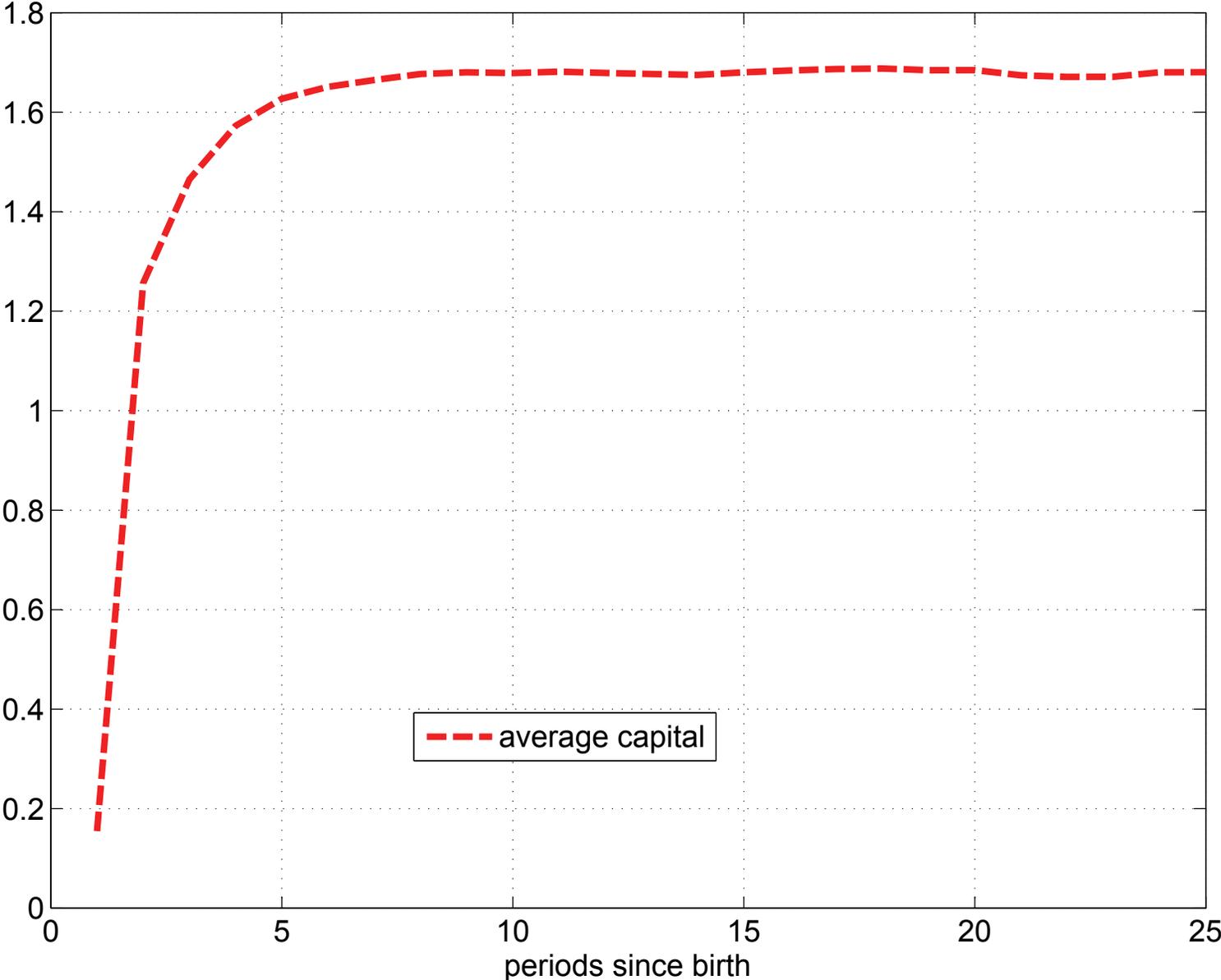


FIGURE 4. Capital choices among unconstrained firms

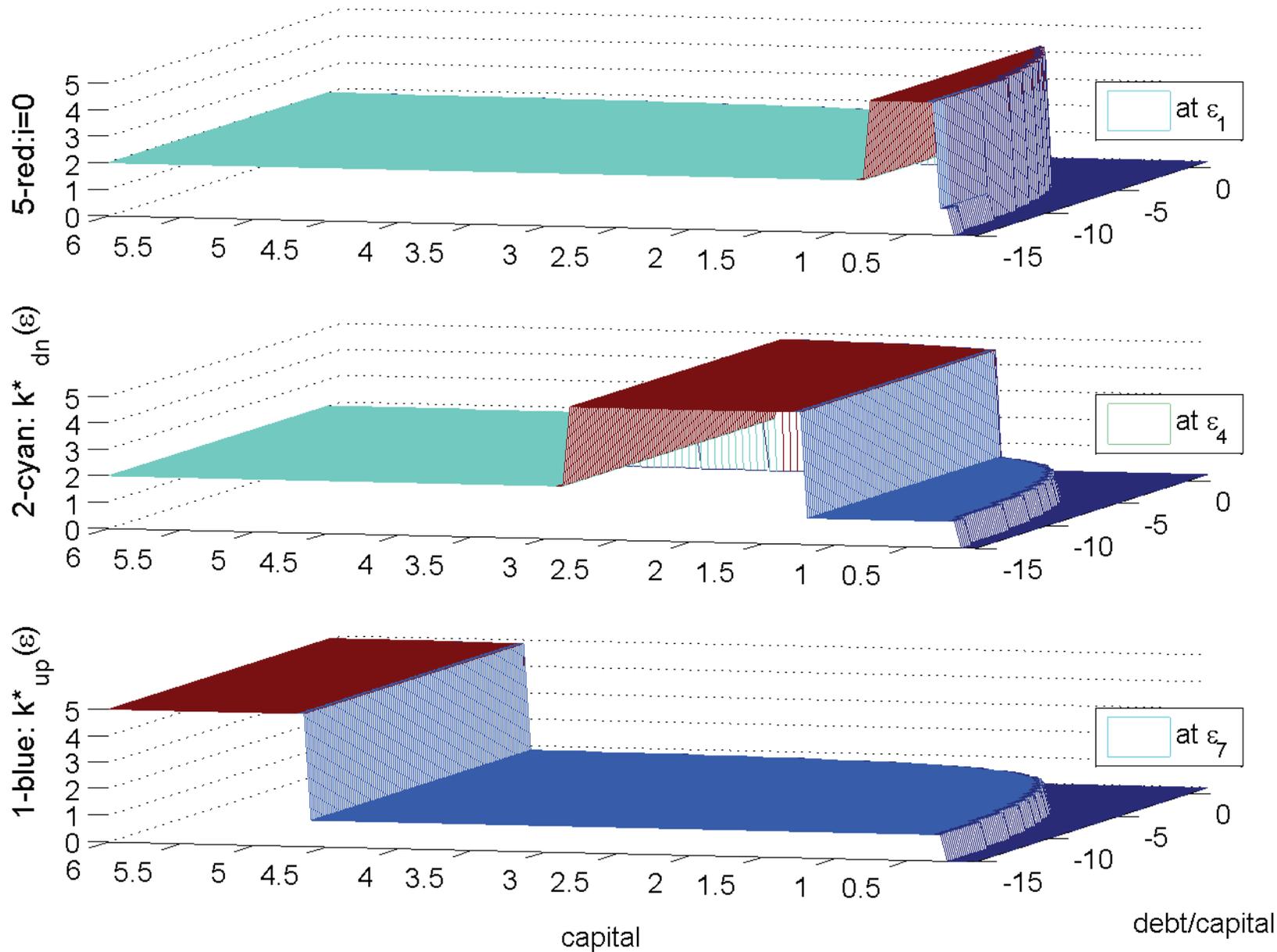


FIGURE 5. Capital choices among constrained firms

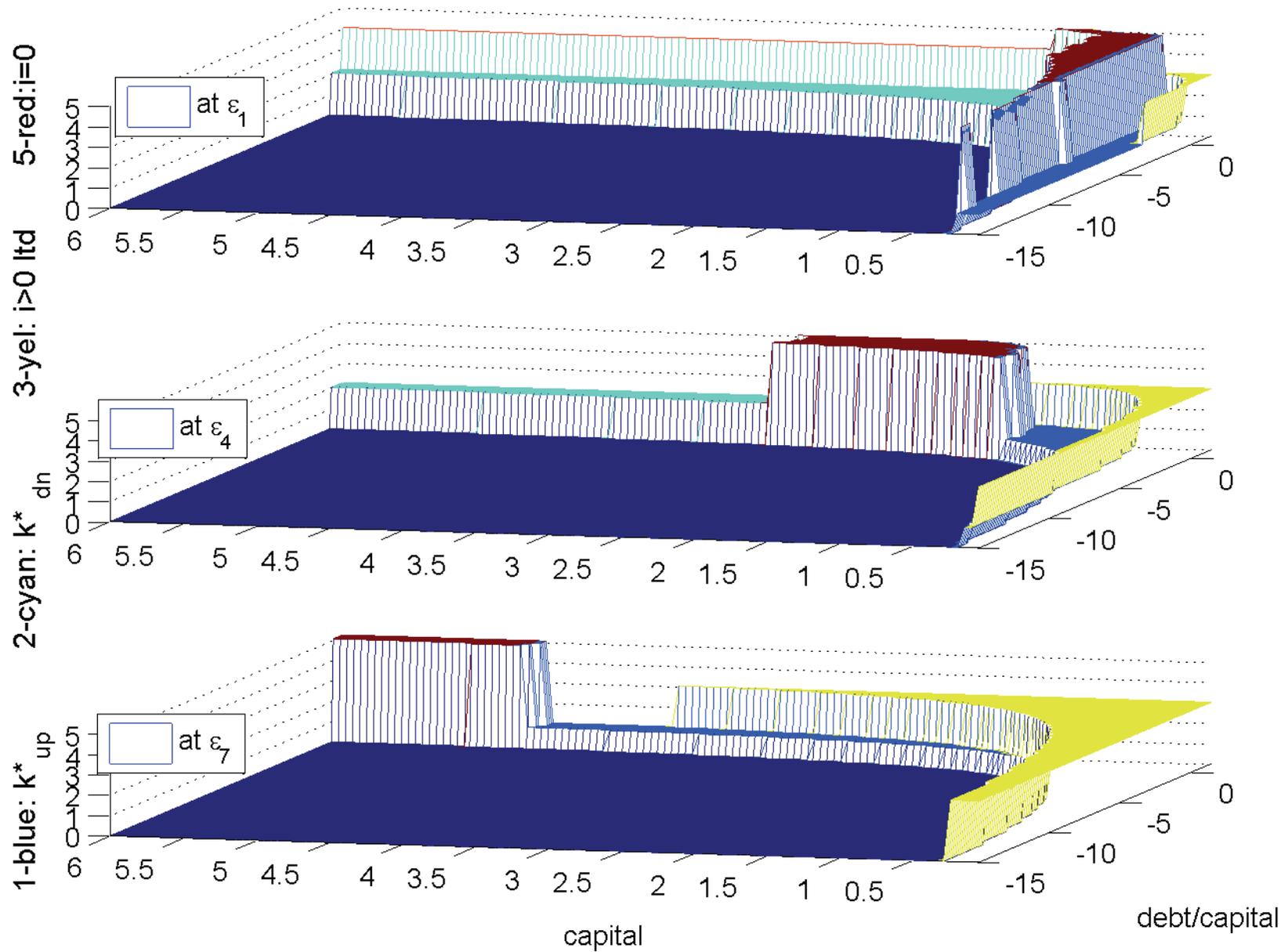


FIGURE 6. Cohort Expected Discounted Return to Capital in Steady State

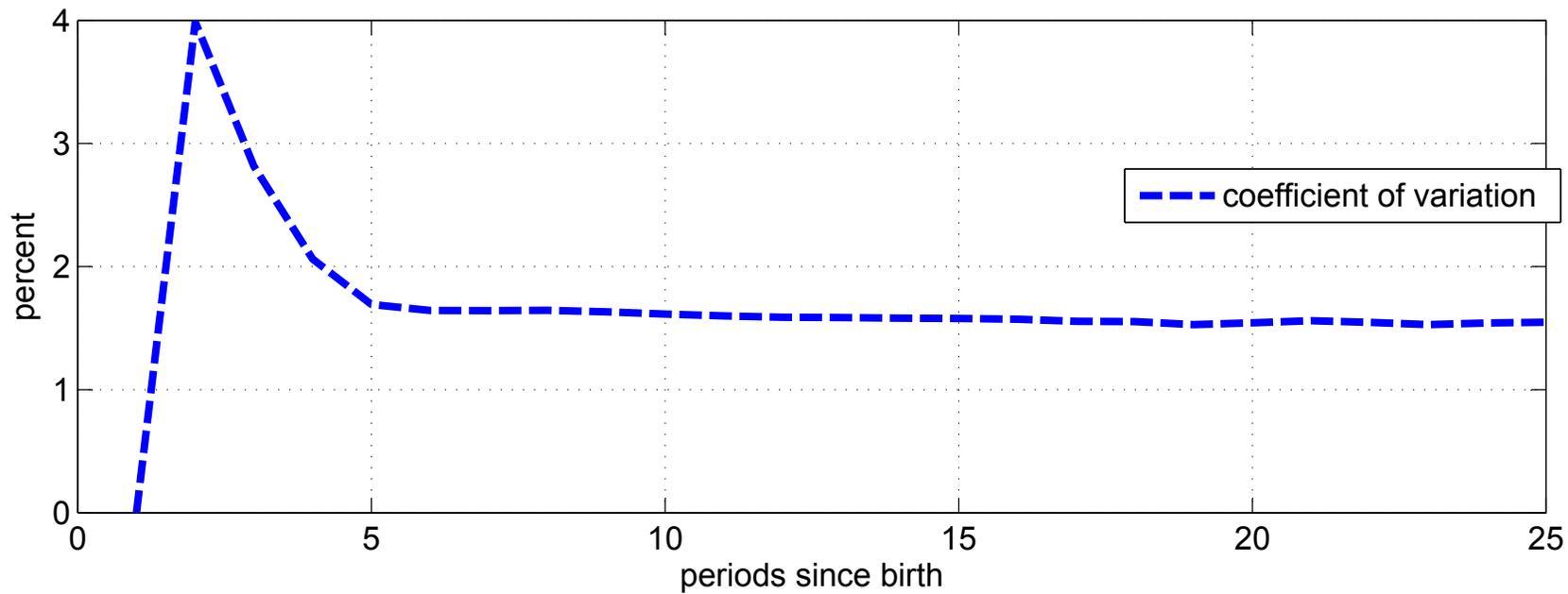
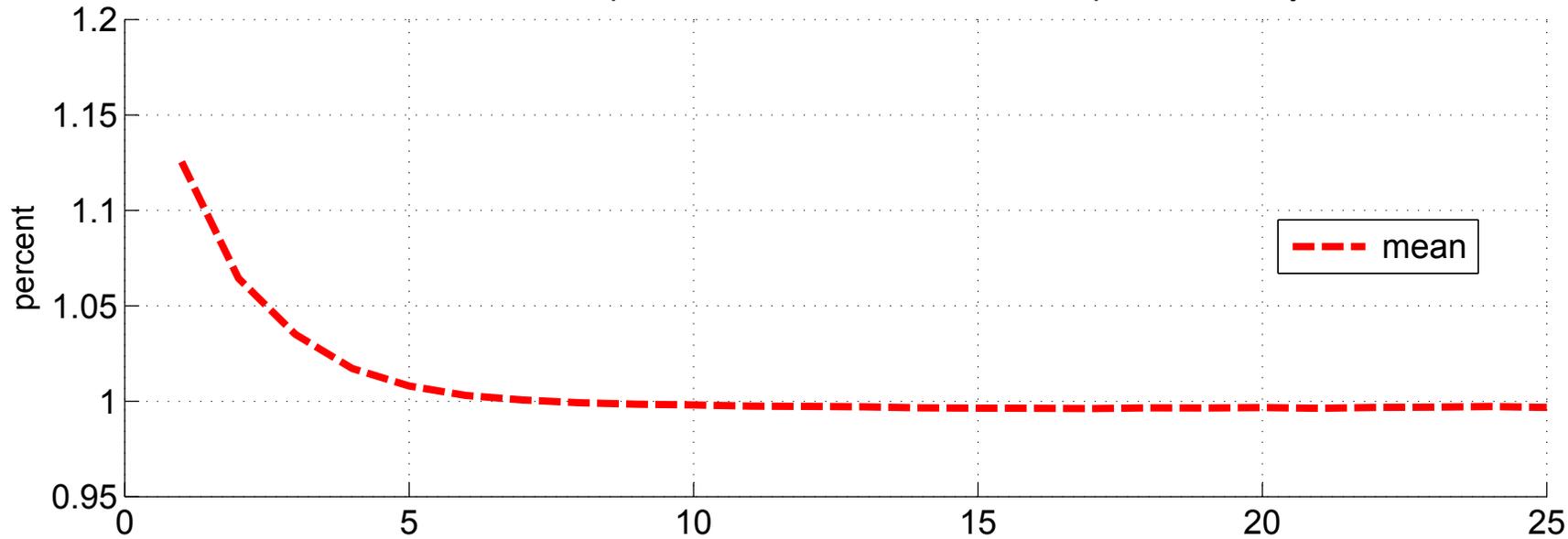


FIGURE 7. Negative technology shock

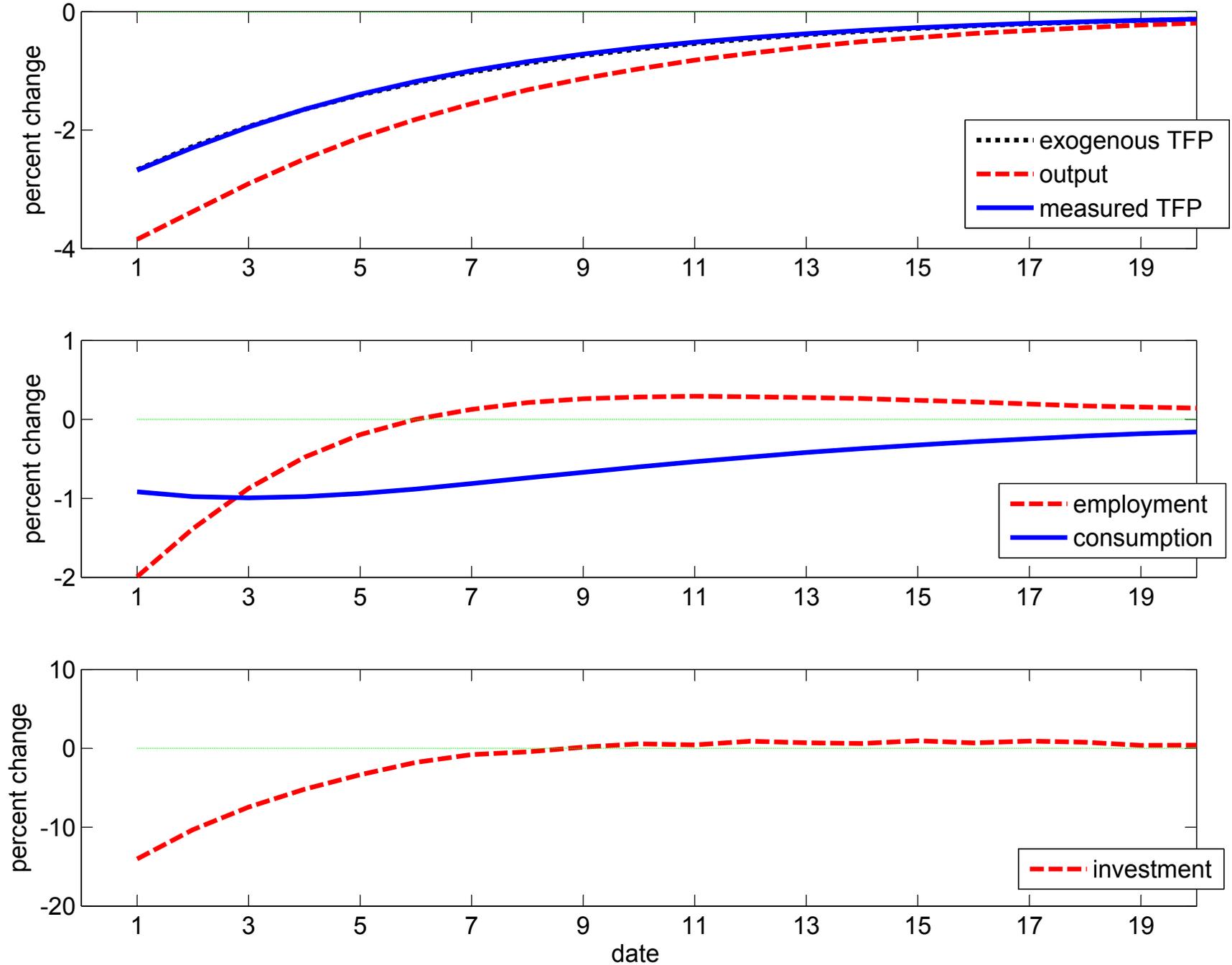


FIGURE 8. The Recent Recession

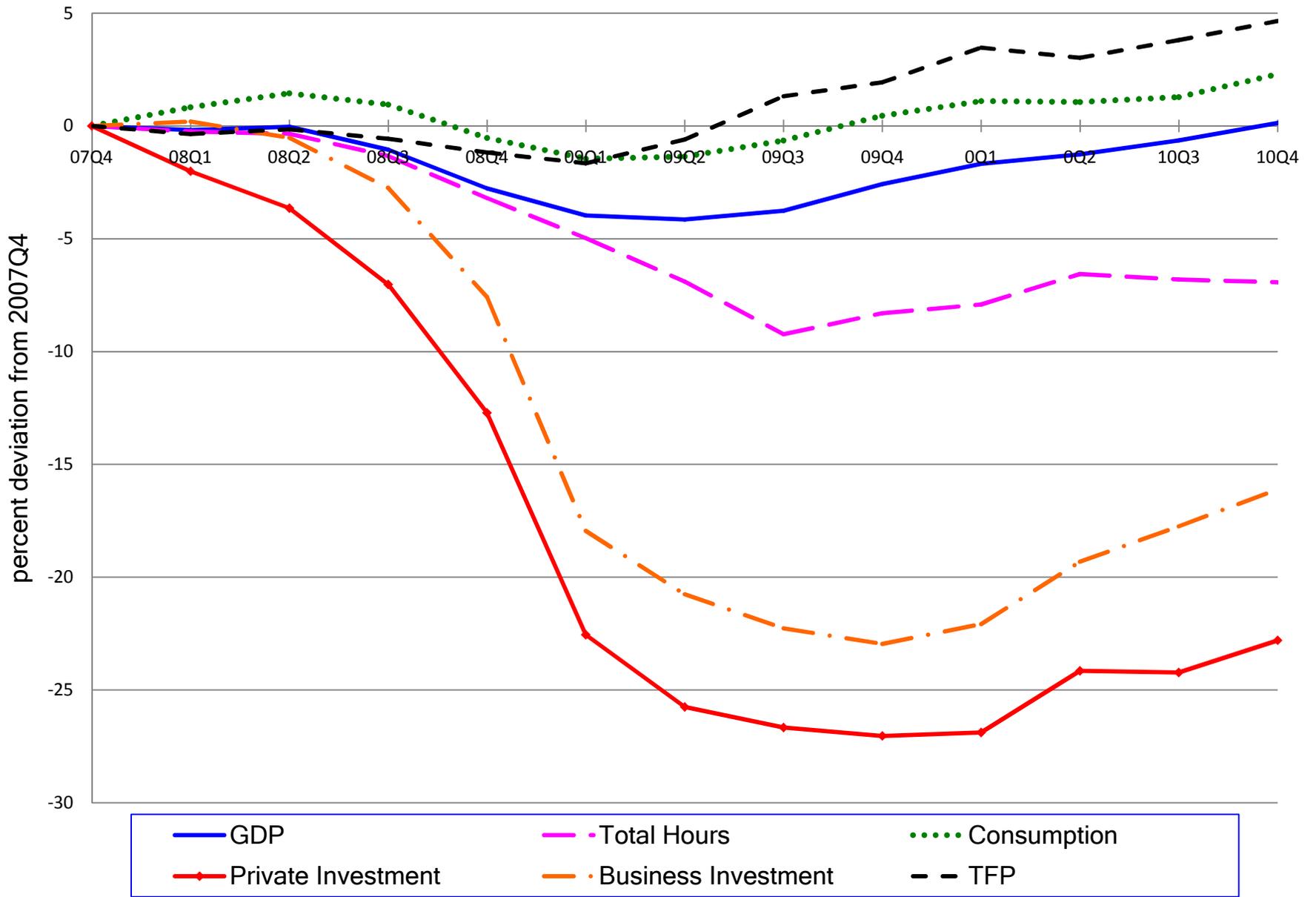


FIGURE 9. Persistent financial crisis

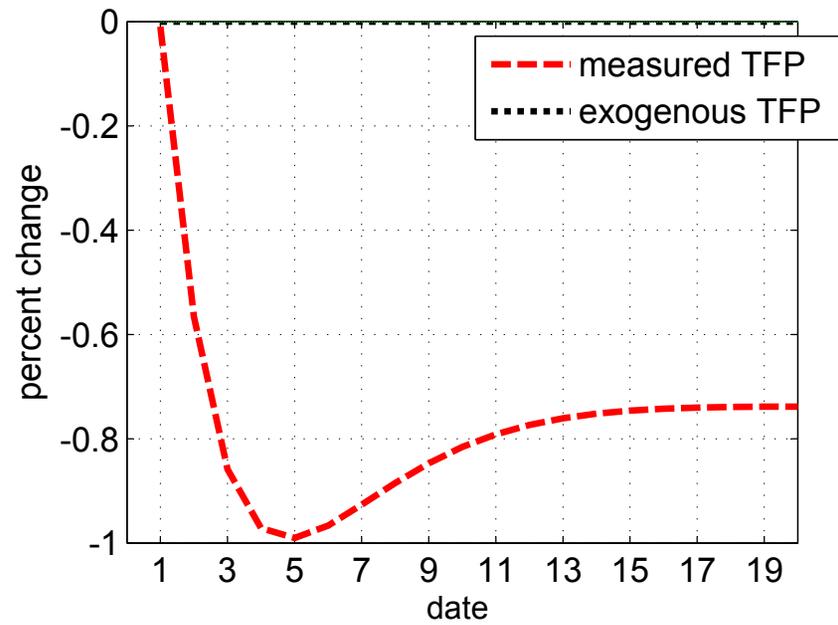
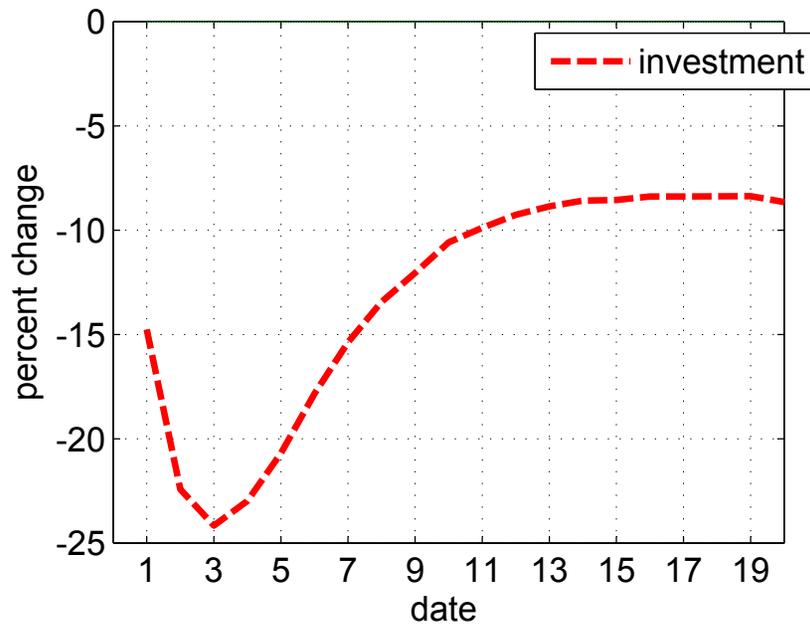
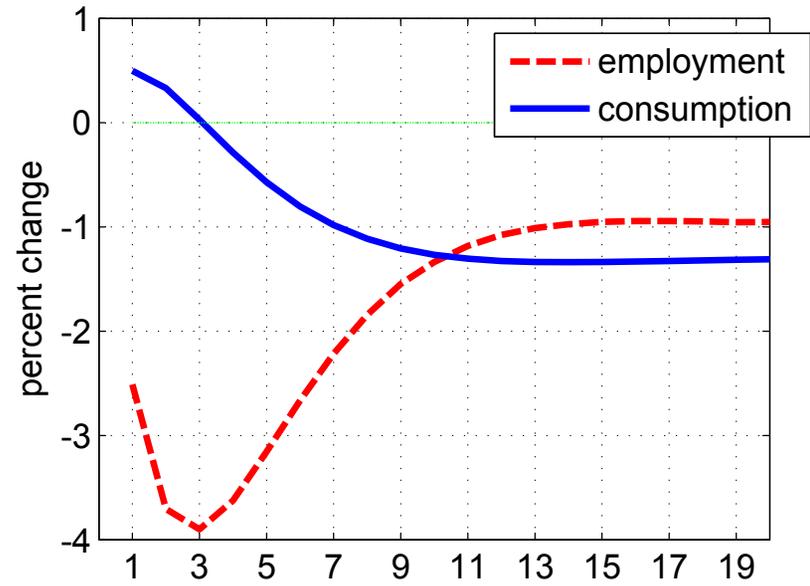
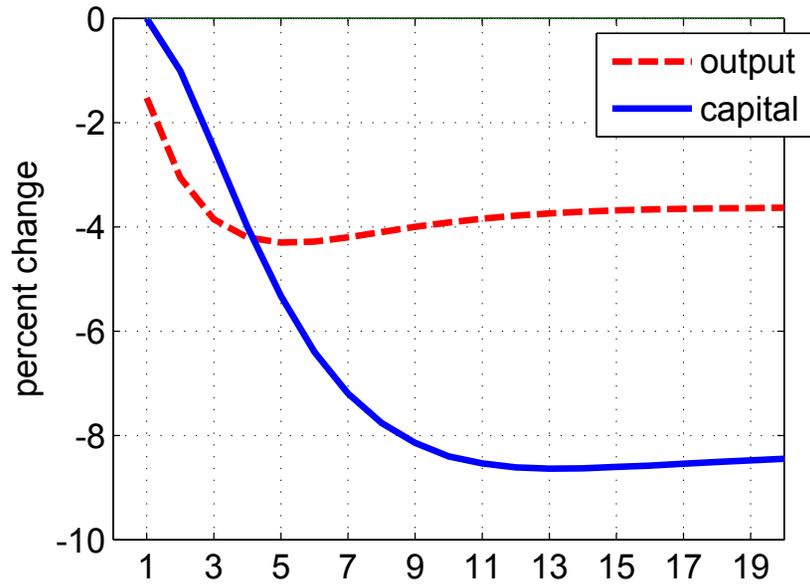


FIGURE 10. Persistent financial crisis: Distribution over (k, ε) at impact and after

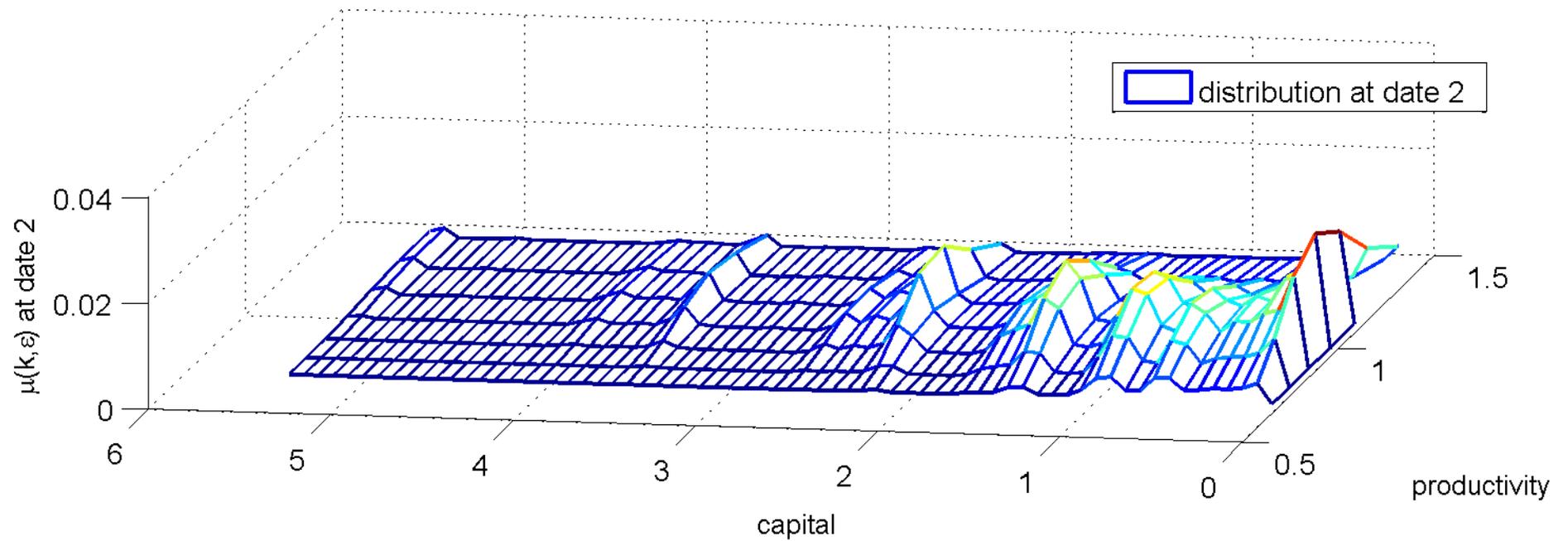
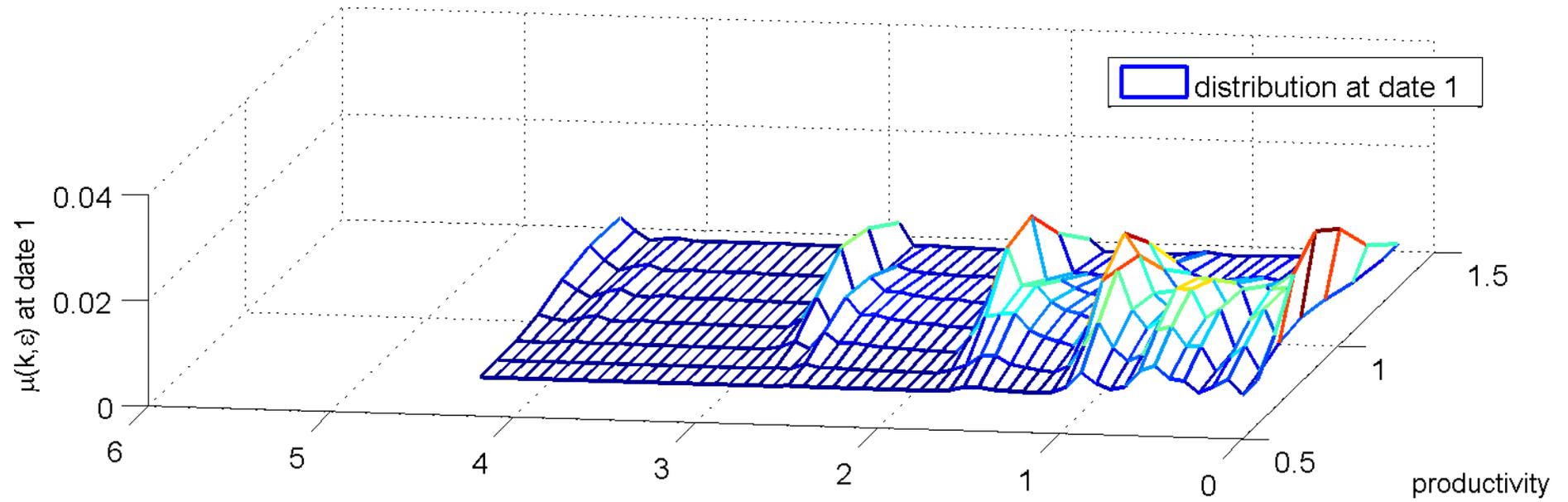


FIGURE 11. Persistent financial crisis: marginal products of capital

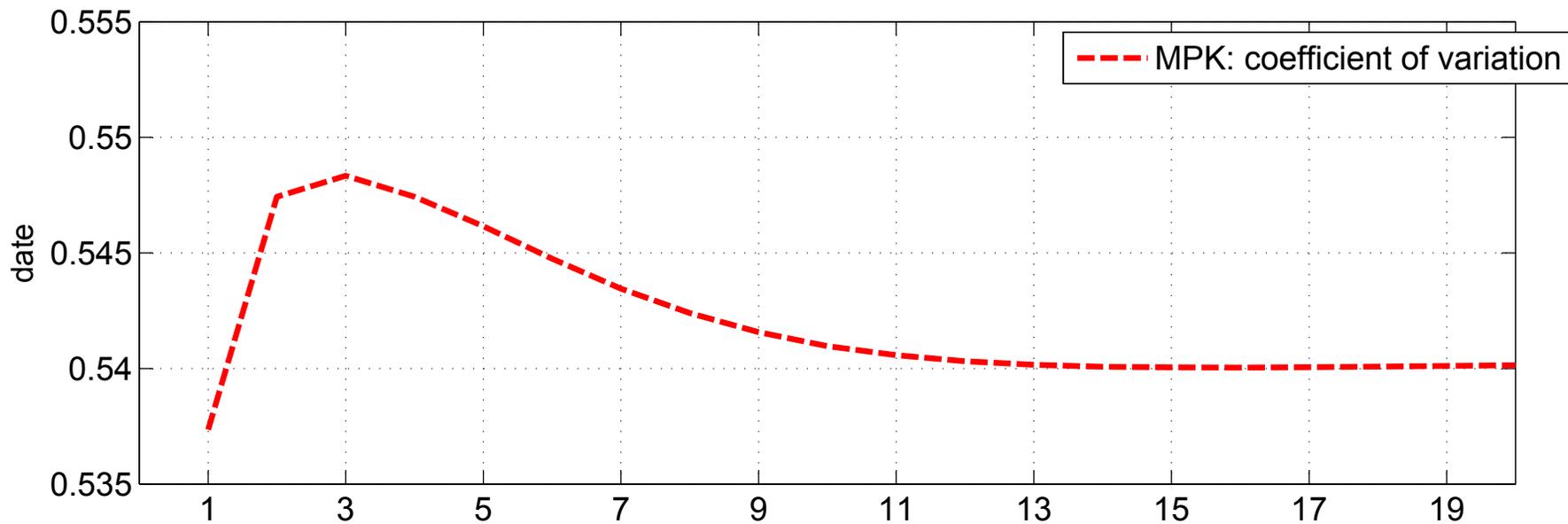
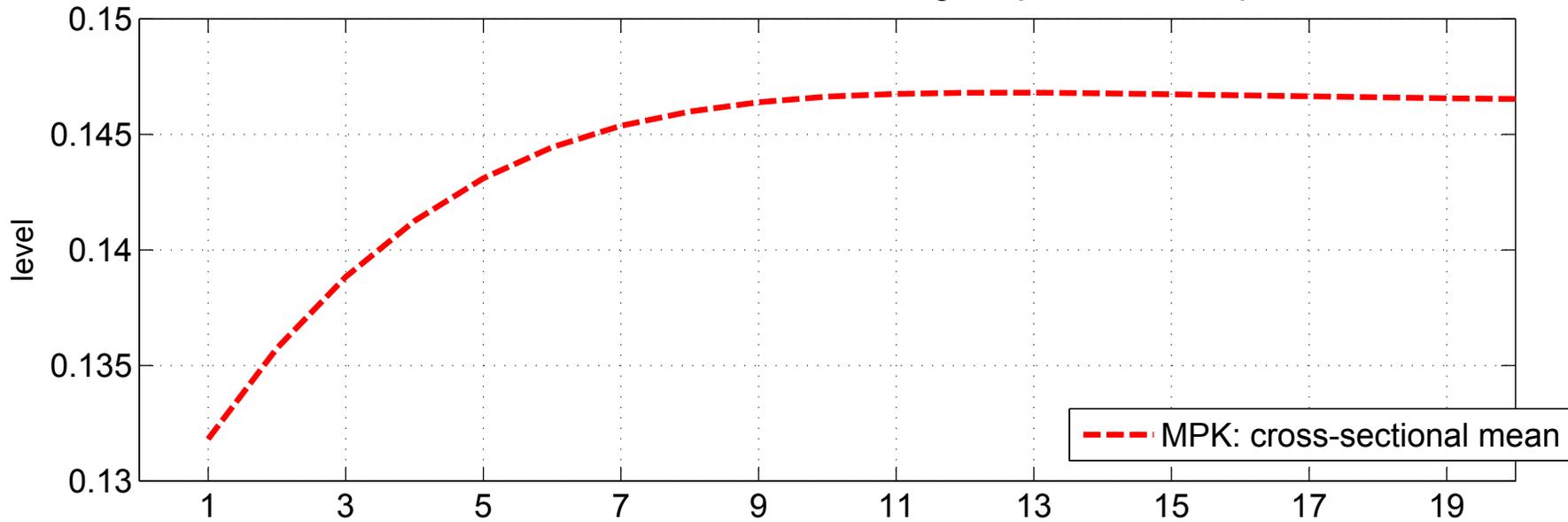


FIGURE 12. Persistent financial crisis with a technology shock

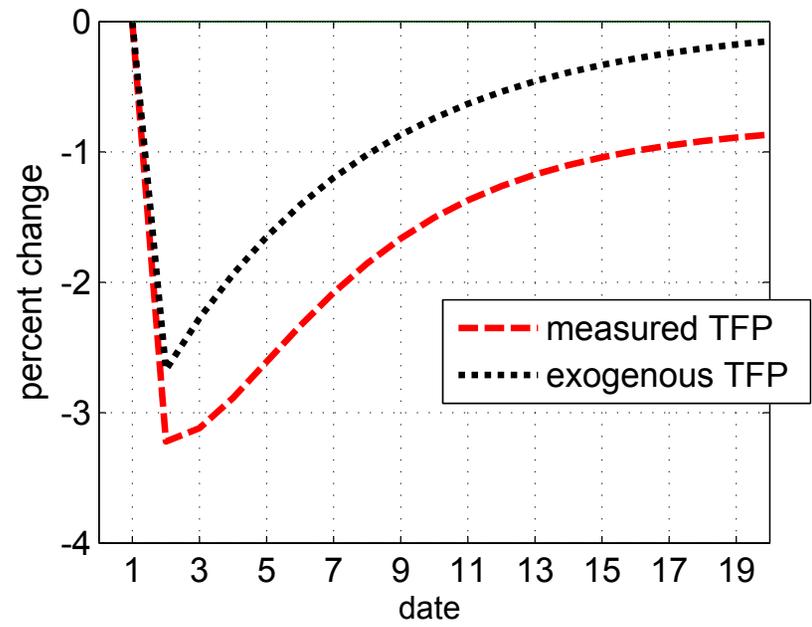
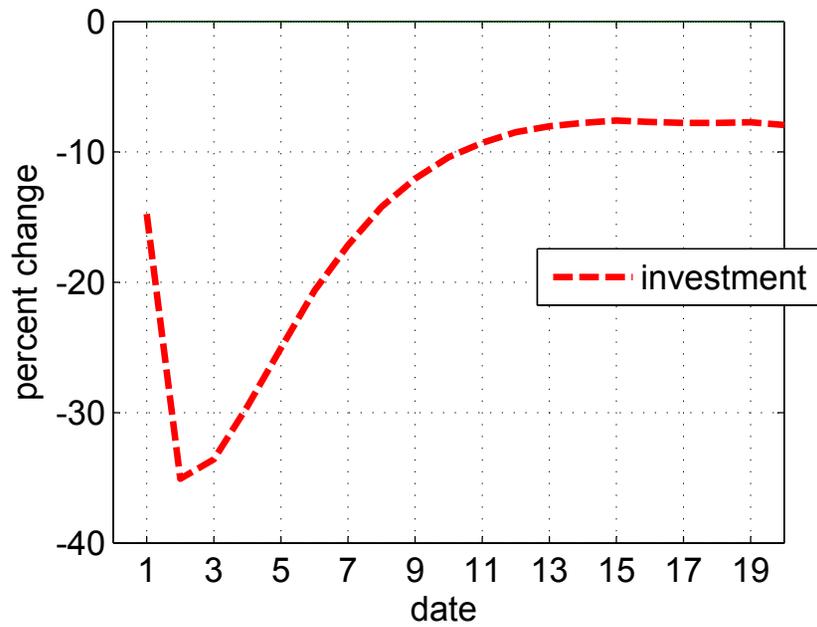
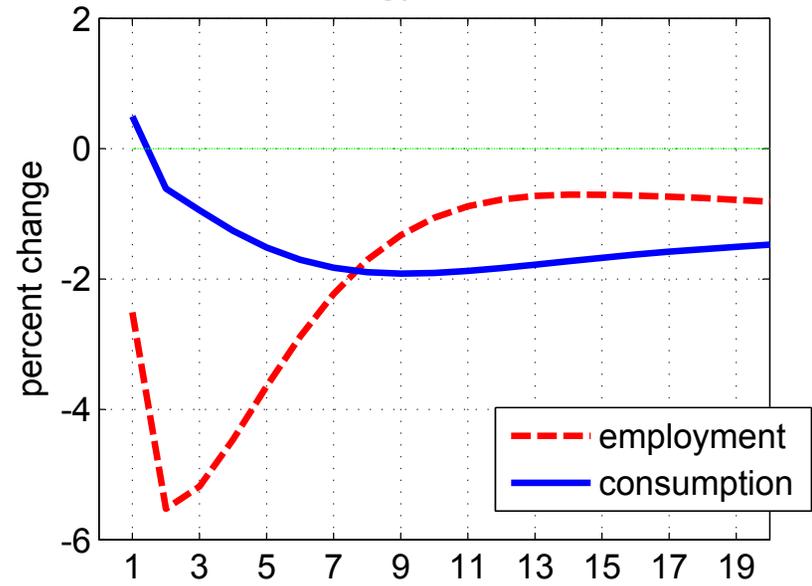
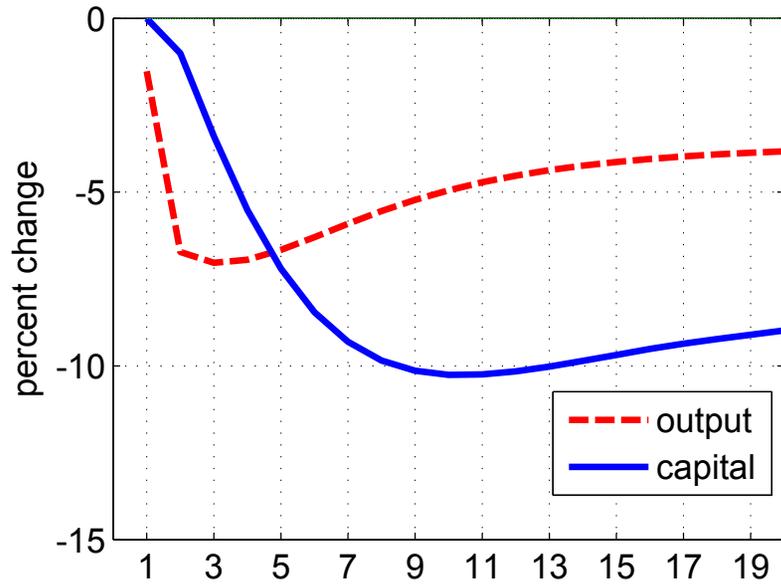


FIGURE 13. Financial crisis and recovery

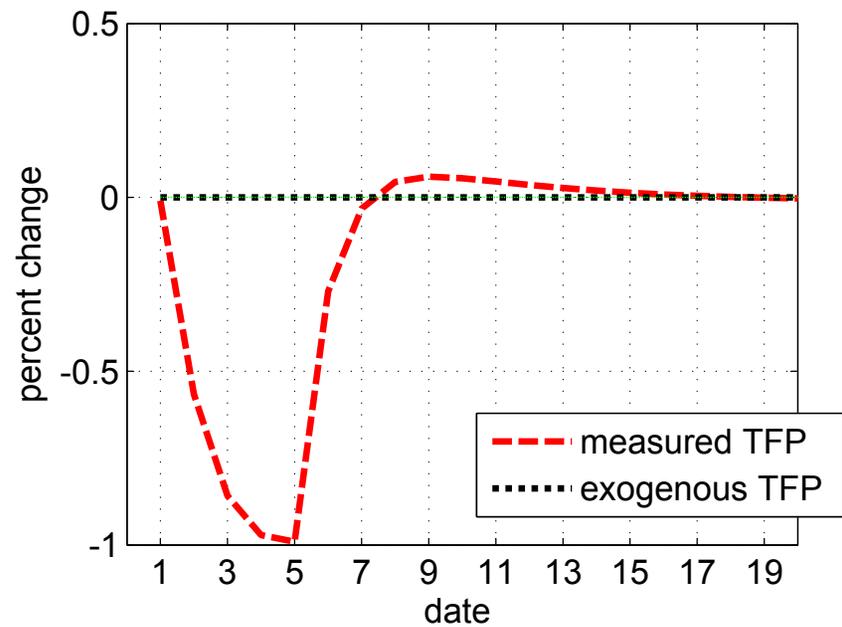
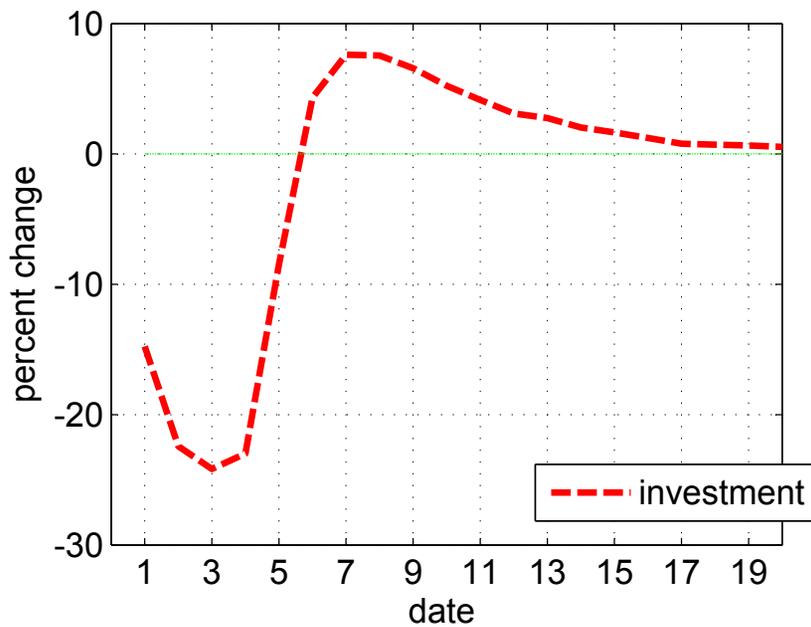
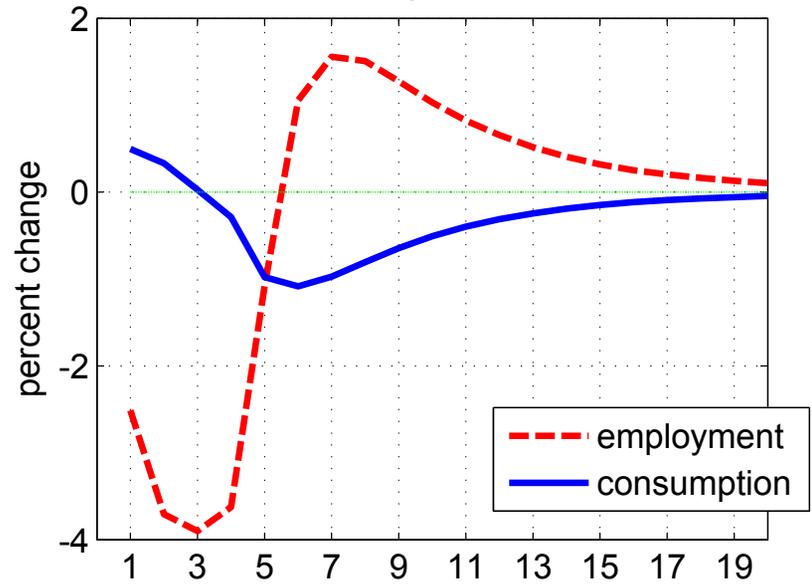
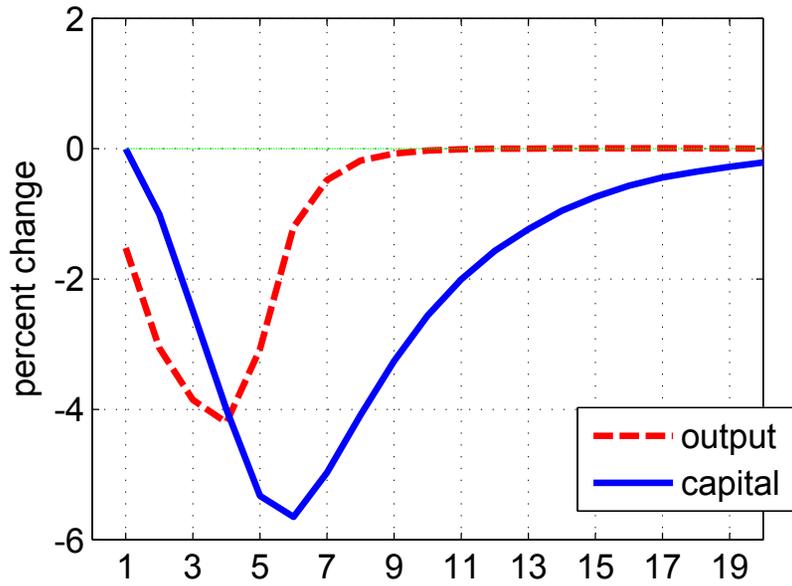


TABLE 1. Business Cycles in the Full Economy

$x =$	Y	C	I	N	K	r
$\text{mean}(x)$	0.581	0.487	0.094	0.333	1.321	0.042
σ_x/σ_Y	(1.919)	0.514	3.834	0.547	0.477	0.476
$\text{corr}(x, Y)$	1.000	0.939	0.968	0.946	0.066	0.665

TABLE 2. Business Cycles Without Financial Frictions

$x =$	Y	C	I	N	K	r
$\text{mean}(x)$	0.605	0.502	0.103	0.336	1.438	0.042
σ_x/σ_Y	(1.955)	0.497	3.768	0.568	0.471	0.454
$\text{corr}(x, Y)$	1.000	0.930	0.969	0.948	0.062	0.675

TABLE 3. Business Cycles Without Financial or Real Frictions

$x =$	Y	C	I	N	K	r
$\text{mean}(x)$	0.619	0.518	0.101	0.333	1.555	0.042
σ_x/σ_Y	(1.972)	0.479	4.037	0.588	0.451	0.440
$\text{corr}(x, Y)$	1.000	0.918	0.968	0.950	0.047	0.682

TABLE 4. Peak-to-Trough Changes: U.S. 2007 Recession and Model

$x =$	GDP	I	N	C	TFP
Data	-4.14	-25.75	-6.89	-1.36	-0.60
Model (fig. 9)	-4.20	-22.98	-3.62	-0.58	-0.97

TABLE 5. Conditional Forecasting Rules

	z realization	β_0	β_1	$S.E.$	adj. R^2
forecasting m'	z_1 (915 obs.)	0.03577	0.79631	0.00040	0.99982
	z_2 (3142 obs.)	0.05680	0.79552	0.00038	0.99975
	z_3 (984 obs.)	0.07924	0.79320	0.00023	0.99994
forecasting p	z_1 (915 obs.)	0.86951	-0.40721	0.00006	0.99998
	z_2 (3142 obs.)	0.83032	-0.39881	0.00009	0.99995
	z_3 (984 obs.)	0.79137	-0.39253	0.00003	0.99999