Inefficient Investment Waves*

Zhiguo He
University of Chicago and NBER

Péter Kondor
Central European University and CEPR

July 2012

Abstract

We develop a dynamic model of trading and investment with limited aggregate resources to study investment cycles. Unverifiable idiosyncratic investment opportunities imply market prices to play a role of rent distribution, distorting private investment incentives from a social point of view. This distortion is price-dependent, leading to two-sided inefficient investment cycles—too much investment in booms with high prices and too little in recessions with low prices. Interventions targeting only the underinvestment in recessions might make all agents worse off. We connect our results to both industry specific and aggregate boom-and-bust patterns.

Key Words: Pecuniary externality, overinvestment and underinvestment, market intervention, Greenspan's put

1 Introduction

The history of modern economies is rich in boom-and-bust patterns. Boom periods with vast resources invested in new projects and low expected returns are followed by downturns when long-run projects are liquidated early, liquid resources are hoarded in safe short-term assets and there is little investment in new projects even if expected returns are high. 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 90s, or the recent housing boom around the mid 2000s and the crisis afterwards. These investment cycles are in the forefront of the academic and policy debate. Can these investment cycles be caused by financing frictions only? Are they inefficient, i.e., is there

*Email addresses: zhiguo.he@chicagobooth.edu, kondorp@ceu.hu. We are grateful to Ulf Axelson, Arvind Krishnamurthy, Guido Lorenzoni, Semyon Malamud, John Moore, Martin Oehmke, Alp Simsek, Balazs Szentes, Jaume Ventura, Rob Vishny, and numerous seminar participants. Péter Kondor acknowledges the financial support of the Paul Woolley Centre at the LSE.

[1] For example, Hoberg and Phillips (2010) statistically identifies 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 nineties.

overinvestment in booms and/or underinvestment in downturns? Relatedly, should the policy maker intervene in booms, in downturns or both?

In this paper, we contribute to this debate as follows. Constrained efficient investment cycles arise naturally in our dynamic economy with limited aggregate resources to fund risky projects. However, introducing unverifiable idiosyncratic investment opportunities leads to inefficient investment behaviors in the market equilibrium. As the main novelty of our paper, we show that this friction may induce a two-sided inefficiency, i.e., overinvestment in booms with high asset prices and underinvestment in downturns with low asset prices. As a mirror image, firms store too little liquid resources in booms and hoard too much of them in downturns. We show that intervention targeted at raising prices in downturns to avoid underinvestment typically make overinvestment in booms worse. What is more, this adverse effect might be so strong that the intervention becomes Pareto inferior compared to the case of no intervention at all.

We present an analytically tractable, stochastic dynamic model of trading and investment. There are two goods; a capital good and cash. *Capital* stands for risky long-term projects that generate stochastic cash flows according to a linear technology. In contrast, *cash* stands for riskless short-term asset which serves as both the consumption good and the input for investment in capital. Firms who operate the capital can invest and disinvest, that is, they can create new capital at a constant unit cost or dismantle the capital for a relatively smaller constant benefit, both in terms of cash. Firms can also trade capital among each other at an equilibrium price. Capital generates risky interim cash flows, and these represent aggregate shocks in our economy. Negative cash flows imply that capital requires costly maintenance in terms of cash; when there is shortage of aggregate cash, firms might need to dismantle the capital. As a result, firms will store cash in order to avoid inefficient liquidation of the project.

The crucial feature of our economy is that ex ante identical firms are subject to idiosyncratic shocks in their investment opportunities. Namely, at a given time some firms experience a high productivity shock on their capital, while others receive an investment opportunity into another new technology. The latter group of firms sells their capital to the former group in a Walrasian market, and after trading the latter group invests all their cash into the new opportunity while the former operates their capital holdings. We consider both the complete market case when the idiosyncratic shock is contractible, and the incomplete market case when the market for such contracts is missing.

The aggregate capital stock and cash are two state variables of our economy. Thanks to scale invariance, we solve our model by keeping track of the aggregate cash-to-capital ratio as our unidimensional state variable. It is also our proxy for the level of aggregate liquidity in our economy. When interim cash flow shocks are negative, the cash-to-capital ratio falls, and so does the equilibrium price of capital, boosting the expected return on buying capital. When the price drops to the level of the liquidation benefit, capital are dismantled back to cash keeping the aggregate cash-to-capital ratio above an endogenous lower threshold. We think of low liquidity states as a downturn. The low price level of capital in downturn represents a liquidity premium, because firms have to be compensated for the increasing probability of forced inefficient liquidation when no cash is available for potential maintenance. As the cash-to-capital ratio rises, this risk is reduced, the capital price increases and the premium decreases. When the price reaches the cost of creating new capital, firms build new capital keeping the aggregate cash-to-capital ratio below an endogenous upper threshold. We think of the high aggregate liquidity state when new projects are created as a boom period.

With the aid of analytical solutions, we study whether the investment threshold at the boom period and disinvestment threshold at the downturn are at their efficient levels. In the complete market benchmark, the market solution and the social planner's choice coincide: Firms dismantle their productive capital only when the cash-to-capital ratio hits zero, and firms invest when the cash-to-capital ratio hits a positive threshold in booms. This upper threshold is determined by a trade-off. On one hand, building capital is a positive net present value project. On the other hand, storing some cash to avoid costly liquidation is valuable for buffering purposes. Hence in the complete market benchmark, although expected returns and economic activities fluctuate with the cash-to-capital ratio, the resulting investment cycle is not a sign of inefficiencies per se.

However, in the incomplete market where the market for idiosyncratic shocks is missing, the investment and disinvestment thresholds are distorted. In particular, firms always dismantle capital at a positive cash-to-capital ratio. Also, under some conditions, they build capital at a lower investment threshold than the social planner would. That is, they invest too little (dismantle too much) in downturns, and overinvest in booms. As a mirror image, they hoard too much cash in a downturn, and hold too little cash in a boom.

The intuition behind our mechanism is as follows. The price of existing capital plays a double role in our economy. First, it determines the investment decision, i.e., firms build new capital (dismantle existing capital) when the price of existing capital is sufficiently high (low). Second, it also determines the terms of exchange when firms are subject to different idiosyncratic shocks. Although the exchange moves all cash and capital to the most efficient hands, the term of this exchange affects the rent distribution across firms with different idiosyncratic shocks. For example, the equilibrium price of capital is high when the cash-to-capital ratio is high, and firms who receive the idiosyncratic new investment opportunity are able to exchange their capital holdings for a large amount of cash, acquiring large rents on their capital. Importantly, though not affecting the total welfare, the ex post distribution of this rent affects firm's ex ante incentives to hold cash versus capital. In other words, capital prices, through its second role of determining the distribution of rent, cause a wedge between the private marginal rate of substitution between capital and cash and that of the social planner. This wedge distorts the price of capital away from its social value, and in turn—because of its first role—distorts the investment and disinvestment decisions.

Novel to the existing literature on pecuniary externalities, we show that the direction of price distortion in our model depends on the state of the economy. The high rent on capital versus cash in a boom implies that the private value of capital is higher than the social value of capital in these states. This is a pecuniary externality inducing even higher price and overinvestment in capital in booms. As a symmetric argument, in downturns the price of the capital is low, inducing a negative wedge between the private and the social value of capital. This implies even lower price and underinvestment in capital in downturns.

We suggest a number of applications for our model. First, we highlight the dynamic consequences of one-sided (in downturn only) interventions, which relates our paper to the debate on asymmetric interest rate policy often referred to as the Greenspan's put. The dynamic structure of our model emphasizes a two-way interaction between decisions in booms and downturns. When a firm decides to build capital, she worries that this capital will have to be dismantled if the state of the economy deteriorates significantly. When she dismantles capital in downturns, she similarly takes into account that the economy might revert to a boom. As a result, if the policy maker taxes cash-holdings in downturns to increase the capital price and avoid inefficient liquidation, this one-sided intervention will typically make the overinvestment problem worse in the boom. This unintended effect of the downturn-intervention in booms may make firms worse off even in downturns.

Second, our model can also be interpreted in a narrower, sectoral level. In particular, a series of papers[2] document that there are relative boom/bust patterns across industries. That is, from time-to-time different industries go through patterns where high returns induce an overinvestment phase relative to other industries which is followed by abnormally low returns inducing an underinvestment phase. Interestingly, Hoberg and Phillips (2010) show that these inefficient investment waves are more pronounced in more competitive industries, where firms do not internalize their own impact on the market price. This evidence is consistent with our mechanism that inefficient investment waves are driven by pecuniary externalities (which requires economic agents to be price takers).

Third, as an application with both aggregate and sectoral elements, we also connect our results to the boom and bust pattern in construction and housing prices. Our mechanism suggests that the volume of real estate development in a boom is inefficiently high,[3] because investors build houses (instead of holding liquid financial assets) as a store of value expecting to be able to sell the real estate for a high price in case they want to invest into a new investment opportunity. In downturns, the resale price of real estate is low, so investors prefer to hold inefficiently high level of liquid assets in case they want to invest into a new investment opportunity. That is, the relative liquidity of real estate compared to short-term assets varies over the business cycle, resulting in a two-sided inefficiency.

Finally, we relate our results to the literature on financial development and growth. For this, we consider our assumption of the missing market for idiosyncratic investment opportunities as a proxy for the lack of financial development. Under this interpretation, our model provides a potential justification for the stylized facts that in less financially developed countries investment in productive technologies is more volatile and exhibits stronger procyclicality.[4] As we show,

[2] See Rhodes-Kropf, Robinson and Viswanathan (2005) and Hoberg and Phillips (2010).

[3] One suggestive sign of the inefficiently high level of real estate development is the frequently observed phenomenon of "overbuilding" (e.g. Wheaton and Torto, 1990; Grenadier, 1996), that is, periods of construction booms in the face of rising vacancies and plummeting demand.

[4] See Aghion et al. (2010).

this excess volatility implies slower capital accumulation and lower consumption in less financially developed countries.

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 for the full joint distribution of states and equilibrium objects.

Literature. To our knowledge, our paper is the first to show that the simple friction of unverifiable idiosyncratic investment opportunities results in investment cycles with two-sided inefficiency, i.e., overinvestment in booms and underinvestment in recessions.

Our work belongs to a growing literature analyzing pecuniary externalities in incomplete markets. All this literature, including our paper, builds on the result in Geanakoplos and Polemarchakis (1985) that when markets are incomplete, the competitive equilibrium may be constrained inefficient. In this setting pecuniary externalities can have a first order effect, because prices fail to equate the marginal rate of substitution of each firm across all goods (or states). A large stream in this literature emphasizes a fire-sale feed-back loop induced, typically, by a collateral constraint (e.g. Kiyotaki and Moore, 1997; Gromb and Vayanos, 2002; Krishnamurthy, 2003; Lorenzoni, 2008; Jeanne and Korinek, 2010; Bianchi, 2010; Bianchi and Mendoza, 2011; Stein, 2011; He and Krishnamurthy, 2012*a*). In these papers, firms do not take into account that the more they invest ex ante, the more they have to dismantle once they hit their constraint, which reduces fire-sale prices tightening the constraint and amplifying the effect. Such an amplification mechanism is absent in our paper. Instead, we follow Shleifer and Vishny (1992), Allen and Gale (1994, 2004, 2005), Caballero and Krishnamurthy (2001, 2003), Farhi, Golosov and Tsyvinski (2009) and Gale and Yorulmazer (2011) where an uninsurable shock creates the dispersion in marginal rate of substitution of ex-ante identical firms.[5] Our main point of departure is that in our paper the sign of the distortion turns out to switch with the state of the economy. Our main innovation leading to this result is that the Geanakoplos and Polemarchakis (1985) mechanism is interacted with a theory of countercyclical liquidity premium.

A group of recent papers investigating the moral hazard problem of incentivizing banks in a macroeconomic context derive related implications to our work. Similar to our work, in Gersbach and Rochet (2012) 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.[6] Furthermore, our result that one-sided interventions can be inferior to no interventions is related to the debate on the pros and cons of asymmetric interest rate policy often referred to as the Greenspan's put (e.g. Diamond and Rajan, 2011; Farhi and Tirole, 2012).[7] In these papers,

[5] For fire-sale externalities explored in the literature see Davila (2011) for a comparative analysis. For uninsurable idiosyncratic liquidity shocks, see Holmstrom and Tirole (2011, chap.7.) for 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.

[6] Namely, if private benefits of banks are increasing in the size of their loans, then it is cheaper to make them exert effort by letting them to increase their loan size in booms compared to paying them sufficient rent to avoid this. This private contract does not take into account the price effect of the resulting procyclicality in aggregate loan size.

[7] In a nutshell, supporting distressed institutions by low interest rates is detrimental to ex ante incentives of

agency frictions and related incentive problems for financial intermediaries are crucial. In contrast, our mechanism is based on the novel observation that a missing market for idiosyncratic investment opportunities can lead to a market price which is biased in the opposite direction in booms and downturns. Thus, whatever policy helps in a boom will typically make firms worse off in downturns and vice-versa. Ex ante welfare in any state is the weighted average of these effects.

From a methodological point of view, as a continuous-time model with investment and trade, the closest papers to ours are Brunnermeier and Sannikov (2011) and He and Krishnamurthy (2012*b*). As their focus is balance sheet amplification rather than pecuniary externality, their model is more complex and less analytically tractable.

The structure of our paper is as follows. Section 2 gives an simple static example to highlight the main intuition. In Section 3 we present our model, and analyze the market equilibrium and the constrained efficient allocations of the social planner. In Section 4 we expose the inefficiencies of the market solution. Section 5 presents our applications and extensions. Finally, we conclude.

2 A simple example

Before we move on to set up our main model, we first illustrate the key insight of our paper by a simple example with the following 2-date-2-good economy.

Endowment and goods. At the beginning of date 0 there is a unit mass of risk-neutral firms. Each firm $i \in [0,1]$ holds one unit of the capital good (also referred to as capital), and c units of the consumption good (also referred to as cash). While this example is fully symmetric in these two goods, this will not be the case in our main model in Section 3.

Transformation technology. At date 0, each firm can invest or disinvest by using the consumption good to create capital or the other way around. The technology is such that each firm can convert two units of cash to a capital, or obtain a unit of cash by liquidating two capital. Thus, given the endowment of one unit of capital and c cash, the individual holding of $\left(K^i, C^i\right)$ at the end of date 0 must satisfy

$$\begin{cases} 2K^i + C^i = 2 + c & \text{if } K^i > 1, \\ \frac{1}{2}K^i + C^i = \frac{1}{2} + c & \text{if } K^i \leq 1. \end{cases} \tag{1}$$

This budget constraint reflects a kinked transformation technology.[8]

Idiosyncratic skill shocks and trading. At date 0, each firm is identical. However, at the beginning of date 1, half of the firms are hit by a high productivity shock, so that they can produce 3 units of the consumption good out of each unit of the capital. The rest of the firms cannot use the capital at all, but receive a new investment opportunity which turns each unit of cash into 3

financial intermediaries (e.g., encouraging their excessive risk-taking ex ante). As a result, ex post intervention to save distressed institutions will be needed more often.

[8] We show in Appendix C that our main results do not depend on the kinky technology frontier only on its convexity.

units of consumption good.[9] There are no aggregate shocks in this simple example, but we will introduce them later in the main model.

After these idiosyncratic skill shocks, firms can trade capital for cash with each other. Finally, firms produce and invest in the new opportunity, and consume the proceeds at the end of date 1. Crucially, neither the returns from production nor the returns from the new opportunity are pledgeable, and the realization of the skill-shock is unverifiable.

The market solution. Recall that $\left(K^i, C^i\right)$ describe the holdings after the adjustment in date 0 but before the trade in date 1. Denote the aggregate counterpart $K = \int K^i di$ and $C = \int C^i di$. After idiosyncratic skill shocks, firms who can produce (exploit the new opportunity) will exchange all their cash (capital) for capital (cash). In the competitive market, there are $C/2$ amount of cash to purchase $K/2$ amount of capital, and thus the equilibrium capital price in date 1, in terms of cash, is

$$p = \frac{C}{K}. \tag{2}$$

For this given price, each firm solves

$$\max_{K^i, C^i} J^i\left(K^i, C^i; p\right) = \frac{3}{2}\left(K^i + \frac{C^i}{p}\right) + \frac{3}{2}\left(K^i p + C^i\right), \tag{3}$$

subject to the budget constraint in (1). For instance, with probability 1/2, the firm becomes specialized at capital, and she can purchase C^i/p units of capital from the market. Then in total she will have $K^i + C^i/p$ units of capital, each of them producing 3 units of final consumption good.

Given the simple linear structure, the individual demand function is

$$\begin{cases} K^i = 1 + \frac{c}{2}, C^i = 0 & \text{if } p > 2; \\ K^i = 1, C^i = c & \text{if } \frac{1}{2} \leq p \leq 2; \\ K^i = 0, C^i = c + \frac{1}{2} & \text{if } p < \frac{1}{2}. \end{cases} \tag{4}$$

This is intuitive: individual firms hold the asset (capital or cash) whose relative price is higher than the marginal rate of transformation, and inaction may be optimal because of the kink in the transformation technology.

We focus on symmetric equilibrium where $K^i = K$ and $C^i = C$. We derive the unique symmetric market equilibrium by combining individual demand functions (4) with the equilibrium condition in (2). It is apparent that the equilibrium price p has to be in the interval $\left[\frac{1}{2}, 2\right]$. We characterize market equilibria based on the relative initial cash endowment c.

Case 1 *Suppose $c > 2$ so that the initial cash endowment is relatively high. Then the market*

[9] We show in Appendix C that our main results do not depend on the extreme assumption that ex post each agent prefers to produce with only one of the goods. In particular, we present an example where the two groups end up to use a different convex combination of cash and capital.

equilibrium has $p = 2$, and individual firms invest in capital to reach the holdings of

$$K^i = 1 + \frac{c-2}{4} > 1, C^i = c - \frac{c-2}{2} < c.$$

Case 2 *Suppose $c < \frac{1}{2}$ so that the initial cash endowment is relatively low. Then the market equilibrium has $p = \frac{1}{2}$, and individual firms disinvest to reach the holdings of*

$$K^i = 1 - \left(\frac{1}{2} - c\right) < 1, C^i = c + \frac{1}{2}\left(\frac{1}{2} - c\right) > c.$$

Case 3 *Otherwise, when $c \in \left[\frac{1}{2}, 2\right]$, the market equilibrium has $p = c$, and individual firms do not invest so that*

$$K^i = 1, C^i = c.$$

Social planner's problem and inefficiency. The planner maximizes the sum of final consumption of firms at the end of date 1. The only difference between the planner and the market is that the market takes prices as given, while the social planner takes into account how individual decisions determine prices. Thus, we can write the problem of the planner as

$$\max_{K,C} \frac{3}{2}\left(K + \frac{C}{C/K}\right) + \frac{3}{2}\left(K\frac{C}{K} + C\right) = \max_{K,C} 3\left(K + C\right)$$

subject to the aggregate budget constraint similar to (1):

$$\begin{aligned} 2K + C &= 2 + c \quad \text{if } K > C, \\ \tfrac{1}{2}K + C &= \tfrac{1}{2} + c \quad \text{if } K \leq C. \end{aligned} \tag{5}$$

The optimal solution is simply the endowment allocation:

$$K = 1, C = c. \tag{6}$$

Intuitively, the marginal rate of substitution for social welfare between capital and cash is 1. Given that this lies within the marginal rates of transformation of $\frac{1}{2}$ and 2, it is socially wasteful to invest or disinvest. However, as shown in the market solution, individual firms invest (disinvest) when the initial endowment of cash is relatively high (low).

Intuition and discussion. Let us highlight the main lessons from this example. Inefficiency can potentially come from two potential sources: ex ante date 0 investment (transforming cash to capital or the other way around), and ex post date 1 resource allocation among heterogeneous firms. In our model, ex post resource allocations is always efficient, as the date 1 trading ensures that capital (cash) go to the right hand—firms who can produce will get all the capital, and firms who has a new opportunity will have all the cash to exploit this opportunity. In fact, under both the planner's solution and the market one, given the fixed aggregate resource pair of (K, C), the representative firm obtains

$$\int \left[\frac{1}{2} \left(K^i + \frac{C^i}{p} \right) 3 + \frac{1}{2} \left(K^i p + C^i \right) 3 \right] di = 3 \left(K + C \right)$$

in expectation. As a result, in our model the inefficiency arises only because of the divergent date 0 private investment incentives compared to the one of the social planner.

To highlight the distortion investment incentives, we study the marginal rate of substitution for both the social planner and individual firms. The social planner's value, given the pair of capital-cash holdings (K, C), is simply given by

$$J_P \left(K, C \right) = 3 \left(K + C \right).$$

This implies that independent of market price p, the social planner's marginal rate of substitution (MRS^S) between cash and capital is always 1. In contrast, the private value of the pair of capital and cash holdings $\left(K^i, C^i \right)$ given the price p is (3), and the marginal rate of substitution between capital and cash for individual price-taking firms is the price p:

$$MRS^i = \frac{\partial J^i \left(K^i, C^i; p \right) / \partial K^i}{\partial J^i \left(K^i, C^i; p \right) / \partial C^i} = \frac{\frac{1}{2} \left(3 + 3p \right)}{\frac{1}{2} \left(\frac{3}{p} + 3 \right)} = p. \tag{7}$$

Interestingly, there is a wedge between the social planner's marginal rate of substitution of 1 and that of individual firms p. The economic force behind this wedge is as follows. Although the ex post (date 1) trading guarantees the efficient resource allocation (which the social planner cares about), it introduces the distribution of economic rents (which the social planner does not care about) that in general distorts the individual firm's ex ante (date 0) marginal rate of substitution. To see this, consider the pair of trading firms so that one prefers to have the capital (a capital-firm), while the other prefers to have the cash (a cash-firm). The capital in the cash-firm's hand delivers the firm a utility of $3p$ (by selling the capital for p cash and investing it to get $3p$ utility at the end of date 1), while the cash in the capital-firm's hand delivers $3/p$ (by buying $1/p$ capital which yields $3/p$ utility at the end of date 1). Hence, when cash is abundant (scarce) relative to capital in aggregate so that the ex post capital price p is higher (lower) than 1, most of the rent from holding cash (capital) goes to the firms holding capital (cash). As a result, compared to the social planner, holding capital (cash) becomes more attractive than holding cash (capital).[10]

[10] In Appendix C we explore this inefficiency in more detail, and we show that our results are robust as long as there are other informational or agency frictions. We highlight three potential routes to deal with this inefficiency: (1) agents could exchange contracts in period 0 conditional on their date 1 individual skill-shocks; (2) agents could be forced to pool their assets into a bank in period 0, to get back a unit of capital or cash in period 1 based on their self-reported type; and (3) an agent who holds an asset useless to him can contract with another agent to exploit the asset and return the proceeds.

Solution (1) and solution (2) would eliminate the inefficiency if the skill-shock was verifiable. However, reminiscent to Diamond and Dybvig (1983) and Jacklin (1987), under (1) agents have incentives to misreport their types, and this problem applies to (2) as well if there exists some Walrasian market for the asset. Solution (3) heavily relies on the assumption of pledgeable output. When output is partially pledgeable, we show that agents would prefer

Finally, note that in resolving the inefficiency, the social planner does not need to identify (or make firms to reveal) which firm is hit by which idiosyncratic shock. In fact, ex post trading will implement the efficient allocation of cash and capital across heterogenous firms, and it is sufficient for the social planner to control the ex-ante investment decisions.

The above intuition that agents overinvest in the scarce good because of a pecuniary externality carries the main driving force of our main results in this paper. Our full dynamic model completes this intuition in two crucial dimensions. First, it connects the relative scarcity of assets with the state of the economic cycle. While in this example there is little significance that we call one of the goods a capital good (and the other as cash), this will not be the case in our main model. Indeed, in our full model, the capital good generates a stochastic flow of cash, and firms deciding to save or invest the produced cash drive the economic cycle. Second, our dynamic structure provides a natural way to analyze how a government intervention in a downturn changes firms' incentives in a subsequent boom and how the fact that firms foresee this effect determines the welfare consequences of this intervention.

3 The Model

3.1 Assets

We model an economy with trade and investment. There is a single capital good representing risky and productive projects. The other asset in this economy is cash which both serve as a consumption good and as an input for building capital. We assume that there is a safe storage technology and capital does not depreciate; thus both capital and cash are perfectly storable.

There is a final date arriving at a stopping time τ with Poisson intensity ξ, where ξ is a positive constant. At the end of the final date, each unit of capital pays a single cash-dividend (to be specified shortly). However, 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 $(\Omega, \mathcal{F}, \mathcal{P})$. When $\sigma dZ_t > 0$ the capital generates cash; when $\sigma dZ_t < 0$, the owner of the capital has to invest this amount (of cash) to the capital as maintenance; otherwise the capital turns unproductive.

Denote by K_t the aggregate quantity of capital. Then given the aggregate cash shock σdZ_t for each unit of capital, the evolution of aggregate cash, without investment or disinvestment (to be introduced shortly), is

$$dC_t = K_t \sigma dZ_t. \tag{8}$$

3.2 Firms and frictions

The market is populated by a unit mass of risk neutral firms who operate the capital. At each time instant firm may decide to build new capital, trade capital for cash at the equilibrium price p_t, or

trading their assets directly in the market (as in our model) to writing contracts in (3) if agents can abscond with a sufficiently large fraction of the output.

dismantle the capital. Building new capital costs h units of cash, while liquidating a unit capital provides l units of cash where $h > l$. This scraping technology ensures the limited liability of the asset owner despite the potentially unbounded losses in (8). Firms can also consume their cash at any moment for a constant marginal utility of 1. Because of linear technologies, in general it is optimal to have threshold strategies for (dis)investment. Thus, we can simply focus on thresholds in comparing different (dis)investment strategies.

The major friction in this economy is that firms are subject to non-verifiable idiosyncratic shocks. While firms are ex ante identical, ex post they differ in their skills. Specifically, in the final date, each firm with probability half experience a firm-specific productivity hike implying that in their hands each capital can produce R unit of cash. Firms not experiencing this shock cannot produce any cash from their capital. However, they are the only ones to receive an investment opportunity which allows them to turn every unit of cash to $u > 1$ units of final consumption. This situation is analogous to the capital-cash example studied in Section 2. Just as in that example, we only need that ex post there is heterogeneity among firms in their valuations of the available assets.

We assume that both R from the capital and u from the new technology are not pledgeable; this extreme assumption is a short-cut for agency and/or informational frictions.[11] Throughout we assume that

$$\frac{R}{h} > u, \tag{9}$$

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

Firms learn which group they belong to only at the beginning of the final date. We refer to the group with the skill to invest in the new technology as firms hit by *skill-shock*. Importantly, after firms learn whether they are hit by the shock but before the final productions are taken place, all firms have a last trading opportunity to trade capital for cash. We refer to the potentially infinitely long time interval before the final date τ as ex ante, and refer to the final date τ (where final trading occurs) as ex post. We denote the ex post price by $\widehat{p}_\tau$ (recall that we denote by p_t ex ante prices).

Figure 1 summarizes the time-line of events in our model. While the dynamic structure of our model might seem unusual, we argue that this structure unifies the advantages of two period models and infinite period models. In particular, this structure renders tractability for analytically showing pecuniary externalities, and moreover allows us to analyze the stationary distribution of ex ante variables.[12]

[11] Appendix C, in the context of our simple example, discusses the potential agency problems in detail.

[12] It is worth emphasizing that our qualitative results are robust to a setting where investment opportunities arrive in every instant to a given fraction of agents (and thus separate from the aggregate productivity hike event); see the analysis in Section 5. We show that the main qualitative results still hold in that less tractable but, in some contexts, more intuitive variant.

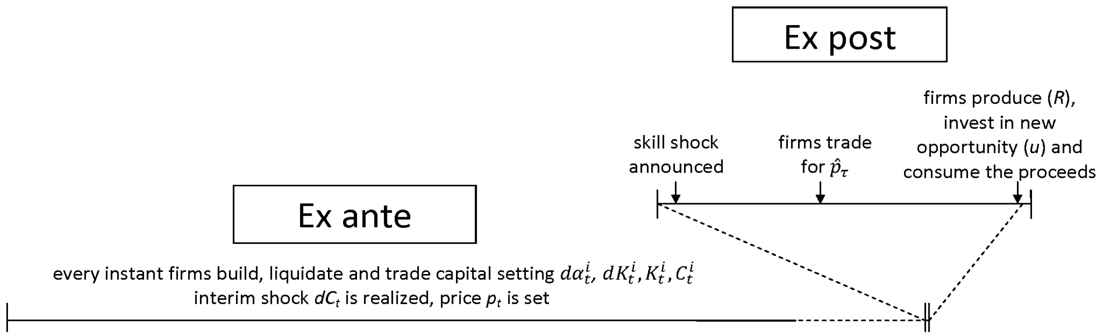

Figure 1: Time line.

3.3 Individual firm's problem

Consider firm i who holds K_t^i units of capital and C_t^i amount of cash, with a wealth of (in terms of cash) $w_t^i \equiv p_t K_t^i + C_t^i$. Then the firm i is solving the following problem:

$$\max_{\{d\alpha^i \geq 0, K^i, C^i, dK^i\}} \mathbb{E}\left\{\int_0^\infty \xi e^{-\xi\tau}\left(\int_0^\tau d\alpha_s + \left[\frac{1}{2}\left(K_\tau^i + \frac{C_\tau^i}{\widehat{p}_\tau}\right)R + \frac{1}{2}\left(K_\tau^i \widehat{p}_\tau + C_\tau^i\right)u\right]\right)d\tau\right\} \tag{10}$$

where α_t^i is the 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 the term in the squared bracket is the consumption at the final date. K_t^i is the amount of capital that she dismantles or builds. In the squared bracket, we also used the fact that those hit by the skill-shock sell their capital for $\widehat{p}_\tau$ to those who are not hit. For instance, when the skill shock hits, the firm sells the capital to receive $K_\tau^i \widehat{p}_\tau$, and then invests them together with C_τ^i in the new technology with productivity u. Note that the expression in the squared bracket is analogous to the objective function (3) in our simple example.

The problem in (10) is subject to the dynamics of individual wealth,

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

where θ is the cost of changing the amount of capital so that

$$\theta = \begin{cases} h & \text{if} \quad dK_t^i \geq 0 \\ l & \text{if} \quad dK_t^i < 0 \end{cases}.$$

Also, wealth cannot be negative at any point, i.e., $w_t^i \geq 0$ for all t.

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

dK_t, (8) implies that the dynamics of aggregate cash level in the economy is[13]

$$dC_t = \sigma K_t dZ_t - \theta dK_t. \tag{11}$$

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:

$$c_t \equiv C_t/K_t,$$

which evolves according to

$$dc_t = \frac{dC_t}{K_t} - \frac{C_t}{K_t}\frac{dK_t}{K_t} = \sigma dZ_t - (\theta + c_t)\frac{dK_t}{K_t}. \tag{12}$$

3.4 Interpretation

At the firm level, C_t^i represents the financial slack of a firm, that is, its cash holdings and other short-term investments, while K_t^i represents the total of its gross property, plants, equipment, inventories and intangible assets. The process $\sigma K_t^i dZ_t$ represents operating cash-flows.[14] In our abstract model firms do not raise funds from outside investors, financing their investment from retained earnings only. This gives a clear motive to build up cash reserves as a buffer for future operational losses. While it is a major simplification, a series of recent papers with strong empirical focus show that this precautionary motive remains the key determinant of firms behavior even when both internal and external finance are considered (e.g. Armenter and Hnatkovska (2011), Bolton, Chen and Wang (2011), Covas and Haan (2011), Eisfeldt and Muir (2012)).[15]

In a domestic, interindustry context, we should expect that the relevance of our main contracting friction of unverifiable investment opportunities is more relevant in industries with a larger role of non-tangible assets such as human capital. In an international context, this friction should be more relevant in countries with less financial development, and lower degree of law enforcement.

3.5 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 (10), and*

[13] To simplify notation we ignore the possibility that at any given point in time some agents create capital while some agents liquidate capital. It will be easy to see that this never happens in equilibrium.

[14] In contexts where firm-level data might not be available, we can think of the net of liquid asset holdings C_t as a metaphor for quantities invested in short-run, low-return projects, while the net of capital holdings as total cumulative investment in long-run, higher-return, riskier projects. For example, Aghion et al. (2010), in an international context, use the share of structural investment compared to total investment as an empirical proxy for the share of investment in long-run productive projects.

[15] For example, Covas and Haan (2011) and Eisfeldt and Muir (2012) document that firms tend to safeguard their investments in downturns by accumulating retained earnings and issuing more equities and debt in booms than their direct investment needs.

2. *markets clear in every instant both ex ante and ex post.*

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. Thus, often it is convenient to pick the particular market equilibrium where all firms follow the same action. We refer to this case as the symmetric equilibrium.

Definition 2 *A symmetric equilibrium is a market equilibrium where*

$$d\alpha_t^i = d\alpha_t,\ K_t^i = K_t,\ C_t^i = C_t,\ \text{and}\ dK_t^i = dK_t.$$

In the rest of the paper we omit the time subscript t or τ whenever it does not cause any confusion.

3.6 Market Equilibrium

We solve for the market equilibrium in this section. As we show, in this economy consumption before Poisson event is strictly suboptimal, thus $d\alpha_t^i = d\alpha_t = 0$ always.

3.6.1 Ex post equilibrium prices

Consider the final date. All firms who are hit by the skill-shock sell their capital, because their marginal valuation of capital drops to zero. And, as long as the capital price $\widehat{p}$ is less than R, all cash holders who are not hit by the shock are happy to exchange all their cash to capital. Appealing to the law of large numbers, the market clearing condition at the final date implies that

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

We still need to ensure that $c \leq R$. Later we show that the full support of c is endogenous as firms build (dismantle) capital whenever the aggregate cash is sufficiently high (low). For simplicity, we restrict the parameter space to ensure that the condition $c \leq R$ holds always in equilibrium.

3.6.2 Ex ante equilibrium values, prices, and investment polices

In the ex-ante interval determining the equilibrium objects is more subtle. As we state in the next lemma, our formalization has a number of useful properties. Namely, the only relevant aggregate state variable is the cash-to-capital ratio, and the value function of any individual firm is linear in their capital and cash holdings.

Lemma 1 *Let $J\left(K^i, C^i, K, C\right)$ be the value function of firm i who 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,*

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

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

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

Firms build capital whenever the capital price p reaches the cash cost h, and dismantle capital whenever the price falls to the liquidation value l. Define c_h^* (c_l^*) as the endogenous threshold of the aggregate cash-to-capital ratio where firms start to build (dismantle) capital, then we must have

$$\frac{v(c_h^*)}{q(c_h^*)} = h, \text{ and } \frac{v(c_l^*)}{q(c_l^*)} = l. \tag{13}$$

Moreover, the linear technology implies that c_h^* and c_l^* are reflective boundaries of the process c. Therefore, based on (12), the aggregate cash-to-capital ratio c must fluctuate in the interval $[c_l^*, c_h^*]$, with a dynamics of

$$dc = \sigma dZ_t - dU_t + dB_t, \tag{14}$$

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. Moreover, the standard properties of reflective boundaries imply the following smooth pasting conditions for our value functions:

$$v'(c_h^*) = q'(c_h^*) = q'(c_l^*) = v'(c_l^*) = 0. \tag{15}$$

3.6.3 Characterizing the market equilibrium

Now we turn to characterizing the value functions $v(c)$ and $q(c)$ in the range $c \in [c_l^*, c_h^*]$. We give here a draft and show the details in the Appendix. Because of Lemma 1, firms are indifferent in the composition of their portfolios, and we can consider the value function of a firm who holds only capital and another firm with cash only. The cash-holding firm gives an ODE for $q(c)$:

$$0 = \frac{\sigma^2}{2} q'' + \frac{\xi}{2}(u - q(c)) + \frac{\xi}{2}\left(\frac{R}{c} - q(c)\right), \tag{16}$$

and the capital-holding firm, given $q(c)$, yields the ODE for $v(c)$:

$$0 = q'(c)\sigma^2 + \frac{\sigma^2}{2} v''(c) + \frac{\xi}{2}(uc - v(c)) + \frac{\xi}{2}(R - v(c)). \tag{17}$$

These ODEs are Hamilton-Jacobi-Bellman (HJB) equations given the dynamics of the state c. We first explain the terms without ξ in both ODEs. For the cash value q equation (16), $\frac{\sigma^2}{2}q''$ captures the impact of changing c; and a similar term shows up in the capital value v equation (17). In addition, we have $q'(c)\sigma^2$ in equation (17) because of the Ito's correction term. To see this intuitively, the capital itself generates random cash flows σdZ_t that are correlated with the aggregate state $c_{t+dt} = c_t + \sigma dZ_t$ (see 14), and the expected value of these cash flows is

$$\mathbb{E}_t \left[q\left(c+\sigma dZ_t\right)\sigma dZ_t\right] = \mathbb{E}_t \left[q'\left(c\right)\sigma^2 \left(dZ_t\right)^2\right] = q'\left(c\right)\sigma^2 dt.$$

The terms multiplied by the intensity ξ describe the change in expected utility once the final date arrives. The first of these terms in equation (16) shows that, once a firm holding a unit of cash is hit by a skill shock, her value jumps to u from $q(c)$. Otherwise, the second term says that she uses the unit of cash to buy $1/\widehat{p} = 1/c$ unit of capital, so her utility jumps to R/c from $q(c)$. The interpretation in equation (17) is analogous.

Define the constant $\gamma \equiv \sqrt{2\xi}/\sigma$ which is important for our analysis. We solve the ODE system in (16)-(17) in closed-form, which admits the following general form:

$$q\left(c\right) = \frac{u}{2} + e^{-c\gamma}A_1 + e^{c\gamma}A_2 + R\frac{\gamma}{2}\frac{-e^{c\gamma}\,\mathrm{Ei}\left(-\gamma c\right) + e^{-c\gamma}\,\mathrm{Ei}\left(c\gamma\right)}{2}, \tag{18}$$

and

$$v\left(c\right) = R + \frac{uc}{2} + e^{c\gamma}\left(A_3 - cA_2\right) - e^{-c\gamma}\left(A_4 + cA_1\right) + cR\frac{\gamma}{2}\frac{\left(e^{\gamma c}\,\mathrm{Ei}\left(-\gamma c\right) - e^{-c\gamma}\,\mathrm{Ei}\left(\gamma c\right)\right)}{2}, \tag{19}$$

where $\mathrm{Ei}(x)$ is the exponential integral function defined as

$$\mathrm{Ei}\left(x\right) \equiv \int_{-\infty}^{x} \frac{e^t}{t}dt,$$

and the constants A_1-A_4 are determined from boundary conditions in (15).

Finally, we determine the endogenous investment/liquidation thresholds c_l^* and c_h^* using (13). 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 describe the basic properties of this equilibrium. We summarize this result below and give formal proof in the Appendix.

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

1. *firms do not consume before the final date;*
2. *each firm in each state $c \in [c_l^*, c_h^*]$ is indifferent in the composition of her portfolio;*
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 iff $c = c_h^*$ and cover the negative cash shocks by liquidating a sufficient fraction of capital iff $c = c_l^*$;

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 ex ante price is $p(c) = v(c)/q(c)$;*

5. *ex post, a firm hit by the skill shock sells all her capital to firms who are not hit by the shock for the price $\widehat{p} = c$; and*

6. *$q(c)$ is monotonically decreasing, $v(c)$ is monotonically increasing, and $p(c)$ is monotonically increasing.*

Because all firms are ex ante indifferent how much cash or capital to hold at the equilibrium prices, the properties of our market equilibrium leave individual portfolios undetermined. The symmetric equilibrium picks the equilibrium where all individual portfolios are the same.[16]

3.6.4 Investment waves

The thick, solid lines on panels A-E of Figure 2 illustrate the properties of the market equilibrium. While panels A-C show the functions $p(c), v(c), q(c)$ describing the price of capital, the value of cash and the value of capital in equilibrium, panels D-E depict the cash-to-capital ratio and the investment/disinvestment activity along one particular sample path.

We can think of the cash-to-capital ratio c as "aggregate liquidity," and the time with high (low) aggregate liquidity in which capital are built (dismantled) as a boom (downturn). In our model investment takes a simple threshold strategy so that investment (disinvestment) occurs only at c_h^* (c_l^*). However, we believe the resulting clustered investment and disinvestment activities depicted on 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 aggregate level of liquidity. This is illustrated in panel D. This particular sample path starts with a series of positive shocks, which increases the marginal value of capital and decreases the marginal value of cash. Thus, the price of capital increases along this path (not shown)[17] and the expected return of holding capital rises. This is so, because in these states the probability that the economy slips into a downturn (so that capital has to be dismantled) is low. Thus, firms are willing to hold capital even if its expected excess return over cash is relatively low. Around period 120 the price hike reaches the cost of building capital, h, which triggers investment (as shown in Panel E) to keep the cash-to-capital ratio at c_h^*. For symmetric reasons, as a series of subsequent negative shocks decrease

[16] Note that there are parameter combinations for which the equilibrium described in Proposition 1 does not exist. However, in those cases, typically we find another type of equilibrium to arise with very similar features. The main difference is that in this alternative equilibrium firms do not dismantle capital at a single point when $c = c_l^*$, but do so gradually as long as the cash-to-capital ratio is within a range $c \in [c_l^*, c_x]$. Within this range, the price is flat at $p(c) = l$. Given that otherwise this alternative equilibrium differs little to the one described in Proposition 1, we relegate its characterization to Appendix B.

[17] As 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^*.

aggregate liquidity, the marginal value of cash and expected returns rise, while the value of capital falls, which depresses prices. Around period 700, these negative shocks trigger disinvestment in capital, keeping the cash-to-capital ratio above c_l^*.

The constant $\gamma = \sqrt{2\xi}/\sigma$ parametrizing functions $v(c)$, $q(c)$ in (19) and (18) plays an important role. Intuitively, γ drives the relative importance of the ex post payoffs for ex ante decisions. A high switching intensity ξ or a low σ reduces the chance of large interim shocks, and hence ex post payoffs are important determinants of ex ante decisions. The following results on the investment and disinvestment thresholds are useful in understanding the intuition behind our results.

Proposition 2 *In the market equilibrium,*

1. $c_h^* > h, c_l^* < l$; *and*
2. *as* $\gamma \to \infty$, $c_h^* \to h$ *and* $c_l^* \to l$.

Consider the last result. When γ grows without bound the firm's ex ante decisions are almost solely determined by ex post payoffs. That is, in that limit our dynamic set up is very close to our simple static example in Section 2. Indeed, Case 1 and Case 3 in our simple example in Section 2 stated that agents build capital whenever the initial endowment of cash relative to capital is larger than $h = 2$ (the cost of building a unit of capital) and dismantle capital whenever the endowment of cash relative to capital is smaller than $l = \frac{1}{2}$ (the benefit of liquidating a unit of capital). This result directly corresponds to the second statement in this proposition. Away from this limit, when building a unit of capital, the firm considers also the risk of reaching the low liquidity state when this unit is dismantled inefficiently. Thus, she decides to build capital only at a higher threshold, i.e., $c_h^* > h$. The analogous argument implies that $c_l^* < l$.

Clearly, the levels of the endogenous investment and disinvestment thresholds, c_h^* and c_l^*, determine the main characteristics of investment waves in our economy. The main question of this paper is whether these thresholds at their welfare maximizing level in the market equilibrium. This is what we study in the next section.

4 Externalities

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

4.1 Constrained efficient benchmarks

We study the constrained efficient allocation where the planner takes into account the technological constraint that the aggregate cash has to be kept non-negative by liquidating capital if necessary.[18]

[18] Without this technological constraint, condition (9) implies that the planner should convert any amount of cash to capital.

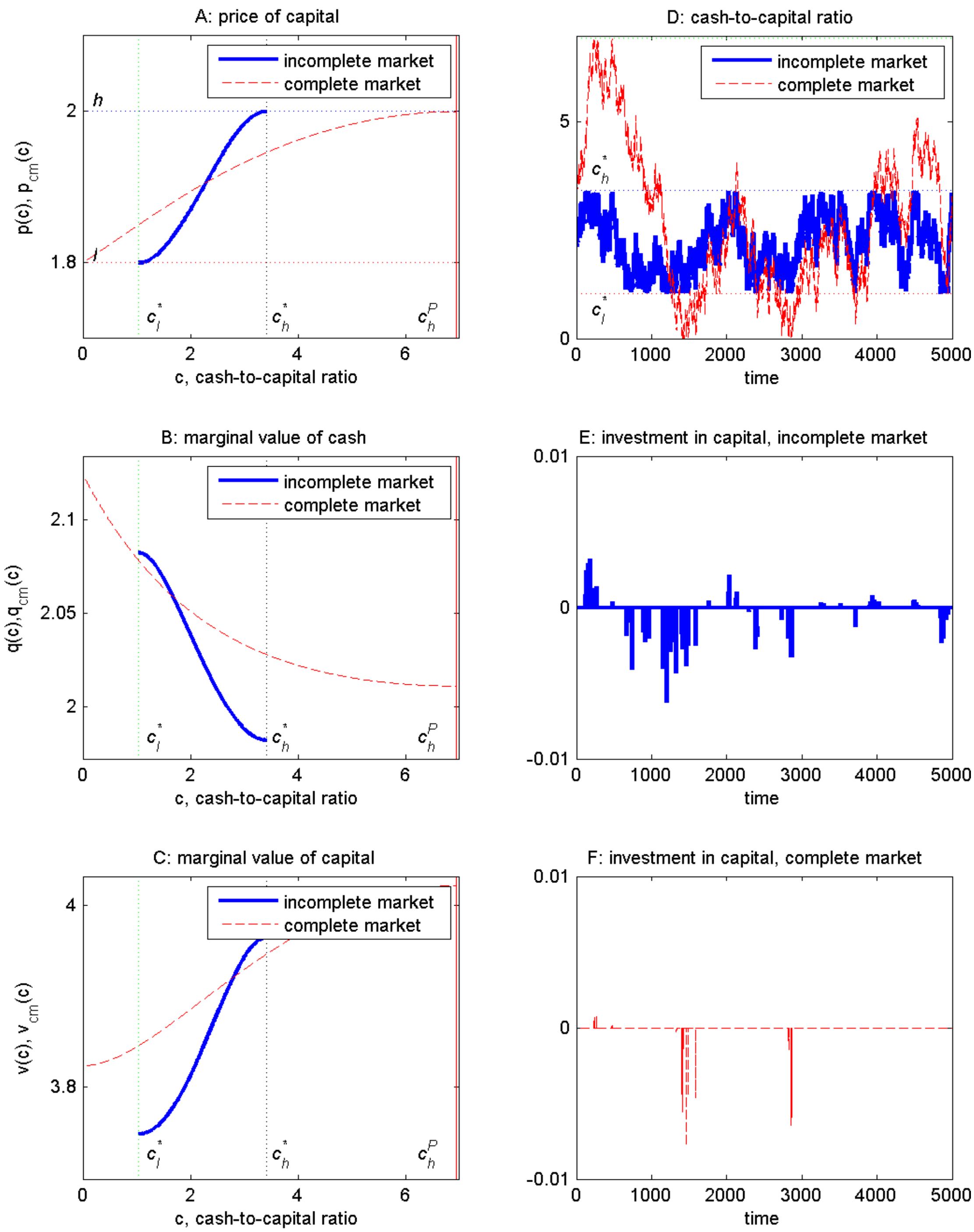

Figure 2: Panels A-C depict the price of capital, value of cash and the value of capital. The solid, vertical line on the right of each graph is at the investment threshold in complete markets and in social planners solution, c_h^P, while the two dashed vertical lines are the disinvestment and investment thresholds in our baseline case, c_l^*, c_h^*. The horizontal lines on the Panel A are at the levels of l and h. Panels D-F depict a simulated sample path. Vertical lines on panel E from top to bottom are c^*, c_h and c_l. Each panel shows objects both for the baseline model with incomplete markets (thick solid curves) and the benchmark of complete markets (thin, dashed curves). Parameter values are $R = 4.1$, $\sigma^2 = 1$, $\xi = 0.1$, $u = 2$, $l = 1.8$ and $h = 2$.

We will consider two benchmark economies which both produces the same constrained efficient outcome.

First, we consider a social planner who can dictate investment policies but without knowing the realization of skill shocks of market participants. 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. Second, we consider a decentralized economy where markets are complete so that each individual firms have access to the same investment opportunities (potentially through contracting), which allows us to characterize the asset prices without inefficiency.

4.1.1 Social planner's problem

Denote by $J_P(K, C)$ the social planner's value function. The planner can decide when to build and dismantle capital. Thus, she optimally regulates the c process subject to the constraint that the cash level C must stay non-negative.

Ex post, cash (capital) always ends up in the firms with (without) skill shock at the market clearing price $\widehat{p} = c$. Therefore, the total value ex post is[19]

$$KR + Cu. \tag{20}$$

Thus, given the aggregate state pair (K, C), the social planner is solving

$$J_P(K, C) = \max_{dK} \mathbb{E}\left[\int_0^\infty \xi e^{-\xi\tau}(K_\tau R + C_\tau u)\, d\tau\right] \equiv K j_P(c) \tag{21}$$

subject to the constraint $C \geq 0$ and (11). In the second equality, we have invoked the scale-invariance to define $j_P(c)$ to be the social planner's value per unit of capital.

Because of the linear technology of building and liquidating capital, regulation with reflective barriers on c is optimal (Dixit (1993)). That is, there exists low and high thresholds c_l^P and c_h^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 $c = c_l^P$ $(c = c_h^P)$ at the lower (upper) threshold.

Consider the social value given the particular policy that c is regulated by an arbitrary reflecting barriers $c_l < c_h$. Define the corresponding (scaled) social value[20] as $j_P(c; c_l, c_h)$ so that

$$K j_P(c; c_l, c_h) \equiv \mathbb{E}\left[\int_0^\infty \xi e^{-\xi\tau}(K_\tau R + C_\tau u)\, d\tau \mid c_l, c_h\right]. \tag{22}$$

[19] Given (K, C), the representative firm hit by the skill-shock gets $u(\widehat{p}K + C) = u(cK + C) = 2Cu$, while if she is not affected she gets $RK + RC/(\widehat{p}) = RK + \frac{R}{c}C = 2KR$. Hence, in expectation the total welfare is $RK + uC$.

[20] In our model the total welfare is state dependent. Thus, we can make a distinction between policies which improve total welfare at some states (e.g. in downturns only), and policies which improve welfare everywhere. This is because for any current c_t the total welfare function factors in the effect of the policy in each other state. For instance, the probability of arriving in a given state depends on the current c_t, i.e., in a downturn, a boom looks less likely than a continuing downturn.

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

$$0 = \frac{\sigma^2}{2} j_P'' + \xi \left(R + uc - j_P\right), \text{ for } c \in (c_l, c_h), \tag{23}$$

where we suppressed the arguments of j_P and $j_P'' = \frac{\partial^2 j_P}{\partial^2 c}$; and at the reflective barriers c_l, c_h the smooth pasting conditions must hold:

$$\frac{\partial \left[K j_P\left(c_l; c_l, c_h\right)\right]}{\partial K} = l \frac{\partial \left[K j_P\left(c_l; c_l, c_h\right)\right]}{\partial C}, \text{ and } \frac{\partial \left[K j_P\left(c_h; c_l, c_h\right)\right]}{\partial K} = h \frac{\partial \left[K j_P\left(c_h; c_l, c_h\right)\right]}{\partial C}. \tag{24}$$

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.[21] The ODE (23) has a closed from solution

$$j_P\left(c; c_l, c_h\right) = R + uc + D_1 e^{-\gamma c} + D_2 e^{\gamma c}. \tag{25}$$

For any fixed c_l, c_h, we solve for the constants D_1 and D_2 based on (24).

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

$$j_P(c) \equiv j_P\left(c; c_l^P, c_h^P\right) = \max_{c_l, c_h} j_P\left(c; c_l, c_h\right). \tag{26}$$

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

$$\left.\frac{\partial^2 \left[K j_P\left(C/K; c_l^P, c_h^P\right)\right]}{\partial K \partial C}\right|_{C = K c_h^P} = h \left.\frac{\partial^2 \left[K j_P\left(C/K; c_l^P, c_h^P\right)\right]}{\partial^2 C}\right|_{C = K c_h^P}, \tag{27}$$

which we can rewrite as

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

For the lower barrier c_l^P, at the optimal choice the constraint $C \geq 0$ might bind. Thus, the supercontact condition is

$$\left.\frac{\partial^2 \left[K j_P\left(C/K; c_l^P, c_h^P\right)\right]}{\partial K \partial C}\right|_{C = K c_l^P} \leq l \left.\frac{\partial^2 \left[K j_P\left(C/K; c_l^P, c_h^P\right)\right]}{\partial^2 C}\right|_{C = K c_l^P}, \text{ for } c_l^P \geq 0 \tag{28}$$

with complementarity. The next proposition shows that the optimal lower threshold is $c_l^P = 0$, and the optimal upper threshold is the unique solution of a simple analytical equation.

[21] See Dixit (1993) for a detailed argument.

Proposition 3 *The social planner dismantles capital whenever c reaches 0 and builds capital whenever c reaches the endogenous investment threshold $c_h^P > 0$. The investment threshold c_h^P is given by the unique solution of*

$$\frac{R - hu}{R - lu}\left(e^{c_h^P \gamma}(1 + l\gamma) - (1 - l\gamma)\, e^{-c_h^P \gamma}\right) - 2\gamma\left(c_h^P + h\right) = 0. \tag{29}$$

Under the optimal policy, the optimal social value $j_P(c)$ is concave over $\left[0, c_h^P\right]$, and $j_P(c) \leq R + uc$.

4.1.2 Properties of constrained efficient solution

To understand the optimal choice by the social planner, it is useful to consider the following comparative statics.

Proposition 4 *The socially optimal investment threshold c_h^P*

1. *is converging to 0 as $\gamma \to \infty$, and decreasing in γ given that $\gamma > \hat{\gamma}$ for a given $\hat{\gamma}$,*
2. *decreasing in l and R and increasing in h,*
3. *approaching to ∞ as $R \to uh$.*

An unboundedly large γ is due to either a large ξ (i.e., the final date arrives very fast) or a small σ (i.e., the interim shocks are small). Both imply that the social planner puts almost zero weight on the possibility that a sequence of negative cash flow shocks force her to dismantle capital at the lower threshold. Thus, as suggested in the first statement in Proposition 4, the social planner does not store any cash (i.e., convert any cash to capital immediately). This is in line with condition (9) which ensures that creating capital socially dominates holding cash. When γ is not that large, the social planner puts positive weight on the possibility of forced liquidation, thus stores some cash for buffering purposes. The higher the l, i.e. the lower the cost of liquidating capital for cash, the less the cash buffer that the social planner is building up. This reduces the upper bound c_h^P of the cash buffer, as stated in the second result in Proposition 4. Finally, when $R - uh$ is sufficiently small, the cash is almost as valuable as capital. Thus, it is optimal to store cash always, so that c_h^P increases without bound.

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. Increasing the lower threshold or decreasing the upper threshold, relative to those of the social planners' solution respectively, unambiguously decreases welfare everywhere.

Proposition 5 *For any $c_h < c_h^P$ and $c_l > 0$, we have*

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

As a result, Proposition 5 implies that, whenever $c_h^* < c_h^P$ and $c_l^* > 0$, firms underinvest in recessions and overinvest in booms in capital in our market equilibrium. Recall that in a symmetric equilibrium, $Kj_P(c; c_l, c_h)$ is also the ex ante value of the representative firm. Hence, a welfare-increasing policy constitutes an ex ante Pareto improvement with respect to the symmetric market equilibrium.

Finally, is 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:

$$T(c; c_l, c_h) \equiv \mathbb{E}\left[\int_0^\tau \frac{|dK_t|}{K_t} dt\right]. \tag{30}$$

Proposition 6 *For any c_h and c_l, we have*

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

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

4.1.3 Solution with complete market

Consider the variant of our decentralized model where markets are complete so that constrained efficient solution is achieved. There are many different ways to model complete markets. For instance, if individual skill states are contractible, then allowing for trading Arrow-Debreu securities will restore the investment incentives.[22] In the context of our model with investment, we simply assume that the ex post proceeds R and u are fully pledgeable so that individual firms can enjoy the investment opportunities of others. More specifically, individual firms with skill-shock can hire firms without skill shock to realize the full marginal return of R from capital, and similarly firms without new investment opportunity can lend their cash and receive the new investment benefit u. Thus, ex post all firms can invest their cash-holdings to the new technology and none of them loses their expertise in operating the capital, which effectively eliminates the ex post heterogeneity among firms.

We refer to this variant of economy as the complete market economy or the subscript cm. By following the same derivation, the value of cash $q_{cm}(c)$ and the value of capital $v_{cm}(c)$ solve the ODE system

$$0 = \frac{\sigma^2}{2} q_{cm}''(c) + \xi (u - q_{cm}(c)), \tag{31}$$

$$0 = q_{cm}'(c)\sigma^2 + \frac{\sigma^2}{2} v_{cm}''(c) + \xi (R - v_{cm}(c)). \tag{32}$$

[22] However, if individual types are not contractible, misreporting will occur. For formal arguments in the two period context, see Appendix C.

Relative to (16) and (17), the difference lies in the ξ term capturing the ex post event: for instance, in the complete market, each unit of capital realizes a value of R in (32), while in (17) with half probability the firm with new investment opportunities sells that unit of capital at a price of $\widehat{p} = c$ and obtain uc. The following statement characterizes the equilibrium in this variant of our model.

Proposition 7 *In the symmetric complete market equilibrium,*

1. *firms do not consume before the final date;*
2. *each firm in each state $c \in \left[0, c_h^P\right]$ is indifferent in the composition of her portfolio;*
3. *each firm use every positive cash shock to build capital iff $c = c_h^P$ and dismantle the capital (to cover negative cash flow shocks) iff $c = 0$;*
4. *we have*

$$q_{cm}(c) = u + e^{-c\gamma} B_1 + e^{c\gamma} B_2, \tag{33}$$
$$v_{cm}(c) = R + e^{c\gamma}(B_3 - cB_2) + e^{-c\gamma}(B_4 - cB_1), \tag{34}$$

where B_1, B_2, B_3, B_4 and c_h^P is given by boundary conditions

$$\frac{v_{cm}\left(c_h^P\right)}{q_{cm}\left(c_h^P\right)} = h, \ \frac{v_{cm}(0)}{q_{cm}(0)} = l, \ v'_{cm}\left(c_h^P\right) = q'_{cm}\left(c_h^P\right) = v'_{cm}(0) = 0, \tag{35}$$

and the welfare of representative firm $v_{cm}(c) + cq_{cm}(c)$ achieves $j_P(c)$ given in (26); and

5. *$v_{cm}(c)$ is increasing in c, $q_{cm}(c)$ is decreasing in c, and $p_{cm}(c) \equiv \frac{v_{cm}(c)}{q_{cm}(c)}$ is increasing in c.*

In this complete market economy, because individual firms have the same objective as the social planner, the market implements the constrained efficient solution. We later also call this complete market equilibrium the constrained-efficient equilibrium. The qualitative properties of the constrained-efficient economy is similar to the market solution of our baseline economy in Proposition 1. In particular, as illustrated by the thin, dashed curves on Figure 2, when the cash-to-capital ratio decreases, the price falls and the capital trades with a significant liquidity premium. Thus, the fluctuation across booms with high prices and high investment and downturns with cheap capital and large liquidity premium is consistent with a constrained efficient economy.

4.2 Two-sided inefficiency

Now we show that there is a large subset of parameters where firms overinvest in productive capital in booms and underinvest in downturns. We refer to this case as *two-sided inefficiency.* Figures 2 illustrates such a case, where the dashed (solid) vertical lines show the thresholds of the market equilibrium (constrained-efficient equilibrium). In particular, in the market equilibrium firms dismantle capital when still some cash is around, $c_l^* > 0$; and firms create new capital at a lower liquidity level than the social planner would do, $c_h^* < c_h^P$. Proposition 6 and 5 imply that in this case the resulting investment waves are too volatile (illustrated by panels E and F of Figures 2) and any policy that raises (decreases) the upper (lower) threshold would unambiguously increase total welfare.

4.2.1 Existence of two-sided inefficiency and intuition

While the social planner would dismantle capital only when all cash in the economy is gone, the market solution in Proposition 1 has $c_l^* > 0$ always. That is, in the market equilibrium firms dismantle capital when the social planner would still avoid it. In this sense there is underinvestment in productive assets or, equivalently, over hoarding of liquidity in a recession. On the other side, in booms firms in the market equilibrium could over or underinvest in capital, i.e., $c_h^* \gtrless c_h^P$, depending on the parameter values.

Proposition 8 *We have the following results.*

1. *In market solution firms dismantle capital before the aggregate liquidity reaches zero, i.e., $c_l^* > 0$. Hence market solution implies underinvestment in capital and over hoarding of cash in downturns.*

2. *Keeping u, l, h, R fixed, there is a threshold $\hat{\gamma}$ that if $\gamma > \hat{\gamma}$, we have $c_h^* > c_h^P$. That is to say, the market solution implies underinvestment in capital and over hoarding of cash in booms as well.*

3. *If the difference between the productivity of capital and that of the new investment opportunity, $R/h-u$, is sufficiently small, then we have $c_h^* < c_h^P$. That is to say, the market solution implies overinvestment in capital in booms.*

The general intuition behind our mechanism is as follows. As we have already observed, the planner's choice of the investment and disinvestment thresholds is driven by a simple trade-off. While capital is more productive than cash, a limited buffer of cash is useful to avoid the inefficient liquidation of capital in the case of a series of adverse shocks requiring maintenance. In the market solution, while the same trade-off is present, there is an additional force as highlighted in the simple example in Section 2. The ex post market clearing price not only moves resources to the most efficient hands but also allocates the rent among different firms, and this distorts the private investment incentives ex-ante. Importantly, the direction of price distortion can change with the state of the economy, leading to underinvestment in booms and overinvestment in recessions. To see the intuition behind the conditions implying a two sided inefficiency, let us compare the private and social value of capital in our economy. Because the representative firm sells the capital at the ex post price $\widehat{p}$ given a skill-shock, the private (ex post) value of a capital is $\frac{1}{2}u\widehat{p}+\frac{1}{2}R$, while the social (ex post) value of a capital is always R. Therefore, whether the representative firm overvalues the capital compared to the planner depends on whether $\widehat{p} > R/u$, i.e., whether the private marginal rate of substitution is above the social marginal rate, as shown in Section 2. Given that $\widehat{p} = c$ fluctuates in the interval $[c_l^*, c_h^*]$ we should expect overinvestment in booms and underinvestment in recessions whenever

$$uc_l^* < R < uc_h^*. \tag{36}$$

Consistent with Proposition 8, the first inequality in (36) is always satisfied because $uc_l^* < ul < uh < R$, whereas the second inequality might or might not hold.

As emphasized earlier in Section 2, in our model trading moves the assets (cash) to the hands with the highest profitability, leading to ex post efficient resource allocation even in the (incomplete) market equilibrium. As a result, given the state (K, C) the social planner does not change the welfare of the representative firm ex post. Instead, by changing the thresholds c_l and c_h, the social planner influences the future distribution of c (or, equivalently, the joint distribution of (K, C)), which improves the representative firm's ex ante welfare according to Proposition 5.

To illustrate the effect of the social planner through changing the distribution of c, on Figure 3 we depict equilibrium objects generated by a simulation of 100 sample paths for both the market equilibrium (thick, solid curves) and the incomplete market benchmark (thin, dashed curves). In each panel we take the average over the 100 sample paths in each period. For the cash-to-capital ratio, c_t, the level of capital, K_t, and cash, C_t, we depict the ex-ante path, that is, the realization conditional on that the capital good has *not* matured before period t. For the consumption, $RK_\tau + uC_\tau$, we depict the ex-post path, that is, the realization conditional on the event that the capital good matures exactly at period t. In this way, we get an approximation of the expected path conditional on any given realization of the Poisson shock determining the final date $\tau = t$;[23] the unconditional objects can be obtained by weighting the conditional objects by the probability distribution imposed by the Poisson structure. In panels A, C and D, we normalize by the realization in time 0. Parameters are as at Figure 2, so the market equilibrium exhibits two-sided inefficiency.

It should not be surprising that our ex-ante economy, while stationary when normalized by the level of capital, is shrinking. It is a simple consequence of no drift in (11) and the costly adjustment of capital.[24] More importantly, in line with our two-sided inefficiency and Proposition 5 , we see that the level of capital, consumption and cash all has lower rate of growth (i.e., higher rate of shrinking) in the market equilibrium. Note also, that the average realizations in the market equilibrium are below their counterpart in the efficient benchmark case for any realization of the Poisson event τ. Therefore, the unconditional expected consumption, level of capital and cash at period zero must also be lower in the market equilibrium. The reason behind this pattern is stated in Proposition 6: under two-sided inefficiency, agents adjust the level of capital too often, losing two much in the process.

[23] For example, suppose the final date is $\tau = 200$. Then an approximation of the expected consumption at the final date is the point corresponding to $\tau = 200$ on Panel A , and an approximation of the expected ex ante path of length 200 of K_t, C_t, and c_t are given by the points corresponding to $t < 200$ on Panels B-D.

[24] One could introduce a positive growth rate in (11) easily, which still gives an analytical solution for the equilibrium. However, this treatment would make some of the proofs cumbersome, without providing additional insights.

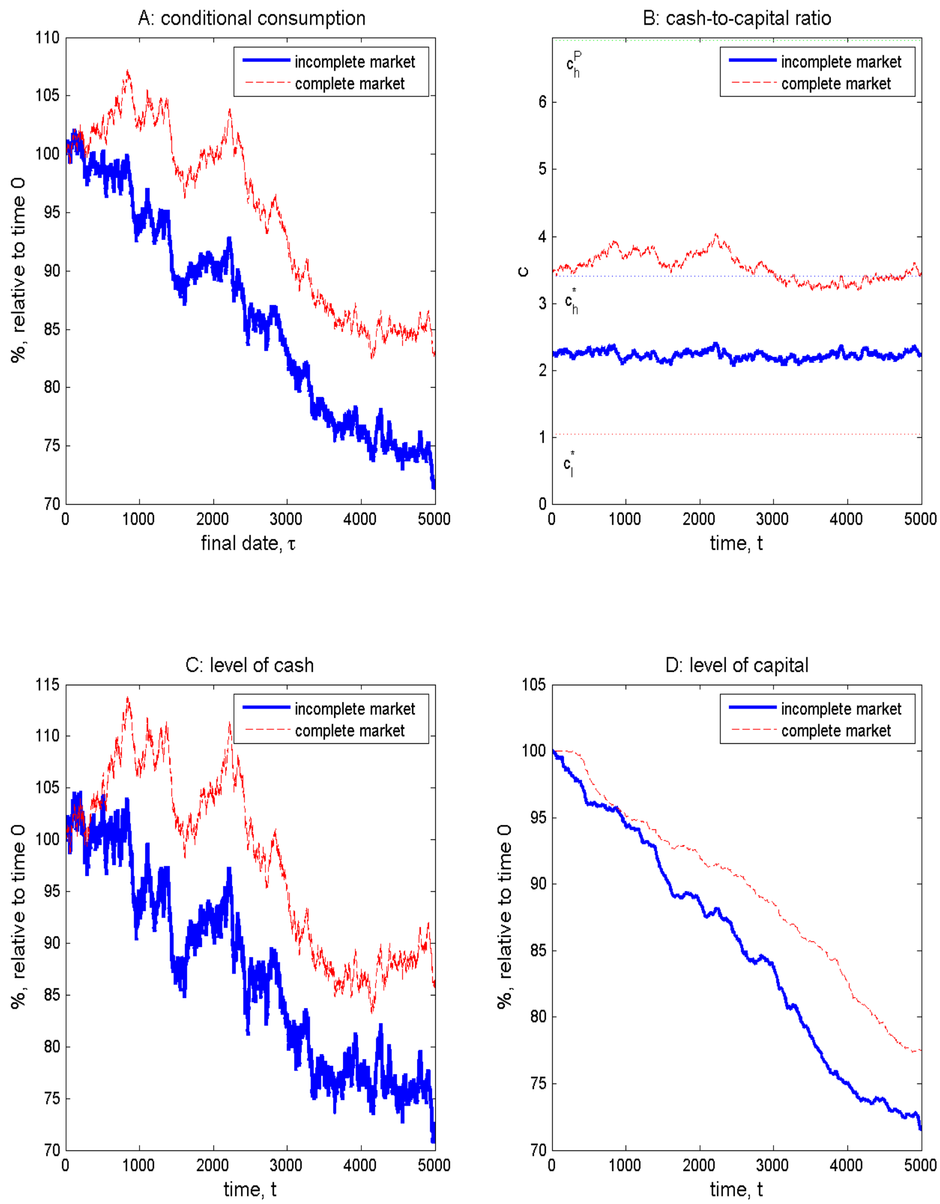

Figure 3: Each panel depicts the mean of 100 sample paths of the equilibrium objects for the baseline model with incomplete markets (thick solid curves) and the benchmark of complete markets (thin, dashed curves). Panel A depicts consumption $RK_\tau + uC_\tau$ conditional on the final date τ, while Panel B-D depicts , c_t, C_t and K_t, respectively, conditional on $t < \tau$. Parameter values are $R = 4.1$, $\sigma^2 = 1$, $\xi = 0.1$, $u = 2$, $l = 1.8$ and $h = 2$.

At a deep level, our mechanism is in line with the welfare effects of pecuniary externalities identified by Geanakoplos and Polemarchakis (1985). That seminal paper shows that when markets are incomplete and, consequently, prices do not equate marginal rate of substitution of firms, then pecuniary externalities might have first-order effects on welfare. Our mechanism works by the same logic: Because of the missing market, price $\widehat{p}$ does not serve its Walrasian function of signalling the relative social value of different goods. This makes firms' ex ante investment decisions socially inefficient. Our main contribution relative to Geanakoplos and Polemarchakis (1985) and the subsequent literature is to point out that the distortion implied by the pecuniary externality can change sign with the state of the economy, because the distortion in the price can change sign.

Proposition 8 translates the above intuition in terms of the deep parameters of the model. Vaguely speaking, Proposition 8 suggests that there are two-sided inefficiencies if $R/h - u$ is small, i.e., the profitability of the existing capital technology is close to that of the new investment opportunity. As pointing out the two-sided inefficiency is the major novelty in our paper, from now on we focus mostly on this case.

5 Applications and extensions

We suggest a number of applications for our model. As a main policy application, in part 5.2 we analyze the dynamic effects of asymmetric interventions relating our results to the debate on asymmetric monetary policy often referred to as the Greenspan's put. Before that, in the next part we suggest three further applications. First, we explore the connections of our findings to industry-specific boom and bust patterns. Second, we relate our findings to the cross-country evidence on growth, financial development and the composition of investment. Finally, we connect our results to the housing cycle. As an extension, in part 5.3, we show that the two-sided inefficiency holds in an alternative (and perhaps more realistic) specification where a random group of agents receive the investment opportunity in each time instant.

5.1 Sectoral and aggregate investment waves

Industrial 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 like the tech-bubble in the nineties and the biotech bubble in the eighties. They also show that only in competitive industries there is a negative correlation between relative valuation or investment and subsequent profits or earnings. That is, only in competitive industries boom patterns are not validated by high subsequent earnings. This is consistent with our model as our inefficiency is driven by a pecuniary externality. If there are only few firms present in an industry who take into account the price effects of their own actions, our inefficiency would attenuate.[25] An additional prediction of our model is that this negative

[25] We show this formally in Appendix C in the context of our two-period example.

correlation should be stronger in those competitive industries where contracting frictions are likely to be more severe.

Real estate and housing cycles We can also apply our results in the context of the boom and bust pattern in real estate development and house prices.[26] Our mechanism suggests that the volume of construction in a boom is inefficiently high: banks/investors invest in real estate developments instead of liquid financial assets expecting to be able to sell the real estate for a high price in case they find a new investment opportunity.[27] One suggestive sign of this inefficiency is the frequently observed phenomenon of "overbuilding," that is, periods of construction booms in the face of rising vacancies and plummeting demand.[28] On the other hand, in recessions, our model suggests that banks/investors hold inefficiently high level of liquid assets, expecting to be able to buy cheap real estate in case a group of distressed investors have to dismantle their holdings. This precautionary behavior is consistent with findings of financial institutions cutting their lendings and overbuilding their cash reserve in the recent financial crisis.[29]

Financial development and investment dynamics Finally, our model also suggests a novel rationale for stylized facts on the connection of financial development and investment dynamics. Aghion et al. (2010) is a useful starting point. They argue that the level of financial development has a first order effect on the composition of investment and its variation over the business cycle. In particular, they decompose aggregate investment to structural and other investment arguing that structural investment is a proxy for investment in longer-term, more productive, but, in the short-term, riskier 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 activity is an important channel how the lack of financial development hinders growth.

[26] Shiller (2007) illustrates this pattern by the cyclicality of the residential investment to GDP ratio. He points out that cycles in this ratio correspond closely to the recessions after 1950, typically peaking few quarters before the start of the recession. This pattern was not observed before the 2000-01 recession but was observed again before the 2007-2009 recession.

[27] Related arguments were made in connection to the development of Japan.

> It took most Japanese banks years to whittle down the tens of billions of dollars in unrecoverable loans left on their books after the collapse of a real estate bubble in Japan's overheated 1980's. They finally succeeded in the last two or three years [...]But analysts criticize most banks for failing to find new, more profitable – and less risky – ways of doing business. Instead, analysts say many have gone back to lending heavily to real estate development companies and investment funds, as the rebounding economy has touched off a construction boom in Tokyo. "If the economy stalled, Japanese banks would have a bad loan problem all over again," said Naoko Nemoto, an analyst for Standard & Poor's in Tokyo. Ms. Nemoto estimates that banks loaned 1.6 trillion yen ($14 billion) to real estate developers in the six months that ended last September – half of all new bank lending in that period." (The New York Times, January 17, 2006, pg.4)

[28] See Wheaton and Torto (1990) and Grenadier (1996) for alternative explanations of overbuilding. Overbuilding was also observed before the 2007-2009 recession in the sense that rental vacancies peaked in 2004, before the peak of the construction boom. (See http://www.census.gov/hhes/www/housing/hvs/historic/index.html.)

[29] See Heider, Hoerova and Holthausen (2009), Ashcraft, Mcandrews and Skeie (2011) and Acharya and Merrouche (2012) for U.S. and international evidence.

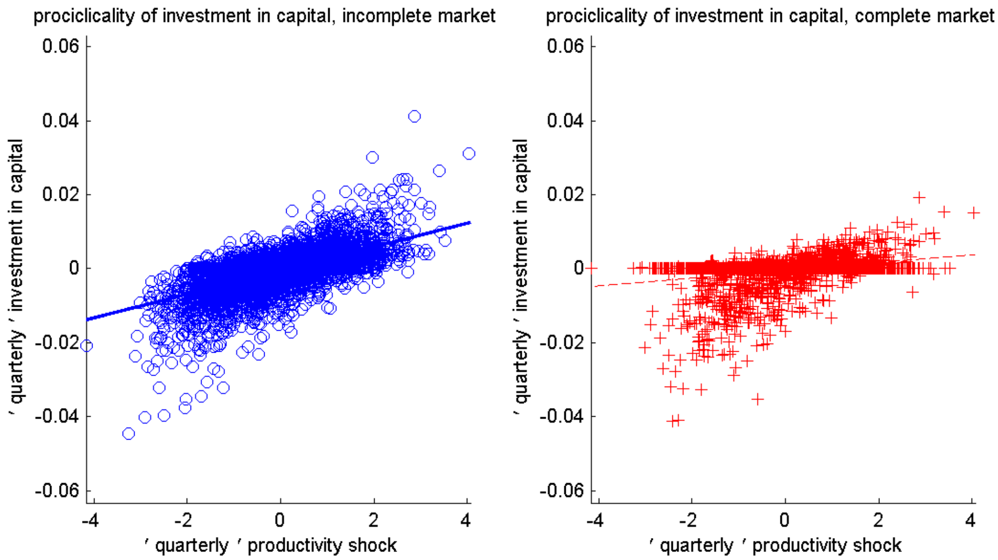

Figure 4: For each panels we used a simulation of 100 paths as illustrated in Figure 3. Here, we formed the means of 100 subsequent observations of σB_t (productivity shock, x-axis) and dK_t (investment in capital, y-axis) in each sample path. The panel show the corresponding scatter plot and a regression line for the baseline case of incomplete market (left) and benchmark case of complete market (right). Parameter values are $R = 4.1$, $\sigma^2 = 1$, $\xi = 0.1$, $u = 2$, $l = 1.8$ and $h = 2$.

Our results are broadly consistent with the stylized facts in Aghion et al. (2010), if we take the lack of contractibility on idiosyncratic investment opportunities as a proxy of low level of financial development, and capital as a proxy for more productive and riskier projects. Our two-sided externality implies more volatile investment in capital (Proposition 6), lower level of expected consumption (Proposition 5) and lower growth rate of the level of capital and cash in the long term (Figure 3) for financially less developed countries. Also, as a counterpart for the procyclicality of structural investment, we form 'quarterly data' from our simulated samples illustrated on Figure 3. That is, thinking of each period as a day, we generate quarterly observations by summing up shocks, σdZ_t, and investment in capital, dK_t, in each subsequent 100 days. We plot the resulting data points on the two panels of Figure 4, and run a linear regression to assess the connection between shocks and investment. As apparent, two-sided inefficiency implies a larger volatility (in this example, app. 2.5 times larger) and a stronger connection between shocks ($R^2 = 0.38$ vs. $R^2_{cm} = 0.10$). Thus, our results are consistent with the empirical facts that in less financially developed countries, structural investment is more volatile and more procyclical, causing slower growth of consumption and of the size of the economy.

5.2 One-sided interventions

An important advantage of the dynamic structure of our model is that in any state firms' decisions are affected by their expectation of economic conditions in all other (future) states. Suppose that the economy is in a downturn and a policy is introduced with the promise that it will be abandoned as soon as the economy recovers. This policy will necessarily influence firms' choices in the downturn. However, it will also affect firms' choices in the boom, as firms foresee downturns during the boom.

Motivated by this idea, we now analyze a class of (suboptimal) policies that we call one-sided interventions. At a state close to c_l^* the policy maker who observe the price falling dangerously close to the disinvestment threshold l may decide to intervene to boost prices. We do not allow the policy maker to regulate prices directly. Instead, the tool we give to the policy maker is a combination of ex ante taxes and subsidies to the cash holders and capital holders subject to a balanced-budget condition. The policy maker can affect the equilibrium prices and the equilibrium investment/disinvestment thresholds through these taxes and subsidies. Since raising prices is unnecessary in a boom, the planner might make the policy conditional on being in a sufficiently low c state.

5.2.1 Tax-subsidy scheme

A one sided intervention lowers the disinvestment threshold by definition. We know from Proposition (5) that if the investment threshold, c_h^*, remained constant, this policy would improve welfare. However, not surprisingly, a one-sided intervention will reduce the investment threshold c_h^*: knowing the distortionary subsidy in downturns, firms in the boom over invest more egregiously. This implies a negative effect of welfare imposed by the one sided intervention. Interestingly, we show that this negative effect can be so strong that the policy reduces welfare everywhere. That is, even if the policy reduces inefficient liquidation in downturns, it might lower firms' welfare even in downturns, because the policy will make overinvestment in future booms much worse.

We first define one-sided intervention and the corresponding intervention equilibrium. We distinguish equilibrium objects under intervention, with the index π, from their counterpart without intervention.

Definition 3 *A one-sided intervention is a tax-subsidy scheme $\pi(c)$ and an intervention-threshold c_0 such that*

1. *Given c, firms pay $\pi(c)$ for each unit of cash holding and receive $c\pi(c)$ for each unit of capital holding, so that the government budget is balanced;*

2. $\pi(c) = 0$ *for any* $c > c_0$,

3. *the disinvestment threshold is reduced by the intervention,* $c_l^\pi < c_l^*$*;*

4. *the equilibrium price is increased at the intervention threshold,* $p_\pi(c_0) > p(c_0)$.

An intervention equilibrium is the market equilibrium under a one-sided intervention.

We emphasize that we only require the one-sided intervention to raise the price at the intervention-threshold c_0. We may think of one-sided interventions as policies which raise prices for *every*

$c \in [c_l^*, c_0]$ by increasing the capital value $v_\pi(c)$ and/or decreasing the cash value $q_\pi(c)$ over the range $[c_l^*, c_0]$. However, for our result we require less. In this sense, we do not restrict the sign of $\pi(c)$ and impose only weaker requirement in part 4 of Definition 3.

In the next proposition, we show that a one-sided intervention typically decreases the investment threshold ($c_h^\pi < c_h^*$), i.e., worsens overinvestment in the boom. After all, the value of the capital in one state is naturally positively related to its value in every other state. Thus, when intervention boosts the capital price in low states, the price tends to increase also in high states, triggering earlier investment thus a lower threshold c_h^π. That is to say, an intervention focusing on improving underinvestment in the downturn will typically make overinvestment worse in the boom.

Proposition 9 *Any one-sided intervention* $(\pi(c), c_0)$ *in which the value of cash decreases at* c_0*, i.e.,* $q_\pi(c_0) \leq q(c_0)$*, reduces the upper investment threshold,* $c_h^\pi < c_h^*$*.*

5.2.2 An example with a Pareto-dominated one-sided intervention

The intuitive result in Proposition 9 opens an interesting question. The price-boosting one-sided intervention alleviates the underinvestment problem in downturns, but also leads to more severe overinvestment in booms. Speaking about welfare, is it possible that the latter negative effect dominates the earlier positive effect, even in downturns where the one-sided intervention is designed for? We provide an affirmative answer to this question by constructing the following example.

Consider a constant tax-subsidy up to c_0, i.e., $\pi(c) = \pi$ for every $c < c_0$ where π now is a positive constant. We solve the intervention equilibrium in Appendix A.10, which gives investment and disinvestment thresholds c_l^π, c_h^π. We plot one particular example in Figure 5, in which while the policy lowers both the investment and disinvestment thresholds (i.e. $c_l^\pi < c_l^*$ and $c_h^\pi < c_h^*$), it also reduces welfare everywhere. Thus, the depicted one-sided intervention is Pareto inferior compared to the symmetric market equilibrium without intervention.

It is instructive to connect this result to the current debate on "Greenspan's put," i.e., the doctrine that it is sufficient if monetary policy intervenes in a recession but stays inactive when the economy is recovered. We can interpret our taxes-and-subsidies schemes as vague representations of an expansionary monetary policy. An interest rate cut decreases incentives to save cash and increases incentives to invest in capital, just as our simple one-sided intervention does. Our result shows that such interventions might be harmful even at the recession.

Recently, several papers proposed arguments against the Greenspan's put including Farhi and Tirole (2012) and Diamond and Rajan (2011). However, their argument is different. In Diamond and Rajan (2011), ex post inefficient bank-runs serve as a disciplining device for banks. Anticipated interest rate cuts in bad times weakens the disciplining device, and banks take on too much leverage ex ante and are subject to runs ex post too often.

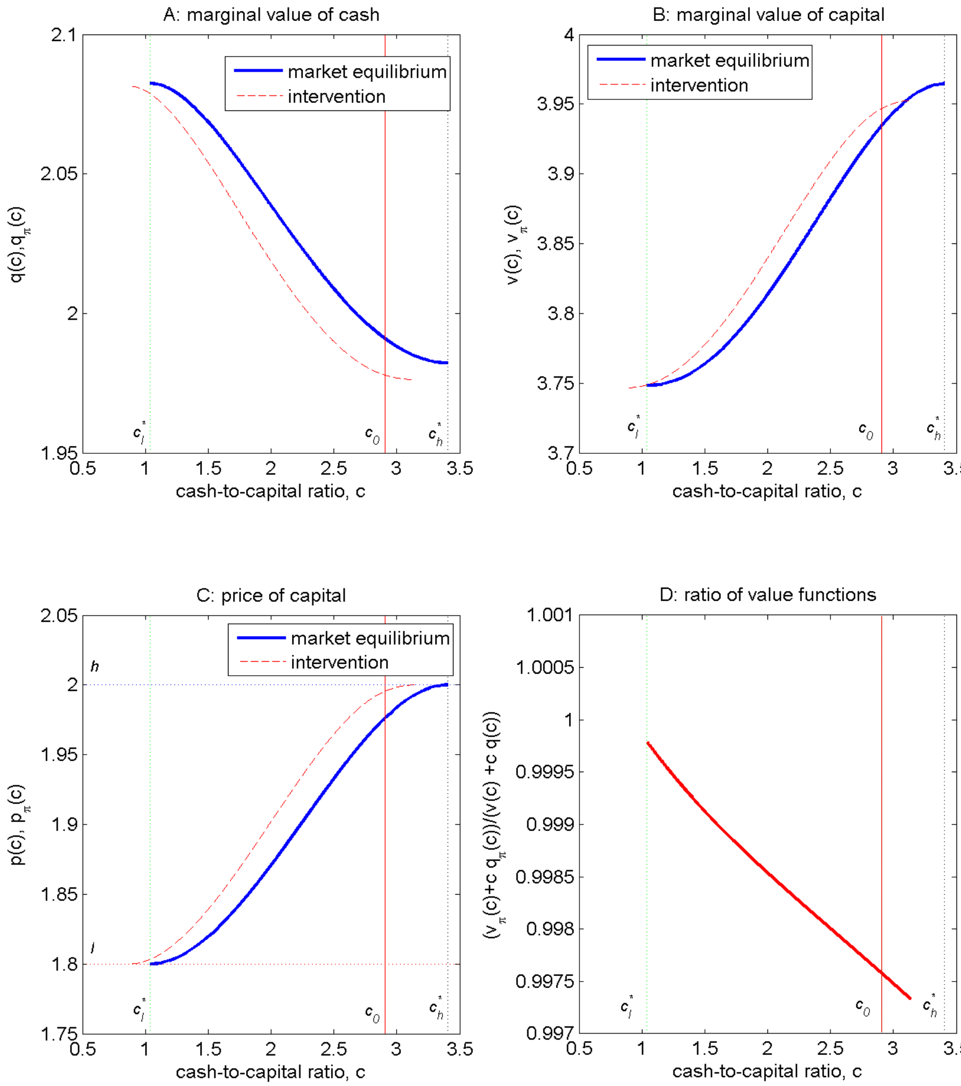

Figure 5: The marginal value of cash, the marginal value of a capital, the price of capital and the ratio of value functions for our baseline model with incomplete markets (thick solid curves) and a particular one-sided intervention (thin, dashed curves). On Panels A, B and C, the solid, vertical line is the thresholds for intervention, c_0,while the two dashed vertical lines are the disinvestment and investment thresholds in our baseline case, c_l^*, c_h^*. Parameter values are $R = 4.1$, $\sigma^2 = 1$, $\xi = 0.1$, $u = 2$, $l = 1.8$ and $h = 2$ and $c_0 = c_h^* - 0.5$ and $\pi = 0.015$.

Farhi and Tirole (2012) show that there is strategic complementarity in choosing higher leverage ex ante, and, consequently, needing a more frequent non-directed bail-out in the form of low interest rates ex post. In both papers, incentive issues inherent in financial intermediation play the pivotal role. In contrast, the pecuniary externality is central to our mechanism.

Until now we have put little emphasis on the parameters that imply a one-sided inefficiency, i.e., underinvestment in capital both in booms and downturns. It is useful to note, however, that with one-sided inefficiency, the price increasing one-sided intervention (at least if it is sufficiently small) improves welfare by pushing the economy closer to the second-best everywhere. Thus, an alternative reading of our results is that the pros and cons of an asymmetric interest rate policy depend on the nature of the externality. In our model, it depends on whether the technology represented by capital is much more productive than the idiosyncratic investment opportunities. Only in the latter case a one-sided intervention could be harmful.

5.3 An alternative specification with flow new opportunities

So far our analysis relies on the specific structure that the idiosyncratic shocks are realized just at the period of capital productivity hike. This section aims to show that this particular timing assumption is immaterial for our main qualitative results of two-sided inefficiency.

5.3.1 Setting and solutions

We consider two major changes. First, the aggregate Poisson event that the capital is subject to a productivity hike (recall that the hike occurs with intensity ξ) and the idiosyncratic shock is separated. In particular, we keep the productivity hike in capital in the baseline model, but in each point of time ϕdt fraction of the firms are hit by the idiosyncratic skill shock. That is, they lose the skill to tender the capital, but can invest in a new opportunity with a constant return of $u > 1$. As a result, in each instant, a group of firms with measure ϕdt sell all their capital to the rest of firms and exit the market (and consume the final consumption goods after investing in the new opportunity). For simplicity, after the final date of productivity hike, the model ends and firms consume their consumption goods.

The second change is about timing of (dis)investment opportunities. Instead of letting the firms to invest and disinvest at any point, we assume that they can do so only irregularly. In particular, with intensity η, an aggregate Poisson event realizes, and at that instant all firms can build capital at cost h or dismantle capital at cost l in any amount they wish. Because firms might not be able to dismantle capital for cash at any time they wish, even in the constrained-efficient solution firms will dismantle capital (if they can) once the aggregate cash-to-capital ratio drops to a sufficiently low but positive threshold, as opposed to $c_l^P = 0$ in our main model shown in Proposition 3. This interior solution structure guarantees exactly zero first-order conditions in setting the upper-investment and lower-liquidation thresholds in the constrained-efficient benchmark. Thus, any (small) divergence between the social marginal rate of substitution and that of private firms can deliver distorted (dis)investment thresholds in the market equilibrium.[30]

Different from the baseline model, in this variant there is no guarantee that the aggregate cash level is kept away from zero by liquidation if necessary. Thus, we assume an infinite pool of outside cash holders who can inject one unit of cash to this market for a total cost of $\lambda > 1$. That is, outside investors can acquire the knowledge of firms, but it is costly. We think of λ to be sufficiently high.

In the Appendix A.11 we write down the ODEs and boundary conditions for the value of capital $v(c)$ and the value of cash $q(c)$ in the market solution, and solve them numerically. As before, we assess whether a social planner could improve welfare by only changing the investment/disinvestment thresholds c_h and c_l, as opposed to leaving their determination to the market. To this end, we further characterize the ODE for the equilibrium value functions with intervention, and solve them numerically. We then search for the optimal dis(investment) thresholds that maximizes the social value.

5.3.2 Two-sided inefficiency

As an illustration, Figure 6 graphs the investment threshold (the left panel, solid line) and disinvestment threshold (the right panel, solid line) in the market solution, with corresponding optimal ones (dashed lines) if the social planner can intervene.

Two sided inefficiency prevails in this alternative setting. In the left panel, the disinvestment threshold of the market solution is above that of the social planner's solution, while in the right panel the opposite holds. The intuition is the same as in our baseline model. Firms who suffer idiosyncratic skill shocks will sell their capital at an equilibrium price; the higher the aggregate cash level, the higher the equilibrium prices. These prices affect the firms' private incentives whenever they can adjust their capital/cash holdings. Take the example of low aggregate level of cash; firms tend to dismantle their capital excessively, worrying that they might be hit by a skill shock and therefore sell their capital (to others who have not hit by the skill shocks) at low prices. In contrast, the social planner should completely ignore this issue of rent distribution. Similarly, when aggregate cash is abundant, firms will invest more than the social planner would like to: firms factor in that their capital can be sold at a high price (above the corresponding social value) once they exit the market.

Figure 6 also gives a comparative static result for two-sided externalities. When we vary the investment cost h, effectively we are varying the productivity advantage of capital $R/h-u$ (and this is the x-axis we are plotting against in Figure 6). Consistent with Proposition 8 which suggests that two-sided inefficiencies are more likely to occur when the productivity advantage of capital $R/h-u$ is small, we observe that the difference between the market solution and that of the social planner goes up for a smaller $R/h-u$. Intuitively, when the profitability of the existing capital technology

[30] If firms can (dis)invest at any time (as in the main model), our numerical solution suggests that although the upper investment threshold is generally distorted, firms do not liquidate capital until the aggregate cash level $c = 0$ (which is the constrained efficient level). This is partly because in the constrained efficient solution, the cornered (at zero) optimal liquidation threshold has a non-zero first-order condition (recall Eq. (28)). Lumpy investment opportunities featuring an interior (strictly positive) disinvestment threshold resolves this technical problem.

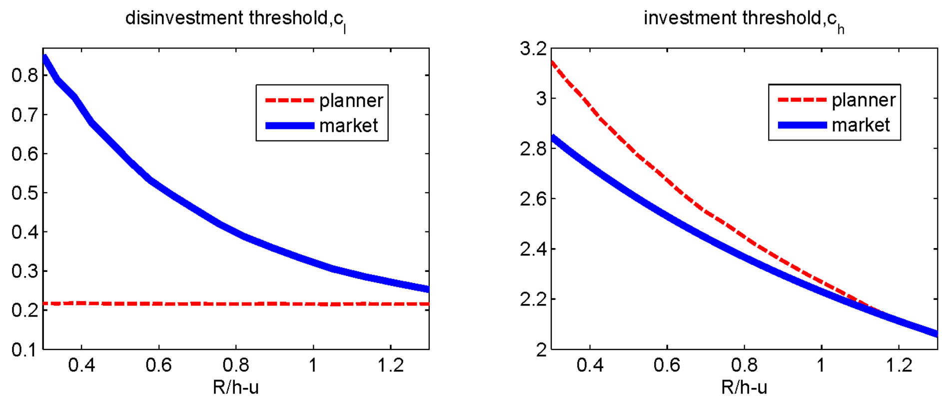

Figure 6: Disinvestment and investment thresholds in the alternative specification under the market solution (solid) and under the planner (dashed) for different h values. We plot these polices as functions of the productivity advantage of capital, $R/h - u$. Parameter values are $R = 4.1$, $\xi = 0.1$, $\sigma^2 = 0.5$, $u = 2$, $l = 0.11$, $\lambda = 5$, $\eta = 0.2$.

is close to that of the new investment opportunity, these two assets have similar productivities. Agents in the market solution tend to move back and forth between these two assets too often for rent extraction, while the social planner realize that most of these activities are wasted due to deadweight loss of transaction costs.

6 Conclusion

We build an analytically tractable, dynamic stochastic model of investment and trade, in which investment cycles, i.e., boom periods with abundant investment and low returns and bust periods with low investment and high returns, arise naturally. In the presence of unverifiable idiosyncratic investment opportunities, a two-sided inefficiency can arise: there are two much investment in the technology and too low buffer in cash in booms, and there are too little investment and too much cash holdings in downturns. We show that in this case a one-sided policy targeting only the underinvestment in downturns might be ex ante Pareto inferior to no intervention in all states (including downturns).

Apart from analyzing two-sided inefficiencies, we also presented a novel way of modelling problems of investment and trade. This method provides analytical tractability in a dynamic stochastic framework for the full joint distribution of states and equilibrium objects. To explore its potential, we use this framework to analyze the role of sovereign wealth funds in financial crises by introducing groups of firms with different level of skills in a parallel project.

References

Acharya, Viral V., and Ouarda Merrouche. 2012. “Precautionary Hoarding of Liquidity and

Interbank Markets: Evidence from the Subprime Crisis." *Review of Finance*, 16: 210–230.

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.

Armenter, Roc, and Viktoria Hnatkovska. 2011. "The macroeconomics of firms' savings." Federal Reserve Bank of Philadelphia Working Papers 12-1.

Ashcraft, Adam, James Mcandrews, and David Skeie. 2011. "Precautionary Reserves and the Interbank Market." *Journal of Money, Credit and Banking*, 43: 311–348.

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.

Bolton, Patrick, Hui Chen, and Neng Wang. 2011. "Market Timing, Investment, and Risk Management." NBER NBER Working Papers 16808.

Brunnermeier, Markus K., and Yuliy Sannikov. 2011. "A macroeconomic model with a financial sector." http://scholar.princeton.edu/markus/publications/macroeconomic-model-financial-sector.

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.

Covas, Francisco, and Wouter J. Den Haan. 2011. "The Cyclical Behavior of Debt and Equity Finance." *American Economic Review*, 101(2): 877–99.

Davila, Eduardo. 2011. "Dissecting Fire Sales Externalities." http://www.people.fas.harvard.edu/ edavila.

Diamond, Douglas W, and Philip H Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, 91(3): 401–19.

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.

Eisfeldt, Andrea, and Tyler Muir. 2012. "The Joint Dynamics of Internal and External Finance." https://sites.google.com/site/andrealeisfeldt/.

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.

Grenadier, Steven R. 1996. "The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets." *The Journal of Finance*, 51(5): pp. 1653–1679.

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.

Heider, Florian, Marie Hoerova, and Cornelia Holthausen. 2009. "Liquidity Hoarding and Interbank Market Spreads: The Role of Counterparty Risk." European Central Bank Working Paper Series 1126.

He, Zhiguo, and Arvind Krishnamurthy. 2012*a*. "Intermediary Asset Pricing." *American Economic Review*, ,: forthcoming.

He, Zhiguo, and Arvind Krishnamurthy. 2012*b*. "A Macroeconomic Framework for Quantifying Systemic Risk." http://faculty.chicagobooth.edu/zhiguo.he/research/systemicmacro.pdf.

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.

Jacklin, Charles J. 1987. "Demand Deposits, Trading Restrictions, and Risk Sharing." , ed. Edward C. Prescott and Neil Wallace, Chapter Contractual Arrangements for Intertemporal Trade, Vol. 1. University of Minnesota Press.

Jeanne, Olivier, and Anton Korinek. 2010. "Managing Credit Booms and Busts: A Pigouvian Taxation Approach." Peterson Institute for International Economics WP10-12.

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." *Jour-*

nal of Economic Theory, 111(2): 277–292.

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.

Shiller, Robert J. 2007. “Understanding Recent Trends in House Prices and Home Ownership.” National Bureau of Economic Research, Inc NBER Working Papers 13553.

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.

Wheaton, William C., and Raymond G. Torto. 1990. “An Investment Model of the Demand and Supply For Industrial Real Estate.” *Real Estate Economics*, 18(4): 530–547.

A Appendix: Proofs and Derivations

A.1 Proof of Lemma 1 and Proposition 1

We construct the proof in steps. In particular, we separate Proposition 1 into the following four Lemmas. These four lemmas are sufficient to prove Proposition 1.

Lemma A.1 *If the equation system (18)-(19), (13)-(15) has a solution where $c_h^* < R$, and both $v(c)$ and $q(c)$ are increasing in the range $c \in [c_l^*, c_h^*]$, then Proposition 1 holds.*

Lemma A.2 *The system (18)-(19), (13)-(15) always has at least one solution.*

Lemma A.3 *If $h - l$ is sufficiently small, then $c_h^* < R$.*

Lemma A.4 *$q(c)$ is decreasing in c. If $h - l$ is sufficiently small, then $v(c)$ is increasing for $c \in [c_l^*, c_h^*]$.*

A.1.1 Step 1: Proof of Lemma 1 and Lemma A.1

Denote the dollar share of capital in the firm's portfolio by ψ_t^i, so that $\psi_t^i = K_t^i p_t / w_t^i$. According to our conjecture, the value function can be written as (recall the aggregate cash-to-capital ratio $c = C/K$)

$$J\left(K_t, C_t, K_t^i, C_t^i\right) = w_t^i \left[\left(1 - \psi_t^i\right) q\left(c_t\right) + \frac{\psi_t^i}{p_t} v\left(c\right)\right] = J\left(K_t, C_t, w_t^i\right),$$

is linear in w_t. This is equivalent to $J\left(C, K, K_t^i, C_t^i\right) = K_t^i v\left(c\right) + C_t^i q\left(c\right)$ stated in the Lemma. Also, we have the wealth dynamics, expressed in terms of portfolio choice ψ_t^i, as

$$dw_t^i = -d\alpha_t^i - \theta dK_t^i + \psi_t^i w_t^i \frac{1}{p_t}\left(dp_t + \sigma dZ_t\right).$$

And, $q(c) \geq 1$ has to hold as firms can consume cash at the final date (and there is no discounting), which implies $d\alpha_t^i = 0$, i.e., firms do not consume ex ante.

As the firm is choosing portfolio share ψ_t^i, and the capital to build or dismantle dK_t^i, the Hamiltonian-Jacobi-Bellman (HJB) of problem (10) can be written as:

$$0 = \max_{d\psi_t^i, dK_t^i} d\alpha_t^i + J_C \mathbb{E}_t\left[dC_t\right] + \frac{1}{2} J_{CC} \mathbb{E}_t\left[dC_t^2\right] + J_w \mathbb{E}_t\left(dw_t\right) + J_K' dK_t^i + J_{w,C} \mathbb{E}_t\left[dw_t dC_t\right].$$

The endogenous price dynamics (using Ito's Lemma) is

$$dp_t = \frac{1}{2}\sigma^2 p''\left(c\right) dt + \sigma p'\left(c\right) dZ_t + dB_t^p - dU_t^p,$$

where dB_t^p (dU_t^p) reflects p at $p\left(c_l^*\right) = l$ ($p\left(c_h^*\right) = h$). This is because in any market equilibrium firms will create (dismantle) capital if $p_t = h$ ($p_t = l$), and keep doing it until the price adjusts. We derived the boundary conditions in the main text. Also, by risk neutrality and ex ante homogeneity of firms, before the final date the price of the capital has to make firms indifferent whether to hold capital or cash. Otherwise markets could not clear. We also explained that $\hat{p}_\tau = c_\tau$.

Thus, inside the reflection boundary (c_l^*, c_h^*) the above HJB is (we drop i from now on)

$$0 = \max_{\psi_t} \left\{ \begin{array}{c} \frac{\sigma^2}{2} w_t q_c'' (c_t) + q(c_t) \psi_t w_t \frac{\frac{1}{2}\sigma^2 p''(c_t)}{p_t} + q'(c_t) \left(\left(\psi_t w_t \frac{\sigma}{p_t} (\sigma + p'(c_t) \sigma) \right) \right) \\ +\xi w_t \left[\frac{1}{2} \left(\frac{\psi_t R}{p_t} + (1-\psi_t) \frac{R}{c_t} \right) + \frac{1}{2} \left(\frac{\psi_t}{p_t} u c_t + (1-\psi_t) u \right) - q(c_t) \right] \end{array} \right\}.$$

Since the problem is linear in ψ_t, in equilibrium firms must be indifferent in their choice of ψ_t. Thus, we can calculate the dynamics of the cash (capital) value by choosing $\psi_t = 0$ ($\psi = 1$). Setting $\psi_t = 0$ directly implies (16). Choosing $\psi_t = 1$ gives

$$0 = \frac{\sigma^2}{2} q''(c) + q(c) \frac{\frac{1}{2}\sigma^2 p''(c)}{p} + q'(c) \left(\frac{1}{p} (\sigma + p'\sigma) \sigma \right) + \frac{1}{p} \left(\frac{\xi}{2} (R + uc) - q(c) p \right).$$

Since $v(c) = p(c) q(c)$, $v' = q'p + p'q$, and $v'' = q''p + 2p'q' + p''q$, we can rewrite the above equation as (17). Given that the ODEs for $v(c)$ and $q(c)$ were derived by substituting in $\psi_t = 1$ and $\psi_t = 0$, it is easy to see that these functions can be interpreted as the value of a capital and that of a unit of cash. This implies that

$$J(C, K, w_t^i) = \left(w_t^i (1 - \psi_t^i) q(c) + \frac{\psi_t^i}{p_t} w_t^i v(c) \right) = q(c) w_t$$

verifying both Lemma 1 and our conjecture on the form of $J(C, K, w_t^i)$.

A.1.2 Step 2: Proof of Lemma A.2

First, note that for any arbitrary c_h and c_l from (15), we can express A_1-A_4 in (18)-(19) as functions of c_h and c_l only. Substituting back to (18)-(19) we get our functions parameterized by c_h and c_l which we denote as $v(c; c_l, c_h)$ and $q(c; c_l, c_h)$. Evaluating these functions at $c = c_l$ and $c = c_h$, we get the following expressions. Define

$$\begin{aligned} f_l(c_l, c_h) &\equiv \frac{e^{-\gamma c_h} (\mathrm{Ei}[c_h\gamma] - \mathrm{Ei}[c_l\gamma]) + e^{\gamma c_h} (\mathrm{Ei}[-c_h\gamma] - \mathrm{Ei}[-c_l\gamma])}{e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}}, \\ g_l(c_l, c_h) &\equiv \frac{e^{-\gamma c_h} (\mathrm{Ei}[c_h\gamma] - \mathrm{Ei}[c_l\gamma]) + e^{\gamma c_h} (\mathrm{Ei}[-\gamma c_l] - \mathrm{Ei}[-\gamma c_h])}{e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}}, \\ f_h(c_l, c_h) &\equiv \frac{e^{-\gamma c_l} (\mathrm{Ei}[c_h\gamma] - \mathrm{Ei}[c_l\gamma]) + e^{\gamma c_l} (\mathrm{Ei}[-\gamma c_h] - \mathrm{Ei}[-\gamma c_l])}{e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}}, \\ g_h(c_l, c_h) &\equiv \frac{e^{-\gamma c_l} (\mathrm{Ei}[c_h\gamma] - \mathrm{Ei}[c_l\gamma]) + e^{\gamma c_l} (\mathrm{Ei}[-\gamma c_l] - \mathrm{Ei}[-\gamma c_h])}{e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}}, \text{ and} \\ m(c_l, c_h) &\equiv \frac{e^{\gamma(c_h - c_l)} - 1}{1 + e^{\gamma(c_h - c_l)}} \in (0, 1). \end{aligned}$$

Then the cash and capital values can be rewritten as

$$\begin{aligned}
q\left(c_l;c_l,c_h\right) &= \frac{u}{2}+\frac{R\xi}{\gamma\sigma^2}f_l\left(c_l,c_h\right), q\left(c_h;c_l,c_h\right)=\frac{u}{2}+\frac{R\xi}{\gamma\sigma^2}f_h\left(c_l,c_h\right),\\
v\left(c_l;c_l,c_h\right) &= R+\frac{c_l u}{2}+\frac{u}{2\gamma}m\left(c_l,c_h\right)+\frac{R\gamma}{2}\left(\frac{g_l\left(c_l,c_h\right)}{\gamma}-c_l f_l\left(c_l,c_h\right)\right), \text{ and}\\
v\left(c_h;c_l,c_h\right) &= R+\frac{c_h u}{2}-\frac{u}{2\gamma}m\left(c_l,c_h\right)+\frac{R\gamma}{2}\left(\frac{g_h\left(c_l,c_h\right)}{\gamma}-c_h f_h\left(c_l,c_h\right)\right).
\end{aligned}$$

For any c_h, define the function $H\left(c_h\right)$ implicitly as the corresponding lower threshold c_l so that at $c=c_h$ the market price is just h, i.e.,

$$p\left(c_h;c_l=H\left(c_h\right),c_h\right)=\frac{v\left(c_h;c_l=H\left(c_h\right),c_h\right)}{q\left(c_h;c_l=H\left(c_h\right),c_h\right)}=h.$$

Similarly, define $L\left(c_h\right)$ is defined implicitly by

$$p\left(c_l;c_l=L\left(c_h\right),c_h\right)\equiv\frac{v\left(c_l;c_l=L\left(c_h\right),c_h\right)}{q\left(c_l;c_l=L\left(c_h\right),c_h\right)}=l,$$

which makes the market price to be l at $c=c_l$. Obviously, once we find such c_h that $H\left(c_h\right)=L\left(c_h\right)$, then this particular c_h and the corresponding $c_l=H\left(c_h\right)=L\left(c_h\right)$ is a solution of (13)-(15), (18)-(19). To show that this solution exists, we first establish properties of $L\left(c_h\right)$ then we proceed to the properties of $H\left(c_h\right)$.

Properties of $L\left(c_h\right)$ It is useful to observe that

$$\begin{aligned}
\frac{\partial f_l}{\partial c_l} &= \frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}\left(\gamma f_l-\frac{1}{c_l}\right), \frac{\partial f_l}{\partial c_h}=2\frac{\frac{1}{c_h}-\gamma f_h}{e^{\gamma(c_h-c_l)}-e^{\gamma(c_l-c_h)}}\\
\frac{\partial g_l}{\partial c_l} &= \frac{1}{c_l}+\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}\gamma g_l, \frac{\partial g_l}{\partial c_h}=-\frac{2\gamma g_h}{e^{\gamma(c_h-c_l)}-e^{\gamma(c_l-c_h)}},\\
\lim_{c_l\to c_h} f_l &= \frac{1}{\gamma c_h}, \quad \lim_{c_l\to c_h} g_l=0, \quad \lim_{c_l\to c_h} m=0.
\end{aligned}$$

1. We show that $f_l\left(c_h,c_l\right)$ is monotonically decreasing in c_l. Its slope in c_l is

$$\frac{\partial f_l}{\partial c_l}=\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}\left(\gamma f_l\left(c_h,c_l\right)-\frac{1}{c_l}\right), \tag{A.1}$$

and the second derivative is

$$\frac{\partial^2 f_l}{\partial^2 c_l} =$$

$$= -\left(4\gamma e^{2\gamma c_h}\frac{e^{2\gamma c_l}}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)^2}-\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)^2}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)^2}\gamma\right)\left(\frac{1}{c_l}-\gamma f_l\left(c_h,c_l\right)\right)-\frac{\left(-\frac{1}{c_l^2}\right)\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)} =$$

$$= \gamma\left(\frac{1}{c_l}-\gamma f_l\left(c_h,c_l\right)\right)+\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}\frac{1}{c_l^2}$$

Note that if the first derivative is zero, then the second derivative is positive implying that $f_l\left(c_h,c_l\right)$ can have only local minima, but no local maxima in c_l. At the limit one can check that

$$\lim_{c_l\to c_h}\frac{\partial f_l}{\partial c_l}=\lim_{c_l\to c_h}\left(\frac{1}{c_l}\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}\left(\gamma c_l f_l\left(c_h,c_l\right)-1\right)\right)=\frac{1}{c_h}\left(-\frac{1}{2\gamma c_h}\right)<0.$$

Thus, $f_l\left(c_h,c_l\right)$ is decreasing at $c_h=c_l$. Suppose that it is not monotonic over the range of $c_l<c_h$ in c_l. Then the largest $\hat{c}_l$ where the first derivative is 0, would be a local maximum. But we have just ruled out the existence of a local maximum. Thus $f_l\left(c_h,c_l\right)$ monotonically decreasing over the whole range of $c_l<c_h$ in c_l. This statement is equivalent to $\gamma f_l\left(c_h,c_l\right)-\frac{1}{c_l}<0$ for $c_l<c_h$, for any fixed c_h.

2. We show that $X\left(c_l\right)\equiv f_l\left(c_h,c_l\right)-\frac{1}{\gamma c_l}$ is increasing in c_l. We would like to show that

$$X'\left(c_l\right)=\gamma\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}X\left(c_l\right)+\frac{1}{\gamma c_l^2}>0. \tag{A.2}$$

Clearly, we have

$$X\left(c_l=c_h\right)=0,\ X'\left(c_l=c_h\right)=f_l'\left(c_h,c_h\right)+\frac{1}{\gamma c_h^2}=\frac{1}{2\gamma c_h^2}>0.$$

We know that when $c_l\to 0$, $f\left(c_h,c_l\right)$ has the order of $\mathrm{Ei}\left(\gamma c_l\right)$ which is $O\left(\ln c_l\right)$; this implies that $X\left(c_l\right)\to-\infty$ when $c_l\to 0$. Then, if $X\left(c_l\right)$ is not monotone, we must have two points $x_1<x_2$ closest to (but below) c_h so that

$$0>X\left(x_1\right)>X\left(x_2\right),\ X'\left(x_1\right)=X'\left(x_2\right)=0.$$

Setting (A.2) to be zero, we have (because $0<x_1<x_2$)

$$X\left(x_1\right)=-\frac{\left(e^{2\gamma c_h}-e^{2\gamma x_1}\right)}{\gamma^2x_1^2\left(e^{2\gamma c_h}+e^{2\gamma x_1}\right)}<-\frac{\left(e^{2\gamma c_h}-e^{2\gamma x_2}\right)}{\gamma^2x_2^2\left(e^{2\gamma c_h}+e^{2\gamma x_2}\right)}=X\left(x_2\right),$$

in contradiction with $X\left(x_1\right)>X\left(x_2\right)$. Thus (A.2) holds always.

3. We show that the function $\frac{g_l(c_h,c_l)}{\gamma} - c_l f_l\left(c_h, c_l\right)$ is monotonically increasing in c_l. Its first derivative is (all the derivatives in this part are with respect to c_l)

$$\begin{aligned}\left(\frac{g_l}{\gamma} - c_l f_l\right)' &= \frac{1}{\gamma c_l} + \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)} g_l\left(c_l, c_h\right) - \left(\frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}\left(c_l \gamma f_l\left(c_l, c_h\right) - 1\right) + f_l\left(c_l, c_h\right)\right)\\ &= \frac{1}{\gamma c_l} + \gamma \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}\left(\frac{g_l}{\gamma} - c_l f_l\right) + \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)} - f_l\end{aligned}$$

Whenever the first derivative is zero, at that point we have

$$\frac{g_l}{\gamma} - c_l f_l = \frac{f_l - \frac{1}{\gamma c_l}}{\gamma \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}} - \frac{1}{\gamma}. \tag{A.3}$$

We also know that

$$\lim_{c_l \to c_h}\left(\frac{g_l}{\gamma} - c_l f_l\right)' = 0, \text{ and } \lim_{c_l \to c_h}\left(\frac{g_l}{\gamma} - c_l f_l\right)'' = -\frac{1}{3\gamma c_h^2} < 0;$$

so for any fixed c_h, $c_l = c_h$ is a local maximum. Thus to show that $\frac{g_l}{\gamma} - c_l f_l$ is monotone, it suffices to rule out the case of a local minimum $\hat{c}_l < c_h$ so that $\left(\frac{g_l}{\gamma} - c_l f_l\right)' = 0$ and $\left(\frac{g_l}{\gamma} - c_l f_l\right)'' > 0$. In general

$$\left(\frac{g_l}{\gamma} - c_l f_l\right)'' = -\frac{1}{\gamma c_l^2} + \gamma \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}\left(\frac{g_l}{\gamma} - c_l f_l\right)' - f_l' + \frac{4e^{2\gamma c_h} e^{2\gamma c_l}}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)^2}\gamma^2\left(\left(\frac{g_l}{\gamma} - c_l f_l\right) + \frac{1}{\gamma}\right).$$

Thus, if there were a $\hat{c}_l$ that $\left(\frac{g_l}{\gamma} - c_l f_l\right)' = 0$, using (A.1) and (A.3) we have $\left(\frac{g_l}{\gamma} - c_l f_l\right)''$ to be equal to

$$-\frac{1}{\gamma \hat{c}_l^2} - f_l' + \frac{4\gamma^2 e^{2\gamma c_h} e^{2\gamma \hat{c}_l}}{\left(e^{2\gamma c_h} - e^{2\gamma \hat{c}_l}\right)^2}\left(\frac{f_l - \frac{1}{\gamma \hat{c}_l}}{\gamma \frac{\left(e^{2\gamma c_h} + e^{2\gamma \hat{c}_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma \hat{c}_l}\right)}} - \frac{1}{\gamma} + \frac{1}{\gamma}\right) = -\frac{1}{\gamma \hat{c}_l^2} - \gamma \frac{\left(e^{2\gamma c_h} - e^{2\gamma \hat{c}_l}\right)}{e^{2\gamma c_h} + e^{2\gamma \hat{c}_l}}\left(f_l - \frac{1}{\gamma \hat{c}_l}\right).$$

But from (A.2) we know the above term is strictly negative, which proves the contradiction.

4. We show that $q\left(c_l; c_l, c_h\right)$ is also decreasing in c_l for any $c_l < c_h$. Given that $\left(\frac{g_l}{\gamma} - c_l f_l\right)' > 0$ and $\partial\left(\frac{c_l u}{2} + \frac{u\left(e^{-\gamma(c_h - c_l)} + e^{\gamma(c_h - c_l)} - 2\right)}{2\gamma\left(e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}\right)}\right)/\partial c_l = \frac{1}{2} u \frac{e^{-2\gamma c_h + 2\gamma c_l} + 1}{\left(e^{-\gamma c_h + \gamma c_l} + 1\right)^2} > 0$, $v\left(c_l; c_l, c_h\right)$ is increasing in c_l. Thus, $p\left(c_l; c_l, c_h\right)$ is increasing in c_l for any $c_l < c_h$. Also one can show that $\lim_{c_l \downarrow 0} = p\left(c_l; c_l, c_h\right) = -\frac{\tanh(\gamma c_h)}{\gamma} < 0$, and

$$\lim_{c_l \to c_h} p\left(c_l; c_l, c_h\right) = \frac{R + c_h \frac{u}{2} + \frac{R\gamma}{2}\left(-c_h \frac{1}{\gamma c_h}\right)}{\frac{u}{2} + \frac{R\xi}{\gamma \sigma^2} \frac{1}{\gamma c_h}} = \frac{R + c_h \frac{u}{2} - \frac{R}{2}}{\frac{u}{2} + \frac{1}{2c_h}},$$

which is larger than l as long as $c_h > l$. Thus, as long as $c_h > l$, $\lim_{c_l \to c_h} p(c_l; c_l, c_h) \geq l$ and there is a unique solution c_l for any c_h of $p(c_l; c_l, c_h) = l$. Therefore $L(c_h)$ exist. From the monotonicity in c_l, and continuity of $p(c_l; c_l, c_h)$ we also know that $L(c_h)$ is continuous.

Properties of $H(c_h)$ First, we show that for any $c_h \in [l, R]$, $H(c_h)$ is a continuous function and $H(c_h) \in [0, c_h]$. Again, the notation $'$ means we are taking the derivative with respect to c_l. We use the following facts:

$$\begin{aligned}
\frac{\partial f_h}{\partial c_l} &= \frac{2\left(\gamma f_l(c_h, c_l) - \frac{1}{c_l}\right)}{\left(e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}\right)}, \frac{\partial g_h}{\partial c_l} = \frac{2\gamma g_l(c_h, c_l)}{\left(e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}\right)} \\
\frac{\partial f_h}{\partial c_h} &= \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}\left(\frac{1}{c_h} - \gamma f_h(c_h, c_l)\right), \frac{\partial g_h}{\partial c_h} = \frac{1}{c_h} - \frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}\gamma g_h(c_l, c_h) \\
\lim_{c_l \to c_h} f_h &= \frac{1}{\gamma c_h}, \quad \lim_{c_l \to c_h} g_h = 0.
\end{aligned}$$

1. The result of $\frac{\partial f_h}{\partial c_l} = \frac{2\left(\gamma f_l(c_h, c_l) - \frac{1}{c_l}\right)}{\left(e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}\right)} < 0$ follows from the step 1 in the previous subsection.

2. We show $\left(\frac{g_h}{\gamma} - f_h c_h\right)' > 0$ for $c_l < c_h$. We have $\left(\frac{g_h}{\gamma} - f_h c_h\right)' = 2\frac{g_l - c_h\gamma f_l + c_h\frac{1}{c_l}}{e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}}$ and

$$\frac{\partial^2\left(\frac{g_h}{\gamma} - f_h c_h\right)}{\partial^2 c_l} = \frac{2g_l' - c_h 2\gamma f_h' - 2\frac{c_h}{c_l^2}}{e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}} + \gamma e^{-\gamma(c_h - c_l)}\frac{e^{2(-\gamma(c_h - c_l))} + 1}{\left(e^{-2\gamma(c_h - c_l)} - 1\right)^2}\left(2g_l - c_h 2\gamma f_l + \frac{2c_h}{c_l}\right).$$

If the first derivative is zero at a point $c_h > c_l$, then the second derivative is

$$\frac{2\frac{1}{c_l} + 2\gamma\frac{\left(e^{2\gamma c_h} + e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h} - e^{2\gamma c_l}\right)}\left(g_l(c_l, c_h) - c_h\gamma f_l(c_h, c_l) + \frac{c_h}{c_l}\right) - c_h 2\frac{1}{c_l^2}}{\left(e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}\right)} = \frac{-2\frac{c_h - c_l}{c_l^2}}{\left(e^{\gamma(c_h - c_l)} - e^{-\gamma(c_h - c_l)}\right)} < 0.$$

for any $c_h > c_l$, which implies that it can have no minimum in that range. Also

$$\lim_{c_l \to c_h}\frac{\partial\left(\frac{g_h}{\gamma} - f_h c_h\right)}{\partial c_l} = 0, \quad \lim_{c_l \to c_h}\frac{\partial^2\left(\frac{g_h}{\gamma} - f_h c_h\right)}{\partial^2 c_l} = -\frac{1}{3\gamma c_h^2}$$

so $c_l = c_h$ must be the unique maximum in the range $c_h \geq c_l$, and the result follows.

3. Consequently, $q(c_h; c_h, c_l)$ is monotonically decreasing and $v(c_h; c_h, c_l)$ is monotonically increasing in c_l. Thus, $p(c_h; c_h, c_l)$ is monotonically increasing in c_l.

4. Observe that the following hold

$$\lim_{c_l \to c_h} p(c_h; c_l, c_h) = \lim_{c_l \to c_h}\frac{v(c_h; c_l, c_h)}{q(c_h; c_l, c_h)} = \frac{Rc_h + c_h^2\frac{u}{2} - \frac{R}{2}c_h}{\frac{u}{2}c_h + \frac{R}{2}} = \frac{c_h^2 u + Rc_h}{uc_h + R} = c_h.$$

Because $\lim_{c_l \to 0} = p(c_h; c_l, c_h) = -c_h$, hence we know that for any $c_h > h$ there is a unique $c_l \in [0, c_h]$ which solves $p(c_h; c_l, c_h) = h$. From the monotonicity of $p(c_h; c_h, c_l)$ in c_l and the continuity in c_h, the resulting function $H(c_h)$ is continuous in c_h.

Intercept of $H(c_h)$ and $L(c_h)$

1. Here we show that $H(h) > L(h)$. We know that $H(h) = h$ because

$$\lim_{c_l \to h} = \frac{v(c_h; c_l, c_h)}{q(c_h; c_l, c_h)} = \frac{R + h\frac{u}{2} + \frac{R\xi}{\gamma\sigma^2}\left(-h\frac{1}{\gamma h}\right)}{\frac{u}{2} + \frac{R\xi}{\gamma\sigma^2}\frac{1}{\gamma h}} = \frac{R + h\frac{u}{2} + \frac{R}{2}\gamma\left(-h\frac{1}{\gamma h}\right)}{\frac{u}{2} + \frac{R}{2}\gamma\frac{1}{\gamma h}} = h.$$

However, note that

$$\lim_{c_l \to h} \frac{v(c_l; c_l, c_h)}{q(c_l; c_l, c_h)} = \frac{R + h\frac{u}{2} + \frac{R\gamma}{2}\left(-h\frac{1}{\gamma h}\right)}{\frac{u}{2} + \frac{R}{2h}} = h,$$

and $\frac{v(c_l;c_l,c_h)}{q(c_l;c_l,c_h)}$ is increasing in c_l. Since $L(h)$ is defined by $\frac{v(c_l;L(h),h)}{q(c_l;L(h),h)} = l < h$, $L(h) < h = H(h)$ must hold.

2. Now we show that $\lim_{c_h \to \infty} H(c_h) = 0 < \lim_{c_h \to \infty} L(c_h)$. It is easy to check that

$$\lim_{c_h \to \infty} f_l = \frac{-\operatorname{Ei}[-c_l\gamma]}{e^{\gamma(-c_l)}}, \lim_{c_h \to \infty} g_l = \frac{\operatorname{Ei}[-\gamma c_l]}{e^{\gamma(-c_l)}}, \lim_{c_h \to \infty} f_h = 0, \lim_{c_h \to \infty} g_h = 0$$

Thus, $\lim_{c_h \to \infty} \frac{v(c_l;c_l,c_h)}{q(c_l;c_l,c_h)}$ takes the value of

$$\lim_{c_h \to \infty} \frac{R + \frac{c_l u}{2} + \frac{u m(c_l,c_h)}{2\gamma} + \frac{R\gamma}{2}\left(\frac{g_l(c_l,c_h)}{\gamma} - c_l f_l(c_l, c_h)\right)}{\frac{u}{2} + \frac{R\gamma}{2} f_l(c_l, c_h)} = \frac{R + \frac{c_l u}{2} + \frac{u}{2\gamma} + \frac{R\gamma}{2}\left(\frac{\operatorname{Ei}[-\gamma c_l]}{\gamma e^{\gamma(-c_l)}} - c_l \frac{-\operatorname{Ei}[-c_l\gamma]}{e^{\gamma(-c_l)}}\right)}{\frac{u}{2} - \frac{\operatorname{Ei}[-c_l\gamma]}{e^{\gamma(-c_l)}}}.$$

Thus, $\lim_{c_h \to \infty} L(c_h)$ is the finite positive solution of

$$\frac{R + \frac{c_l u}{2} + \frac{u}{2\gamma} + \frac{R\gamma}{2}\left(\frac{\operatorname{Ei}[-\gamma c_l]}{\gamma e^{\gamma(-c_l)}} - c_l \frac{-\operatorname{Ei}[-c_l\gamma]}{e^{\gamma(-c_l)}}\right)}{\frac{u}{2} - \frac{\operatorname{Ei}[-c_l\gamma]}{e^{\gamma(-c_l)}}} = l.$$

In contrast, $\lim_{c_h \to \infty} \frac{v(c_h;c_l,c_h)}{q(c_h;c_l,c_h)}$ takes the value of

$$\begin{aligned} &\lim_{c_h \to \infty} \frac{R + \frac{c_h u}{2} - \frac{u}{2\gamma} m(c_l, c_h) + \frac{R\gamma}{2}\left(\frac{g_h(c_l,c_h)}{\gamma} - c_h f_h(c_l, c_h)\right)}{\frac{u}{2} + \frac{R\gamma}{2} f_h(c_l, c_h)} \\ = &\lim_{c_h \to \infty} \frac{\frac{R}{c_h} + \frac{u}{2} - \frac{u}{c_h 2\gamma} + \frac{R\gamma}{2}\left(\frac{g_h(c_l,c_h)}{c_h\gamma} - f_h(c_l, c_h)\right)}{\frac{u}{2c_h} + \frac{R\gamma}{2}\frac{f_h(c_l,c_h)}{c_h}} = \lim_{c_h \to \infty} \frac{\frac{u}{2} + \frac{R\gamma}{2}\left(\frac{g_h(c_l,c_h)}{c_h\gamma}\right)}{\frac{R\gamma}{2}\frac{f_h(c_l,c_h)}{c_h}} = \infty, \end{aligned}$$

Hence, $\frac{v(c_h;c_l,c_h)}{q(c_h;c_l,c_h)}$ grows without bound for any fixed c_l, and $\frac{v(c_h;c_l,c_h)}{q(c_h;c_l,c_h)}$ is monotonically increasing in c_l. As a result, in order to have a solution of $\lim_{c_h\to\infty}\frac{v(c_h;c_l,c_h)}{q(c_h;c_l,c_h)} = l$, c_l has to go to zero, implying $\lim_{c_h\to\infty} H(c_h) = 0$.

The two results imply that there is always an intercept $c_h \in (h,\infty)$ that $H(c_h) = L(c_h)$. This concludes the step proving that (13)-(15), (18)-(19) has a solution

A.1.3 Step 3: Proof of Lemma A.3

We have shown that $H(h) = h$. Note also that if $c_h = c_l$ then $\frac{v_h}{q_h} = \frac{v_l}{q_l}$. This, and the continuity of $H(\cdot)$ and $L(\cdot)$ in l, implies that at the limit $l \to h$, there is a solution of the system (13)-(15), (18)-(19) that $c_l^* - c_h^* \to 0$ and $c_h^*, c_l^* \to h$. Then, the statement comes from $h < hu < R$.

A.1.4 Step 4: Proof of Lemma A.4

First we show that $q(c)$ is always deceasing, and there exists a critical value $\widehat{c} \in (c_l, c_h)$ so that $q''(c) < 0$ for $c \in (c_l, \widehat{c})$ and $q''(c) > 0$ for $c \in (\widehat{c}, c_h)$. Moreover, for $c \in (c_l, \widehat{c})$ where $q''(c) < 0$, we have that $q'''(c) > 0$.

1. To show that $q' < 0$, we differentiate the ODE $0 = \frac{\sigma^2}{2}q'' + \frac{\xi}{2}\left(u + \frac{R}{c}\right) - \xi q$ again to reach
$$0 = \frac{\sigma^2}{2}q''' - \frac{\xi}{2}\frac{R}{c^2} - \xi q'. \tag{A.4}$$
Due to boundary conditions, we have at both ends c_l^* and c_h^*, the function $q'(c)$ equals zero and its second derivative $\frac{\sigma^2}{2}q''' = \frac{\xi}{2}\frac{R}{c^2} > 0$. Suppose to the contrary that $q'(\widetilde{c}) > 0$ for some point $\widetilde{c} \in (c_l, c_h)$; then we can pick $\widetilde{c}$ so that $q'(\widetilde{c}) > 0$ and $q'''(\widetilde{c}) = 0$ (otherwise the function $q'(\cdot)$ is zero at one end, is convex globally, and thus never comes back to zero at the other end). But because $\frac{\sigma^2}{2}q'''(\widetilde{c}) = \frac{\xi}{2}\frac{R}{\widetilde{c}^2} + \xi q'(\widetilde{c}) > 0$, contradiction. This proves that $q' < 0$.

2. We know that $q''(c_l) < 0$ and $q''(c_h) > 0$, and therefore there exists $\widehat{c}$ so that $q''(\widehat{c}) = 0$. We show this point is unique. Because $0 = \frac{\sigma^2}{2}q'' + \frac{\xi}{2}\left(u + \frac{R}{c}\right) - \xi q$, we have $0 = \frac{\sigma^2}{2}q''' - \frac{\xi}{2}\frac{R}{c^2} - \xi q'$, and
$$0 = \frac{\sigma^2}{2}q'''' + \frac{\xi R}{c^3} - \xi q''. \tag{A.5}$$
Suppose we have multiple solutions for $q''(\widehat{c}) = 0$. Clearly, it is impossible to have $q''(\widehat{c}) = 0$ but $q''(\widehat{c}-) > 0$ and $q''(\widehat{c}+) > 0$; otherwise $q''''(\widehat{c}) > 0$ which contradicts with (A.5). Then there must exist two points $c_1 > \widehat{c}$ and $c_2 > c_1 > \widehat{c}$ that $q''(c_1) = 0$, $q''(c_2) < 0$ and $q''''(c_2) > 0$, but $q''(c) < 0$ for $c \in (c_1, c_2)$. This implies that $\frac{\sigma^2}{2}q''''(c_1) = -\frac{\xi R}{c_1^3} + \xi q''(c_1) < 0$. As a result, there exists another point $c_3 \in (c_1, c_2)$ so that $q''''(c_3) = 0$ with $q''(c_3) < 0$. But this contradicts with (A.5).

3. Now we show that for $c \in (c_l, \widehat{c})$ with $q''(c) < 0$, we have $q'''(c) > 0$, i.e., $q''(c)$ is increasing. Suppose not. Since $q'''(c_l) > 0$ so that $q''(c)$ is increasing at the beginning, there must exist

some reflecting point c_4 for the function q'' so that $q''''(c_4) = 0$. But because $q''(c_4) < 0$, it contradicts with (A.5).

Second, we show that $v(c)$ is increasing if $h - l$ is sufficiently small.

1. We show that if $v''(c_l) > 0$, then $v(c)$ is increasing in c. Let $F(c) \equiv v'(c)$, so that

$$0 = q''\sigma^2 + \frac{\sigma^2}{2}F'' + \frac{\xi}{2}u - \xi F$$

with boundary conditions that $F(c_l) = F(c_h) = 0$. The assumption $v''(c_l) > 0$ implies that $F'(c_l) > 0$. Thus, if there are some points with $F(c) < 0$ in the range of (c_l, c_h) then we can find two points c_1 and c_2 (a maximum and a minimum) so that $c_1 < c_2$ but $F''(c_1) < 0$ $F''(c_2) > 0$, $F'(c_1) = F'(c_2) = 0$ and $F(c_1) > 0 > F(c_2)$. We can apply the ODE to these two points:

$$\begin{aligned} 0 &= q''(c_1)\sigma^2 + \frac{\sigma^2}{2}F''(c_1) + \frac{\xi}{2}u - \xi F(c_1), \\ 0 &= q''(c_2)\sigma^2 + \frac{\sigma^2}{2}F''(c_2) + \frac{\xi}{2}u - \xi F(c_2). \end{aligned}$$

The second equation implies that $q''(c_2) < 0$, which implies that $c_1 < c_2 < \widehat{c}$. However, the above two equations also imply that

$$q''(c_1)\sigma^2 > \frac{\xi}{2}u > q''(c_2)\sigma^2$$

contradiction with the previous lemma which shows that q'' is increasing over $[c_l, \widehat{c}]$.

2. Now we show that if $h-l$ is sufficiently small, then $v''(c_l) > 0$; with the first result we obtain our claim. From our ODE,

$$v''(c_l) = -\frac{\xi}{\sigma^2}2\left(\frac{(uc_l + R)}{2} - v(c_l)\right) = \frac{\xi}{\sigma^2}2\left(\frac{R}{2} + \frac{u}{2\gamma}h(c_l, c_h) + \frac{R\xi}{\gamma\sigma^2}\left(\frac{g_l(c_l, c_h)}{\gamma} - c_l f_l(c_l, c_h)\right)\right).$$

We know that as $h - l \to 0$, $c_h - c_l \to 0$. We will prove the statement by showing that (1) $\lim_{c_l \to c_h}\left(\frac{(uc_l+R)}{2} - v(c_l)\right) = 0$, because $\lim_{c_l \to c_h}\left(\frac{(uc_l+R)}{2} - v(c_l)\right)$ equals

$$\lim_{c_l \to c_h}\left(\frac{R}{2} + \frac{u}{2\gamma}h(c_l, c_h) + \frac{R\gamma}{2}\left(\frac{g_l(c_l, c_h)}{\gamma} - c_l f_l(c_l, c_h)\right)\right) = \frac{R}{2} + 0 + \frac{R\xi}{\gamma\sigma^2}\left(0 - \frac{1}{\gamma}\right) = 0$$

and (2) $\lim_{c_l \to c_h}\frac{\partial\left(\frac{(uc_l+R)}{2} - v(c_l)\right)}{\partial c_l} = \lim_{c_l \to c_h}\frac{\partial\left(\frac{u}{2\gamma}h(c_l, c_h) + \frac{R\gamma}{2}\left(\frac{g_l(c_l, c_h)}{\gamma} - c_l f_l(c_l, c_h)\right)\right)}{\partial c_l} < 0$, because it equals

$$\lim_{c_l\to c_h}\left(-\frac{ue^{\gamma(c_h-c_l)}}{\left(e^{\gamma(c_h-c_l)}+1\right)^2}+\frac{R\gamma}{2}\left(\frac{1}{\gamma c_l}+\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}g_l-\frac{\left(e^{2\gamma c_h}+e^{2\gamma c_l}\right)}{\left(e^{2\gamma c_h}-e^{2\gamma c_l}\right)}\left(c_l\gamma f_l-1\right)\right)\right)$$
$$= \quad -u\frac{1}{(1+1)^2}+\frac{R\gamma}{2}\left(\frac{1}{\gamma c_h}-\frac{1}{2\gamma c_h}-\frac{1}{2\gamma c_h}\right)=-\frac{u}{4}<0.$$

These two statements imply that when c_h-c_l is sufficiently small then $v''(c_l)>\lim_{c_l\to c_h}v''(c_l)=0$.

A.2 Proof of Proposition 2

The result $c_h^*>h$ is a consequence of the fact that we defined $H(c_h)$ as the unique c_l solving $\frac{v_h(c_l,c_h)}{q_h(c_l,c_h)}=h$ when $c_h>h$. (see part 4 in section A.1.2.)

For the result $c_l^*\le l$, consider the possibility that $c_l^*>l$. The following lemma states that in this case $p''(c_l^*)<0$. This implies that this is not an equilibrium. To see this, we have $p'(c_l^*)=0$ by the boundary conditions $v'(c_l^*)=q'(c_l^*)=0$. Thus $p''(c_l^*)<0$, combined with $p(c_l^*)=l$ and $p'(c_l^*)=0$, would imply that $p(c)<l$ for c sufficiently close to c_l^*.

Lemma A.5 *The sign of $p''(c_l^*)$ is the same as that of $l-c_l^*$.*

Proof. Simple algebra implies that

$$\begin{aligned}
p''(c_l^*) &= \left(\frac{v'q-q'v}{q^2}\right)'=\frac{\left(v''q+v'q'-\left(q''v+v'q'\right)\right)}{q^2}-2q^{-3}\left(v'q-q'v\right)\\
&= \frac{v''q-q''v}{q^2}=\frac{\left(-\frac{\xi}{2}\left(uc_l^*+R\right)+\xi lq\left(c_l^*\right)\right)\frac{2}{\sigma^2}q-\left(-\frac{\xi}{2}\left(uc_l^*+R\right)+\xi c_l^*q\left(c_l^*\right)\right)\frac{2}{\sigma^2c_l^*}v}{q^2}\\
&= \frac{\left(-\frac{\xi}{2}\left(uc_l^*+R\right)+\xi lq\left(c_l^*\right)+\xi c_l^*q\left(c_l^*\right)-\xi c_l^*q\left(c_l^*\right)\right)\frac{2}{\sigma^2}q-\left(-\frac{\xi}{2}\left(uc_l^*+R\right)+\xi c_l^*q\left(c_l^*\right)\right)\frac{2}{\sigma^2c_l^*}v}{q^2}\\
&= \frac{\left(-\frac{\xi}{2}\left(uc_l^*+R\right)+\xi c_l^*q\left(c_l^*\right)\right)\frac{2}{\sigma^2}\left(q-\frac{v}{c_l^*}\right)+\left(l-c_l^*\right)\xi q\left(c_l^*\right)\frac{2}{\sigma^2}q}{q^2}\\
&= \left(l-c_l^*\right)\frac{\frac{1}{c_l^*}\left(\frac{\xi}{2}\left(uc_l^*+R\right)-\xi c_l^*q\left(c_l^*\right)\right)\frac{2}{\sigma^2}+\xi q\left(c_l^*\right)\frac{2}{\sigma^2}}{q}
\end{aligned}$$

which gives the Lemma by noticing that q is decreasing in c and the boundary $q'(c_l^*)=0$ implies that

$$-\frac{\xi}{2}\left(uc_l^*+R\right)+\xi c_l^*q\left(c_l^*\right)\propto q''\left(c_l^*\right)<0.$$

■

The third statement is a consequence of the following Lemma.

Lemma A.6 *We have the following limiting results:*

$$\begin{aligned}\lim_{\gamma\to\infty}\gamma f_l &= \frac{1}{c_l},\ \lim_{\gamma\to\infty}\gamma f_h=\frac{1}{c_h},\ \lim_{\gamma\to\infty}g_h=0,\ \lim_{\gamma\to\infty}g_l=0;\\ \textit{and } \lim_{\gamma\to\infty}c_h^* &= h,\ \lim_{\gamma\to\infty}c_l^*=l.\end{aligned}$$

Proof. The first four results are based on L'Hopital rule. Take the first result for illustration:

$$\begin{aligned}\lim_{\gamma\to\infty}\gamma f_l &= \lim_{\gamma\to\infty}\frac{\gamma\left(\mathrm{Ei}[-c_h\gamma]-\mathrm{Ei}[-c_l\gamma]\right)}{e^{\gamma(-c_l)}}=\lim_{\gamma\to\infty}\frac{\mathrm{Ei}[-c_h\gamma]-\mathrm{Ei}[-c_l\gamma]}{\frac{1}{\gamma}e^{\gamma(-c_l)}}\\ &= \lim_{\gamma\to\infty}\frac{\frac{e^{-c_h\gamma}}{\gamma}-\frac{e^{-c_l\gamma}}{\gamma}}{-\frac{1}{\gamma^2}e^{\gamma(-c_l)}+\frac{(-c_l)}{\gamma}e^{\gamma(-c_l)}}=\lim_{\gamma\to\infty}\frac{-e^{-c_l\gamma}/\gamma}{\frac{(-c_l)}{\gamma}e^{\gamma(-c_l)}}=\frac{1}{c_l}.\end{aligned}$$

These four results imply that

$$\lim_{\gamma\to\infty}\frac{v_h}{q_h}=\lim_{\gamma\to\infty}\frac{R+\frac{c_hu}{2}-\frac{u}{2\gamma}m\left(c_l,c_h\right)+R\frac{\gamma}{2}\left(\frac{g_h(c_l,c_h)}{\gamma}-c_hf_h\left(c_l,c_h\right)\right)}{\frac{u}{2}+R\frac{\gamma}{2}f_h\left(c_l,c_h\right)}=\frac{R+\frac{c_hu}{2}-R\frac{1}{2}}{\frac{u}{2}+R\frac{1}{2c_h}}$$

Thus, in the limit the solution of $\frac{v_h}{q_h}=h$ is the solution for the equation of

$$\frac{R+\frac{c_hu}{2}-R\frac{1}{2}}{\frac{u}{2}+R\frac{1}{2c_h}}=h,$$

which gives $\lim_{\gamma\to\infty}c_h^*=h$. Similarly, the following calculation implies that $\lim_{\gamma\to\infty}c_l^*=l$:

$$\lim_{\gamma\to\infty}\frac{v_l}{q_l}=\lim_{\gamma\to\infty}\frac{R+\frac{c_lu}{2}+\frac{u}{2\gamma}m\left(c_l,c_h\right)+R\frac{\gamma}{2}\left(\frac{g_l(c_l,c_h)}{\gamma}-c_lf_l\left(c_l,c_h\right)\right)}{\frac{u}{2}+R\frac{\gamma}{2}f_l\left(c_l,c_h\right)}=\frac{R+\frac{c_lu}{2}+R\frac{1}{2}}{\frac{u}{2}+R\frac{1}{2c_l}}.$$

■

A.3 Proof of Proposition 3

The solution for D_1-D_2 is obvious. To verify that $c_l^P=0$, we have to show that

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

Now we show that the solution exists and unique. Define a function $G(c)$:

$$G(c)\equiv\frac{R-hu}{R-lu}\left(e^{c\gamma}\left(1+l\gamma\right)-\left(1-l\gamma\right)e^{-c\gamma}\right)-2\gamma\left(c+h\right), \tag{A.6}$$

with $G(0) = 2R\gamma\frac{l-h}{R-lu} < 0$ (recall $R - hu > R - lu > 0$) and $G(\infty) = \infty$. We have $G'(c) = \gamma\left(\frac{(R-hu)}{(R-lu)}\left((l\gamma+1)e^{c\gamma} + e^{-c\gamma}(1-l\gamma)\right) - 2\right)$, $G'(0) = 2u\gamma\frac{l-h}{R-lu} < 0$, 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 $G\left(c_h^P\right) = 0$, verifying the equation (29).

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

$$0 = \frac{\sigma^2}{2}j_P''(c) + \xi\left(R + uc - j_P(c)\right) \tag{A.7}$$

with boundary conditions $j_P(0) = lj_P'(0)$, $j_P\left(c_h^P\right) = \left(h + c_h^P\right)j_P'\left(c_h^P\right)$, and $j_P''\left(c_h^P\right) = 0$. Note that the boundary conditions imply that $j_P\left(c_h^P\right) = R + uc_h^P$. For later reference, we show that $j_P(c)$ is concave and increasing over $[0, c^*]$, and $j_P(c) < R + uc$. First, from smooth pasting condition at c_h^P we have

$$u - j_P'\left(c_h^P\right) = u - \frac{j_P\left(c_h^P\right)}{h + c_h^P} = u - \frac{R + uc_h^P}{h + c_h^P} = \frac{uh - R}{h + c_h^P} < 0.$$

Then, taking derivative again on (A.7) and evaluate at the optimal policy point c_h^P, we have

$$j_P'''\left(c_h^P\right) = -\frac{2\xi}{\sigma^2}\left(u - j_P'\left(c_h^P\right)\right) = \frac{2\xi}{\sigma^2}\frac{R - uh}{h + c_h^P} > 0,$$

and as a result $j_P''\left(c_h^P-\right) < 0$. Suppose that j_P fails to be globally concave over $\left[0, c_h^P\right]$. Then there exists some point $j_P'' > 0$, and pick the largest one $\widehat{c}$ so that j_P'' is concave over $[\widehat{c}, c^*]$ with $j_P''(\widehat{c}) = 0$ and $j_P'''(\widehat{c}) < 0$. But since j_P'' is concave over $\left[\widehat{c}, c_h^P\right]$, $j_P'(\widehat{c}) > j_P'\left(c_h^P\right) > u$, therefore $\frac{\sigma^2}{2}j_P'''\left(c_h^P\right) = \xi\left(j_P'\left(c_h^P\right) - u\right) > 0$, contradiction. Therefore j_P is globally concave over $\left[0, c_h^P\right]$, which also implies that $j_P(c) < R + uc$ due to (A.7).

We may also need to evaluate the social value $j_P(c)$ for $c > c_h^P$. Because the optimal policy is investing, if $C > Kc_h^P$ so that $c > c_h^P$, then immediately the economy should build $K\frac{c - c_h^P}{h + c_h^P}$ capital to keep the cash-to-capital ratio at c_h^P. This implies a social value of

$$j_P(c) = \frac{1}{K}\left(K + K\frac{c - c_h^P}{h + c_h^P}\right)j_P\left(c_h^P\right) = \left(\frac{h + c}{h + c_h^P}\right)j\left(c_h^P\right) \text{ for } c > c_h^P.$$

A.4 Proof of Proposition 4

Recall $G(c)$ defined in (A.6). Since $\lim_{\gamma\to\infty}\frac{\frac{R-hu}{R-lu}\left(e^{c\gamma}(1+l\gamma) - (1-l\gamma)e^{-c\gamma}\right)}{\gamma c} = \infty$, to ensure that $G\left(c_h^P\right) = 0$ as $\gamma \to \infty$ we must have $c_h^P \to 0$. This is the first part of the first statement. In addition,

$$\frac{\partial G(c)}{\partial\gamma} = \frac{R - hu}{R - lu}\left(ce^{c\gamma}(1 + l\gamma) + le^{c\gamma} - c(l\gamma - 1)e^{-c\gamma} + le^{-c\gamma}\right) - 2(c + h),$$

which is positive for sufficiently large γ. Finally, from the proof of Proposition 3 we know that $G'\left(c_h^P\right) > 0$. Hence, for sufficiently large γ, we have $\frac{\partial c_h^p}{\partial\gamma} = -\frac{\partial G\left(c_h^p\right)/\partial\gamma}{G'\left(c_h^p\right)} < 0$ which concludes the first part. The second part follows because $\frac{R-hu}{R-lu}$ is increasing in R, and

$$\begin{aligned}
\frac{\partial G(c)}{\partial h} &= \frac{-u}{R-lu}\left(e^{c\gamma}(1+l\gamma)-(1-l\gamma)e^{-c\gamma}\right)-2\gamma<0,\\
\frac{\partial G(c)}{\partial l} &= (R-hu)\frac{u\left(1-e^{-2c\gamma}\right)+R\gamma+R\gamma e^{2(-c\gamma)}}{e^{-c\gamma}(R-lu)^2}>0.
\end{aligned}$$

Finally, fixing any c we have $\lim_{R\to hu} G(c) = -2\gamma(c+h) < 0$ always. This implies that for $\lim_{R\to hu} G\left(c_h^P\right)=0$ to hold, it must be that $c_h^p\to\infty$ so that $\lim_{R\to hu}\frac{\frac{R-hu}{R-lu}\left(e^{c\gamma}(1+l\gamma)-(1-l\gamma)e^{-c\gamma}\right)}{c}\to 2\gamma$. This concludes the proposition.

A.5 Proof of Proposition 5

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 $j_P''\left(c_h^P;0,c_h^P\right)=0$. 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 $j_P\left(c;0,c_h^P\right)$ by $j_P(c)$. We need to show that

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

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

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 suboptimal, we must have $j(c;c_l,c_h)<j_P(c)\le R+uc$ (recall Proposition 3). Then $0=\frac{\sigma^2}{2}j''(c)+\xi(R+uc-j(c))$ 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 social planner's solution), the smooth pasting condition (not optimality condition!) at the regulated ends implies that

$$\begin{aligned}
j(c_h;c_l,c_h)-(c_h+h)j'(c_h;c_l,c_h) &= 0, &\text{(A.8)}\\
j(c_l;c_l,c_h)-(c_l+l)j'(c_l;c_l,c_h) &= 0. &\text{(A.9)}
\end{aligned}$$

Now we start proving the properties for the top 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 top investment policy on the social value. Differentiating the basic ODE 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

$$\frac{\sigma^2}{2}F_h''(c;c_l,c_h)-\xi F_h(c;c_l,c_h)=0. \tag{A.10}$$

Moreover, take the total derivative with respect to c_h on the equality (A.8), i.e., take derivative that affects both the policy c_h and the state $c = c_h$, we have

$$\begin{aligned}
&\frac{\partial}{\partial c_h} j\left(c_h; c_l, c_h\right) + j'\left(c_h; c_l, c_h\right) = j'\left(c_h; c_l, c_h\right) + \left(c_h + h\right)\left(\frac{\partial}{\partial c_h} j'\left(c_h; c_l, c_h\right) + j''\left(c_h; c_l, c_h\right)\right) \\
\Rightarrow\quad &\frac{\partial}{\partial c_h} j\left(c_h; c_l, c_h\right) - \left(c_h + h\right)\frac{\partial}{\partial c_h} j'\left(c_h; c_l, c_h\right) = \left(c_h + h\right) j''\left(c_h; c_l, c_h\right) < 0 \\
\Rightarrow\quad &F_h\left(c_h; c_l, c_h\right) - \left(c_h + h\right) F_h'\left(c_h; c_l, c_h\right) < 0. \qquad \text{(A.11)}
\end{aligned}$$

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.9), we have the boundary condition of $F_h(\cdot)$ at c_l:

$$\frac{\partial}{\partial c_h} j\left(c_l; c_l, c_h\right) = \left(c_l + l\right)\frac{\partial}{\partial c_h} j'\left(c_l; c_l, c_h\right) \Rightarrow F_h\left(c_l; c_l, c_h\right) - \left(c_l + l\right) F_h'\left(c_l; c_l, c_h\right) = 0. \qquad \text{(A.12)}$$

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

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

1. *$F_h(c)$ cannot change sign over $[c_l, c_h]$. Suppose that $F_h(c_l) > 0$; then from (A.12) we know that $F_h'(c_l) > 0$. Then simple argument based on ODE (A.10) 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 $\widehat{c}$ so that $F_h'(\widehat{c}) = 0$, $F_h(\widehat{c}) > 0$ and $F_h''(\widehat{c}) < 0$. But this contradicts with the ODE (A.10).*

2. *Define $W_h(c) \equiv F_h(c) - (l + c) F_h'(c)$ so that $W_h'(c) = -(l + c) F_h''(c) = -\frac{2\xi(l+c)}{\sigma^2} F_h(c)$. 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 $W_h'(c) > 0$. Step 2 implies that $W_h(c) > 0$, and $F_h'(c_h) = \frac{h-l}{l+c}(F_h - W_h) < 0$. But we then have*

$$W_h\left(c_h\right) = F_h\left(c_h\right) - (l + c) F_h'\left(c_h\right) = F_h\left(c_h\right) - (h + c) F_h'\left(c_h\right) + (h - l) F_h'\left(c_h\right) < 0,$$

where we have used (A.11), contradiction. Thus we have shown that $F_h(c) > 0$.

■

A.6 Proof of Proposition 6

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

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

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

$$\frac{\sigma^2}{2}T''\left(c; c_l\right) = \xi T\left(c; c_l\right) \Rightarrow \frac{\sigma^2}{2}F''\left(c; c_l\right) = \xi F\left(c; c_l\right). \tag{A.13}$$

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

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

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

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

Here we used the fact that $T''\left(c = c_l; c_l\right) > 0$ due to (A.13) and $T > 0$ by definition.

Now we show that $F\left(c\right) > 0$ so that the total investment activity goes up for a higher c_l. To see this, first note that $F\left(c\right)$ never changes sign. Otherwise, suppose that there exists some c_1 so that $F\left(c_1\right) = 0$. If $F'\left(c_1\right) > 0$ then it must be that F is convex and positive for $c > c_1$, which contradicts with $F'\left(c_h\right) = 0$. Similarly we rule out $F'\left(c_1\right) < 0$. If $F'\left(c_1\right) = 0$, then combining with $F\left(c_1\right)$ we can solve for $F\left(c\right) = 0$ for all c, contradicting with $F'\left(c_l\right) < 0$. Now since $F\left(c\right)$ never changes sign, it suffices to rule out $F\left(c\right) < 0$ always. If it were true, then F is concave always due to (A.13). This contradicts with $F'\left(c_l\right) < 0 = F'\left(c_h\right)$. As a result, $F\left(c\right) > 0$.

A.7 Proof of Proposition 7

It is easy to check that the general forms (33)-(34) are indeed solutions of (31)-(32). Now we show that equations (35) have a solution $B_1, B_2, B_3, B_4, c_h^{cm}$ where $c_h^{cm} = c_h^P$, and that $j_P\left(c\right) = v_{cm}\left(c\right) + cq_{cm}\left(c\right)$. Observe first that

$$v_{cm}\left(c\right) + cq_{cm}\left(c\right) = R + uc + e^{-c\gamma}B_3 + e^{c\gamma}B_4. \tag{A.14}$$

Also, (35) can be written as

$$\begin{aligned}
\frac{R+B_3+B_4}{\frac{u}{2}+B_1+B_2} &= l, && \text{(A.15)}\\
\frac{u}{2}+\gamma B_3-B_2-\gamma B_4-B_1 &= 0, && \text{(A.16)}\\
-\gamma e^{-c_h^{cm}\gamma}B_1+\gamma e^{c_h^{cm}\gamma}B_2 &= 0, && \text{(A.17)}\\
\frac{u}{2}+\gamma e^{c_h^{cm}\gamma}\left(B_3-c_h^{cm}B_2\right)-B_2e^{c_h^{cm}\gamma}-\gamma e^{-c_h^{cm}\gamma}\left(B_4-c_h^{cm}B_1\right)-B_1e^{-c_h^{cm}\gamma} &= 0, && \text{(A.18)}\\
\frac{R+uc_h^{cm}+e^{c_h^{cm}\gamma}B_3+e^{-c_h^{cm}\gamma}B_4}{\frac{u}{2}+e^{-c_h^{cm}\gamma}B_1+e^{c_h^{cm}\gamma}B_2} &= h+c_h^{cm}. && \text{(A.19)}
\end{aligned}$$

Adding c_h^{cm} times (A.17) to (A.18) gives

$$\frac{u}{2}+e^{c_h^{cm}\gamma}\left(\gamma B_3-B_2\right)-e^{-c_h^{cm}\gamma}\left(\gamma B_4+B_1\right)=0. \tag{A.20}$$

Together with (A.16) this implies

$$\gamma B_3=B_2,\ \text{and}\ -B_1=\gamma B_4. \tag{A.21}$$

Substituting this into (A.17) gives

$$e^{-c_h^{cm}\gamma}B_4+e^{c_h^{cm}\gamma}B_3=0. \tag{A.22}$$

Also, expressing (B_1+B_2) from (A.16) and plugging into (A.15) gives

$$R+B_3+B_4=l\left(u+\gamma B_3-\gamma B_4\right) \tag{A.23}$$

and by (A.21), (A.19) is equivalent to

$$R+uc_h^{cm}+e^{c_h^{cm}\gamma}B_3+e^{-c_h^{cm}\gamma}B_4=\left(h+c_h^{cm}\right)\left(u-\gamma B_4e^{-c_h^{cm}\gamma}+\gamma B_3e^{c_h^{cm}\gamma}\right). \tag{A.24}$$

Observe that the system (A.22)-(A.24) is equivalent to the system of boundary conditions (24), thus $B_3=D_2$, $B_4=D_1$ and $c_h^{cm}=c_h^P$. Given (A.14) and the fact that (A.21), we proved the statement.

Finally, $v'_{cm}(c)>0$ because

$$\begin{aligned}
v'_{cm}(c) &= \frac{u}{2}+\gamma e^{c\gamma}\left(B_3-cB_2\right)-B_2e^{c\gamma}-\gamma e^{-c\gamma}\left(B_4-cB_1\right)-B_1e^{-c\gamma}\\
&= \frac{u}{2}-c\gamma^2D_2e^{c\gamma}-c\gamma^2D_1e^{-c\gamma}=\frac{u}{2}+c\gamma^2\left(R-lu\right)e^{c\gamma}\frac{e^{2\gamma\left(c_h^P-c\right)}-1}{e^{2\gamma c_h^P}+l\gamma\left(e^{2\gamma c_h^P}-1\right)+1}>0;
\end{aligned}$$

and $q'_{cm}(c) < 0$ because

$$\begin{aligned} q'_{cm}(c) &= -\gamma e^{-c\gamma} B_1 + \gamma e^{c\gamma} B_2 = \gamma e^{-c\gamma}\gamma B_4 + \gamma e^{c\gamma}\gamma B_3 \\ &= \gamma e^{-c\gamma}\gamma D_1 + \gamma e^{c\gamma}\gamma D_2 = \gamma^2 (R - lu) e^{c\gamma} \frac{1 - e^{2\gamma\left(c_h^P - c\right)}}{e^{2\gamma c_h^P} + l\gamma\left(e^{2\gamma c_h^P} - 1\right) + 1} < 0. \end{aligned}$$

These two results imply that the price is monotonically increasing.

A.8 Proof of Proposition 8

The first statement comes from the construction of the Proof of Proposition 1. In particular, from the fact that c_h^* and c_l^* are constructed as the intercept of continuous functions $H(c_h)$ and $L(c_h)$, with both mapping $[h, \infty) \rightarrow R^{++}$, $H(h) = h > L(h) > 0$, and $0 < \lim_{c_h \to \infty} L(c_h)$ $\lim_{c_h \to \infty} H(c_h) = 0 < \lim_{c_h \to \infty} L(c_h) < \infty$. Thus, both $c_h^* \in (h, \infty)$ and $c_l^* \in (0, c_h^*)$.

The second statement is the consequence of Lemma A.6 and the first result in Proposition 4.

For the last statement, note that the proof of Proposition 1 goes through without any modification for the case when $R = uh$. That is, even in the limit $R \to uh$, c_h^* is finite. However, Proposition 4 states that for any parameters, in the limit $R \to uh$, $c_h^P \to \infty$. This gives the result.

A.9 Proof of Proposition 9

Consider the functions $\tilde{q}(c; q_0, v_0, c_h)$ and $\tilde{v}(c; q_0, v_0, c_h)$ of c parameterized by q_0, v_0, and c_h:

$$0 = \frac{\sigma^2}{2}\tilde{q}''(c) + \frac{\xi}{2}(u - \tilde{q}(c)) + \frac{\xi}{2}\left(\frac{R}{c} - \tilde{q}(c)\right) \tag{A.25}$$

$$0 = \tilde{q}'(c)\sigma^2 + \frac{\sigma^2}{2}\tilde{v}''(c) + \frac{\xi}{2}(uc - \tilde{v}(c)) + \frac{\xi}{2}(R - \tilde{v}(c)). \tag{A.26}$$

and the boundary conditions

$$\tilde{v}'(c_h) = \tilde{q}'(c_h) = 0, \tag{A.27}$$

$$\tilde{q}(c_0) = q_0, \tilde{v}(c_0) = v_0. \tag{A.28}$$

The general solution is

$$\widetilde{q}(c) = \frac{u}{2} + e^{-c\gamma} A_1 + e^{c\gamma} A_2 + \frac{R\gamma}{2} \frac{-e^{c\gamma}\,\mathrm{Ei}(-\gamma c) + e^{-c\gamma}\,\mathrm{Ei}(c\gamma)}{2} \tag{A.29}$$

$$\widetilde{v}(c) = R + \frac{cu}{2} + e^{c\gamma}(A_3 - cA_2) - e^{-c\gamma}(A_4 + cA_1) + \frac{cR\gamma}{2} \frac{e^{\gamma c}\,\mathrm{Ei}(-\gamma c) - e^{-c\gamma}\,\mathrm{Ei}(\gamma c)}{2} \tag{A.30}$$

$$= R + uc + e^{c\gamma} A_3 - e^{-c\gamma} A_4 - c\widetilde{q}(c). \tag{A.31}$$

where A_1-A_4 (may differ from those in (18) and (19)) are pinned down by (A.27)-(A.28). We have

$$\begin{aligned}
\tilde{q}'(c) &= -\gamma e^{-c\gamma}A_1 + \gamma e^{c\gamma}A_2 - \frac{R\gamma^2\left(e^{-c\gamma}\operatorname{Ei}[c\gamma] + e^{c\gamma}\operatorname{Ei}[-c\gamma]\right)}{2}, \\
\tilde{v}'(c) &= \frac{u}{2} + \frac{R\gamma\left(-e^{-c\gamma}\operatorname{Ei}[c\gamma] + e^{c\gamma}\operatorname{Ei}[-c\gamma]\right)}{2} + \frac{Rc\gamma^2\left(e^{-c\gamma}\operatorname{Ei}[c\gamma] + e^{c\gamma}\operatorname{Ei}[-c\gamma]\right)}{2} \\
&\quad + e^{c\gamma}\left(\left(-\gamma c - 1\right)A_2 + \gamma A_3\right) + e^{-c\gamma}\left(\left(\gamma c - 1\right)A_1 + \gamma A_4\right).
\end{aligned}$$

Define the function $c_h(q_0, v_0)$ implicitly by $\tilde{v}(c_h; q_0, v_0, c_h) = h\tilde{q}(c_h; q_0, v_0, c_h)$, and we are interested in the derivatives

$$\frac{\partial c_h}{\partial q_0} = -\frac{\tilde{v}'_{q_0} - h\tilde{q}'_{q_0}}{\tilde{v}'_{c_h} - h\tilde{q}'_{c_h}}, \frac{\partial c_h}{\partial v_0} = -\frac{\tilde{v}'_{v_0} - h\tilde{q}'_{v_0}}{\tilde{v}'_{c_h} - h\tilde{q}'_{c_h}}.$$

For this, consider the following Lemmas.

Lemma A.8 *We have*

$$\frac{\partial \tilde{q}(c_h; q_0, v_0, c_h)}{\partial q_0} = \frac{2}{e^{c_h\gamma}e^{-c_0\gamma} + e^{-c_h\gamma}e^{\gamma c_0}} > 0, \tag{A.32}$$

$$\frac{\partial \tilde{v}(c_h; q_0, v_0, c_h)}{\partial v_0} = \frac{2}{e^{-\gamma(c_h - c_0)} + e^{\gamma(c_h - c_0)}} > 0, \frac{\partial \tilde{q}(c_h; q_0, v_0, c_h)}{\partial v_0} = 0. \tag{A.33}$$

Proof. We show (A.32) first. We know that $\tilde{q}(c_0) = q_0$, which based on (A.29) can be written as $e^{-c_0\gamma}A_1 + e^{\gamma c_0}A_2 + l_q = q_0$ (where l_q is independent of q_0) which implies

$$A_1 = \frac{-l_q - e^{\gamma c_0}A_2 + q_0}{e^{-c_0\gamma}}. \tag{A.34}$$

and $\tilde{q}'(c_h) = 0$ which can be rewritten as $-e^{-c_h\gamma}\gamma A_1 + e^{c_h\gamma}\gamma A_2 + s_q = 0$ (where s_q is independent of q_0) which implies

$$A_2 = \frac{e^{-c_h\gamma}\gamma A_1 - s_q}{e^{c_h\gamma}\gamma} = \frac{e^{-c_h\gamma}\gamma\frac{-l_q - e^{\gamma c_0}A_2 + q_0}{e^{-c_0\gamma}} - s_q}{e^{c_h\gamma}\gamma} \Rightarrow A_2 = \frac{e^{-c_h\gamma}\gamma\frac{-l_q + q_0}{e^{-c_0\gamma}} - s_q}{\left(1 + e^{-2c_h\gamma}e^{\gamma 2c_0}\right)e^{c_h\gamma}\gamma}. \tag{A.35}$$

Thus, (A.35) and (A.34) imply that

$$\frac{\partial A_2}{\partial q_0} = \frac{e^{-c_h\gamma}}{e^{c_h\gamma}e^{-c_0\gamma} + e^{-c_h\gamma}e^{\gamma c_0}}, \tag{A.36}$$

$$\frac{\partial A_1}{\partial q_0} = \frac{1}{e^{-c_0\gamma}} - e^{\gamma 2c_0}\frac{e^{-c_h\gamma}}{e^{c_h\gamma}e^{-c_0\gamma} + e^{-c_h\gamma}e^{\gamma c_0}} = \frac{e^{c_h\gamma}}{e^{c_h\gamma}e^{-c_0\gamma} + e^{-c_h\gamma}e^{\gamma c_0}}. \tag{A.37}$$

Using (A.29) we obtain our result.

The first result in (A.33) follows similarly. The second result $\frac{\partial \tilde{q}(c_h; q_0, v_0, c_h)}{\partial v_0} = 0$ comes from the fact that (A.25) and the boundary conditions $\tilde{q}'(c_h) = 0$ and $\tilde{q}(c_0) = q_0$ are independent of v_0. ■

Lemma A.9 *We have*

$$
\begin{aligned}
\frac{\partial \widetilde{v}\left(c_h;q_0,v_0,c_h\right)}{\partial q_0} &= 2\frac{e^{\gamma(c_h-c_0)}-e^{-\gamma(c_h-c_0)}-\gamma\left(c_h-c_0\right)\left(e^{-\gamma(c_h-c_0)}+e^{\gamma(c_h-c_0)}\right)}{\gamma\left(e^{\gamma c_0}e^{-\gamma c_h}+e^{-\gamma c_0}e^{\gamma c_h}\right)^2}<0,\\
\frac{\partial \widetilde{v}\left(c_h;q_0,v_0,c_h\right)}{\partial q_0}-h\frac{\partial \widetilde{q}\left(c_h;q_0,v_0,c_h\right)}{\partial q_0} &= 2\frac{e^{\gamma(c_h-c_0)}-e^{-\gamma(c_h-c_0)}-\gamma\left(c_h+h-c_0\right)\left(e^{-\gamma(c_h-c_0)}+e^{\gamma(c_h-c_0)}\right)}{\gamma\left(e^{\gamma c_0}e^{-\gamma c_h}+e^{-\gamma c_0}e^{\gamma c_h}\right)^2}<0
\end{aligned}
$$

Proof. We show the first result. We rewrite $\widetilde{v}\left(c_0\right)$ and $\widetilde{v}'\left(c_h\right)$ as (as before here l_{vq} and s_{vq} are independent of q_0)

$$
\begin{aligned}
\widetilde{v}\left(c_0\right) &= e^{c_0\gamma}\left(A_3-c_0A_2\right)-e^{-c_0\gamma}\left(A_4+c_0A_1\right)+l_{vq},\\
\widetilde{v}'\left(c_h\right) &= s_{vq}+e^{c_h\gamma}\left(\left(-\gamma c_h-1\right)A_2+\gamma A_3\right)+e^{-c_h\gamma}\left(\left(\gamma c_h-1\right)A_1+\gamma A_4\right)
\end{aligned}
$$

Thus, the boundary conditions $\widetilde{v}\left(c_0\right)=v_0$ and $\widetilde{v}'\left(c_h\right)=0$ imply that

$$
\begin{aligned}
A_3 &= c_0A_2+e^{-c_0\gamma}v_0-e^{-c_0\gamma}l_{vq}+e^{-2c_0\gamma}\left(A_4+c_0A_1\right),\\
A_4 &= -\frac{\begin{array}{c}\left(-e^{\gamma c_h}\left(\gamma c_h-\gamma c_0+1\right)\right)A_2+\left(e^{-\gamma c_h}\left(\gamma c_h-1\right)+\gamma c_0e^{-2\gamma c_0}e^{\gamma c_h}\right)A_1\\ +\left(\gamma e^{-\gamma c_0}e^{\gamma c_h}\right)v_0+\left(s_{qv}-\gamma e^{-\gamma c_0}e^{\gamma c_h}l_{qv}\right)\end{array}}{\gamma e^{-\gamma c_h}+\gamma e^{-2\gamma c_0}e^{\gamma c_h}}
\end{aligned}
$$

Thus, using the result in (A.36) and (A.37) one can derive that

$$
\frac{\partial A_4}{\partial q_0}=e^{\gamma c_h}\frac{2e^{\gamma c_0}e^{-\gamma c_h}-\gamma c_0\left(e^{\gamma c_0}e^{-\gamma c_h}+e^{-\gamma c_0}e^{\gamma c_h}\right)}{\gamma\left(e^{\gamma c_0}e^{-\gamma c_h}+e^{-\gamma c_0}e^{\gamma c_h}\right)^2}.
$$

Similarly it implies that

$$
\frac{\partial A_3}{q_0}=\frac{\partial A_1}{q_0}e^{-2c_0\gamma}c_0+\frac{\partial A_2}{q_0}c_0+\frac{\partial A_4}{q_0}e^{-2c_0\gamma}=\frac{2e^{-\gamma c_0}+\gamma c_0\left(e^{\gamma c_0}e^{-2\gamma c_h}+e^{-\gamma c_0}\right)}{\gamma\left(e^{\gamma c_0}e^{-\gamma c_h}+e^{-\gamma c_0}e^{\gamma c_h}\right)^2}
$$

Consequently, using (A.31), we have (where we have used (A.32))

$$
\begin{aligned}
\frac{\partial \widetilde{v}\left(c_h\right)}{\partial q_0} &= e^{c_h\gamma}\frac{\partial A_3}{q_0}-e^{-c_h\gamma}\frac{\partial A_4}{q_0}-c_h\frac{\partial \widetilde{q}\left(c_h\right)}{\partial q_0}\\
&= 2\frac{e^{\gamma(c_h-c_0)}-e^{-\gamma(c_h-c_0)}-\gamma\left(c_h-c_0\right)\left(e^{-\gamma(c_h-c_0)}+e^{\gamma(c_h-c_0)}\right)}{\gamma\left(e^{\gamma c_0}e^{-\gamma c_h}+e^{-\gamma c_0}e^{\gamma c_h}\right)^2}<0.
\end{aligned}
$$

The last inequality comes from the fact that the function $e^x-e^{-x}-x\left(e^{-x}+e^x\right)$ is negative and monotonically decreasing for all $x>0$. The second statement comes directly from the expression for $\frac{\partial \widetilde{q}(c_h)}{\partial q_0}$. ■

Lemma A.10 *If* $\frac{v_0}{q_0}<h$, *then* $\tilde{v}'_{c_h}-h\tilde{q}'_{c_h}>0$.

Proof. We parameterize c_h by y. The idea is that if the function $\tilde{v}\left(y;q_0,v_0,y\right)-h\tilde{q}\left(y;q_0,v_0,y\right)$ is negative at $y=c_0$ and positive as $y\to\infty$, then there is a $y=c_h$ so that this function is zero (satisfying the definition of c_h) and where the slope of this function is positive, which is the claim of our lemma.

The function $\tilde{v}(y; q_0, v_0, y) - h\tilde{q}(y; q_0, v_0, y)$ can be solved by imposing the boundary conditions

$$\tilde{v}'(y) = \tilde{q}'(y) = 0, \tilde{q}(c_0) = q_0, \tilde{v}(c_0) = v_0. \tag{A.38}$$

for all $y \geq c_0$. Thus, by setting $y = c_0$, we must have

$$\tilde{v}(c_0; q_0, v_0, c_0) - h\tilde{q}(c_0; q_0, v_0, c_0) = v_0 - hq_0 < 0,$$

by the condition of the proposition.

Now we show that $\tilde{v}(y; q_0, v_0, y) - h\tilde{q}(y; q_0, v_0, y) \to \infty$ as $y \to \infty$. We first show calculate $\lim_{y\to\infty} \tilde{q}(y; q_0, v_0, y)$ in (A.29). For this, we solve for $e^{-y\gamma}A_1$ and $e^{y\gamma}A_2$ from (A.29)-(A.30) and (A.38):

$$e^{-y\gamma}A_1 = \frac{q_0 - \frac{u}{2} + e^{(c_0-y)\gamma}\frac{RM'(y)}{2} - \frac{R\gamma}{2}M(c_0)}{e^{(y-c_0)\gamma} + e^{\gamma(c_0-y)}}, e^{y\gamma}A_2 = \frac{q_0 - \frac{u}{2} - e^{(y-c_0)\gamma}\frac{R\gamma M'(y)}{2} - \frac{R\gamma}{2}M(c_0)}{e^{(y-c_0)\gamma} + e^{\gamma(c_0-y)}}.$$

where $M(y) \equiv -e^{\gamma y}\,\mathrm{Ei}[-\gamma y] + e^{-\gamma y}\,\mathrm{Ei}[y\gamma]$. Using $\lim_{y\to\infty} M'(y) = 0$, it is easy to show that $\lim_{y\to\infty} e^{y\gamma}A_2 = \lim_{y\to\infty} e^{-y\gamma}A_1 = 0$, which implies that $\lim_{y\to\infty} \tilde{q}(y; q_0, v_0, y) = \frac{u}{2}$ in (A.29). A similar argument implies that $\lim_{c\to\infty} \tilde{v}(c; q_0, v_0, c) = \infty$. Thus, $\tilde{v}(c; q_0, v_0, c) - h\tilde{q}(c; q_0, v_0, c) = \infty$. This completes our proof. ■

Putting together the above three lemmas, we have

$$\frac{\partial c_h}{\partial q_0} = -\frac{\tilde{v}'_{q_0} - h\tilde{q}'_{q_0}}{\tilde{v}'_{c_h} - h\tilde{q}'_{c_h}} > 0, \text{ and } \frac{\partial c_h}{\partial v_0} = -\frac{\tilde{v}'_{v_0} - h\tilde{q}'_{v_0}}{\tilde{v}'_{c_h} - h\tilde{q}'_{c_h}} < 0$$

This implies that $c_h^\pi < c_h^*$ whenever $q_\pi(c_0) \leq q(c_0)$ and $v_\pi(c_0) \geq v(c_0)$ (which automatically implies that $p_\pi(c_0) \geq p(c_0)$).

The claim in the proposition is stronger which says that $c_h^\pi < c_h^*$ holds even if $q_\pi(c_0) \leq q(c_0)$ and $v_\pi(c_0) < v(c_0)$, but $p_\pi(c_0) = \frac{v_\pi(c_0)}{q_\pi(c_0)} \geq p(c_0) = \frac{v(c_0)}{q(c_0)}$. Because $\frac{\partial c_h}{\partial v_0} = -\frac{\tilde{v}'_{v_0} - h\tilde{q}'_{v_0}}{\tilde{v}'_{c_h} - h\tilde{q}'_{c_h}} < 0$, it suffices to show that this result holds for the worst v_0 drop to maintain p_0, i.e., v_0 and q_0 decrease proportionally so v_0/q_0 remains at constant.

To this end, we consider increasing q_0 to $\bar{q}_0 = q_0 + \varepsilon$ where ε is very small. To make sure that $\frac{\bar{v}_0}{\bar{q}_0} = \frac{v_0}{q_0}$, we need that $\bar{v}_0 = v_0 + a\varepsilon$ where $a = \frac{v_0}{q_0}$. Let us refer to all the objects after the change with the bar. Our goal is to show that $\bar{v}(c_h)/\bar{q}(c_h)$ would decrease; then $\tilde{v}'_{c_h} - h\tilde{q}'_{c_h} > 0$ implies that c_h^π increases. Using the first two Lemmas above, we have (denoting $x \equiv (c_h - c_0)\gamma$)

$$\begin{aligned}
\bar{q}(c_h) &= \tilde{q}(c_h) + \varepsilon\frac{2}{e^x + e^{-x}} \\
\bar{v}(c_h) &= \tilde{v}(c_h) + \varepsilon 2\frac{e^x - e^{-x} - x(e^{-x} + e^x)}{\gamma(e^x + e^{-x})^2} + \frac{v_0}{q_0}\frac{2\varepsilon}{e^x + e^{-x}}
\end{aligned}$$

Hence for sufficiently small ε we have (up to the first order)

$$\begin{aligned}\frac{\bar{v}(c_h)}{\bar{q}(c_h)} &= \frac{\tilde{v}(c_h)}{\tilde{q}(c_h)} + \frac{2\varepsilon}{\tilde{q}(c_h)}\left(\frac{e^x - e^{-x} - x\left(e^{-x}+e^x\right)}{\gamma\left(e^x+e^{-x}\right)^2} + \frac{v_0}{q_0}\frac{1}{e^x+e^{-x}}\right) - \frac{\tilde{v}(c_h)}{\tilde{q}^2(c_h)}\frac{2\varepsilon}{e^x+e^{-x}} \\ &< \frac{\tilde{v}(c_h)}{\tilde{q}(c_h)} + \frac{\varepsilon}{\tilde{q}(c_h)}\frac{2\left(\frac{v_0}{q_0} - \frac{\tilde{v}(c_h)}{\tilde{q}(c_h)}\right)}{e^x+e^{-x}}. \end{aligned} \tag{A.39}$$

Here, the third inequality in (A.39) is because the term $e^x - e^{-x} - x\left(e^{-x}+e^x\right) < 0$ for all $x > 0$. Note that the last term in (A.39) is strictly negative because $\frac{v_0}{q_0} < \frac{\tilde{v}(c_h)}{\tilde{q}(c_h)} = h$; hence the first order impact of ε is always negative. Because the above argument holds for any v_0 and q_0, tracing out the first-order effect implies that any intervention which lowers cash value but keeps capital price unchanged will lower $\frac{\bar{v}(c_h)}{\bar{q}(c_h)}$. This concludes our proof.

A.10 ODEs for one sided intervention

Following our derivation of the market equilibrium, value functions in the intervention equilibrium are defined by the ODEs

$$0 = \frac{\sigma^2}{2}q_\pi'' - \mathbf{1}_{c<c_0}\pi + \xi\left(\frac{u+R/c}{2} - q_\pi\right) \tag{A.40}$$

$$0 = \frac{\sigma^2}{2}v_\pi'' + q_\pi'\sigma^2 + \mathbf{1}_{c<c_0}\pi c + \xi\left(\frac{uc+R}{2} - v_\pi\right) \tag{A.41}$$

where $\mathbf{1}$ is the indicator function, subject to the boundary conditions

$$\frac{v_\pi\left(c_h^\pi\right)}{q_\pi\left(c_h^\pi\right)} = h, \quad \frac{v_\pi\left(c_l^\pi\right)}{q_\pi\left(c_l^\pi\right)} = l, \tag{A.42}$$

$$v_\pi'\left(c_h^\pi\right) = q_\pi'\left(c_h^\pi\right) = q_\pi'\left(c_l^\pi\right) = v_\pi'\left(c_l^\pi\right) = 0. \tag{A.43}$$

Besides, each function has continuous first order derivative at c_0. It is simple to check that the following general solution satisfies the system

$$\begin{aligned} q_\pi(c) &= -\frac{\pi\mathbf{1}_{c<c_0}}{\xi} + \frac{u}{2} + e^{-c\gamma}M_{1,5} + e^{c\gamma}M_{2,4} + \frac{R\gamma}{2}\frac{e^{-c\gamma}\,\mathrm{Ei}(c\gamma) - e^{c\gamma}\,\mathrm{Ei}(-\gamma c)}{2}, \\ v_\pi(c) &= \frac{\pi c\mathbf{1}_{c<c_0}}{\xi} + R + \frac{uc}{2} + e^{c\gamma}\left(M_{3,7} - cM_{1,5}\right) - e^{-c\gamma}\left(M_{4,8} + cM_{2,6}\right) + \frac{cR\gamma}{2}\frac{\left(e^{\gamma c}\,\mathrm{Ei}(-\gamma c) - e^{-c\gamma}\,\mathrm{Ei}(\gamma c)\right)}{2}, \end{aligned}$$

where $M_{i,j} \equiv \mathbf{1}_{c>c_0}M_i + \mathbf{1}_{c<c_0}M_j$. The constants $M_1, ..., M_8$ are determined by (A.42)-(A.43) and the smooth-pasting conditions at c_0 for $v(c)$ and $q(c)$.

A.11 ODEs for the alternative specification

In the alternative setting, the value of capital $v(c)$ and the value of cash $q(c)$ satisfy the following ODEs:

$$
\begin{aligned}
0 &= -v'(c)(c+p(c))\phi + q'(c)\sigma^2 + \frac{\sigma^2}{2}v''(c) + \phi(pu - v(c)) + \xi(R - v(c)) + \\
&\quad \eta\left[\left(\mathbf{1}_{c>c_h}v(c_h) + \mathbf{1}_{c<c_l}v(c_l) + \mathbf{1}_{c_h>c>c_l}v(c)\right) - v(c)\right], \qquad \text{(A.44)}\\
0 &= -q'(c)(c+p(c))\phi + \frac{\sigma^2}{2}q''(c) + \phi(u - q(c)) + \xi(1 - q(c)) \\
&\quad +\eta\left[\left(\mathbf{1}_{c>c_h}q(c_h) + \mathbf{1}_{c<c_l}q(c_l) + \mathbf{1}_{c_h>c>c_l}q(c)\right) - q(c)\right]. \qquad \text{(A.45)}
\end{aligned}
$$

where $\mathbf{1}$ is the indicator function and $p(c) = v(c)/q(c)$. There are two main changes compared to our baseline model. First, the first term in each equation is due to the additional drift term in the dynamics of the aggregate cash-to-capital ratio c (when firms cannot invest/disinvest):

$$dc_t = -\phi(c_t + p_t)\,dt + \sigma dZ_t.$$

It is because there are ϕdt fraction of firms exiting the market with cash $c_t + p_t$, i.e., they sell their capital holdings at a price of p_t and leave the market with these proceeds plus their cash holdings c_t. Second, the last bracketed term in each equation captures the event in which firms can invest or disinvest (which occurs with intensity η). For instance, if $p_t > h$ (i.e. $c_t > c_h$) firms build new capital until the point where the aggregate liquidity falls to c_h where $p(c_h) = h$, and accordingly their capital value $v(c)$ and the cash value $q(c)$ jump to $v(c_h)$ and $q(c_h)$.

We have the following six boundary conditions:

$$
\begin{aligned}
v'(0) &= 0 \text{ and } q(0) = \lambda, && \text{(A.46)}\\
p(c_l) &= \frac{v(c_l)}{q(c_l)} = l \text{ and } p(c_h) = \frac{v(c_h)}{q(c_h)} = h, && \text{(A.47)}\\
\lim_{c\to\infty} v'(c) &= 0, \text{ and } \lim_{c\to\infty} q'(c) = 0. && \text{(A.48)}
\end{aligned}
$$

Conditions (A.46) holds because $c = 0$ is a reflective barrier where outside cash is injected whenever the value of cash is larger than λ. Conditions (A.47) are determined by the adjustment of capital explained above. The last two conditions in (A.48) are standard. These six boundary conditions allow us to solve for $v(c)$ and $q(c)$ and two endogenous thresholds c_h and c_l numerically.

Now we consider the equilibrium with intervention. The specific intervention rule, i.e., taxation policies that finance building/dismantling capital, affects the market price p. In contrast to the baseline model where the ex ante price plays no role, in this alternative setting it affects the amount of cash that goes to the firms hit by new investment opportunities, which in turn affects welfare. We hence consider the following taxation policy which balances the government budget any time. Suppose that the government builds capital through taxing cash only, and distribute the newly built capital back to firms immediately. For $c > c_h^g$, to achieve the desired upper cash-to-capital ratio c_h^g, the government needs to build $K\frac{c-c_h^g}{t+c_h^g}$ units of capital, and thus requires $K\frac{c-c_h^g}{h+c_h^g}h$ units of cash. Given existing cash $C = Kc$, the government needs to tax $\frac{c-c_h^g}{h+c_h^g}\frac{h}{c}$ per unit of existing cash. In the meantime, each firm receives newly built capital $K\frac{c-c_h^g}{h+c_h^g}$ which can be sold in the market at a price of $p\left(c_h^g\right) = v\left(c_h^g\right)/q\left(c_h^g\right)$. Combining both pieces, the net taxation (outflows) per unit of cash is

$$\frac{c-c_h^g}{h+c_h^g}\frac{h}{c} - \frac{c-c_h^g}{h+c_h^g}\frac{v\left(c_h\right)}{q\left(c_h\right)c} = \frac{c-c_h^g}{ch+cc_h^g}\left(h - \frac{v\left(c_h\right)}{q\left(c_h\right)}\right) = \frac{c-c_h^g}{ch+cc_h^g}\left(h - p\left(c_h\right)\right). \tag{A.49}$$

Similarly, when $c < c_l^g$ is low, the government dismantles $K\frac{c_l^g-c}{l+c_l^g}$ units of capital for $K\frac{c_l^g-c}{l+c_l^g}l$ units of cash. So per unit of capital, the government taxes $\frac{c_l^g-c}{l+c_l^g}$ units of capital, and redistributes back $\frac{c_l^g-c}{l+c_l^g}l$ amount of cash. Converting to utilities, effectively each capital is taxed at $\frac{c_l^g-c}{l+c_l^g}\left(v\left(c_l^g\right) - lq\left(c_l^g\right)\right)$.

The next proposition gives the system of ODEs determining the social welfare for an arbitrarily given c_h and c_l in our alternative setting.

Proposition 10 *The total welfare in the alternative specification for arbitrarily given thresholds $c_l < c_h$ is $Kj_P\left(c; c_l, c_h\right) = K\left(v_P\left(c\right) + q_P\left(c\right)c\right)$ where v_P and q_P is given by the system*

$$\begin{aligned} 0 &= -q_P'\left(c+p_P\right)\phi + \frac{\sigma^2}{2}q_P'' + \phi\left(u - q_P\right) + \xi\left(1 - q_P\right) \\ &\quad + \eta\left(1_{c>c_h}\left(-\frac{c-c_h}{ch+cc_h}\left(hq_P\left(c_h\right) - v_P\left(c_h\right)\right) + q_P\left(c_h\right) - q_P\left(c\right)\right) + 1_{c<c_l}\left(q_P\left(c_l\right) - q_P\left(c\right)\right)\right), \\ 0 &= q_P'\left(c_t\right)\sigma^2 - v_P'\left(c+p_P\right)\phi + \frac{\sigma^2}{2}v'' + \phi\left(p_P u - v_P\right) + \xi\left(R - v\right) \\ &\quad + \eta\left(1_{c>c_h}\left(v_P\left(c_h\right) - v_P\left(c\right)\right) + 1_{c<c_l}\left(-\frac{c_l-c}{l+c_l}\left(v_P\left(c_l\right) - lq_P\left(c_l\right)\right) + v_P\left(c_l\right) - v_P\left(c\right)\right)\right). \end{aligned}$$

with the boundary conditions (A.46) and (A.48).

Proof. The expression of total welfare is obvious. With given intervention policy, for q equation, for $c > c_h^g$ we have (we need to multiply the taxation (A.49) by $q\left(c_h^g\right)$ to get back to utilities)

$$0 = -q_P'\left(c+p_P\right)\phi + \frac{\sigma^2}{2}q_P'' + \phi\left(u - q_P\right) + \xi\left(1 - q_P\right) + \eta\left(-\frac{c-c_h^g}{ch+cc_h^g}\left(hq_P\left(c_h^g\right) - v_P\left(c_h^g\right)\right) + q_P\left(c_h^g\right) - q_P\left(c\right)\right)$$

When $c < c_l^g$ the cash is free of taxation, and thus the adjustment in the η event is simply $q_P\left(c_l^g\right) - q_P\left(c\right)$. For capital, taxation occurs when $c < c_l^g$. In this situation, since each capital is taxed with a utility equivalent of $\frac{c_l^g-c}{l+c_l^g}\left(v_P\left(c_l^g\right) - lq_P\left(c_l^g\right)\right)$, we have

$$\begin{aligned} 0 = q_P'\left(c\right)\sigma^2 - v_P'\left(c+p_P\right)\phi + \frac{\sigma^2}{2}v_P'' + \phi\left(p_P u - v\right) + \xi\left(R - v_P\right) + \\ + \eta\left(-\frac{c_l^g-c}{l+c_l^g}\left(v_P\left(c_l^g\right) - lq_P\left(c_l^g\right)\right) + v_P\left(c_l^g\right) - v_P\left(c\right)\right). \end{aligned}$$

■

B Appendix: An alternative equilibrium

In the main text, we showed that an equilibrium exist when $h - l$ is sufficiently small. While our condition is only sufficient, and not necessary, it is possible that the type of equilibrium presented in the main text does not exist. In this Appendix, we provide some insights on the type of equilibrium that arises instead. We argue that the main properties of this alternative equilibrium are very similar to the one presented.

While the equation system (18)-(19), (13)-(15) always have a solution, for some parameters this solution implies that for a c sufficiently close to c_l^*, the price is below the threshold l. This obviously cannot be an equilibrium—because firms would dismantle the first instant when the price drops below the liquidation value l. For that set of parameters we can construct the equilibrium as follows. There is a $c_x \in (c_l^*, c_h^*)$ that for every $c \in [c_l^*, c_x]$

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

and an endogenous fraction of capital are dismantled at every instant. That is, in this range the price is constant in c and firms dismantle an increasing fraction of their capital as c drops further from c_x. The following Proposition describes this equilibrium.

Proposition B.1 *Suppose that there is a $c_h^* < R$, $c_x \in (l, c_h^*)$, q_0, A_1, A_2, A_3, A_4 solving (18)-(19), (13)*

$$\begin{aligned}
\frac{\xi}{2\sigma^2}\left(u + \frac{R}{c_x}\right)(l - c_x) &= q'(c_x) \\
l\frac{\xi}{2\sigma^2}\left(u + \frac{R}{c_x}\right)(l - c_x) &= v'(c_x) \\
\frac{v(c_x)}{q(c_x)} = l, \frac{v(c_h^*)}{q(c_h^*)} &= h, v'(c_h^*) = q'(c_h^*) = 0.
\end{aligned}$$

Then there is a market equilibrium with partial liquidation where

1. *firms do not consume before the final date,*

2. *each firm in each state $c \in [l, c_h^*]$ is indifferent in the composition of her portfolio*

3. *firms do not build or dismantle capital when $c \in (c_x, c_h^*)$ and, in aggregate, firms spend every positive cash shock to build capital iff $c = c_h^*$ and cover the negative cash shocks by liquidating a fraction of capital iff $c \in [l, c_x]$. When $c = l$, firms finance every negative cash shock by liquidating capital.*

4. *the value of holding a unit of cash and the value of holding a unit of capital are described by* $q(c)$ *and* $v(c)$ *and the ex ante price is* $p = \frac{v(c)}{q(c)}$ *when* $c \in [c_x, c_h^*]$*, and by*

$$\begin{aligned} q(c) &= q_0 + \frac{\xi}{2\sigma^2}\left[(ul - R)(c - l) - \frac{u}{2}\left(c^2 - l^2\right) + lR(\ln c - \ln l)\right] \\ v(c) &= lq_m(c) \end{aligned}$$

and the ex ante price is $p = l$ *when* $c \in [l, c_x]$.

5. *Ex post, each firms hit by the shock sells all her capital to the firms who are not hit by the shock for the price* $\hat{p}_\tau = c$.

Proof. Under the conditions of the Proposition, firms start to disinvest whenever $p(c) = l$. Given the liquidation rate $y(c)\,dt = -dK/K$, then its impact on the aggregate cash-to-capital ratio c is

$$x(c)\,dt \equiv \frac{dC}{K} - \frac{C}{K}\frac{dK}{K} = -\frac{ldK}{K} - \frac{C}{K}\frac{dK}{K} = (l + c)\,y(c)\,dt,$$

so the c evolves as $dc = x(c)\,dt + \sigma dZ_t$. We must have $v(c) = lq(c)$ as firms are always indifferent in liquidating the capital, and v and q satisfies:

$$\begin{aligned} 0 &= x(c)\,q'(c) + \frac{\sigma^2}{2}q''(c) + \frac{\xi}{2}\left(u + \frac{R}{c}\right) - \xi q(c) \\ 0 &= x(c)\,v'(c) + q'(c)\,\sigma^2 + \frac{\sigma^2}{2}v''(c) + \frac{\xi}{2}(uc + R) - \xi v(c) \end{aligned}$$

Using $v(c) = lq(c)$, we obtain

$$\begin{aligned} 0 &= x(c)\,lq'(c) + \frac{\sigma^2}{2}lq''(c) + \frac{\xi l}{2}\left(u + \frac{R}{c}\right) - \xi lq(c) \\ 0 &= x(c)\,lq'(c) + q'(c)\,\sigma^2 + \frac{\sigma^2}{2}lq''(c) + \frac{\xi}{2}(uc + R) - \xi lq(c) \end{aligned}$$

Eliminating identical terms, we get

$$q'(c) = \frac{\xi}{2\sigma^2}\left(u + \frac{R}{c}\right)(l - c) = 0.$$

As $q'(c_l) = 0$ has to hold, $c_l = l$. The closed-form solution is

$$q(c) = q_0 + \frac{\xi}{2\sigma^2}\left[(ul - R)(c - l) - \frac{u}{2}\left(c^2 - l^2\right) + lR(\ln c - \ln l)\right]$$

And, we have $q''(c) = -\frac{\xi}{2\sigma^2}\left(u + \frac{lR}{c^2}\right) < 0$. We know that for $c \in [l, c_x]$ we have $v(c) = lq(c)$ which allows us to back out the endogenous drift of c:

$$x(c) = \frac{-\frac{\sigma^2}{2}q''(c) - \frac{\xi}{2}\left(u + \frac{R}{c}\right) + \xi q(c)}{q'(c)},$$

and thus the endogenous liquidation rate $y(c) = \frac{x(c)}{l+c}$. For $c > c_x$ we have the ODE as usual. We then search for the c_x, c_h pair that satisfies the conditions of the proposition. ■

Plotting v, q and p give very similar graphs to Figure 2 with the main difference that at the range $c \in [l, c_x]$ the price is flat at the level l. In the same range $q(c)$ is decreasing implying that $v(c) = lq(c)$ is also decreasing.

C Appendix: Extensions of the two-period example

In this appendix, we extend our two period example into several directions to shed more light on some of the complexities we touched upon in the main text. In particular, we focus on three main topics.

First, we highlight the exact nature of frictions we implicitly or explicitly assume in the main text. As a main issue, we discuss under what conditions the efficiency can be restored in our two period example. In particular, we consider the following three routes.

1. If the realization of the skill shock is verifiable, so there are contracts contingent on these shocks, then the planner's solution is attainable. However, we argue that if firms can misreport their shocks, then they indeed will misreport, and these contracts will not help.

2. If the output of harvesting the capital and investing in new technology is fully pledgeable, then the planner's solution is attainable. However, if firms can steal a sufficiently large fraction of these outputs, then the inefficiency we present in the main text reemerges.

3. As a third alternative, in the spirit of Diamond and Dybvig (1983), we argue that there might be a banking solution to the inefficiency. That is, if all firms are forced to keep their assets at a bank which can produce, invest the cash into the new technology, and distribute the proceeds equally among the participants, then the first best is attainable. However, we also argue that each firm will have a strong motivation to not to participate in the bank, but invest on her own and trade on the ex post market at the equilibrium price instead. In this sense, the banking solution is fragile. This argument is closely related to the Jacklin (1987) critique to the Diamond and Dybvig (1983) model.

As a second topic, we generalize the preferences and the technology of our two-period example to show that neither the assumed risk-neutrality nor the kinky technological frontier is critical for our intuition.

Finally, if there were only two firms in our economy with perfectly negatively correlated skill-shocks, so they would bargain with each other ex post instead of trading on a Walrasian market, then our inefficiency would not arise. This is an argument which connects our result to the finding of Hoberg and Phillips (2010), who find that competitive industries are much more subject to inefficient investment waves than non-competitive industries.

C.1 Frictions

In this part, we use a slightly more general version of the two-period model than the one in the main text, which is closer to our full model in the main text. In particular, just as in the full model, we assume that harvesting the capital gives R, investing in the new technology gives u, building a capital costs h cash, and an additional cash needs $1/l$ capital to be destroyed. We assume that the initial endowment is K_i^0 of capital and C_i^0 of cash. (In the example of the main-text, $R = u = 3$, $K_i^0 = 1$ and $C_i^0 = c$, $h = 2$, $l = \frac{1}{2}$.) In the spirit of the example in the main-text we restrict $l < \frac{R}{u} < h$ implying that the social planner always prefers firms not to change their allocation. That is, $K_i = K_i^0, C_i = C_i^0$ is the planner's solution.

C.1.1 Contractibility

Fully verifiable skill-shock Consider the following problem

$$\begin{aligned} &\max_{x_{KK},x_{KC},x_{CK},x_{CC},K_i,C_i} \frac{R}{2}\left(K_i + x_{KK} + \frac{C_i + x_{KC}}{p_1}\right) + \frac{u}{2}\left(p_1\left(K_i + x_{CK}\right) + C_i + x_{CC}\right) \\ &s.t. \quad \pi_{KK}x_{KK} + \pi_{KC}x_{KC} + \pi_{CK}x_{CK} + \pi_{CC}x_{CC} + p_0 K_i + C_i = p_0 K_i^0 + C_i^0 \\ &\qquad \begin{cases} hK_i + C_i = hK_i^0 + C_i^0 & \text{if } K_i > K_i^0, \\ lK_i + C_i = lK_i^0 + C_i^0 & \text{if } K_i < K_i^0. \end{cases} \end{aligned} \tag{C.1}$$

Here, $x_{s_1 s_2}$ is the amount purchased or sold from the Arrow-Debrue security paying 1 unit of s_2 good when the firm has the skill to use the s_1 good, and $\pi_{s_1 s_2}$ is the price of the Arrow-Debrue security, where $s_1, s_2 = K, C$. The prices p_0, p_1 are the prices for a unit of capital in terms of cash in period 0 and 1 respectively.

The first order conditions are

$$\begin{aligned} \frac{R}{2} + \frac{u}{2}p_1 - \mu p_0 &= 0 \\ \frac{R}{2p_1} + \frac{u}{2} - \mu &= 0 \\ \frac{R}{2p_1} &= \pi_{KC}\mu \\ \frac{R}{2} &= \pi_{KK}\mu \\ \frac{u}{2}p_1 &= \mu\pi_{CK} \\ \frac{u}{2} &= \mu\pi_{CC} \end{aligned}$$

subject to the technology constraint, implying that

$$\begin{aligned}
\pi_{KC} + \pi_{CC} &= 1 \\
\pi_{KK} + \pi_{CK} &= p_0 \\
\frac{\pi_{KK}}{\pi_{CC}} &= \frac{R}{u} \\
\frac{\pi_{CK}}{\pi_{CC}} &= p_1, \frac{\pi_{KK}}{\pi_{KC}} = p_1.
\end{aligned}$$

Given our assumption on the ex post price formation and that exactly the same mass of firms receives both shocks, the market clearing conditions are

$$\begin{aligned}
x_{CC} &= -x_{KC}, x_{KK} = -x_{CK}, \\
\text{and } p_1 &= \frac{C_i + x_{KC}}{K_i + x_{CK}}
\end{aligned}$$

and $p_0 \leq h$ implies that $K_i \leq K_i^0$ and $\frac{1}{p_0} \leq l$ implies that $C_i \leq C_i^0$. Simple substitution shows that the following Lemma holds, and the planner's solution is achievable.

Lemma B.1 *Any x_{KC}, x_{CK} satisfying $\frac{R}{u} = \frac{C_i^0 + x_{KC}}{K_i + x_{CK}}$ constitutes an equilibrium with*

$$\begin{aligned}
x_{CC} &= -x_{KC}, x_{KK} = -x_{CK}, \pi_{KK} = \pi_{CK} = \frac{R}{2u}, \pi_{CC} = \pi_{KC} = \frac{1}{2} \\
p_0 &= p_1 = \frac{R}{u}, C_i = C_i^0, K_i = K_i^0.
\end{aligned}$$

Misreporting In this part, we show that given the planer's solution achieved by the Arrow-Debreu securities, some firms will always want to misreport their type. An firm who can produce will misreport her type if

$$\frac{x_{CC}}{p_1} R + x_{CK} R > \frac{x_{KC}}{p_1} R + x_{KK} R,$$

implying $2\left(x_{CC} u + x_{CK} R\right) > 0$ by Lemma B.1. On the other hand, an firm who can invest in the new opportunity will misreport her type if

$$x_{KC} u + x_{KK} p_1 u > x_{CC} u + x_{CK} p_1 u$$

implying that $2\left(x_{CC} u + x_{CK} R\right) < 0$. Hence, the only possibility that no firms misreport is $\left(x_{CC} u + x_{CK} R\right) = 0$, or

$$\frac{x_{CK} R}{u} = x_{KC}.$$

In the planner's solution, $\frac{R}{u} = \frac{C_i^0 + x_{KC}}{K_i + x_{CK}}$, which implies that

$$\frac{R}{u} = \frac{C_i^0 + x_{KC}}{K_i + x_{CK}} = \frac{C_i^0 + \frac{x_{CK} R}{u}}{K_i + x_{CK}} \Rightarrow \frac{R}{u} = \frac{C_i^0}{K_i}.$$

Hence, firms do not misreport only when the original endowment ratio is $\frac{R}{u}$. But this implies that the market price R/u have to coincide with the marginal rate of substitution of social planner.

C.1.2 Pledgeability

If the final output for both technologies are fully pledgeable, then a firm with the new opportunity would offer a labour contract to other firms who can produce. According to this contract, the firm who can produce are hired by the other firm, and does so for zero wage. (Given that each firm can harvest any number of capital, given that she has the skill to harvest the first, in equilibrium zero wage emerges.) So each firm can get R for each capital, and, by the symmetric argument, each firm can get u for each cash, regardless of their skill-shock. In this case, each firm maximizes $\frac{1}{2}K^i R + \frac{1}{2}C^i u$, and the decentralized solution coincides with the planner's solution.

Suppose now that the worker can abscond with λ fraction of the goods she produces. Then the optimal contract will be to give her λR of the output of the capital and λu of the output of the cash. It is easy to see that as long as firms

$$(1-\lambda)\,h < \frac{R}{u} < \frac{l}{1-\lambda},$$

firms will prefer to use the market as opposed to contracting (as p_1 is always in between l and h). Thus, as long as the worker can steal a sufficiently large part of the output, our inefficiency reemerges.

C.1.3 Banking solution

Suppose that there is a bank which holds the assets of all firms. In period one, firms report their type and receive y_K capital, if they report that they can produce and y_C cash, if they report that they can invest the cash into the new technology. Then the banking solution is

$$\begin{aligned}&\max_{y_K, y_C} \frac{1}{2}y_K R + \frac{1}{2}y_C u\\&s.t.\ \text{budget constraint of technology}\end{aligned}$$

which is identical to the social planner.

Do we have a Diamond and Dybvig (1983) type run where firms want to misreport?

1. Without market then clearly no one wants to misreport. It is because misreporting gives the firm the type of good which they cannot operate.

2. As it was pointed out by Jacklin (1987), when there is a market, firms might misreport and sell their obtained good on the market.

Just for this exercise, suppose that K firm can still consume the cash with a marginal utility of 1. This assumption would not change the previous analysis. Then the K firm determines the price of one capital in terms of cash. It is because she can either get $y_K R$ by claiming K and get y_C by claiming C, so the price of capital would be

$$p = y_C / \left(y_K R\right).$$

Suppose that this price is high. Similar to Jacklin's idea, it might be possible that an individual firm may want to convert cash to capital, and sell the capital in the market if it turns out to be u firm.

C.2 Generalizing technology

In this part, we argue that neither our extreme technology shocks nor the kinky ex ante technology frontier is critical for our results.

C.2.1 Ex post technology shocks

In contrast to the main text model where the ex post technology is linear, we now will specify a general ex post technology shock and show that the market failure still holds. To illustrate this point, we first assume away the linear investment technology at date 0, and solve for the equilibrium market price at date 0. If the equilibrium market price differs from social planner's marginal rate of substitution, then individual firms, once equipped with investment technology, will invest inefficiently.

The ex post idiosyncratic shock of date 1 is as follows. In each idiosyncratic state, s, agents can produce by a general CES technology, but in one of the states the marginal product of capital is larger $(s = K)$, while in the other one the marginal product of cash is larger $(s = C)$. That is, if we denote (K_i^s, C_i^s) their final position of capital and cash in state s, agent i knows that at date 1 she can produce Q_i consumption good with the CES technology

$$Q_i = \left\{ \begin{array}{lll} \left(\kappa \left(K_i^K\right)^r + (1-\kappa)\left(C_i^K\right)^r\right)^{\frac{1}{r}} & w.p. & \frac{1}{2} \\ \left((1-\kappa)\left(K_i^C\right)^{\prime r} + \kappa\left(C_i^C\right)^{\prime r}\right)^{\frac{1}{r}} & w.p. & \frac{1}{2} \end{array} \right\}$$

where $\kappa > \frac{1}{2}$ and $0 < r \leq 1$ are constants. Thus, if at the end of period 0, each of them is endowed by K units of capital and C units of cash, then, after they learn their skill-shocks, they adjust their position by solving the following problem.

$$\begin{aligned} \max_{K_i^K, C_i^K, K_i^C, C_i^C} \; & \frac{1}{2}\left(\kappa\left(K_i^K\right)^r + (1-\kappa)\left(C_i^K\right)^r\right)^{\frac{1}{r}} + \frac{1}{2}\left((1-\kappa)\left(K_i^C\right)^r + \kappa\left(C_i^C\right)^r\right)^{\frac{1}{r}} \\ s.t. \quad & pK_i^K + C_i^K = pK + C \qquad\qquad \text{Lagrange multiplier } \lambda \\ & pK_i^C + C_i^C = pK + C \qquad\qquad \text{Lagrange multiplier } \lambda' \end{aligned}$$

The FOCs are

$$\begin{aligned}
\frac{1}{2}\kappa\left(\kappa+(1-\kappa)\left(\frac{C_i^K}{K_i^K}\right)^r\right)^{\frac{1-r}{r}} &= p\lambda \\
\frac{1}{2}(1-\kappa)\left(\kappa\left(\frac{K_i^K}{C_i^K}\right)^r+(1-\kappa)\right)^{\frac{1-r}{r}} &= \lambda \\
\frac{1}{2}(1-\kappa)\left((1-\kappa)+\kappa\left(\frac{C_i^C}{K_i^C}\right)^r\right)^{\frac{1-r}{r}} &= p\lambda' \\
\frac{1}{2}\kappa\left((1-\kappa)\left(\frac{K_i^C}{C_i^C}\right)^r+\kappa\right)^{\frac{1-r}{r}} &= \lambda'
\end{aligned}$$

and the market clearing condition is

$$C_i^C + C_i^K = 2C.$$

Together with the constraints, this is a system of 7 equations and 7 unknowns (4 decision variables, 2 Lagrange multipliers and the price).

Instead, the social planner solves

$$\begin{aligned}
&\max_{K^K,C^K,K^C,C^C} \frac{1}{2}\left(\kappa\left(K^K\right)^r+(1-\kappa)\left(C^K\right)^r\right)^{\frac{1}{r}}+\frac{1}{2}\left((1-\kappa)\left(K^C\right)^r+\kappa\left(C^C\right)^r\right)^{\frac{1}{r}} \\
&s.t. \quad C^C+C^K=2C \qquad \text{Lagrange multiplier } \lambda_C \\
&\qquad\;\; K^C+K^K=2K \qquad \text{Lagrange multiplier } \lambda_K
\end{aligned}$$

The FOCs are

$$\begin{aligned}
\frac{1}{2}\kappa\left(\kappa+(1-\kappa)\left(\frac{C^K}{K^K}\right)^r\right)^{\frac{1-r}{r}} &= \lambda_C \\
\frac{1}{2}(1-\kappa)\left(\kappa\left(\frac{K^K}{C^K}\right)^r+(1-\kappa)\right)^{\frac{1-r}{r}} &= \lambda_K \\
\frac{1}{2}(1-\kappa)\left((1-\kappa)+\kappa\left(\frac{C^C}{K^C}\right)^r\right)^{\frac{1-r}{r}} &= \lambda_C \\
\frac{1}{2}\kappa\left((1-\kappa)\left(\frac{K^C}{C^C}\right)^r+\kappa\right)^{\frac{1-r}{r}} &= \lambda_K
\end{aligned}$$

If the optimal choices in state C are described by $C^C(C,K)$ and $K^C(C,K)$ as functions of the endowment C and K, then, by the envelope theorem the marginal rate of substitution for the planner (with value V) is

$$\frac{V_K}{V_C}=\frac{(1-\kappa)\left(2K-K^C(C,K)\right)^{r-1}}{\kappa\left(2C-C^C(C,K)\right)^{r-1}}.$$

While the analytical analysis is cumbersome, we illustrate by a simple numerical exercise that our main result, that agents have private incentives to overinvest in the scarce asset, still holds. In particular, in Figure 7, we plot the ratio of marginal valuation of capital versus cash in the market solution, p, and the social valuation $\frac{V_K}{V_C}$ as a function of r and the initial cash to capital ratio C/K. The figure illustrates that agents privately overvalue the capital good exactly when cash is abundant, that is, when $C/K > 1$.

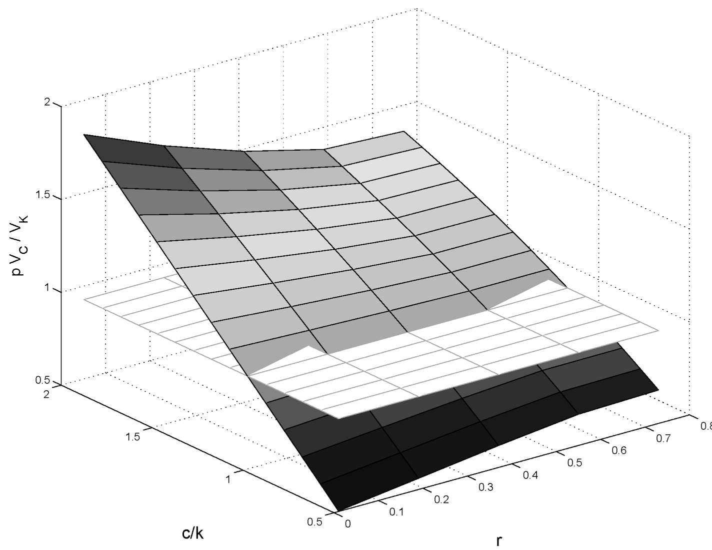

Figure 7: Ratio of price of capital versus social value of capital to cash. Agents are subject to idiosyncratic CES parametrized by $r = \frac{1}{1-s}$ where s is the elasticity of substitution. The fixed parameters are $k = 1$, $\alpha = \frac{2}{3}$.

C.2.2 General concave ex ante technology

In this exercise, we keep the ex post shocks as in the main text, but suppose that instead of (C.1), the ex ante technology is given by a smooth concave function

$$F(K_i, C_i) \equiv 0.$$

Then, in the decentralized economy, an individual firm's first order condition is given by

$$\frac{1}{2} R \left(1 - \frac{\frac{F_K}{F_C}}{p}\right) + \frac{1}{2} u \left(p - \frac{F_K}{F_C}\right) = 0$$

where F_K, F_C are the partial derivatives of the technology frontier at the optimum values. Similarly, the social planner's first order condition is

$$\frac{1}{2}R\left(1-\frac{\frac{F_K}{F_C}}{p}-C\frac{1}{p^2}\frac{dp}{dK}\right)+\frac{1}{2}u\left(p-\frac{F_K}{F_C}+K_i\frac{dp}{dK}\right)=0$$

as the social planner takes into account the effect of changing dK on price. Clearly, the difference between the two conditions is

$$K\left(u-R\frac{K}{C}\right)\frac{1}{2}\frac{dp}{dK}=0$$

with a sign which clearly depends on

$$\frac{C}{K}\lesseqgtr\frac{R}{u}.$$

That is, the difference between two FOCs depends on the relative scarcity of the two goods compared to their relative social value. This is the intuition we describe in the main text.

C.3 A non-competitive industry

Instead of competitive market with a continuum of firms, we consider the following variation of model with two firms. To eliminate aggregate uncertainty, one firm will be hit by R shock and the other will be hit by u shock, with equal probabilities. The key difference we would like to capture is that with two firms, individuals are no longer infinitesimal so that they are not price takers. We show that the non-competitive nature can alleviate the pecuniary externalities that are prevailing at the competitive market.

Without Walrasian market at the second period after the skill shock, we need to specify a bargaining protocol governing the transaction. We simply assume that it is equally likely to have anyone to make a take-it-or-leave-it offer to the other firm. More specifically, suppose that at the beginning of period 2 the endowment of firm i is (K_i, C_i), where $i = 1, 2$. If the firm i gets the chance to make the offer, she can propose the new allocation $\left(K_i', C_i'; K_j', C_j'\right)$ to firm j, which satisfies the budget constraints $K_i' + K_j' = K_i + K_j$ and $C_i' + C_j' = C_i + C_j$. For instance, if the firm i is hit by R shock, then the equilibrium offers that are proposed by firm i and will be accepted (by firm j) can be written as

$$\left(K_i + K_j, (1-\alpha)\, C_i; 0, \alpha C_i + C_j\right)$$

where $\alpha \in [0, 1]$. Obviously, $\alpha = 1$ reaches the socially optimal allocation; we allow for any α for the generality of our result. Thus the firm i's utility is $(K_i + K_j)\, R$ while the firm j's utility is $(\alpha C_i + C_j)\, u$.

At period 1, the firm i is solving the following problem:

$$\max_{K_i,C_i}\frac{1}{2}\left[\frac{1}{2}(K_i+K_j)\,R+\frac{1}{2}(K_i+\alpha K_j)\,R\right]+\frac{1}{2}\left[\frac{1}{2}(C_i+C_j)\,u+\frac{1}{2}(C_i+\alpha C_j)\,u\right]$$

subject to the technology constraint $F(K_i, C_i) = 0$. The objective can be rewritten as

$$\begin{aligned} & \frac{1}{2}\left[K_i + \frac{1+\alpha}{2}K_j\right]R + \frac{1}{2}\left[C_i + \frac{1+\alpha}{2}C_j\right]u \\ = & \frac{1}{2}\left[K_i R + C_i u\right] + \frac{1+\alpha}{2}\frac{1}{2}\left[K_j R + C_j u\right] \end{aligned}$$

Since the firm i will take the firm j's decision as given, she is maximizing $\frac{1}{2}(K_i R + C_i u)$, which is equivalent to the social planner's objective. This result is independent of α, and to be consistent with ex post efficiency in the main text we can set $\alpha = 1$. Also, as the argument suggests, the result goes through as long as the following holds: in the event that the other firm j is making the offer, the firm j's behavior of giving his worthless asset to firm i is not affected by the firm i's endowment.

This extension illustrates that, in line with Hoberg and Phillips (2010), our inefficiency arises only in competitive markets.