

LSE

THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

LSE Research Online

He Zhiguo and Peter Kondor Inefficient investment waves

Article (Accepted version)
(Refereed)

Original citation:

Zhiguo, He and Kondor, Peter (2016) Inefficient investment waves. *Econometrica*, 84 (2). pp. 735-780. ISSN 0012-9682
DOI: [10.3982/ECTA11788](https://doi.org/10.3982/ECTA11788)

© 2016 The Econometric Society

This version available at: <http://eprints.lse.ac.uk/64412/>
Available in LSE Research Online: May 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Inefficient Investment Waves*

Zhiguo He

Péter Kondor

University of Chicago and NBER

London School of Economics and CEPR

August 2015

forthcoming in Econometrica

Abstract

We show that firms' individually optimal liquidity management results in socially inefficient boom-and-bust patterns. Financially constrained firms decide on the level of their liquid resources facing cash-flow shocks and time-varying investment opportunities. Firms' liquidity management decisions generate simultaneous waves in aggregate cash holdings and investment, even if technology remains constant. These investment waves are not constrained efficient in general, because the social and private value of liquidity differs. The resulting pecuniary externality affects incentives differentially depending on the state of the economy, and often overinvestment occurs during booms and underinvestment occurs during recessions. In general, policies intended to mitigate underinvestment raise prices during recessions, making overinvestment during booms worse. However, a well-designed price-support policy will increase welfare in both booms and recessions.

Key Words: Pecuniary externality, overinvestment and underinvestment, market intervention

*Email addresses: zhiguo.he@chicagobooth.edu, kondorp@ceu.hu. We are grateful to the editor, the referees, Ulf Axelson, Hans Gersbach, Arvind Krishnamurthy, Guido Lorenzoni, Semyon Malamud, John Moore, Tyler Muir, Martin Oehmke, Alp Simsek, Balazs Szentes, Jaume Ventura, Rob Vishny, and numerous seminar participants. We thank Miklós Farkas for excellent research assistance. Zhiguo He acknowledges the financial support from the Center for Research in Security Prices at the University of Chicago Booth School of Business. Péter Kondor acknowledges the financial support of the Paul Woolley Centre at the LSE and of the European Research Council (Starting Grant #336585).

1 Introduction

The history of modern economies is rich with boom-and-bust patterns. Boom periods during which vast resources are invested in new projects are followed by downturns during which long-run projects are liquidated early, liquid resources are hoarded in safe short-term assets, and there is little investment in new projects. While some of these patterns affect only certain industries,¹ others affect the aggregate economy—e.g., the emerging market boom and bust at the end of 1990s, or the recent investment boom around the mid-2000s and the crisis afterwards. These investment cycles are in the forefront of the academic and policy debate.

In this paper, we show that firms' individually optimal liquidity management results in socially inefficient boom-and-bust patterns. Financially constrained firms choose what level of liquid resources required to absorb cash-flow shocks and to take advantage of time-varying investment opportunities. Firms hold liquid resources both to avoid inefficient liquidation of productive capital in case of adverse cash-flow shocks and to be prepared for potentially cash-intensive future investment opportunities. Our focus is on the implications for the aggregate economy when cash-flow shocks are correlated across firms.

Our first observation is that firms' liquidity management decisions generate simultaneous waves in firms' aggregate holdings of liquid assets and investment and waves of the opposite phase in market value of liquidity, even if technology remains constant. We argue that the emerging picture partially rationalizes evidence on liquidity holdings of non-financial firms and the time variation in the market value of liquidity.

The main result of this paper is that we show that such investment waves are not constrained efficient when future investment opportunities are non-contractible. The social and private value of liquidity differs in general. In particular, the incentive to turn liquid resources into illiquid capital, which affects individual firms but not the planner, is stronger during booms (i.e. after a series of favorable cash-flow shocks so that the capital price is relatively high) than during recessions. We show that the externality is often two-sided depending on the aggregate state: there is overinvestment in capital during booms and underinvestment in capital during recessions. As a result, firm investment is too volatile.

The presence of a two-sided externality radically changes the outcome of policy interventions. In general, policies targeted on raising prices in recessions help mitigate underinvestment, but make overinvestment in booms worse. As an example, consider a transfer scheme that does not allow the price of capital to fall below a certain level during recessions. We show that setting the appropriate price level for such a policy is critical. If the set price for the recession is not sufficiently low, it may decrease welfare during both booms and recessions, as agents foresee the induced overinvestment in booms. We show how a specific price-floor policy can change incentives through all states of the economy in order to increase welfare during both booms and recessions.

¹For example, Hoberg and Phillips (2010) document a large number of examples of industry-specific boom-and-bust patterns beyond the well-known examples such as the boom and bust of the semi-conductor industry in the 1990s. (See also Rhodes-Kropf, Robinson and Viswanathan (2005) for related findings.)

For our analysis, we integrate a novel, analytically tractable, stochastic dynamic model of liquidity management into a macroeconomic context. Our model focuses on non-financial firms. We call their long-term risky asset *capital*, and their liquid asset holdings *cash*. Capital stands for certain fixed investment in long-term risky technology, which produces stochastic flows in cash. Cash can be stored safely, exchanged for consumption goods, or used to build new capital at a constant proportional cost. Capital can also be liquidated for a relatively smaller constant proportional benefit in terms of cash. Thus, aggregate cash holdings represent non-financial firms' liquid financial claims on the rest of the economy. The risky cash flows generated by capital (which can also be interpreted as short-lived TFP shocks) represent aggregate shocks in our economy, and negative cash-flows imply that capital requires costly maintenance in terms of cash.

The economy is initially in the aggregate stage where identical firms facing the aggregate cash-flow shocks trade, build or liquidate capital, or consume. With Poisson intensity, firms move to the idiosyncratic stage. In that stage, some firms find a productive project, which uses the existing capital (capital firms), while others get a new idea for a project, which requires cash to be exploited (cash firms). Then, cash firms sell their capital to capital firms in a Walrasian market. After trading, cash firms invest all their cash into the new opportunity, whereas capital firms operate their capital holdings more productively. Finally, firms consume all their obtained wealth.

A crucial equilibrium implication of our setup is that the aggregate stage features simultaneous waves in investment, cash-holding of firms, and the price of capital in terms of cash, even with constant technology. Firms store the cash as a buffer in order to avoid inefficient liquidation of capital. As cash-flow shocks are perfectly correlated, a series of positive cash-flow shocks raise the aggregate level of cash holding. The larger buffer decreases the chance of a series of adverse shocks forcing firms to liquidate productive capital, and as a result raises the equilibrium price of capital. When the price of capital reaches the fixed cost of investment, firms decide to build new capital. Analogously, as a result of a series of negative cash-flow shocks the price of capital might drop to the level of the liquidation benefit, leading firms to liquidate capital. This process keeps the aggregate cash-to-capital ratio within the implied liquidation and investment thresholds. We think of the state when new capital is built as a boom period and the state when capital is liquidated as a recession.

We show that the equilibrium liquidation and investment thresholds do not coincide with a planner's choice if the investment opportunities in the idiosyncratic stage are not contractible. In the planner's solution, firms liquidate their productive capital only when the cash-to-capital ratio hits zero, and invest during booms when the cash-to-capital ratio hits a positive threshold, which is the socially optimal cash buffer in this economy. However, in the decentralized equilibrium, the investment and disinvestment thresholds are distorted. In particular, firms always liquidate capital at a strictly positive cash-to-capital ratio, implying that firms always underinvest in downturns. Interestingly, under some conditions firms invest in capital when the cash buffer is lower than the one the planner would choose. That is, they underinvest in capital (liquidate too much) in downturns and overinvest during booms. As a mirror image, they hoard too much cash during a

downturn, and hold too little cash during a boom.

Here is the economic intuition: As we noted, firms' incentive to build liquidity buffers against cash-flow shocks generates procyclicality in aggregate liquidity holdings and countercyclicality in the value of liquidity, implying that the value of capital relative to cash, i.e. the capital price, has to be procyclical. Once investment opportunities arrive, cash firms can sell the capital they have, and capital firms buy the capital at the prevailing market price in terms of cash. Therefore, in booms, when the price of capital is higher, firms value their capital more than cash. That is, preparing for investment opportunities aggravates procyclicality in capital prices. However, this additional effect that influences private incentives is absent from social incentives, because one firm's gain from trading capital to cash is the other firm's loss. Therefore, there is a state-dependent wedge between the private and social valuation of capital (relative to cash), creating the possibility of overinvestment in booms and underinvestment in recessions.

This argument holds because we assume that certain markets are missing. For example, firms writing contracts *ex ante* on investment opportunities would insure each other against the gains and losses from *ex post* trading. Similarly, firms able to pledge the output of their investment opportunities, would exchange capital to cash at terms determined by the (fixed) output of these opportunities. These possibilities eliminate the wedge between the market price and the social value of capital, restoring the constrained efficiency for the decentralized economy.

As an extension of our model, we allow firms to pledge capital to obtain external credit by collateralized borrowing. This makes capital more valuable from both the private and social perspectives. We show that collateralized borrowing tends to push up the private benefit of capital more than it does on the capital's social value. Therefore, collateralized borrowing could be excessive, in the sense that a sufficiently large borrowing capacity of capital brings a no-borrowing economy from "underinvestment always" to two-sided inefficiency, with overinvestment during the boom.

As an illustration of the potential of our mechanism to provide new explanations for existing problems in various contexts, we connect our results to the observed phenomenon of relative boom-and-bust patterns across industries, and to stylized facts that in less financially developed countries investment in productive technologies is more volatile and exhibits stronger procyclicality.

As a methodological contribution, we develop a novel dynamic model to analyze the effect of aggregate liquidity fluctuations on asset prices and real activity, with analytical tractability of the full joint distribution of states and equilibrium objects.

Literature. In our model, firms' individually optimal liquidity management decisions generate aggregate waves in investment, market value of liquidity, and aggregate liquidity holdings. As a main contribution, we show that if future investment shocks are non-contractible, firms often have too much incentives to invest during booms and too little incentives to invest during recessions.

Ours is not the first paper to emphasize that firm-level constraints can generate inefficient investment waves. The literature with perhaps the largest influence on current policy discussions emphasizes the fire-sale feedback loops induced by a price-sensitive collateral constraint (e.g., Kiy-

otaki and Moore, 1997; Gromb and Vayanos, 2002; Krishnamurthy, 2003; Jeanne and Korinek, 2010; Bianchi, 2010; Bianchi and Mendoza, 2011; Stein, 2011; He and Krishnamurthy, 2012; Jermann and Quadrini, 2012; Brunnermeier and Sannikov, 2014). In these models, firms fail to internalize that the more they borrow and invest during booms, the more they have to deleverage and disinvest during recessions, which depresses fire-sale prices and tightens the constraint faced by other firms as well. Compared to a social planner facing the same constraints, in these models firms' incentives to borrow and/or invest are always too strong. Our research differs from this literature in two crucial dimensions. First, our mechanism is unrelated to any form of collateral-based or net-worth-based amplification mechanism. Second, and more important, the externality in our model changes sign with the state of the economy. As a result, policy measures limiting overexpansion in booms, which are unambiguously beneficial in an economy with collateral constraints, cause inefficient hoarding of liquidity in our economy and potentially decrease welfare everywhere.²

Like the literature on fire-sale feedback loops, our work also belongs to the literature analyzing the welfare effects of pecuniary externalities. This literature is based on the seminal papers of Stiglitz (1982), Greenwald and Stiglitz (1986), and Geanakoplos and Polemarchakis (1985), which, like the recent work of Farhi and Werning (2013), establish general conditions implying welfare-changing pecuniary externalities. Our application of this general principle is closest to the vein of research in which market incompleteness hinders the equalization of firms' marginal utility of wealth across states or time (e.g., Shleifer and Vishny (1992), Allen and Gale (1994, 2004, 2005), Caballero and Krishnamurthy (2001, 2003), Lorenzoni (2008), Farhi, Golosov and Tsyvinski (2009) and Gale and Yorulmazer (2011)). Compared to a planner, this mechanism can imply that incentives to invest are either too strong or too weak, depending on the exact specification.³ Our main innovation is that we highlight the effect of interacting these types of pecuniary externalities with varying incentives to hold liquid assets over the cycle. This interaction leads to our main result that the sign of the distortion in investment incentives switches with the state of the economy.

A few recent papers cast in two-period settings investigate two-sided inefficiency and derive implications related to our work. Gersbach and Rochet (2012) study the moral hazard problem of incentivizing banks in a macroeconomic context, and show that banks extend too much credit in booms and too little in recessions. Their mechanism relies on the difference between the private and social solution of bank's moral hazard problem. Additionally, in their two-period setting which models booms and recessions separately as two different states in period 1, the period 0 intervention can resolve the two-sided efficiency at once. In contrast, in our dynamic model booms and recessions

²This paper contributes to the discussion on the optimal mix of ex ante regulation and ex post intervention (e.g., Diamond and Rajan, 2011; Farhi and Tirole, 2012; Jeanne and Korinek, 2013), to the extent that we emphasize that a policy of intervention during a recession will also affect incentives during a boom. We characterize economies when, because of the two-sided externality, this fact has crucial consequences on the welfare effects of these policies.

³See Davila (2014) for a comparative analysis of different mechanisms connected to pecuniary externalities and the argument that collateral constraints always imply overinvestment ex ante. For uninsurable idiosyncratic liquidity shocks, see Holmstrom and Tirole (2011, chap.7.) of simplified versions and excellent discussion of Shleifer and Vishny (1992) and Caballero and Krishnamurthy (2003). Finally, a recent paper by Hart and Zingales (2011) studies the excessive supply of private money based on the idea of special pledgeability of certain assets. This friction always results in overinvestment in such assets, in contrast to our model.

occur in cycles, and the potentially inferior one-sided interventions emphasize the interconnected incentives between booms and recessions for forward-looking economic agents. Eisenbach (2013) studies banks financed with short-term debt in a general equilibrium setting, and show that in good (bad) times banks face too little (much) market discipline imposed by rolling over short-term debt. In contrast to our paper, in which idiosyncratic investment opportunities drive inefficiency, that paper emphasizes aggregate risk, and the fact that short-term debt lacks aggregate-state contingency.

In our model, firms hold liquid assets to avoid adverse effects of cash-flow shocks and to prepare for future investment opportunities. This is consistent with a large body of previous work on liquidity management (e.g., Almeida, Campello and Weisbach (2004), Bates, Kahle and Stulz (2009), Denis and Sibilkov (2010), Ivashina and Scharfstein (2010), Lins, Servaes and Tufano (2010), Eisfeldt and Muir (2013), Acharya, Almeida and Campello (2013)). This argument goes back to Keynes, who calls this the precautionary motive.⁴ However, instead of aiming for a detailed picture of firms' individual saving and investment decisions, we focus on the consequences of such decisions to the aggregate economy.

The structure of our paper is as follows. In Section 2 we present the setup and the equilibrium of our model. In Section 3 we expose the inefficiencies of the market solution. Section 4 presents our findings on economic policy and other applications. We discuss the robustness of our mechanism in Section 5. We conclude in Section 6. All proofs are in Appendix, Online Appendix, or Additional Material available on the author's website.⁵

2 A Dynamic Model of Saving and Investment

2.1 Assets

We model an economy where firms facing cash-flow shocks and time-varying investment opportunities make saving and investment decisions. There is a single capital good representing risky and productive projects. The other asset in this economy is cash which serves both as a consumption good and as an input for building capital. We assume that there is a safe storage technology and that capital does not depreciate; thus both capital and cash are perfectly storable.

For each firm, there is a final date arriving at a stopping time τ with Poisson intensity ξ , where ξ is a positive constant. At this final date, firms receive potentially different investment opportunities (to be specified shortly), and any unused capital depreciates fully. For now, we think of the arrival of the final date as an aggregate shock (we offer an alternative interpretation in Section 2.4). Before the final date, each unit of capital generates random cash flows. This shock is common across capital units and driven by σdZ_t , where σ is a positive constant and $Z \equiv \{Z_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is a standard Brownian-motion on a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$. One can interpret the

⁴Others proposed the tax motive, the transaction motive, and the agency motive as alternative explanations (see Bates, Kahle and Stulz (2009) for detailed arguments and references).

⁵See http://faculty.chicagobooth.edu/zhiguo.he/research/hekondor_additionalmaterial.pdf.

aggregate cash-flow shocks σdZ_t as short-lived TFP shocks. When $\sigma dZ_t > 0$ the capital generates cash. When $\sigma dZ_t < 0$, the firm needs to spend $|\sigma dZ_t|$ amount of cash on this capital as maintenance cost; otherwise the capital turns unproductive.

Denote by K_t the aggregate quantity of capital. Given the aggregate cash shock σdZ_t of each unit of capital, when firms do not invest or disinvest (to be introduced shortly), the aggregate level of cash accumulated in storage, C_t , would follow the evolution of

$$dC_t = K_t \sigma dZ_t. \quad (1)$$

2.2 Firms and frictions

The market is populated by a unit mass of risk-neutral firms who operate the capital. At each time instant, firms may decide to build new capital, trade capital for cash at the equilibrium price p_t , or liquidate the capital. Building new capital costs h units of cash, while liquidating a unit of capital provides l units of cash, where $h > l > 0$. Firms can also consume their cash at any moment of a constant marginal utility of 1. Because of linear technologies, in general it is optimal to have threshold strategies of (dis)investment. Thus, we can simply focus on thresholds in comparing different (dis)investment strategies.

The major friction in this economy is that firms can neither write contracts on the different investment opportunities they face, nor they can pledge the future return on these opportunities. Although firms are initially identical, they receive different investment opportunities. Specifically, in the random final date, each firm with probability half finds a project which uses the existing capital productively generating $R_K > 0$ unit of final consumption per each unit of used capital. The other group of firms find a new idea requiring liquid resources. Hence, this latter group have a superior use for liquid resources, and we assume that they receive $R_C > 1$ unit of final consumption per unit of cash invested. These shocks are independent across firms, and we refer to the earlier group as capital firms and the latter group as cash firms. R_K and R_C are positive constants. Our extreme assumption that neither group's project returns are pledgeable is a short-cut for agency and/or informational frictions.⁶ We partially relax this assumption in Section 5.2. Throughout we assume that

$$\frac{R_K}{R_C} > h, \quad (2)$$

which ensures that building capital is socially efficient when the economy has sufficient cash.⁷

Firms learn which group they belong to only at the beginning of the final date. In the final date the conversion technology between capital and cash is no longer available, but firms have a last trading opportunity to trade capital for cash before final production. We refer to the potentially infinitely long interval before the final date τ as the aggregate stage of the economy, as at this stage

⁶ Appendix C of He and Kondor (2012), in the context of a simple two-period example, discusses the potential agency problems in detail.

⁷ Allowing for $h > \frac{R_K}{R_C} > l$ would leave the derivation and the characterization of the market equilibrium untouched. Although the comparison to the planner's case remains similar, the derivation is more cumbersome. Hence, for easier readability, we discuss this case in Remark 1 in Section 3.2.

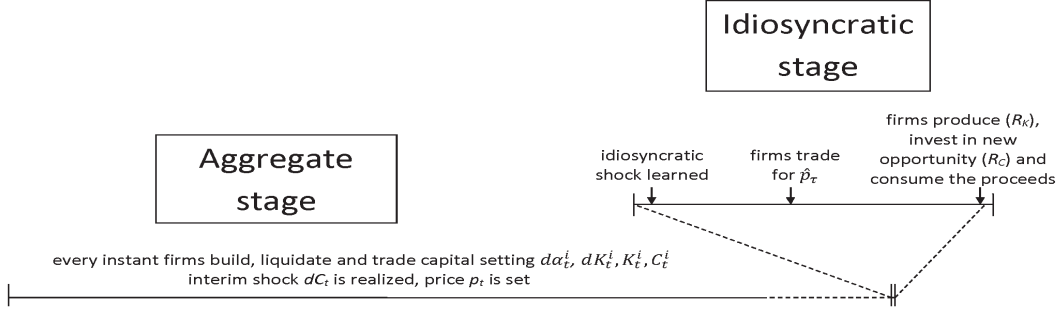


Figure 1: Time line.

all shocks affect each agent the same way. By similar logic, we refer to the final date τ (in which final trading occurs) as the idiosyncratic stage. We denote the price in the idiosyncratic stage by \hat{p}_τ (recall that we denote by p_t the prices in the aggregate stage). Figure 1 summarizes the time line of events in our model. We expand on the interpretation of the two stages in Section 2.4.

2.3 Individual firm's problem

Consider firm i , which holds K_t^i units of capital and C_t^i amount of cash, with a wealth (in terms of cash) of $w_t^i \equiv p_t K_t^i + C_t^i$. Since the idiosyncratic stage arrives according to an exponential distribution with density $\xi e^{-\xi\tau}$, firm i is solving the following problem:

$$\max_{\{d\alpha^i \geq 0, K^i \geq 0, C^i \geq 0, dK^i\}} \mathbb{E} \left\{ \int_0^\infty \xi e^{-\xi\tau} \left(\int_0^\tau d\alpha_t^i + \left[\frac{1}{2} \left(K_\tau^i + \frac{C_\tau^i}{\hat{p}_\tau} \right) R_K + \frac{1}{2} (K_\tau^i \hat{p}_\tau + C_\tau^i) R_C \right] \right) d\tau \right\} \quad (3)$$

where α_t^i is firm i 's cumulative consumption before the final date τ (so it is non-decreasing with $d\alpha_t^i \geq 0$; later we see that it is zero in equilibrium), and dK_t^i is the amount of capital that it dismantles or builds. The term in the squared bracket is the consumption at the idiosyncratic stage. For instance, if the firm turns out to be cash-type, it will sell its capital holding K_τ^i at the price of \hat{p}_τ to receive $K_\tau^i \hat{p}_\tau$, and then invest its cash together with C_τ^i in exploiting new cash-intensive projects with return R_C .

The problem in (3) is subject to the dynamics of individual wealth,

$$dw_t^i = -d\alpha_t^i - \theta dK_t^i + K_t^i (dp_t + \sigma dZ_t), \quad (4)$$

where θ is the cost of changing the amount of capital, so that $\theta = h\mathbf{1}_{\{dK_t^i \geq 0\}} + l\mathbf{1}_{\{dK_t^i < 0\}}$. Also, wealth cannot be negative at any point, i.e. $w_t^i \geq 0$ of all t .

Recall $K_t = \int_i K_t^i di$ is the aggregate capital. Combining the investment/disinvestment policy

dK_t , (1) implies that the dynamics of aggregate cash level in the economy is⁸

$$dC_t = \sigma K_t dZ_t - \theta dK_t. \quad (5)$$

The scale-invariance implied by the linear technology suggests that it is sufficient to keep track of the dynamics of the cash-to-capital ratio:

$$c_t \equiv \frac{C_t}{K_t},$$

which evolves according to

$$dc_t = \frac{dC_t}{K_t} - \frac{C_t}{K_t} \frac{dK_t}{K_t} = \sigma dZ_t - (\theta + c_t) \frac{dK_t}{K_t}. \quad (6)$$

2.4 Interpretation of the aggregate and idiosyncratic stages

We stress that thinking of the arrival of the idiosyncratic stage as an aggregate shock and the resulting separation of the two stages is a didactic tool. It helps show how the incentives related to the idiosyncratic investment opportunities affect the incentives for saving and investing in the aggregate stage. In the real world, some firms might be in the idiosyncratic stage while others are still in the aggregate stage. Therefore, the final date does not correspond to an observable time point in the economy. Instead, we will think of recessions and booms and economic policies affecting saving and investment in these states within the aggregate stage of the economy. With this structure we can analyze the dynamic fluctuation of our economy without sacrificing analytical tractability.

Indeed, there is a formally equivalent economy where the arrival of the final date is idiosyncratic to individual firms. Under this interpretation, in each time interval dt a ξdt fraction of firms randomly receive heterogenous investment opportunities as above (i.e., $\frac{\xi}{2} dt$ fraction are capital firms while the other $\frac{\xi}{2} dt$ fraction cash firms), enter the idiosyncratic stage and trade cash for capital among themselves on a separate market, while the remaining firms continue to operate in the aggregate stage. Thus, under this interpretation the economy never terminates.

In this alternative economy, the individual firm's problem (3) and the evolution of aggregate state (6) remain the same. Because at each instant there are equal fractions of cash and capital flowing out from the economy, the aggregate cash-to-capital ratio in the remaining economy is not affected; but the size of the remaining economy shrinks. The trading price also remains \hat{p}_τ in the separate market, while all incumbent firms face a trading price of p_t .

To further emphasize that this separation is a technical innovation, in Section 5.1 we present and analyze a version of our model where a fraction of firms learn about new investment opportunities in each time instant and trade cash and capital in a single market together with the firms who remain in the aggregate stage. That is, the aggregate stage and the idiosyncratic stage are not

⁸To simplify notation we ignore the possibility that at any given point in time some firms create capital while some firms liquidate capital. It is easy to see that this never happens in equilibrium.

separated. While that version is not analytically tractable, we will illustrate by numerical analysis that our main result goes through.

2.5 Interpretation of cash and capital

Cash holding in the aggregate stage, C_t^i , represents the financial slack of a firm—cash holdings; other short-term, liquid investments; or credit lines. It can be used either to cover any operating losses, or invest in any new opportunities (even outside the industry). As we illustrate in Figure 3, it is possible to map C_t^i to data by thinking of it as the liquid financial asset holdings of non-financial firms of the economy. Note that in reality these assets represent claims on the government, households, or foreigners: entities which we do not explicitly model.

K_t^i represents firms' total gross property, plants, equipment, inventories, and intangible assets, which is much more specific to each industry and thus much less liquid. The process $\sigma K_t^i dZ_t$ might represent cash flows from both operating and financing activities. In our abstract model without external financing, firms finance their investment from retained earnings only. However, in Section 5.2 we show that allowing for collateralized borrowing could make our main results more pronounced.

Importantly, R_C in the idiosyncratic stage should not be interpreted as the return from liquid investments. Instead, it is a reduced-form representation of the expected return from the cash-intensive development of a new idea. We follow a reduced-form treatment. In reality, the cash might have to be used to hire labor, or purchase specific capital for the new idea. R_K can be interpreted similarly, but for an idea that uses the same type of capital as the existing technology. Cash firms are the ones with comparative advantage in exploiting the former, whereas capital firms have comparative advantage in exploiting the latter.

2.6 Definition of equilibrium

Definition 1 *In the market equilibrium,*

1. *each firm chooses $d\alpha_t^i$, K_t^i , C_t^i , and dK_t^i to solve (3), and*
2. *markets clear in every instant, during both the aggregate and the idiosyncratic stages.*

As we will see, in our framework, the equilibrium only pins down the aggregate variables: prices, net trade, and net investment and disinvestment. Typically, any combination of individual actions consistent with the aggregate variables is an equilibrium. It is convenient to pick the particular market equilibrium where all firms follow the same action, which we refer to as the *symmetric equilibrium*.

Henceforth, we omit the time subscript t or τ whenever it does not cause any confusion.

2.7 Market equilibrium

We solve for the market equilibrium in this section. As we show, in this economy consumption (of cash) before the idiosyncratic stage is strictly suboptimal, thus $d\alpha_t^i = d\alpha_t = 0$ always.

2.7.1 Equilibrium price in the idiosyncratic stage

Consider the idiosyncratic stage. The law of large numbers implies that there exists a half measure of capital (cash) firms. All capital firms use their cash holdings to buy capital holdings from cash firms, and the market clearing condition implies that

$$\frac{1}{2}C = \frac{1}{2}K\hat{p} \Rightarrow \hat{p} = c.$$

We still need to ensure that $R_K \geq \hat{p} = c$: This is because capital firms have the option of consuming their cash holdings instead of purchasing capital, which puts an upper bound on \hat{p} . Later we show that the full support of c is endogenous, because firms build (dismantle) capital whenever the aggregate cash is sufficiently high (low). For simplicity, we restrict the parameter space to ensure that the condition $c \leq R_K$ holds always in equilibrium.

2.7.2 Equilibrium values, prices, and investment policies in the aggregate stage

Now we determine equilibrium objects in the aggregate stage. The next lemma states two useful features of our formalization: First, the only relevant aggregate state variable is the cash-to-capital ratio. Second, the value function of any individual firm is linear in its capital and cash holdings.

Lemma 1 *Let $J(K^i, C^i, K, C)$ be the value function of firm i which holds capital K^i and cash C^i in an economy with aggregate capital K and aggregate cash C . Then, for aggregate cash-to-capital ratio $c = C/K$, there are functions $v(c)$ and $q(c)$ that,*

$$J(C, K, K^i, C^i) = K^i v(c) + C^i q(c).$$

That is, regardless of the firm's composition of asset holdings, the value of every unit of capital is $v(c)$, and the value of every unit of cash is $q(c)$. Both functions depend only on the aggregate cash-to-capital ratio. Because of linearity, the equilibrium price has to adjust in a way such that firms are indifferent to whether they hold capital or cash. That is, the equilibrium price of capital $p(c)$ in the aggregate stage must satisfy that

$$p(c) = \frac{v(c)}{q(c)}.$$

Firms build capital whenever the capital price p reaches the cash cost h , and they dismantle capital whenever the price falls to the liquidation benefit l . Define c_h^* and c_l^* as the endogenous thresholds of the aggregate cash-to-capital ratio where firms start to build and dismantle capital,

respectively. These thresholds satisfy

$$\frac{v(c_h^*)}{q(c_h^*)} = h, \text{ and } \frac{v(c_l^*)}{q(c_l^*)} = l. \quad (7)$$

Moreover, the linear technology implies that c_h^* and c_l^* are reflective boundaries of the process c . Therefore, based on (6), the aggregate cash-to-capital ratio c must fluctuate in the interval $[c_l^*, c_h^*]$, with a dynamics of

$$dc = \sigma dZ_t - dU_t + dB_t, \quad (8)$$

where $dU_t \equiv (h + c_h^*) \frac{dK_t}{K_t}$ reflects c at c_h^* from above, while $dB_t \equiv (l + c_l^*) \frac{dK_t}{K_t}$ reflects c at c_l^* from below. The standard properties of reflective boundaries imply the following smooth-pasting conditions for our value functions (Dixit (1993)):

$$v'(c_h^*) = q'(c_h^*) = q'(c_l^*) = v'(c_l^*) = 0. \quad (9)$$

2.7.3 Characterizing the market equilibrium

Now we turn to characterizing the value functions $v(c)$ and $q(c)$ in the range $c \in [c_l^*, c_h^*]$. Here we give a sketch; full details are available in the Online Appendix. Because of Lemma 1, firms are indifferent to the composition of their asset holdings, and we can consider the value function of a firm that holds only capital or only cash. The value function of a firm holding only cash gives an Ordinary Differential Equation (ODE) of $q(c)$:

$$0 = \underbrace{\frac{\sigma^2}{2} q''(c)}_{\text{volatility of } dc_t} + \underbrace{\frac{\xi}{2} (R_C - q(c))}_{\text{becoming a cash firm}} + \underbrace{\frac{\xi}{2} \left(\frac{R_K}{c} - q(c) \right)}_{\text{becoming a capital firm}}, \quad (10)$$

and the value function of a firm holding only capital, given $q(c)$, yields the ODE for $v(c)$:

$$0 = \underbrace{\frac{\sigma^2}{2} v''(c)}_{\text{volatility of } dc_t} + \underbrace{q'(c) \sigma^2}_{\text{expected value of dividends}} + \underbrace{\frac{\xi}{2} (R_C c - v(c))}_{\text{becoming a cash firm}} + \underbrace{\frac{\xi}{2} (R_K - v(c))}_{\text{becoming a capital firm}}. \quad (11)$$

These ODEs are Hamilton-Jacobi-Bellman (HJB) equations for cash and capital given the dynamics of the state c . We first explain the terms unrelated to ξ in each ODE. For (10), the Ito correction term $\frac{\sigma^2}{2} q''(c)$ captures the impact of the evolution of the state variable c ; a similar term shows up in (11). In addition, we have $q'(c) \sigma^2$ in (11) because the capital generates random cash flows σdZ_t which are perfectly correlated with the aggregate state $c_{t+dt} = c_t + \sigma dZ_t$ (see (8)).⁹

Multiplied by the intensity ξ , the terms describe the change in expected utility once the idiosyncratic stage arrives. The first of these terms in (10) captures that, once a firm holding a unit of cash learns to be a cash firm, its value jumps to R_C from $q(c)$. The second term says that it

⁹Heuristically, given $q(\cdot)$ as the marginal value of cash, the expected value of the cash flows σdZ_t standing at time t is $\mathbb{E}_t [q(c + \sigma dZ_t) \sigma dZ_t] = \mathbb{E}_t [q'(c) \sigma^2 (dZ_t)^2] = q'(c) \sigma^2 dt$.

uses the unit of cash to buy $1/\hat{p} = 1/c$ unit of capital, so its value jumps to R_K/c from $q(c)$. The interpretation in (11) is analogous.

Define the constant $\gamma \equiv \sqrt{2\xi}/\sigma$. The ODE system in (10)-(11) has the closed-form solution:

$$q(c) = \frac{R_C}{2} + e^{-c\gamma} A_1 + e^{c\gamma} A_2 + R_K \frac{\gamma}{2} \frac{-e^{c\gamma} \text{Ei}(-\gamma c) + e^{-c\gamma} \text{Ei}(c\gamma)}{2}, \quad (12)$$

and

$$v(c) = R_K + \frac{R_C c}{2} + e^{c\gamma} (A_3 - c A_2) - e^{-c\gamma} (A_4 + c A_1) + c R_K \frac{\gamma}{2} \frac{(e^{\gamma c} \text{Ei}(-\gamma c) - e^{-c\gamma} \text{Ei}(\gamma c))}{2}, \quad (13)$$

where $\text{Ei}(x) \equiv \int_{-\infty}^x t^{-1} e^t dt$ is the exponential integral function, and the constants A_1 - A_4 are determined from boundary conditions in (9).

Finally, we determine the endogenous investment/liquidation thresholds c_l^* and c_h^* using (7). The functions $v(c)$, $q(c)$ and the thresholds constitute an equilibrium if the resulting price $p(c) = \frac{v(c)}{q(c)}$ falls in the range of $[l, h]$ when $c \in [c_l^*, c_h^*]$. The following proposition gives sufficient conditions for such a market equilibrium to exist and describes the basic properties of this equilibrium.¹⁰

Proposition 1 *If the difference between the benefit of liquidation, l , and the cost of building capital, h , is sufficiently small, then the market equilibrium exists with the following properties:*

1. *firms do not consume before the final date;*
2. *each firm in each state $c \in [c_l^*, c_h^*]$ is indifferent to the composition of its asset holdings and $0 < c_l^* < c_h^* < R_K$;*
3. *firms do not build or dismantle capital when $c \in (c_l^*, c_h^*)$ and, in aggregate, firms spend every positive cash shock to build capital if and only if $c = c_h^*$, and they cover negative cash shocks by liquidating a sufficient fraction of capital if and only if $c = c_l^*$;*
4. *the value of holding a unit of cash and the value of holding a unit of capital are described by $v(c)$ and $q(c)$, and the price in the aggregate stage is $p(c) = v(c)/q(c)$;*
5. *in the idiosyncratic stage, a capital firm sells all its capital to cash firms for the price $\hat{p}(c) = c$;*
6. *$q(c)$ is monotonically decreasing, $v(c)$ is monotonically increasing, and $p(c)$ is monotonically increasing. Furthermore, $q(c)$ has exactly one inflection point: there is a $c_q \in (c_h^*, c_l^*)$ such that $q''(c) < 0$ for $c \in (c_l^*, c_q)$ and $q''(c) > 0$ for $c \in (c_q, c_h^*)$.*

2.7.4 Investment waves

The thick, solid curves on panels A-E of Figure 2 illustrate the properties of the market equilibrium. In panels A-C, the functions $p(c)$, $v(c)$, $q(c)$ describe the price of capital, the value of

¹⁰When $h - l$ is not sufficiently small, a variant of this equilibrium often prevails. Because this variant has very similar features, we relegate the discussion of it to Additional Material, available on the author's website.

cash, and the value of capital, respectively. Panels D-E depict the cash-to-capital ratio and the investment/disinvestment activity along one particular sample path.

The cash-to-capital ratio, c , represents the relative scarcity of liquid assets in the economy compared to illiquid capital. Thus, we refer to this ratio as “aggregate liquidity.” We also think of intervals with a large increase (drop) of capital as a boom (downturn). In our model, investment takes a simple threshold strategy, in such a way that investment (disinvestment) occurs only at c_h^* (c_l^*). However, we believe the resulting clustered investment and disinvestment activities depicted in panel E captures the essence of boom-and-bust patterns observed in reality.

The economy fluctuates across states because the aggregate cash-flow shocks drive the level of aggregate liquidity. This is illustrated in panel D. This particular sample path starts with a series of positive shocks, which increase the capital value v and decrease the cash value q . Thus, the price of capital increases along this path (not shown),¹¹ because in these states the probability that the economy will slip into a downturn (and capital must be dismantled) is low. When the price hike reaches the cost of building capital, h , investment is triggered (as shown in Panel E). This keeps the cash-to-capital ratio below c_h^* . For symmetric reasons, as a series of negative shocks decrease aggregate liquidity, rising cash values and falling capital values lead to lower capital prices. When the price of capital drops to l , disinvestment in capital is triggered. This keeps the cash-to-capital ratio above c_l^* .

Figure 3 shows our first step in mapping our model to data. Based on FED Flow of Funds data, we construct a series of aggregate liquid financial assets for non-financial US-firms, normalized by the nominal GDP, and showing NBER recessions as shaded areas. Based on the FRED database, we also plot the CD/T-bill spread as a proxy for the market value of liquidity; this spread is often used to measure the liquidity premium as CD is relatively less liquid compared to T-Bills. We also show the cyclical component of both series. These two series correspond to aggregate liquidity, c_t , and the value of a unit of liquidity, $q(c_t)$ in our model. In the data, the cyclical components of the two series are negatively correlated, with a coefficient of -0.3 .

Note that in recessions, liquid financial assets tend to be low but the value of liquidity tends to be high. Indeed, the correlation between the cyclical component of liquid financial assets and the recession dummy is -0.4 . These observations support our interpretation that recessions are associated with relatively low aggregate holdings of liquid assets and high valuations for liquidity.

As we will explain in the rest of the paper, the general pattern of investment waves, procyclical liquidity holdings, and countercyclical valuation for liquidity are a robust pattern in our economy. These features are present regardless of whether the economy is constrained efficient. It turns out that the efficiency properties of our economy are determined by whether the investment thresholds c_l^* and c_h^* are at their welfare-maximizing level. We examine this issue in the next section.

¹¹As a monotonically increasing function of c , the path of $p(c)$ looks qualitatively similar to the path of c , except that it fluctuates between h and l instead of c_h^* and c_l^* .

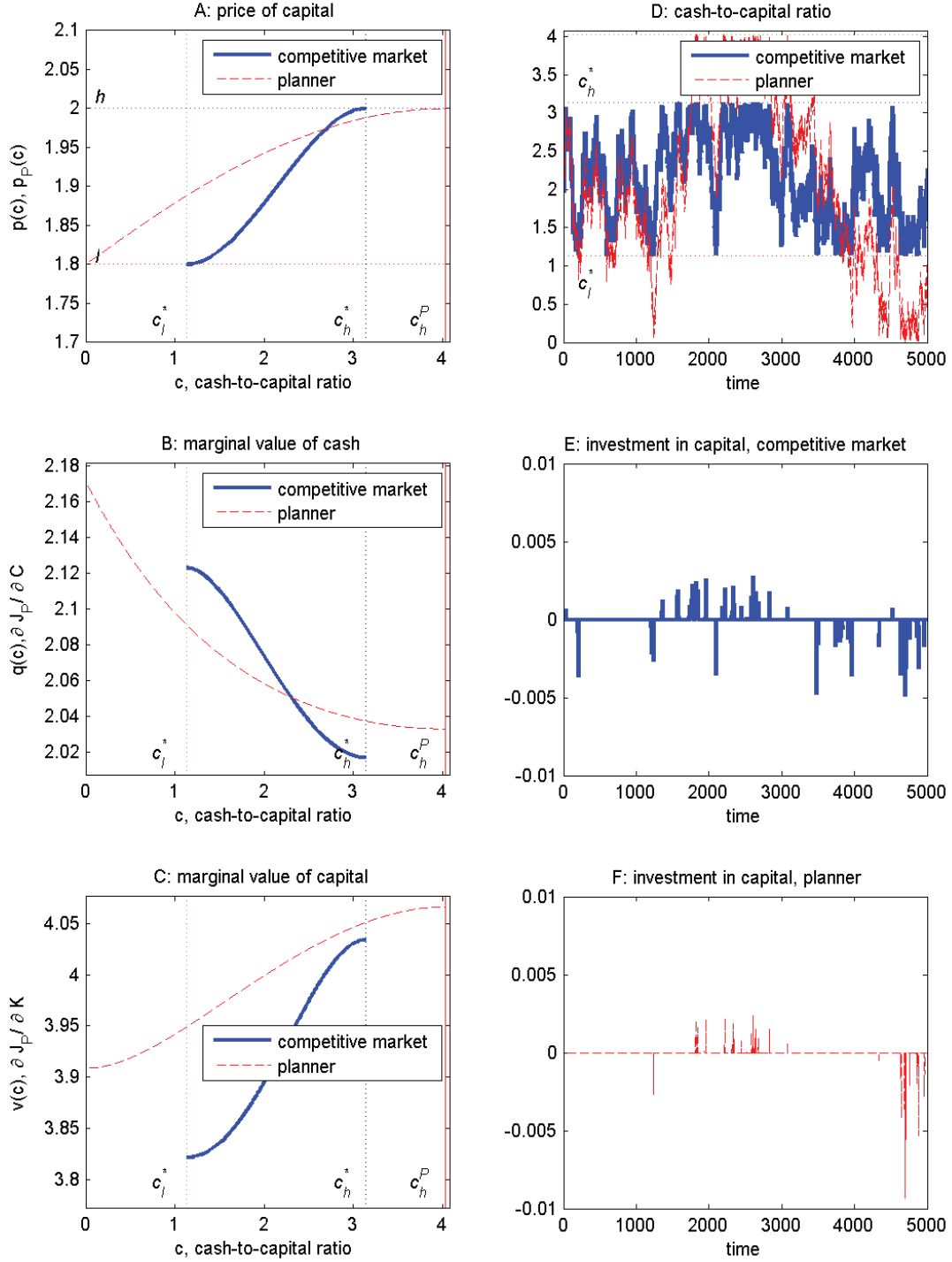


Figure 2: Panels A-C depict the price of capital, the value of cash, and the value of capital. The solid vertical line on the right of each graph is at the investment threshold in the planner's solution, $c_h^P = 4.03$, while the two dashed vertical lines are the disinvestment and investment thresholds in our baseline case, $c_l^* = 1.13$, $c_h^* = 3.14$. The horizontal lines on Panel A are at h and l . Panels D-F depict a simulated sample path. Horizontal lines on panel D from top to bottom are c_h^* , c_h^P and c_l^* . Each panel shows objects of both the baseline model with competitive market (thick solid curves) and the planner's solution (thin, dashed curves). Parameter values are $R_K = 4.2$, $R_C = 2$, $\sigma^2 = 0.6$, $\xi = 0.1$, $l = 1.8$ and $h = 2$.

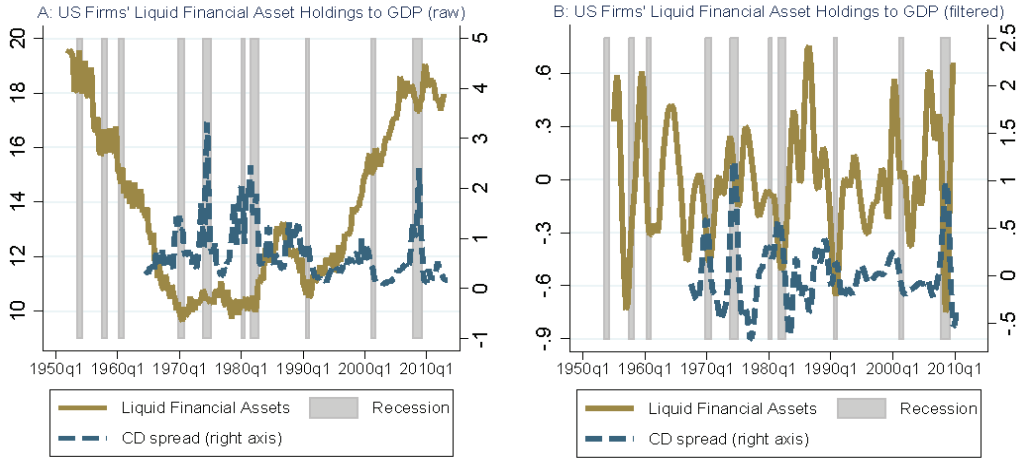


Figure 3: Quarterly aggregate liquid financial assets for non-financial US-firms normalized by nominal GDP (calculated as the sum of items included in rows 2-14 in Flow of Funds Tables L.102, with NBER recessions as shaded areas); and CD/T-bill spread as a proxy for the market value of liquidity (CD6M/TB6M series in FRED database). Panel A plots the raw series, and Panel B plots the cyclical component applying the Baxter-King filter.

3 Welfare

To study pecuniary externalities, we first solve for constrained efficient allocation in this economy as a benchmark. We then show that our model features a two-sided inefficiency on investment waves: Firms underinvest in capital during downturns and often overinvest during booms.

3.1 Constrained efficient benchmark

We study the constrained efficient allocation with the technological constraint that the aggregate cash has to be kept non-negative by liquidating capital if necessary. Without this technological constraint, condition (2) implies that the planner should convert any amount of cash to capital.

We consider a social planner who can dictate investment policies but cannot know the realization of the idiosyncratic shock. Compared to the market equilibrium, the only difference is that in the market equilibrium investment and disinvestment are driven by the market price of capital. In contrast, the social planner ignores market prices and directly decides when to build or dismantle capital. The resulting outcome corresponds to the solution of the planner's problem when he controls both investment in the aggregate stage and allocation in the idiosyncratic stage, given self-reporting (see He and Kondor (2012) for a detailed argument).

3.1.1 Social planner's problem

Denote by $J_P(K, C)$ the planner's value function which decides when to build and dismantle capital. By the end of the idiosyncratic stage, at least as long as $c_t \leq R_K$, due to linearity all cash ends up in the hands of cash firms and all capital ends up in the hands of capital firms.¹² Therefore, the

¹²This result relies on the linearity of technology and can be formally shown by the mechanism design approach (see Additional Material). Also, the conditions of Proposition 1 ensure that in the decentralized case $\hat{p}_\tau \leq c_h^* < R_K$,

total value in the idiosyncratic stage is¹³

$$KR_K + CR_C. \quad (14)$$

Thus, given the aggregate state pair (K, C) , since the final date τ arrives with exponential distribution with intensity ξ , the social planner is solving

$$J_P(K, C) = \max_{dK} \mathbb{E} \left[\int_0^\infty \xi e^{-\xi\tau} (K_\tau R_K + C_\tau R_C) d\tau \mid K_0 = K, C_0 = C \right] \equiv K j_P \left(\frac{C}{K} \right) = K j_P(c) \quad (15)$$

subject to the constraint $C_t \geq 0$ and (5). In the second equality in (15), we have invoked the scale-invariance to define $j_P(c)$ as the planner's value per unit of capital.

Because of the linear technology, regulation with reflective barriers on c is optimal (Dixit (1993)). That is, there exists lower and upper thresholds $c_l^P \geq 0$ and $c_h^P > c_l^P$, so that it is optimal to stay inactive whenever $c \in (c_l^P, c_h^P)$, and dismantle (build) just enough capital to keep $c = c_l^P$ ($c = c_h^P$).

Consider a given policy $\{c_l, c_h\}$ in which c is regulated by reflecting barriers $c_l < c_h$. Given initial state $K_0 = K$ and $C_0 = cK$, define the corresponding (scaled) social value as $j_P(c; c_l, c_h)$, so that

$$K \cdot j_P(c; c_l, c_h) \equiv \mathbb{E} \left[\int_0^\infty \xi e^{-\xi\tau} (K_\tau R_K + C_\tau R_C) d\tau \mid K_0 = K, C_0 = cK; c_l, c_h \right]. \quad (16)$$

Using standard results in regulated Brownian motions, $j_P(c)$ must satisfy

$$0 = \frac{\sigma^2}{2} j_P''(c) + \xi (R_K + R_C c - j_P(c)), \text{ for } c \in (c_l, c_h), \quad (17)$$

and at the reflective barriers c_l, c_h the smooth pasting conditions must hold:

$$\frac{\partial [K j_P(c_l; c_l, c_h)]}{\partial K} = l \frac{\partial [K j_P(c_l; c_l, c_h)]}{\partial C}, \text{ and } \frac{\partial [K j_P(c_h; c_l, c_h)]}{\partial K} = h \frac{\partial [K j_P(c_h; c_l, c_h)]}{\partial C}. \quad (18)$$

We emphasize that these conditions are not optimality conditions. They hold for any arbitrarily chosen barriers $c_l < c_h$ as a consequence of forming expectations on a regulated Brownian-motion (see Dixit (1993)). The ODE (17) has a closed-form solution

$$j_P(c; c_l, c_h) = R_K + R_C c + D_1 e^{-\gamma c} + D_2 e^{\gamma c}. \quad (19)$$

therefore capital firms are willing to use all their cash to buy capital, instead of consuming their cash. However, in the planner's solution, even for the same parameter values, it might be the case that the support of c_t is not a subset of $[0, R_K]$. Then, the planner who does not know idiosyncratic firm types cannot ensure that only cash firms are the end users of all cash. While our Propositions 2-6 are stated for the general case, we limit the discussion in the main text to the simpler case when $c_t \in [0, R_K]$ in the planner's solution. We show that the Propositions hold in the remaining cases in Additional Material by explicitly solving the planner's problem based on the mechanism design approach when $c_t > R_K$ has a positive support.

¹³ Given (K, C) , the representative cash firm gets $R_C(\hat{p}K + C) = R_C(cK + C) = 2CR_C$, while the representative capital firm gets $R_K K + R_K C / (\hat{p}) = R_K K + \frac{R_K}{c} C = 2KR_K$. As both types are equally likely the expected total welfare is $KR_K + CR_C$.

For any fixed $\{c_l, c_h\}$, we solve for the constants D_1, D_2 based on (18).

Denote by $\{c_l^P, c_h^P\}$ the social planner's optimal barrier pair. With a slight abuse of notation, we denote the planner's optimal value, $j_P(c; c_l^P, c_h^P)$, simply by $j_P(c)$:

$$j_P(c) \equiv j_P(c; c_l^P, c_h^P) = \max_{c_l, c_h} j_P(c; c_l, c_h). \quad (20)$$

Following Dumas (1991), we impose super-contact conditions to determine the optimal barrier pair. For the upper barrier c_h^P , this is

$$\left. \frac{\partial^2 [K j_P(C/K; c_l^P, c_h^P)]}{\partial K \partial C} \right|_{C=K c_h^P} = h \left. \frac{\partial^2 [K j_P(C/K; c_l^P, c_h^P)]}{(\partial C)^2} \right|_{C=K c_h^P}. \quad (21)$$

For the lower barrier c_l^P , at the optimal choice the constraint $C \geq 0$ might bind. Thus, the super-contact condition is a complementarity slackness condition¹⁴

$$\left. \frac{\partial^2 [K j_P(C/K; c_l^P, c_h^P)]}{\partial K \partial C} \right|_{C=K c_l^P} \geq l \left. \frac{\partial^2 [K j_P(C/K; c_l^P, c_h^P)]}{(\partial C)^2} \right|_{C=K c_l^P}, \text{ with equality if } c_l^P > 0 \quad (22)$$

The next proposition shows that the optimal lower threshold is $c_l^P = 0$. However, the optimal upper threshold is characterized by the unique solution to an analytical equation. We explain the intuition in Section 3.1.3.

Proposition 2 *The planner dismantles capital whenever c reaches $c_l^P = 0$ and builds capital whenever c reaches a finite, strictly positive investment threshold c_h^P . When the unique solution to the following equation*

$$\frac{R_K - hR_C}{R_K - lR_C} \left(e^{c_h^P \gamma} (1 + l\gamma) - (1 - l\gamma) e^{-c_h^P \gamma} \right) - 2\gamma (c_h^P + h) = 0 \quad (23)$$

lies in $[0, R_K]$, this solution is the socially optimal investment threshold. The optimal social value $j_P(c)$ is concave over $[0, c_h^P]$.

While the market price in the aggregate stage is undefined in an economy where the social planner sets the investment and disinvestment thresholds, we can define the shadow price of capital, $p_P(c)$, as the ratio of the planner's marginal valuation of capital, $\frac{\partial J_P(K, C)}{\partial K}$, over that of cash, $\frac{\partial J_P(K, C)}{\partial C}$, where

$$\frac{\partial J_P(K, C)}{\partial C} = j'_P(c), \quad \frac{\partial J_P(K, C)}{\partial K} = j_P(c) - c j'_P(c), \quad p_P(c) = \frac{j_P(c) - c j'_P(c)}{j'_P(c)}. \quad (24)$$

¹⁴Heuristically, we can understand the super-contact condition as follows. Converting capital to cash at a cost of l brings a marginal gain of $-J_K(c_l^P K, K) + l J_C(c_l^P K, K)$, and the social planner is considering the marginal impact of reducing c_l^P on this marginal gain, i.e., $J_{KC}(c_l^P K, K) - l J_{CC}(c_l^P K, K)$. At the optimal policy this marginal impact is zero. If the optimal policy is binding at $c_l^P = 0$, then this marginal benefit of reducing c_l^P remains strictly positive.

We plot these objects in Figure 2 along with market equilibrium counterparts.

3.1.2 Investment thresholds, welfare, and expected investment volatility

As a preparation for our welfare analysis, we show that (scaled) social welfare, $j_P(c; c_l, c_h)$, is monotonic in thresholds in the following sense: It is welfare improving to decrease the lower threshold (increase the upper threshold), whenever it is above (below) the choice of the the social planner. This is a strong global result: First, it holds for any policy pair as long as $c_l > 0$ and $c_h < c_h^P$. Second, the sign of welfare impact by changing investment thresholds is unambiguous everywhere.

Proposition 3 *For any $c_h < c_h^P$ and $c_l > 0$, we have*

$$\frac{\partial j_P(c; c_l, c_h)}{\partial c_l} < 0, \text{ and } \frac{\partial j_P(c; c_l, c_h)}{\partial c_h} > 0 \text{ for all } c \in [c_l, c_h].$$

It is also useful to define a measure of the volatility of our investment waves. For this purpose, we define the expected total adjustment of capital, parameterized by the thresholds c_l, c_h :

$$T(c; c_l, c_h) \equiv \mathbb{E} \left[\int_0^\tau \frac{|dK_t|}{K_t} \right]. \quad (25)$$

Proposition 4 *For any c_h and c_l , we have*

$$\frac{\partial T(c; c_l, c_h)}{\partial c_l} > 0, \text{ and } \frac{\partial T(c; c_l, c_h)}{\partial c_h} < 0.$$

This proposition states that the expected investment volatility increases in the disinvestment threshold, c_l , and decreases in the investment threshold, c_h . Thus, if in the market equilibrium $c_h^* < c_h^P$ and $c_l^* > 0$, then the economy exhibits more volatile investment compared with that in the constrained efficient benchmark.

3.1.3 Investment thresholds in market equilibrium and in the planner's solution: intuition and comparative statics

As the welfare properties of our economy can be traced back to the investment thresholds, it is useful to understand the economic forces that determine them. As we have established in Propositions 1 and 2, the disinvestment threshold in the market equilibrium, $c_l^* > 0$, is strictly positive, whereas the planner disinvests only when it is unavoidable, $c_l^P = 0$. In the next proposition, we state further results and then proceed to the intuition.

Proposition 5 *The following results hold.*

1. *The solution of equation (23) determining the planner's investment threshold, c_h^P*
 - (a) *is converging to 0 as $\gamma \rightarrow \infty$, and decreasing in γ given that $\gamma > \hat{\gamma}$ for a given $\hat{\gamma}$,*

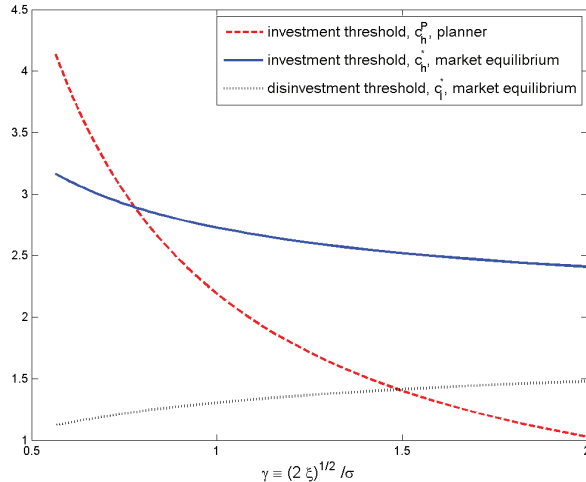


Figure 4: Investment and disinvestment thresholds for the planner ($c_l^P = 0$, c_h^P in dashed line) and for the market (c_l^* in dotted line, c_h^* in solid line). Parameters are $R_K = 4.2$, $R_C = 2$, $\sigma^2 = 0.6$, $\xi = 0.1$, $l = 1.8$ and $h = 2$.

(b) is decreasing in l and R_K , and increasing in h and R_C .

2. In contrast, in the market equilibrium determined in Proposition 1, we have

(a) $c_h^* > h$ and $c_l^* < l$,

(b) $c_h^* \rightarrow h$ and $c_l^* \rightarrow l$ as $\gamma \rightarrow \infty$.

The planner starts disinvesting only when he is forced to, i.e., $c_l^P = 0$. Intuitively, the planner does not want to dismantle capital as long as he has not run out of cash yet. A positive lower threshold would imply that a part of the cash buffer is never used for maintenance. Because capital is more productive than cash, that would be a waste. His choice of the investment threshold $c_h^P > 0$ is driven by a simple trade-off. While capital is more productive than cash, a cash buffer is useful to avoid the inefficient liquidation of capital in the case of a series of adverse cash-flow shocks.

Consider the role of the constant $\gamma = \sqrt{2\xi}/\sigma$. This parameter enters (23), which characterizes the constrained efficient solution, as well as the functions $q(c)$, $v(c)$ in (12) and (13), which characterize the market equilibrium. Figure 4 plots the planner's investment threshold c_h^P (dashed) as a function of γ .

Intuitively, γ measures the relative importance of aggregate cash-flow shocks to idiosyncratic investment opportunities. When γ is large, aggregate shocks are less important, either because their volatility is low, or because the idiosyncratic shock arrives with high intensity. Regardless of the particular reason, a larger γ implies that the planner puts less weight on the possibility that a sequence of negative cash-flow shocks force him to dismantle capital at the lower threshold. In fact, as Proposition 5 states, as γ increases without bound, c_h^P converges to zero as the planner decides not to store any cash (i.e., he will immediately convert any cash to capital) given that capital is relatively more productive $R_K > hR_C$. Figure 4 illustrates that the smaller the γ (say, the larger

the cash-flow volatility σ), the more weight the planner puts on the possibility of forced liquidation, and the larger cash buffer the planner wants to keep.

Figure 4 also plots the investment thresholds c_h^* (solid) and the disinvestment threshold c_l^* (dotted) for the market equilibrium. In the market solution, the same trade-off is present, which is behind the fact that c_h^* is decreasing in γ (just as c_h^P does). However, there is an additional force: The firm in the market equilibrium knows that in the idiosyncratic stage the price of capital will be $\hat{p}_\tau = c_\tau$. A c_t close to 0 implies that holding on to a bit of cash is a good idea because a small amount of cash can be exchanged for a large amount of capital in case the economy enters the idiosyncratic stage. (From the social perspective, the losses and gains from trade in the idiosyncratic stage are a wash.) Hence, the firm liquidates capital well before negative cash-flow shocks deplete all the capital stock, implying that c_l^* is bounded away from 0. That is, the reason to liquidate capital before all cash is depleted is to turn this unit of capital to l units of cash in the aggregate stage, instead of $\hat{p}_\tau = c_\tau$ units of cash in the idiosyncratic stage. Clearly, this logic makes sense only if $\hat{p}_\tau < l$, implying that c_l^* must be smaller than l . Symmetric argument implies that c_h^* must be above h . In fact, we show that in the limit $\gamma \rightarrow \infty$ so that aggregate shocks are unimportant, $c_l^* = l$ and $c_h^* = h$. As firms understand that the price of capital in the idiosyncratic stage is $\hat{p}_\tau = c_\tau$, when only that stage matters, they decide to (dis)invest exactly when that price reaches the cost of (dis)investing.

Turning to the other parameters, the higher l and R_K , and the lower h and R_C (i.e., the lower the adjustment cost and the higher the relative benefit of capital to cash), the less the cash buffer that the planner is willing to build up. This reduces the upper threshold c_h^P , as stated in the second result in Proposition 5.

These results immediately imply that the disinvestment threshold is too high in the market equilibrium. Proposition 5 (or see Figure 4) suggests that the investment threshold c_h^* in the market equilibrium can be either higher or lower than c_h^P in the planner's solution, depending on the parameters. That is, our economy might feature underinvestment always, or underinvestment during recessions but overinvestment during booms. In the next subsection, we identify the subset of parameters for the latter case, call it a two-sided inefficiency, and further explore the underlying mechanism.

3.2 Two-sided inefficiency

The following proposition states the main result of our paper.

Proposition 6 *Under the conditions of Proposition 1, the following statements hold:*

1. *Firms dismantle capital before the cash-to-capital ratio reaches zero, i.e., $c_l^* > 0$. Hence the market equilibrium implies underinvestment in capital and over hoarding of cash in recessions.*
2. *If the difference between the productivity of capital and that of the new investment opportunity, $R_K/h - R_C$ is sufficiently small, then we have $c_h^* < c_h^P$. That is to say, the market equilibrium implies overinvestment in capital during booms.*

Figure 2 illustrates a case of two-sided inefficiency. The thin, dashed curves on panels A-D of Figure 2 illustrate the properties of the solution of the planner's problem. Panels B and C show the planner's marginal valuation of cash and capital in the aggregate stage, while panel A shows the ratio of the two, which is the shadow price of capital as defined in (24). The dashed (solid) vertical lines show the thresholds of the market equilibrium (planner's problem). As explained, in the market equilibrium firms dismantle capital when some cash is still around, $c_l^* > 0$. In this example, firms create new capital at a lower liquidity level than the social planner would, $c_h^* < c_h^P$. Panel D contrasts the resulting evolution of cash-to-capital ratio in the planner's solution and in the market equilibrium under the same sample path of shocks $\{dZ_t\}$. Proposition 4 implies that, in the case of two-sided inefficiency, the resulting investment waves are too volatile in the market equilibrium (illustrated by Panels E and F) in the sense that the expected adjustment intensity of capital is too high.¹⁵ Finally, Proposition 3 implies that any policy that raises (decreases) the upper investment (lower disinvestment) threshold would unambiguously increase total welfare in this case.

The reason for the difference between the planner and the market is a wedge between the private and social valuation of capital relative to cash. To see this, consider firms' marginal rate of substitution (MRS) between capital and cash in the idiosyncratic stage. As the value of capital is R_K and $\hat{p}_\tau R_C$ and the value of cash is $\frac{R_K}{\hat{p}_\tau}$ and R_C for capital firms and cash firms respectively, the MRS is

$$MRS \equiv \frac{\frac{1}{2}(\hat{p}_\tau R_C + R_K)}{\frac{1}{2}\left(R_C + \frac{R_K}{\hat{p}_\tau}\right)} = \hat{p}_\tau. \quad (26)$$

Intuitively, when capital price, \hat{p}_τ , is higher, firms value more the capital they own, because cash firms can sell their capital and capital firms must buy the capital they lack at that higher price. In contrast, the relative social value of capital to cash is always $\frac{R_K}{R_C}$. From the social perspective, the main function of the idiosyncratic stage is that it allocates cash and capital to the highest-value user. The corresponding transfer across the two type of firms in the market equilibrium, pinned down by \hat{p}_τ , is immaterial for the planner!

The wedge between the social and private valuation in the aggregate stage naturally follows from the wedge in the idiosyncratic stage. Because the price in the decentralized economy guides each individual firm's investment decisions, it is this wedge that drives the inefficiency of the investment waves in the aggregate stage. What is more, in our model the valuation wedge fluctuates with the aggregate liquidity state. When γ is finite, because firms are worried about cash-flow shocks, the aggregate liquidity holding of firms is procyclical. Our contracting frictions imply that \hat{p}_τ positively depends on aggregate liquidity holdings. Therefore, from (26), the private incentive to hold capital instead of cash decreases during recessions with low prices and increases during booms with higher prices, compared with its social counterpart.

¹⁵When comparing Panels E and F, recall Proposition 4 and the definition of $T(c; c_h, c_l)$. This excess volatility is in terms of the expected total adjustment intensity. Note that we cannot say whether, conditional on investment, the size of the adjustment is larger or smaller in the market equilibrium than in the planner's solution.

Figure 2 shows our mechanism in action. In Panel A, the solid line shows the market value of capital relative to cash in the aggregate stage, and the dashed line shows its social counterpart. The difference between them comes from our wedge.

Although underinvestment in recession is independent of the parameters, whether there is over- or underinvestment during booms depends on the parameter values. For example, consider again the case $\gamma \rightarrow \infty$. Recall that $R_K > hR_C$ implies $c_h^P = 0$ in this limit, while, as \hat{p}_τ drives investment and disinvestment in the decentralized case, $c_h^* = h$ and $c_l^* = l$. That is, we have underinvestment both during booms and recessions. To generate overinvestment during booms, we need to make capital less attractive relative to cash for the planner.

As Proposition 5 describes, starting from a very large γ , decreasing γ and/or increasing R_C does exactly that. In fact, as Proposition 6 shows, by decreasing $R_K - hR_C$, we can raise c_h^P from zero to any positive level within $[0, R_K]$. While a smaller $R_K - hR_C$ makes cash more attractive in both the market and planner's solution, its effect on the market solution is much smaller because of the additional private incentive force we described above. Loosely speaking, this additional force keeps c_h^* and c_l^* close to h and l in the market equilibrium. Therefore, as Proposition 6 states, when $R_K - hR_C$ is smaller than a given threshold, $c_h^P > c_h^*$ i.e., we have overinvestment during booms.

We conclude this analysis with three remarks.

Remark 1 *We can push the foregoing point further. So far, our analysis is performed under the parameter restriction of $\frac{R_K}{R_C} > h$. What if $l < \frac{R_K}{R_C} < h$, which says that capital is more attractive given the relatively small liquidation benefit and that cash is more attractive given the relatively large capital building cost? In Additional Material we show that, in this case, we always have a two-sided inefficiency. The intuition of the limiting case $\gamma \rightarrow \infty$, which corresponds to a two-period static model, is rather simple. Given the relatively high adjustment cost, the planner would never want to convert capital to cash or vice versa, implying that $c_l^P = 0$ and $c_h^P = \infty$. But in the market solution with $c_l^* = l$, $c_h^* = h$ remains the same as in our base setting, as firms still make their (dis)investment decisions in response to the market price $\hat{p}_\tau = c_\tau$. It follows that $0 = c_l^P < c_l^* < c_h^* < c_h^P = \infty$, hence the two-sided inefficiency.*

Remark 2 *We emphasize that our market inefficiency result comes from non-contractible idiosyncratic investment opportunities. Without contracting frictions, say if R_K and R_C were pledgeable, $\hat{p}_\tau = \frac{R_K}{R_C}$ would always hold and there would be no wedge between the private and social relative value of capital to cash. Just as in our baseline model, in the absence of contracting frictions firms still build up cash buffers against negative cash-flow shocks, because the precautionary motive to hold cash is still present. However, in this variant the investment and disinvestment thresholds are the same in both the planner's solution and the market equilibrium (for a formal proof, see Additional Material). That is, the precautionary motive alone does not create inefficiency. Nevertheless, in our model this precautionary motive to hold cash interacts with the externality.*

Remark 3 *Our mechanism is closely related to the main intuition behind the seminal papers on welfare affecting pecuniary externalities of Stiglitz (1982), Geanakoplos and Polemarchakis (1985),*

and Greenwald and Stiglitz (1986), which are followed by the more recent work of Shleifer and Vishny (1992), Allen and Gale (1994, 2004, 2005), Caballero and Krishnamurthy (2001, 2003), Lorenzoni (2008), Farhi, Golosov and Tsyvinski (2009), Farhi and Tirole (2012) and Gale and Yorulmazer (2011). The critical observation in these papers is that, because of frictions, agents' marginal utility of wealth might not be equalized across time or states in the decentralized equilibrium. In this case, a price change can work as a transfer from low marginal utility states to high marginal utility states, creating *ex ante* welfare improvement. Indeed, there is a parallel argument in our model, as the marginal utility of wealth in the idiosyncratic stage is $\frac{R_K}{\hat{p}_\tau}$ for capital firms and R_C for cash firms. Whenever $\hat{p}_\tau < \frac{R_K}{R_C}$, then the marginal value of wealth is higher for a capital firm. Therefore, if an intervention were to push down the threshold c_i^* to $c_i^* - \varepsilon$ at that state, the delayed disinvestment would lower the capital price at the idiosyncratic stage \hat{p}_τ , which would be a transfer from the cash firms (sellers of capital) to capital firms (buyers of capital), increasing *ex ante* utility. However, note that, unlike in many other models in this literature, in our case it is not the transfer *per se* that is the source of the welfare effect, but the more efficient investment in the aggregate stage.¹⁶

4 Applications

In the first part of this section, as a main application, we discuss the role and limitations of economic policies in our context. In the second part we offer further applications connecting our findings to sectoral investment cycles and financial development.

4.1 Economic policies

Proposition 3 shows how social welfare in our economy changes as the investment and disinvestment thresholds change. However, in a market economy the policymaker cannot set these thresholds directly. Instead, the policymaker might be able to influence the investment/disinvestment threshold by changing the relative incentives of holding cash and capital, that is, by affecting the market price. In this section, we are interested in how various types of economic policies can serve this purpose. We first make the following definition:

Definition 2 *A balanced (budget-neutral) policy is an intervention that changes the marginal value of capital only by the transfer scheme $\pi(c) \geq 0$, such that, given c , $c\pi(c)$ is the effective transfer for each unit of capital held, and $-\pi(c)$ is the effective transfer for each unit of cash held. An intervention equilibrium is a market equilibrium where the wealth dynamics in (4) is adjusted by transfers.*

¹⁶To see this, fix the aggregate stage investment but consider an unexpected intervention in the beginning of the idiosyncratic stage. This unexpected intervention transfers ε cash from cash firms to capital firms and then allows them to trade, produce, and consume. As cash firms would still exchange all their capital for cash and vice versa, the *ex post* allocation of the given (K, C) across firms would remain the same, hence this intervention would not affect *ex ante* welfare. A similar intervention in the bad state of the interim period of Lorenzoni (2008), i.e., transferring cash from consumers to firms, would change the amount of disinvestment, hence affecting welfare.

We refer to the equilibrium objects in an intervention equilibrium by the superscript π . In an intervention equilibrium, the policymaker influences the outcome only through the effect of $\pi(c)$ on the price in the aggregate stage.

The family of balanced policies is rich, because $\pi(c)$ might be defined and implemented in various ways. The simplest case is to impose a particular transfer between cash holders and capital holders. But the policymaker, to avoid inefficient liquidation, might also target a certain price path, $p^\pi(c)$, which will implicitly define $\pi(c)$. If $\pi(c)$ is positive in some range of c , the policymaker might implement $\pi(c)$ by buying a fraction of capital above market prices and selling it back to the market at some point. That is, in our abstract world, it is immaterial to the welfare effect whether policymakers choose to provide subsidies or bailouts to certain industries, to impose measures which affect the collateral value of assets, or to implement asset purchase programs, as long as the implied marginal transfers $\pi(c)$ in these programs are the same.

Note that by the argument derived in Section 3, the (scaled) value of the representative firm in an intervention equilibrium is still $j_P(c; c_l^\pi, c_h^\pi)$ as defined in (20), where c_l^π, c_h^π are the implied investment/disinvestment thresholds. Therefore, Proposition 3 continues to hold: the welfare effect of a policy can be traced back to its effect on the thresholds.

In the rest of this section, we analyze interventions concentrated on certain stages of our investment waves. Since the 2008 financial crisis, there has been an ongoing debate on the potential adverse effects of interventions during recessions on incentives during booms and, relatedly, on the optimal mix of ex ante regulation and ex post intervention (e.g., Diamond and Rajan, 2011; Farhi and Tirole, 2012; Jeanne and Korinek, 2013). Our modelling approach emphasizes that a policy that, say, makes capital more attractive in a recession, affects the relative value of capital in every other state. What is more, the effect of that policy on the investment threshold in booms feeds back to agents' welfare in recessions, too. As we demonstrate, when a two-sided externality is present, this interaction adds an interesting layer to this discussion.

We start our analysis with the following definition.

Definition 3 *A balanced policy is concentrated on low (high) states if $\pi(c) = 0$ for any $c > c_0$ ($c < c_0$).*

The next proposition specifies a simple criterion to decide how such a concentrated policy affects welfare.

Proposition 7 *The following statements hold.*

1. *An intervention concentrated on low states to decrease the disinvestment threshold $c_l^\pi < c_l^*$, also reduces the investment threshold $c_h^\pi < c_h^*$, if $p^\pi(c_0) > p(c_0)$ and $q^\pi(c_0) \leq q(c_0)$. It increases the investment threshold $c_h^\pi > c_h^*$, if $p^\pi(c_0) < p(c_0)$ and $q^\pi(c_0) \geq q(c_0)$.*
2. *An intervention concentrated on high states to increase the investment threshold $c_h^\pi > c_h^*$, also increases the disinvestment threshold $c_l^\pi > c_l^*$, if $p^\pi(c_0) < p(c_0)$ and $q^\pi(c_0) \geq q(c_0)$. It decreases the disinvestment threshold $c_l^\pi < c_l^*$, if $p^\pi(c_0) > p(c_0)$ and $q^\pi(c_0) \leq q(c_0)$.*

Proposition 7 states that, to understand a policy’s welfare consequences, it is sufficient to check the effect of the policy at the single state c_0 , where the intervention stops. Together with Proposition 3 it also provides clear guidelines to the policymaker. For example, suppose that the economy features two-sided inefficiency. The policymaker might want to avoid inefficient liquidation by implementing a policy that increases the price of capital and decreases the value of cash in recessions. As long as the policy has the same effect at c_0 , it unambiguously worsens the overinvestment problem during the boom. However, if the policymaker manages to find an alternative that partially avoids inefficient liquidation and decreases the price of capital (without decreasing the value of cash) at c_0 at the same time, then the intervention improves welfare everywhere.

To illustrate the usefulness of these guidelines, consider a particular family of policies. There, the policymaker chooses a price floor for capital $l + \delta$ with $\delta \geq 0$ together with a lower disinvestment threshold c_l^π with $c_l^\pi < c_l^*$, and designs a policy that does not allow the price to fall below $l + \delta$ as long as $c \geq c_l^\pi$. With this intervention, the policymaker ensures that capital is liquidated only at c_l^π . As we show in Online Appendix B.4, the choice of δ and c_l^π endogenously determines the corresponding transfer scheme $\pi(c)$ and the intervention threshold c_0 . The policymaker can implement the transfer $\pi(c)$ as a direct subsidy to capital holders, or, for example, initiates a tax-financed program of buying assets at a markup ψ above the market price $p^\pi(c)$ with some state-dependent intensity $\chi(c)$, where¹⁷

$$\pi(c) = \chi(c) ((p^\pi(c) + \psi) q^\pi(c) - v^\pi(c)).$$

The dashed and dotted curves in Figure 5 show the main equilibrium objects of the intervention equilibrium for a $\delta = 0$ and a $\delta > 0$ price-floor policy with the same liquidation thresholds c_l^π . The intervention with $\delta = 0$ slightly decreases the price and increases the upper investment threshold c_h^π , thereby increasing the welfare everywhere. In contrast, the intervention with $\delta > 0$ has the opposite effect. The following proposition states the general result for the $\delta = 0$ case when γ is large.

Proposition 8 *A price-floor policy with $\delta = 0$ and any $c_l^\pi < c_l^*$ improves welfare at every state $c \in [c_l^*, c_h^*]$ as long as γ is sufficiently large.*

The intuition behind the opposite effect of the two interventions is instructive. The high price floor close to the recession increases the value of capital during booms, encouraging more overinvestment. While there is less inefficient liquidation during the recession due to the support for capital holders, firms—even during the recession—foresee the resulting stronger overinvestment during booms. In the numerical example in Figure 5 with $\delta > 0$, the latter effect dominates the earlier and decreases welfare.

¹⁷The policymaker can resell the purchased assets to firms at the market price immediately, or after holding them for a given time interval. The latter case might be a reasonable description of the TARP program implemented by the US government in 2008.

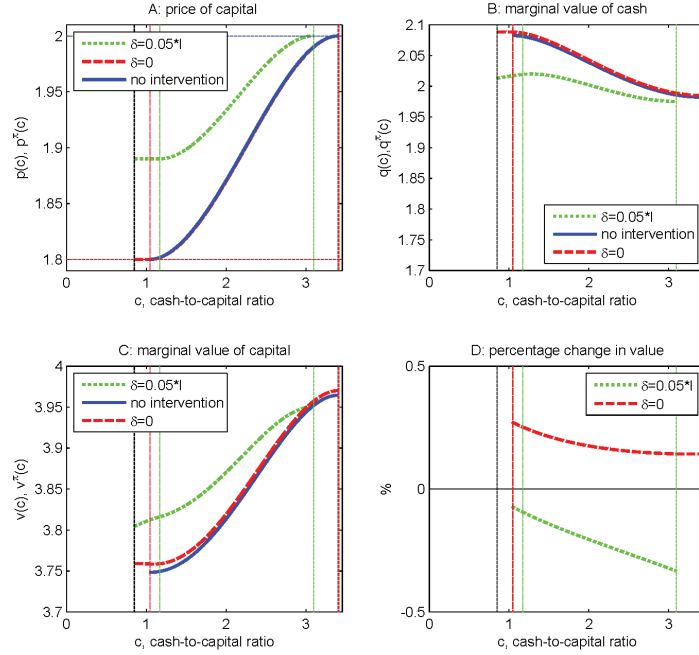


Figure 5: Panels A-C depict the price of capital, the value of cash, and the value of capital in the baseline case (solid), under price floor policies $\delta = 0.05l$ (dotted) and $\delta = 0$ (dashed). (In Panel A, the solid and dashed curves are on top of each other.) Panel D depicts the percentage change in the social value due to the intervention. The vertical lines in each panel from left to right are the liquidation threshold c_l^π of both interventions, the intervention thresholds c_0 of $\delta = 0$ and of $\delta = 0.05l$, and the investment thresholds of $\delta = 0.05l$, the baseline case, and $\delta = 0$. The horizontal lines in Panel A are at the levels of l and h . Parameter values are $R_K = 4.1$, $\sigma^2 = 1$, $\xi = 0.1$, $R_C = 2$, $l = 1.8$, $h = 2$, and $c_l^\pi = 0.85$.

In contrast, when the price floor is sufficiently low, the policymaker prolongs the near-recession state of the economy by keeping the price very close to its minimal value through states of mild recovery. This can decrease the relative value of capital when the economy is booming. Thus, even if the policymaker manages to avoid some inefficient liquidation, the intervention can still decrease the value of capital, which makes the overinvestment problem less severe.

One take-away from this experiment is that setting the appropriate price is critical. If the set price for the recession is not sufficiently low, economic agents foresee the induced overinvestment in booms, thus decreasing welfare during both booms and recessions. If the set price is too low, then it does not stop inefficient liquidation. Hence, the policymaker should set a price floor that just discourages firms from selling the assets for lower-user value agents. This policy endogenously keeps the price of capital low through mild states of recovery, which helps curb incentives for overinvestment during booms.

4.2 Sectoral and aggregate investment waves

In this section we flesh out two further applications that relate our model to sectoral investment waves and financial development. We keep the discussion here brief, and expand on these applications and on the related literature in He and Kondor (2012).

Sectoral investment waves It is well known that certain industries go through boom-and-bust patterns. Hoberg and Phillips (2010) argue that these patterns are widespread in the data, well beyond the handful of well-known episodes such as the 1990s tech-bubble and the 1980s biotech bubble. Interpreting our economy as one sector, our model implies that such cycles arise naturally, even if the technology does not change. The main implication of our mechanism is that in sectors with more non-contractible investment opportunities (e.g., sectors with a larger share of intangible input), other things being equal, these cycles are less efficient. That is, in these sectors too many resources would be spent on frequent adjustment of the capital level, reducing profitability.

Financial development and investment dynamics Our model also suggests a novel rationale for stylized facts on the connection of financial development and investment dynamics. Aghion et al. (2010) provide a useful starting point. The authors decompose aggregate investment to structural and other investment, arguing that structural investment is a proxy for investment in longer-term, riskier, but more productive projects. Then they show that in less financially developed countries structural investment is much more sensitive to productivity shocks, implying a more volatile and more procyclical pattern. They suggest that this difference in the dynamics of the composition of investment activities is an important way in which the lack of financial development hinders growth.

Our results are broadly consistent with the stylized facts in Aghion et al. (2010), if we take capital as a proxy for more productive and riskier projects, and the lack of contractibility on future investment opportunities as a proxy for a low level of financial development. Our two-sided inefficiency implies more volatile investment in capital (Proposition 4), a lower level of expected consumption (Proposition 3), and a lower growth rate of the economy in the long term for less financially developed countries.

5 Robustness

In this section, we present variants of our baseline model to show that the presence of a two-sided inefficiency is not linked to particular technical features of our model. First, we present a variant where the aggregate and idiosyncratic stages are not separated. We show that, due to a similar intuition in the baseline model, two-sided externalities are present. Second, we show that allowing for collateralized borrowing could make the presence of two-sided externalities even more prevalent.

5.1 Contemporaneous aggregate and idiosyncratic shocks and a single market

We argue in Section 2.2 that the separation of the market into two segments, that in which firms in the aggregate stage trade and that in which firms in the idiosyncratic stage trade, is only a technical device. In this part, we build a variant in which we eliminate this segmentation of markets.

Just as in the alternative interpretation of our baseline case in Section 2.4, here we think of firms facing i.i.d. chances of being hit by idiosyncratic investment opportunities occurring with

intensity ξ . Applying the law of large numbers, every instant there are a ξdt fraction of firms hit by the idiosyncratic skill shocks. These firms face the same investment opportunity sets as in our baseline model: with half probability each firm becomes either a capital or a cash firm. Again, as in our baseline, before final production firms can trade their holdings. However, in this variant, both firms with the investment opportunity and those without it trade at the same Walrasian capital price p_t . That is, there is no separate price \hat{p}_τ for the idiosyncratic stage. Firms with investment opportunity, after exploiting it, exit and consume, but the economy goes on forever with the remaining firms who have not got any investment opportunity yet.

As the baseline model, incumbent firms face aggregate cash-flow shocks according to (1), and they can transfer cash to capital or vice versa at the same linear investment technology as in our base model. We further assume that in the aggregate stage there is another sector that combines capital and cash to produce perishable final consumption goods at the Cobb-Douglas technology $\phi K^\alpha C^{1-\alpha}$, where $\alpha \in (0, 1)$ and $\phi > 0$ are positive constants. As we will discuss shortly, the Cobb-Douglas technology with Inada condition is only for ease of illustrating the welfare effect of small policy interventions.

We can solve the model by keeping track of the same state variable, cash-to-capital ratio $c_t = C_t/K_t$. There will be an upper (lower) boundary c_h^* (c_l^*) so that firms start investing in (liquidating) capital when c_t hits the boundary from below (above). Inside the inaction region $c_t \in (c_l^*, c_h^*)$, given the endogenous capital price $p(c_t)$, the cash-to-capital ratio follows

$$dc_t = \underbrace{-\frac{\xi}{2}(p(c_t) + c_t) dt}_{\text{cash firms exiting w. cash}} + \underbrace{\frac{\xi c_t}{2} \left(1 + \frac{c_t}{p(c_t)}\right) dt}_{\text{capital firms exiting w. capital}} + \sigma dZ_t = \frac{\xi}{2} \left(-p(c_t) + \frac{c_t^2}{p(c_t)}\right) dt + \sigma dZ_t. \quad (27)$$

Here, we have labeled the extra drift terms relative to (6). For example, $\frac{\xi}{2} dt$ fraction of capital firms causes an outflow of $\frac{\xi}{2} \left(1 + \frac{c_t}{p(c_t)}\right) dt$ on the scaled (by capital K_t) aggregate capital, which translates to a positive drift of $\frac{\xi c_t}{2} \left(1 + \frac{c_t}{p(c_t)}\right) dt$ for the cash-to-capital ratio.

Although the closed-form solution is no longer available once the drift of the state variable depends on the endogenous capital price $p(c)$, we can study the market equilibrium by numerically solving a system of ODEs. For the marginal value of cash $q(c)$, we have

$$0 = \underbrace{q'(c) \frac{\xi}{2} \left(-p(c) + \frac{c^2}{p(c)}\right)}_{\text{effect of drift}} + \underbrace{\frac{\sigma^2}{2} q''(c) + \xi \left(\frac{1}{2} \left(R_C + \frac{R_K}{p(c)}\right) - q(c)\right)}_{\text{idiosyncratic stage}} + \underbrace{\phi(1-\alpha)c^{-\alpha}}_{\text{aggregate stage}}. \quad (28)$$

We highlight three terms in (28). The first term captures the drift of the state variable c_t as firms are exiting the economy. The second term gives the marginal value of cash when hit by idiosyncratic shocks: Each unit of cash either yields R_C if the firm becomes cash type, or $R_K/p(c)$ if capital type, each with half probabilities.

The third term, $\phi(1-\alpha)c^{-\alpha} = \partial[\phi K^\alpha C^{1-\alpha}]/\partial C$, which is new, gives the marginal value of cash at the aggregate stage before being hit by idiosyncratic investment opportunities. Due to

Inada condition of the Cobb-Douglas technology, this marginal benefit $\phi(1-\alpha)c^{-\alpha}$ soars when c falls towards zero, guaranteeing that in equilibrium firms start liquidating capital for cash before c hits 0. As a result, the equilibrium disinvestment threshold $c_l^* > 0$ always takes an interior solution, which helps us illustrate the effect of a small distortionary tax scheme that we consider later. For details, see Additional Materials.

Similarly the marginal value of cash $v(c)$ satisfies

$$0 = \underbrace{q'(c)\sigma^2 + v'(c)\frac{\xi}{2}\left(-p(c) + \frac{c^2}{p(c)}\right)}_{\text{effect of drift}} + \underbrace{\frac{\sigma^2}{2}v''(c) + \xi\left(\frac{R_{CP}(c) + R_K}{2} - v(c)\right)}_{\text{idiosyncratic stage}} + \underbrace{\phi\alpha c^{1-\alpha}}_{\text{aggregate stage}}. \quad (29)$$

Finally, to pin down the market equilibrium, we have the same boundary conditions as in the base model

$$v'(c_h^*) = q'(c_h^*) = q'(c_l^*) = v'(c_l^*) = 0, \quad (30)$$

and (dis)investment optimality conditions are

$$\frac{v(c_h^*)}{q(c_h^*)} = h, \quad \text{and} \quad \frac{v(c_l^*)}{q(c_l^*)} = l. \quad (31)$$

For the planner's solution, we consider a policy from the family of balanced tax/subsidy considered in Section 4.1:

$$c\pi(c) = \begin{cases} -\pi_h & \text{if } c > (1 - \kappa_h)c_h^* \\ \pi_l & \text{if } c < (1 + \kappa_l)c_l^* \\ 0 & \text{otherwise} \end{cases} \quad (32)$$

where π_h , π_l , κ_h and κ_l are positive constants. This policy taxes capital during booms (specifically, when the economy is close to investment boundary c_h^* , or, a κ_h fraction below c_h^*) and/or subsidizes capital during recessions (specifically, when the economy is close to disinvestment boundary c_l^* , or, a κ_l fraction above c_l^*). Budget-neutral policies imply that cash receives transfers of $\pi(c)$. We obtain the new market equilibrium with intervention by numerically solving the new investment/disinvestment thresholds (c_l^g, c_h^g) , joint with new marginal capital and cash values $v^g(\cdot)$ and $q^g(\cdot)$, respectively.¹⁸ As before, the (scaled) social welfare in the economy is measured as $j^g(c) = v^g(c) + cq^g(c)$.

5.1.1 Two-sided inefficiency

As an illustration, the solid curves in Panels A, B, and C on Figure 6 show the price of capital, the value of capital, and the value of cash, all as a function of c in the alternative setting. We observe that, in the alternative setting without segmented markets, the equilibrium objects are qualitatively similar to those in the base model. The dashed curves on the same panels show the

¹⁸We have the same boundary conditions $p(c_l^g) = l$ and $p(c_h^g) = h$, but modify the ODEs slightly to reflect transfers: For capital, we need to add $c\pi(c)$ in (28); for cash we need to subtract $\pi(c)$ in (29).

corresponding objects under the policy intervention specified in (32). We observe that $c_l^g < c_l^*$ and $c_h^g > c_h^*$. The intuition is clear: As individual firms adjust their real decisions based on market prices, taxing (subsidizing) capital during booms (recessions) postpones the market investment (disinvestment) during that time. Because the policy intervention is small (we set $\pi_h = \pi_l = 1.5\%$ and $\kappa_h = \kappa_l = 7\%$ in this numerical example), the quantitative effect of this policy in Panels A, B, and C is only slightly visible.

Panel D shows the resulting improvement in social surplus for this policy (“two-sided intervention,” solid, U-shaped curve). To show the two-sided inefficiency more convincingly, we also calculate the change in social welfare under two different one-sided intervention policies, i.e., either only taxing capital during booms (“upper intervention” sets $\pi_l = 0$, dotted, increasing curve) or only subsidizing capital during recessions (“lower intervention” sets $\pi_h = 0$, dashed, decreasing curve). Each policy generates a strictly positive value improvement, and the curve of the two-sided intervention is the upper envelope of the one-sided interventions. These results imply the two-sided externality with underinvestment during recessions and overinvestment during booms in this economy.

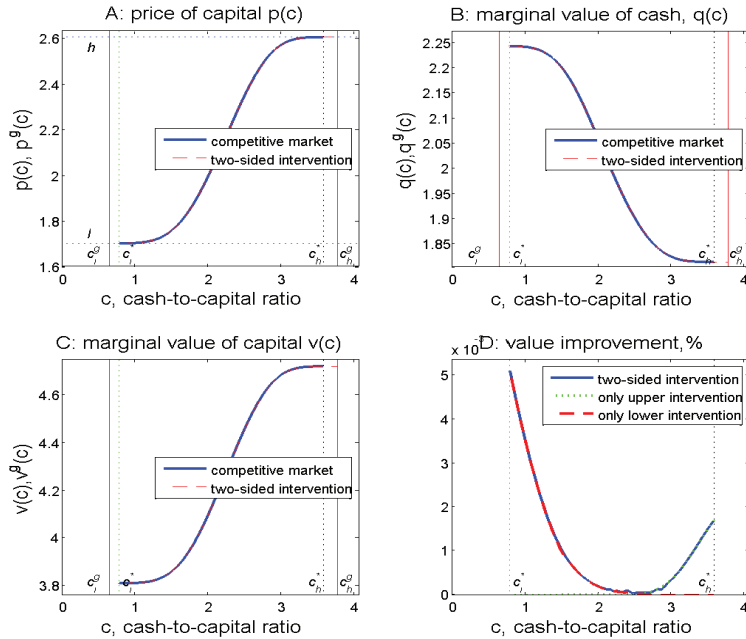


Figure 6: Panels A-C depict the price of capital, the value of cash, and the value of capital in the alternative setting both in the market equilibrium (solid) and when the planner decreases the lower threshold and increases the upper threshold (dashed). Panel D depicts the percentage change in value due the intervention. The vertical lines in each panel from left to right are the disinvestment threshold of the intervention and in the market equilibrium, $c_l^g = 0.6457$, $c_l^* = 0.7857$, and the investment threshold in the market equilibrium and of the intervention, $c_h^* = 3.5729$, $c_h^g = 3.7620$. Parameter values are $R_K = 4.2$, $\sigma^2 = 1$, $\xi = 5$, $R_C = 2$, $l = 1.7$, $h = 2.6$, $\alpha = 0.5$, $\kappa_h = \kappa_l = 7\%$, and $\pi_h = \pi_l = 1.5\%$. We use MATLAB built-in ODE solver `bvp4c` to solve the model, with a convergence criterion of 10^{-7} .

For this example, we chose parameters satisfying $l < \frac{R_K}{R_C} < h$ and picked a large γ . Then, the intuition for two-sided externality follows from the argument in Remark 1 in Section 3.2. In particular, the parameters imply that the firm’s private relative valuation of capital to cash, i.e.,

its private MRS (which is close to $p_t \in [l, h]$ according to the arguments in Section 3.2) will vary around its social counterpart (which is close to $\frac{R_K}{R_C}$ according to the arguments in Section 3.2). Hence, in booms, when p_t is close to h , the private valuation of capital is higher than its social counterpart. This implies that the investment threshold is lower (so the firm invests earlier) in the market equilibrium. The opposite argument holds during recessions, implying two-sided inefficiency.

5.2 Collateralized borrowing

We now go back to the base model and introduce collateralized borrowing. We present here only a sketch of this extension and the main results. A detailed analysis can be found in Additional Material, which is available on the the author's website.

We assume that installing one unit of capital allows the firm to borrow a constant $b \in [0, l)$ units of cash from external creditors, during both the aggregate and the idiosyncratic stages, and there are no other borrowing technologies allowed.¹⁹ The welfare accounting remains the same, because external creditors obtain zero rent always. We can characterize both the market equilibrium and the planner's solution in closed form for the economy with collateralized borrowing, using the fact that, in equilibrium, firms always maximize out their borrowing capacity.

Our main observation is that for a set of economies with small $h - l$, in the absence of borrowing firms underinvest even during booms, but, once allowing for collateralized borrowing, firms will start overinvesting during booms. This result indicates the potential social cost of excessive collateralized borrowing. Formally, define $\varepsilon \equiv h - l$, and imagine improved borrowing technologies modeled as an increasing sequence $b(k) > 0$ for $k \in (0, 1)$, so that $b(k) \equiv l - O(\varepsilon^k)$. The higher the k , the smaller the distance $O(\varepsilon^k)$, hence the higher the borrowing capacity $b(k)$ fixing l .

Proposition 9 *We have the following results:*

1. *When $\varepsilon \rightarrow 0$, in the market equilibrium both c_h^* and c_l^* converges to l . However, in the social planner's solution we have $c_l^P = 0$ and $c_h^P \rightarrow 0$. Hence, there is underinvestment during both booms and recessions.*
2. *For sufficiently small ε , there is overinvestment during the boom $c_h^*(k) < c_h^P(k)$ for $k > 1/3$.*

For the first part, from the planner's view, cash-flow shocks pose little risk given little adjustment cost (as h is close to l). Since capital is more productive ($R_K > hR_C$), the planner should hold almost no cash (i.e., $c_h^P \rightarrow 0$). However, the wedge between social and private incentives of holding capital remains. If $c_h^* \rightarrow h$, $c_l^* \rightarrow l$, and $h \rightarrow l$, then the price in the idiosyncratic stage is close to l , which suggests that in the competitive market firms liquidate capital as c_t falls below l and build as it rises above l .

¹⁹For example, firms cannot borrow against the new investment opportunities at the idiosyncratic stage, potentially because of the non-contractible nature of new opportunities. Here, b can be interpreted as the inefficient recovery that external creditors can obtain if the borrower reneges. This is in the tradition of Kiyotaki and Moore (1997), but to avoid complication we do not link the borrowing capacity to endogenous capital prices.

The second result of Proposition 9 gives the main point of this subsection. A higher k implies a higher $b(k)$, which increases the value of capital for both the planner and the market. We show that the positive effect on the market is stronger and, as $b(k)$ gets close to l , at some point $c_h^* < c_h^P$, i.e., overinvestment during booms. Intuitively, the planner's solution is mainly determined by the adjustment cost $h - l = \varepsilon$. For the market, as k increases, a higher $b(k) = l - O(\varepsilon^k)$ drives up l and $h = l + \varepsilon$, which raises the capital price in the idiosyncratic stage and hence also strengthens the private incentives to hold capital in the aggregate stage. Therefore the increase of the borrowing capacity leads to a faster decrease in c_h^* than in c_h^P .

6 Conclusion

We build an analytically tractable, dynamic stochastic model of investment and savings, in which investment cycles, i.e., boom periods with abundant investment and bust periods with low investment, arise naturally. In the presence of non-contractible idiosyncratic investment opportunities, a two-sided inefficiency can arise: there is too much investment in risky capital and cash buffers are too low during booms, and there is too little investment and too much cash hoarding during recessions. We show that in this case a one-sided policy targeting only the underinvestment during downturns might be ex ante Pareto inferior to the absence of intervention in all states (including downturns).

We acknowledge that there are standard ways to eliminate the inefficiency studied in this paper. In the working paper of this article (He and Kondor (2012)), we investigate this question in a simplified two-period setting. From an ex ante perspective, we show that the market can be completed by introducing Arrow-Debreu securities contingent on the realization of idiosyncratic opportunities, which restore investment incentives of individual firms in the competitive market. However, if idiosyncratic investment opportunities are not verifiable, so that the enforcement of Arrow-Debreu securities requires self-reporting, in general private investment incentives are still distorted away from the social ones. From an ex post perspective, if the final production output is fully pledgeable, then there will be no wedge between the social and private value of capital to cash, even without Arrow-Debreu securities.²⁰ Nevertheless, this result breaks down once we introduce imperfect delegation, which can be further microfounded by lack of commitment, hidden effort, or even information processing. In sum, just as in our two-period setting the social value ratio is independent of the state of cash-to-capital, our key result holds as long as market imperfections lead the price of capital to increase with the cash-to-capital ratio (in our model the idiosyncratic stage price $\hat{p}_\tau = c_\tau$ takes its extreme form due to the cash-in-the-market pricing).

Apart from analyzing two-sided inefficiencies, we also presented a novel way of modeling dynamic investment and savings. This method provides analytical tractability in a dynamic stochastic

²⁰To understand this, note that at the idiosyncratic stage, a cash firm will attach a marginal value of R_K to its capital, because such a cash firm, instead of selling its capital at the market, can hire another capital firm to operate its capital and extract all the output from production. Similarly, a capital firm values its cash by R_C , and the private value of capital to cash is always the social value ratio R_K/R_C .

framework for the full joint distribution of states and equilibrium objects. Exploring the potential of the developed framework in various other contexts is a task for future research.

References

- Acharya, Viral V, Heitor Almeida, and Murillo Campello.** 2013. “Aggregate risk and the choice between cash and lines of credit.” *The Journal of Finance*, 68(5): 2059–2116.
- Aghion, Philippe, George-Marios Angeletos, Abhijit Banerjee, and Kalina Manova.** 2010. “Volatility and growth: Credit constraints and the composition of investment.” *Journal of Monetary Economics*, 57(3): 246–265.
- Allen, Franklin, and Douglas Gale.** 1994. “Limited Market Participation and Volatility of Asset Prices.” *American Economic Review*, 84(4): 933–55.
- Allen, Franklin, and Douglas Gale.** 2004. “Financial Intermediaries and Markets.” *Econometrica*, 72(4): 1023–1061.
- Allen, Franklin, and Douglas Gale.** 2005. “From Cash-in-the-Market Pricing to Financial Fragility.” *Journal of the European Economic Association*, 3(2-3): 535–546.
- Almeida, Heitor, Murillo Campello, and Michael S Weisbach.** 2004. “The cash flow sensitivity of cash.” *The Journal of Finance*, 59(4): 1777–1804.
- Bates, Thomas W., Kathleen M. Kahle, and Rene M. Stulz.** 2009. “Why Do U.S. Firms Hold So Much More Cash than They Used To?” *The Journal of Finance*, 64(5): 1985–2021.
- Bianchi, Javier.** 2010. “Credit Externalities: Macroeconomic Effects and Policy Implications.” *American Economic Review*, 100(2): 398–402.
- Bianchi, Javier, and Enrique G. Mendoza.** 2011. “Overborrowing, Financial Crises and Macro-prudential Policy.” International Monetary Fund IMF Working Papers 11/24.
- Brunnermeier, Markus K., and Yuliy Sannikov.** 2014. “International Credit Flows, Pecuniary Externalities and Capital Controls.” Princeton University.
- Caballero, Ricardo J., and Arvind Krishnamurthy.** 2001. “International and domestic collateral constraints in a model of emerging market crises.” *Journal of Monetary Economics*, 48(3): 513–548.
- Caballero, Ricardo J., and Arvind Krishnamurthy.** 2003. “Excessive Dollar Debt: Financial Development and Underinsurance.” *Journal of Finance*, 58(2): 867–894.
- Davila, Eduardo.** 2014. “Dissecting Fire Sales Externalities.” <http://www.people.fas.harvard.edu/~edavila>.
- Denis, David J., and Valeriy Sibilkov.** 2010. “Financial Constraints, Investment, and the Value of Cash Holdings.” *Review of Financial Studies*, 23(1): 247–269.
- Diamond, Douglas W., and Raghuram Rajan.** 2011. “Illiquid Banks, Financial Stability, and Interest Rate Policy.” NBER Working Papers 16994.
- Dixit, A.K.** 1993. *The art of smooth pasting*. Vol. 2, Routledge.

- Dumas, Bernard.** 1991. “Super contact and related optimality conditions.” *Journal of Economic Dynamics and Control*, 15(4): 675–685.
- Eisenbach, Thomas M.** 2013. “Rollover risk as market discipline: A two-sided inefficiency.” Staff Report, Federal Reserve Bank of New York.
- Eisfeldt, Andrea, and Tyler Muir.** 2013. “Aggregate Issuance and Savings Waves.” <https://sites.google.com/site/andrealeisfeldt/>.
- Farhi, Emmanuel, and Iván Werning.** 2013. “On the Inefficiency of Financial Market Equilibria in Macroeconomics Models with Nominal Rigidities.” Harvard University.
- Farhi, Emmanuel, and Jean Tirole.** 2012. “Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts.” *American Economic Review*, 102(1): 60–93.
- Farhi, Emmanuel, Mikhail Golosov, and Aleh Tsyvinski.** 2009. “A Theory of Liquidity and Regulation of Financial Intermediation.” *Review of Economic Studies*, 76(3): 973–992.
- Gale, Douglas, and Tanju Yorulmazer.** 2011. “Liquidity hoarding.” Federal Reserve Bank of New York Staff Reports 488.
- Geanakoplos, John, and Heracles M. Polemarchakis.** 1985. “Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete.” Cowles Foundation for Research in Economics Discussion Papers 764.
- Gersbach, Hans, and Jean-Charles Rochet.** 2012. “Aggregate Investment Externalities and Macroprudential Regulation.” C.E.P.R. Discussion Papers CEPR Discussion Papers 8764.
- Greenwald, Bruce C, and Joseph E Stiglitz.** 1986. “Externalities in Economies with Imperfect Information and Incomplete Markets.” *The Quarterly Journal of Economics*, 101(2): 229–64.
- Gromb, Denis, and Dimitri Vayanos.** 2002. “Equilibrium and welfare in markets with financially constrained arbitrageurs.” *Journal of Financial Economics*, 66(2-3): 361–407.
- Hart, Oliver D., and Luigi Zingales.** 2011. “Inefficient Provision of Liquidity.” NBER Working Papers 17299.
- He, Zhiguo, and Arvind Krishnamurthy.** 2012. “Intermediary Asset Pricing.” *American Economic Review*, ;: forthcoming.
- He, Zhiguo, and Péter Kondor.** 2012. “Inefficient Investment Waves.” NBER Working Papers 18217.
- Hoberg, Gerard, and Gordon Phillips.** 2010. “Real and Financial Industry Booms and Busts.” *Journal of Finance*, 65(1): 45–86.
- Holmstrom, Bengt, and Jean Tirole.** 2011. *Inside and Outside Liquidity*. MIT Press.
- Ivashina, Victoria, and David Scharfstein.** 2010. “Bank lending during the financial crisis of 2008.” *Journal of Financial economics*, 97(3): 319–338.
- Jeanne, Olivier, and Anton Korinek.** 2010. “Managing Credit Booms and Busts: A Pigouvian Taxation Approach.” Peterson Institute for International Economics WP10-12.
- Jeanne, Olivier, and Anton Korinek.** 2013. “Macroprudential Regulation Versus Mopping Up After the Crash.” National Bureau of Economic Research, Inc NBER Working Papers 18675.

- Jermann, Urban, and Vincenzo Quadrini.** 2012. “Macroeconomic Effects of Financial Shocks.” *American Economic Review*, 102(1): 238–71.
- Kiyotaki, Nobuhiro, and John Moore.** 1997. “Credit Cycles.” *Journal of Political Economy*, 105(2): 211–48.
- Krishnamurthy, Arvind.** 2003. “Collateral constraints and the amplification mechanism.” *Journal of Economic Theory*, 111(2): 277–292.
- Lins, Karl V, Henri Servaes, and Peter Tufano.** 2010. “What drives corporate liquidity? An international survey of cash holdings and lines of credit.” *Journal of Financial Economics*, 98(1): 160–176.
- Lorenzoni, Guido.** 2008. “Inefficient Credit Booms.” *Review of Economic Studies*, 75(3): 809–833.
- Rhodes-Kropf, Matthew, David T. Robinson, and S. Viswanathan.** 2005. “Valuation waves and merger activity: The empirical evidence.” *Journal of Financial Economics*, 77(3): 561–603.
- Shleifer, Andrei, and Robert W Vishny.** 1992. “Liquidation Values and Debt Capacity: A Market Equilibrium Approach.” *Journal of Finance*, 47(4): 1343–66.
- Stein, Jeremy C.** 2011. “Monetary Policy as Financial-Stability Regulation.” NBER Working Papers 16883.
- Stiglitz, Joseph E.** 1982. “The Inefficiency of the Stock Market Equilibrium.” *Review of Economic Studies*, 49(2): 241–61.

A Appendix

In this Appendix we provide proofs for Propositions 2, 3, 4, the first part of Proposition 5 and Proposition 6. The proof of Proposition 9 is given in Additional Material available on the author’s website. Proofs for all the other results are provided in Online Appendix.

A.1 Proof of Proposition 2

Based on boundary conditions $R_K + D_1 + D_2 = l(R_C - \gamma D_1 + \gamma D_2)$ and $D_1 e^{-\gamma c_h^P} + D_2 e^{\gamma c_h^P} = 0$, the solutions for D_1 and D_2 are given by

$$D_1 = -\frac{(R_K - lR_C) e^{2\gamma c_h^P}}{(1 + l\gamma) e^{2\gamma c_h^P} - (1 - l\gamma)}, D_2 = \frac{R_K - lR_C}{(1 + l\gamma) e^{2\gamma c_h^P} - (1 - l\gamma)}. \quad (\text{A.1})$$

To verify that $c_t^P = 0$, we need to show that $j_P''(0) < 0$:

$$\frac{1}{\gamma^2} j_P''(0) = D_1 + D_2 = -(R_K - lR_C) \frac{e^{2\gamma c_h^P} - 1}{e^{2\gamma c_h^P} + l\gamma (e^{2\gamma c_h^P} + 1) - 1} < 0.$$

The super-contact condition at the optimal upper investment threshold c_h^P is

$$0 = \frac{\partial^2 j_P(c; c_l^P, c_h^P)}{\partial c} \Big|_{c=c_h^P} = \gamma^2 (D_1 e^{-\gamma c_h^P} + D_2 e^{\gamma c_h^P}). \quad (\text{A.2})$$

To show c_h^P exists and unique, define a function $G(c)$ ($G(c)$ is proportional to (A.2) if we plug D_1 and D_2 in (A.1) into (A.2))

$$G(c) \equiv \frac{R_K - hR_C}{R_K - lR_C} (e^{c\gamma} (1 + l\gamma) - (1 - l\gamma) e^{-c\gamma}) - 2\gamma(c + h), \quad (\text{A.3})$$

with $G(0) = 2R_K\gamma \frac{l-h}{R_K - lR_C} < 0$ (recall $R_K - hR_C > R_K - lR_C > 0$) and $G(\infty) = \infty$. We have

$$G'(c) = \gamma \left(\frac{(R_K - hR_C)}{(R_K - lR_C)} ((l\gamma + 1) e^{c\gamma} + e^{-c\gamma} (1 - l\gamma)) - 2 \right),$$

$G'(0) = 2R_C\gamma \frac{l-h}{R_K - lR_C} < 0$, and $G'(c)$ changes sign only once. Consequently, there is a unique \hat{c} that $G'(\hat{c}) = 0$, implying that $G(c)$ is decreasing for $c < \hat{c}$ and increasing for $c > \hat{c}$. As $G(0) < 0$ and $G(\infty) = \infty$, there must be a unique c_h^P that $G(c_h^P) = 0$, verifying the equation (23).

The social planner's value function $j_P(c)$ satisfies

$$0 = \frac{\sigma^2}{2} j_P''(c) + \xi (R_K + R_C c - j_P(c)) \quad (\text{A.4})$$

with boundary conditions $j_P(0) = l j_P'(0)$, $j_P(c_h^P) = (h + c_h^P) j_P'(c_h^P)$, and $j_P''(c_h^P) = 0$. Note that the boundary conditions imply that $j_P(c_h^P) = R_K + R_C c_h^P$. For later reference, we show that $j_P(c)$ is concave and increasing over $[0, c_h^P]$, and $j_P(c) < R_K + R_C c$ for $c \in [0, c_h^P]$. First, from smooth pasting condition at c_h^P we have

$$R_C - j_P'(c_h^P) = R_C - \frac{j_P(c_h^P)}{h + c_h^P} = R_C - \frac{R_K + R_C c_h^P}{h + c_h^P} = \frac{R_C h - R_K}{h + c_h^P} < 0.$$

Then, taking derivative again on (A.4) and evaluate at the optimal policy point c_h^P , we have

$$j_P'''(c_h^P) = -\frac{2\xi}{\sigma^2} (R_C - j_P'(c_h^P)) = \frac{2\xi}{\sigma^2} \frac{R_K - R_C h}{h + c_h^P} > 0,$$

and as a result $j_P''(c_h^P) < 0$. Suppose that j_P fails to be globally concave over $[0, c_h^P]$. Then there exists some point $j_P'' > 0$, and pick the largest one \hat{c} so that j_P'' is concave over $[\hat{c}, c_h^P]$ with $j_P''(\hat{c}) = 0$ and $j_P'''(\hat{c}) < 0$. But since j_P'' is concave over $[\hat{c}, c_h^P]$, $j_P'(\hat{c}) > j_P'(c_h^P) > R_C$, therefore $\frac{\sigma^2}{2} j_P'''(\hat{c}) = \xi (j_P'(\hat{c}) - R_C) > 0$, contradiction. Therefore j_P is globally concave over $[0, c_h^P]$, which also implies that $j_P(c) < R_K + R_C c$ for $c \in [0, c_h^P]$ due to (A.4).

A.2 Proof of Proposition 3

Suppose that we are given the policy pair (c_l, c_h) with $0 < c_l < c_h < c_h^P$ where c_h^P satisfies the super-contact condition $j_P''(c_h^P; 0, c_h^P) = 0$. To avoid cumbersome notation we denote the social value $j_P(c; c_l, c_h)$ given the policy pair (c_l, c_h) by $j(c; c_l, c_h)$, and denote the social value under the optimal policy $j_P(c; 0, c_h^P)$ by $j_P(c)$. We need to show that

$$\frac{\partial j_P(c; c_l, c_h)}{\partial c_l} < 0 \text{ and } \frac{\partial j(c; c_l, c_h)}{\partial c_h} > 0.$$

This result further implies that for $0 < c_l^2 < c_l^1 < c_h^1 < c_h^2 < c_h^P$, we have $j(c; c_l^1, c_h^1) < j(c; c_l^2, c_h^2)$.

As preparation, we first show that $j''(c_h; c_l, c_h) < 0$ and $j''(c_l; c_l, c_h) < 0$. Because (c_l, c_h) is sub-optimal, we must have $j(c; c_l, c_h) < j_P(c) \leq R_K + R_C c$ (recall Proposition 2). Then $0 = \frac{\sigma^2}{2} j''(c) + \xi(R_K + R_C c - j(c))$ implies that $j(c)$ is strictly concave at both ends. Second, for any policy pair (c_l, c_h) (including the market solution or the planner's solution), the smooth pasting condition (not optimality condition!) at the regulated ends implies that

$$j(c_h; c_l, c_h) - (c_h + h) j'(c_h; c_l, c_h) = 0, \quad (\text{A.5})$$

$$j(c_l; c_l, c_h) - (c_l + l) j'(c_l; c_l, c_h) = 0. \quad (\text{A.6})$$

Now we start proving the properties for the upper investment policy c_h . Define $F_h(c; c_l, c_h) \equiv \frac{\partial}{\partial c_h} j(c; c_l, c_h)$, which is the marginal impact of changing the investment policy on the social value. Differentiating the ODE (A.7) by the policy c_h , we have $\frac{\sigma^2}{2} \frac{\partial}{\partial c_h} j''(c; c_l, c_h) - \xi \frac{\partial}{\partial c_h} j'(c; c_l, c_h) = 0$, or

$$\frac{\sigma^2}{2} F_h''(c; c_l, c_h) - \xi F_h'(c; c_l, c_h) = 0. \quad (\text{A.7})$$

Moreover, take the total derivative with respect to c_h on the equality (A.5), i.e., take derivative that affects both the policy c_h and the state $c = c_h$, we have

$$\begin{aligned} & \frac{\partial}{\partial c_h} j(c_h; c_l, c_h) + j'(c_h; c_l, c_h) = j'(c_h; c_l, c_h) + (c_h + h) \left(\frac{\partial}{\partial c_h} j'(c_h; c_l, c_h) + j''(c_h; c_l, c_h) \right) \\ \Rightarrow & \frac{\partial}{\partial c_h} j(c_h; c_l, c_h) - (c_h + h) \frac{\partial}{\partial c_h} j'(c_h; c_l, c_h) = (c_h + h) j''(c_h; c_l, c_h) < 0 \\ \Rightarrow & F_h(c_h; c_l, c_h) - (c_h + h) F_h'(c_h; c_l, c_h) < 0. \end{aligned} \quad (\text{A.8})$$

which gives the boundary condition of $F_h(\cdot)$ at c_h . At c_l we can take total derivative with respect to c_h on the equality (A.6), we have the boundary condition of $F_h(\cdot)$ at c_l :

$$\frac{\partial}{\partial c_h} j(c_l; c_l, c_h) = (c_l + l) \frac{\partial}{\partial c_h} j'(c_l; c_l, c_h) \Rightarrow F_h(c_l; c_l, c_h) - (c_l + l) F_h'(c_l; c_l, c_h) = 0. \quad (\text{A.9})$$

With the aid of these two boundary conditions, the next lemma shows that $F_h(\cdot)$ has to be positive always. Because of the definition of $F_h(c; c_l, c_h) \equiv \frac{\partial}{\partial c_h} j(c; c_l, c_h)$, it implies that raising c_h given any state c and any lower policy c_l improves the social value. The argument for the effect of c_l is similar and thus omitted.

Lemma A.1 *We have $F_h(c) > 0$ for $c \in [c_l, c_h]$.*

Proof. *We show this result in two steps.*

1. $F_h(c)$ cannot change sign over $[c_l, c_h]$. Suppose that $F_h(c_l) > 0$; then from (A.9) we know that $F_h'(c_l) > 0$. Then simple argument based on ODE (A.7) implies that $F_h(\cdot)$ is convex and always positive. Now suppose that $F_h(c_l) < 0$; then the similar argument implies that F_h is concave and negative always. Finally, suppose that $F_h(c_l) = 0$ but F_h changes sign at some point. Without loss of generality, there must exist some point \hat{c} so that $F_h'(\hat{c}) = 0$, $F_h(\hat{c}) > 0$ and $F_h''(\hat{c}) < 0$. But this contradicts with the ODE (A.7).

2. Define $W_h(c) \equiv F_h(c) - (l + c) F_h'(c)$ so that

$$W_h'(c) = -(l + c) F_h''(c) = -\frac{2\xi(l + c)}{\sigma^2} F_h(c). \quad (\text{A.10})$$

As a result, $W_h'(c)$ cannot change sign. Because we have $W_h(c_l) = 0$, $W_h(c) = 0$ cannot change sign either.

3. Now suppose counterfactually that $F_h(c) < 0$ so that $W_h'(c) > 0$. Then (A.10) in Step 2 implies that $W_h(c) > 0$, and hence $F_h'(c_h) = \frac{1}{l+c} (F_h(c_h) - W_h(c_h)) < 0$. But we then have

$$W_h(c_h) = F_h(c_h) - (l+c_h)F_h'(c_h) = \underbrace{F_h(c_h) - (h+c_h)F_h'(c_h)}_{(A.8), \text{ negative}} + (h-l)F_h'(c_h) < 0,$$

where we have used (A.8). This contradicts with $W_h(c) > 0$. Thus we have shown that $F_h(c) > 0$.

■

A.3 Proof of Proposition 4

The expected total investment activity $T(c)$ solves $\frac{\sigma^2}{2}T''(c) = \xi T(c)$ with boundary conditions $T'(c_l) = \frac{1}{l+c_l}$ and $T'(c_h) = \frac{1}{h+c_h}$. For example, at $c = c_h$, a positive shock hits with $c = c_h + \epsilon$. To get back to the upper cash-to-capital ratio c_h , the economy builds new capital of $dK = \frac{K}{h+c_h}\epsilon$; thus, we have

$$T(c_h + \epsilon) = \frac{dK}{K} + T(c_h) = \frac{\epsilon}{h+c_h} + T(c_h) \Leftrightarrow T'(c_h) = \frac{1}{h+c_h}.$$

Now we study the impact of policies c_h and c_l on $T(\cdot; c_l, c_h)$. For illustration we analyze c_l only; a similar argument applies to c_h . Define $F(c) \equiv \frac{\partial}{\partial c_l} T(c; c_l, c_h)$; we have

$$\frac{\sigma^2}{2}T''(c; c_l) = \xi T(c; c_l) \Rightarrow \frac{\sigma^2}{2}F''(c; c_l) = \xi F(c; c_l). \quad (\text{A.11})$$

To determine boundaries for F , at c_h we have $T'(c = c_h; c_l) = \frac{1}{h+c_h}$ which implies that

$$F'(c = c_h) = \frac{\partial}{\partial c_l} T'(c = c_h; c_l) = 0.$$

On the other end, $T'(c = c_l; c_l) = \frac{1}{l+c_l}$ implies that $F'(c = c_l) + T''(c = c_l; c_l) = -\frac{1}{(l+c_l)^2}$ or

$$F'(c = c_l) = -\frac{1}{(l+c_l)^2} - T''(c = c_l; c_l) < 0;$$

Here we used the fact that $T''(c = c_l; c_l) > 0$; this fact is implied by (A.11) together with $T(c) > 0$ by definition.

Now we show that $F(c) > 0$ so that the total investment activity goes up for a higher c_l . To see this, first note that $F(c)$ never changes sign. Otherwise, suppose that there exists some c_1 so that $F(c_1) = 0$. If $F'(c_1) > 0$ then it must be that F is convex and positive for $c > c_1$, which contradicts with $F'(c_h) = 0$. Similarly we rule out $F'(c_1) < 0$. If $F'(c_1) = 0$, then combining with $F(c_1) = 0$ we can solve for $F(c) = 0$ for all c , contradicting with $F'(c_l) < 0$. Now since $F(c)$ never changes sign, it suffices to rule out $F(c) < 0$ always. If it were true, then F is concave always due to (A.11). This contradicts with $F'(c_l) < 0 = F'(c_h)$. As a result, $F(c) > 0$.

A.4 Proof of Proposition 5

Recall $G(c)$ defined in (A.3). Since fixing c we have

$$\lim_{\gamma \rightarrow \infty} \frac{\frac{R_K - hR_C}{R_K - lR_C} (e^{c\gamma} (1 + l\gamma) - (1 - l\gamma) e^{-c\gamma})}{\gamma c} = \infty,$$

to ensure that $G(c_h^P) = 0$ as $\gamma \rightarrow \infty$ we must have $c_h^P \rightarrow 0$. This is the first part of the first statement.

For the second part of the first statement,

$$\frac{\partial G(c)}{\partial \gamma} = \frac{R_K - hR_C}{R_K - lR_C} (ce^{c\gamma} (1 + l\gamma) + le^{c\gamma} - c(l\gamma - 1)e^{-c\gamma} + le^{-c\gamma}) - 2(c + h),$$

which is positive for sufficiently large γ . Finally, from the proof of Proposition 2 we know that $G'(c_h^P) > 0$. Hence, for sufficiently large γ , we have $\frac{\partial c_h^P}{\partial \gamma} = -\frac{\partial G(c_h^P)/\partial \gamma}{G'(c_h^P)} < 0$ which concludes the first part. The second part follows because $\frac{R_K - hR_C}{R_K - lR_C}$ is increasing in R_K and decreasing in R_C , and

$$\begin{aligned} \frac{\partial G(c)}{\partial h} &= \frac{-R_C}{R_K - lR_C} (e^{c\gamma} (1 + l\gamma) - (1 - l\gamma)e^{-c\gamma}) - 2\gamma < 0, \\ \frac{\partial G(c)}{\partial l} &= (R_K - hR_C) \frac{R_C (1 - e^{-2c\gamma}) + R_K \gamma + R_K \gamma e^{2(-c\gamma)}}{e^{-c\gamma} (R_K - lR_C)^2} > 0. \end{aligned}$$

Finally, fixing any c we have $\lim_{R_K \rightarrow hR_C} G(c) = -2\gamma(c + h) < 0$ always. This implies that for $\lim_{R_K \rightarrow hR_C} G(c_h^P) = 0$ to hold, it must be that $c_h^P \rightarrow \infty$ so that $\lim_{R_K \rightarrow hR_C} \frac{\frac{R_K - hR_C}{R_K - lR_C} (e^{c\gamma} (1 + l\gamma) - (1 - l\gamma)e^{-c\gamma})}{c} \rightarrow 2\gamma$. This concludes the first statement of the proposition. The second statement is in Online Appendix.

A.5 Proof of Proposition 6

The first statement comes from point 2 of Proposition 1. For the second statement, note that the proof of Proposition 1 goes through without any modification for the case when $R_K = R_C h$. That is, even in the limit $R_C h \rightarrow R_K$, the investment threshold in the market equilibrium c_h^* is finite, and under the parameter restriction of Proposition 1 we have $c_h^* < R_K$. However, note that given any parameters, in the limit $R_C h \rightarrow R_K$ the solution to equation (23) diverges to ∞ . Due to continuity, we can find $R_K - R_C h$ appropriately small so that c_h^P sufficiently close to R_K and hence $c_h^* < c_h^P$. And, even if the solution to (23) is above R_K , we can show that the resulting optimal investment threshold c_h^P lies above R_K and hence $c_h^* < R_K < c_h^P$. The details of this argument are in the Additional Material available on the author's website. This completes the proof.