Inefficient liquidity provision

John Geanakoplos[1,2] · **Kieran James Walsh**[3]

Abstract We prove that in competitive market economies with no insurance for idiosyncratic risks, agents will always overinvest in illiquid long-term assets and underinvest in short-term liquid assets. We take as our setting the seminal model of Diamond and Dybvig (J Polit Econ 91(3):401–419, 1983), who first posed the question in a tractable model. We reach such a simple conclusion under mild conditions because we stick to the basic competitive market framework, avoiding the banks and intermediaries that Diamond and Dybvig (1983) and others introduced.

Keywords Liquidity · Constrained inefficiency · Diamond–Dybvig models · Fire sales

JEL Classification E44 · D5 · E43 · E6 · G18

1 Introduction

The Emergency Economic Stabilization Act of 2008 stated that its goal was "to restore liquidity and stability to the financial system of the United States." To this end, the act originally gave the Treasury authority to purchase up to $700 billion in illiquid "troubled assets." Indeed, it now seems clear that before the late 2000s crisis, financial institutions systematically overinvested in illiquid assets and underinvested in safe,

✉ Kieran James Walsh
  walshk@darden.virginia.edu

1 Department of Economics, Yale University, New Haven, CT, USA

2 Santa Fe Institute, Santa Fe, NM, USA

3 University of Virginia Darden School of Business, Charlottesville, VA, USA

liquid assets. Should this pattern surprise us? In general, do agents in competitive markets systematically choose to provide too little liquidity?

Diamond and Dybvig (1983) provide the seminal analysis on the trade-off between safe, short-term assets and higher-yielding, long-term assets that may be illiquid in the short run. The tension is that while investors want high average returns, they may experience liquidity shocks and require cash before the maturity of their long-term investments. With access to a sufficiently rich set of assets, competitive equilibrium in such economies is efficient, and investors achieve optimal liquidity insurance. However, as Diamond and Dybvig (1983) observe, this set of assets would need to contain derivatives that condition payments on the needs of individual investors. As liquidity needs often are private information or difficult to observe, competitive asset markets of this kind will likely be limited.

The basic Diamond and Dybvig (1983) framework consists of three periods and two technologies, a high-yielding one that takes two periods to mature and a one period storage technology that is "liquid" in the sense that in the middle period its resources are available to consume. In the first period, the ex ante identical agents divide their resources between the two technologies. In the second period, agents may consume and trade bonds, but they cannot liquidate long-term investments. In the third period, returns are realized, and the agents consume. Uncertainty and the need for insurance stem from the fact that in the second period agents face random liquidity shocks. In particular, agents randomly differ in how they discount future utility; that is, they differ in impatience. The sudden impatience represents, for example, the possibility that an investor might experience a random medical problem or that a financial firm might have to honor a credit default swap. We extend this idiosyncratic risk to allow for aggregate risk: Both the fraction of the population at each impatience level and the long-term return are stochastic. Consequently, the interest rate that emerges in the middle period is also stochastic.

At this point in their analysis, Diamond and Dybvig (1983) took a sharp turn. They moved away from competitive and anonymous markets and introduced banks and financial intermediaries. Their paper shows that such institutions may improve upon stark incomplete markets by offering incentive compatible contracts to investors. In the original Diamond and Dybvig (1983) model, a standard bank deposit contract is incentive compatible and sufficient for optimal insurance, even when liquidity needs are private information. The Diamond and Dybvig (1983) analysis is famous for a nasty side effect of bank deposit contracts: Their model exhibits multiple equilibria. While one equilibrium yields optimal insurance, the other consists of a bank run, which is worse than autarky. So, in overcoming the private information problem, the market becomes fundamentally fragile.

While most related papers follow the lead of Diamond and Dybvig (1983) and study bank runs or intermediary contracting problems, we return to the simpler setting consisting of just individual investors and competitive asset markets. We answer our original question by showing that in our Diamond–Dybvig setting, agents always underinvest in short-term, liquid assets.

We require two mild assumptions: (1) Absolute risk aversion is nonincreasing. (2) Ex ante, investors place more weight on their future impatient selves than on their future patient selves. Almost all commonly used utility functions satisfy (1); Arrow

([1965](#)) wrote that violations of decreasing absolute risk aversion are problematic. Assumption (2) is simply the embodiment of liquidity shocks being emergencies, which are associated with high marginal utility, all else equal.

Our main result (Theorem 1) is that under (1) and (2), all incomplete markets competitive equilibria are constrained inefficient: agents overinvest in the illiquid, long-term technology. The investors would all be better off ex ante if they each decided to shift some resources from the longer illiquid investment to the liquid investment. That welfare rises is surprising: weighting their impatient selves more, that is, realizing liquidity shocks are emergencies, the investors were already allocating extra to the short-term technology as insurance against liquidity shocks. However, in equilibrium, they do not allocate enough.

A pecuniary externality is at play. When every individual shifts resources from long-term to short-term, they depress the middle period interest rates, reallocating resources from middle period lenders (individuals who turned out to be patient) to borrowers (individuals who turned out to be impatient). When an individual shifts his own resources from long-term to short-term, he is shifting money from the last period to the middle period whether or not he turns out to be impatient.[1]

A key step in deriving the inefficiency of competitive investment is proving that the interest rate between the last two periods falls when agents get richer in the penultimate period and poorer in the final period. To sign this static, it suffices to show that the resulting two period economies always display stability, so we are able to employ the results of Geanakoplos and Walsh (2016), who show that in two good, I agent economies with common Bernoulli utilities and common endowments but arbitrary heterogeneous discounts, nonincreasing absolute risk aversion guarantees stability.

Our analysis contributes to the existing literature on liquidity provision in three main ways. First, most papers following Diamond and Dybvig (1983) study the intermediary contracting problem, the interaction of intermediaries, or security design. We show that the welfare and policy results of this literature arise in a simple competitive equilibrium setting. We suggest that the main implications and lessons stem simply from market incompleteness, and not from private information mechanism design. Second, we employ weaker assumptions in proving overinvestment in the long-term technology. While our main assumption is nonincreasing absolute risk aversion, previous authors impose some combination of CRRA u, homothetic u, numerical bounds on relative risk aversion, and bounds on the long-term return. Third, unlike previous studies, we allow for aggregate risk in the fraction of the population that turns out to be impatient (as well as in the long-term return), which makes the interest rate random.

The paper concludes with an extension in which the investors may in the middle period liquidate the long-term technology. In states where the fraction of impatient

---

[1] We prove in Theorem 1 that the *over*investment stems from the argument in the last paragraph, that when impatience is caused by emergencies, the utility from impatient selves is weighted more than the utility from patient selves. We prove in Theorem 3 that if the patient selves are weighted sufficiently more than the impatient selves, then there is *under*investment in all equilibria. Finally, in Theorem 4 we prove that when there is no aggregate risk about the fraction of the population that turns out to be impatient, then equilibrium is unique, and for almost all ex ante weights individuals place on their future patient and impatient selves, equilibrium is constrained inefficient: Everybody could be made better off if everybody shifted investment one way or the other.

agents is particularly high, the long-term return is low, or the liquidation return is high, the interest rate rises to a level that induces the agents to liquidate part of the long-term technology. This fire sale of productive assets is a tangible manifestation of underinvestment in the short-term technology. Theorem 2 shows that the fire sale economy also exhibits constrained inefficient liquidity underprovision. In a numerical example, forcing agents to provide more liquidity initially and thus to curtail investment in the more productive long-term asset actually *increases* total production in crisis states of the world: Fire sales decline at a rate faster than the increase in liquidity.[2]

2 Related literature

Economists have long discussed the tension financial institutions face in deciding between safe, liquid reserves and more profitable but illiquid long-term investments. In their Yale College economics textbook, Fairchild et al. (1937) write,

> The banker is impelled by two counteracting motives, profits and safety. The bank derives its profits principally from the making of loans and discounts, and the larger its portfolio of loans and discounts the larger in general will its profits be. But as loans and discounts are made, cash is immediately withdrawn or deposits are created or the bank's note issues are increased. Thus, the reserve ratio falls, and the bank's condition becomes proportionately less safe. To the banker's desire for profits is thus opposed the necessity of keeping a safe ratio between the reserve and the demand liabilities. (p. 439)

Diamond and Dybvig (1983) provide the most famous mathematical treatment of this topic. They observe, as do we, that standard competitive markets will be inefficient without state contingent assets. However, rather than addressing constrained efficiency or possible beneficial market interventions, these authors analyze the ability of bank deposit contracts to insure investors and famously observe that such intermediary contracts introduce the possibility of a bank run equilibrium, which is worse than autarky.

Also, the original Diamond–Dybvig preference structure is quite particular. They assume that impatient types have the utility function $u(c_1)$ and that patient types have the utility function $u(c_1 + c_2)$, where c_1 and c_2 are first and second period consumption. We suppose that both types have the utility function $u(c_1) + \beta u(c_2)$, where β is lower for the impatient agents. Furthermore, unlike Diamond and Dybvig (1983), we require neither a quantitative bound on relative risk aversion nor a cross-restriction on the long-term asset return and the size of the preference shock.

Jacklin (1987) extends Diamond and Dybvig (1983) along two dimensions. First, he considers both the original Diamond–Dybvig preferences and a preference specification similar to the one we employ. Second, he compares the risk sharing properties of standard bank deposit contracts with those of an economy in which agents trade

---

[2] Of course in low interest rate states, the curtailment of long-term investments reduces production. Ex ante expected production also may fall. But as our theorems show expected utility increases for everyone when illiquid investment is curtailed a little.

equity shares in a firm that makes the initial long-run/short-run portfolio decision. When agents cannot trade bank deposits, deposit contracts welfare dominate allocations from competitive equity markets because second-best insurance calls for a wedge between the marginal rate of substitution and the technological rate of return. With competitive, laissez-faire asset markets, investor trade erodes such wedges. Indeed, when agents can retrade bank deposits, allocations from the two mechanisms coincide. For this reason, Jacklin (1987) policy suggestion is the "prohibition of a frictionless credit market," in the presence of an existing bank deposit arrangement.

Farhi et al. (2009) effectively argue, however, that the characterization of Jacklin (1987) is too stark. In their model, intermediaries compete to design liquidity insurance mechanisms for investors. The catch is that the resulting contracts are non-exclusive in the sense that the intermediaries cannot stop their customers from trading bonds among themselves. With more preference generality than in Jacklin (1987), these authors order by welfare a taxonomy of insurance arrangements. First-best (SP^1) allocations maximize ex ante utility subject to resource constraints. Second-best (SP^2) allocations honor both resource and incentive compatibility constraints. With the original Diamond–Dybvig preferences, SP^1 and SP^2 coincide. When profit maximizing intermediaries compete to insure the agents, Farhi et al. (2009) refer to the situation as CE^2. As competition drives profits to zero, CE^2 coincides with SP^2. That is, subject to the private information constraint, the government cannot improve on competitive insurance markets. CE^3 arises when price-taking intermediaries compete *and* investors can privately trade in competitive bond markets. The observation of Jacklin (1987) is basically that CE^3 is worse than SP^2 (and CE^2) and that preventing private retrading improves welfare. Market forces are effectively an additional constraint.

In contrast, Farhi et al. (2009) show that CE^3 is inefficient beyond its distinction from CE^2. Specifically, they characterize the third best (SP^3), in which the planner maximizes utility subject to resource constraints, incentive constraints (regarding liquidity preferences), and constraints that prevent investor side trades at the corresponding market prices. SP^3 is worse than SP^2 because it entails more constraints. SP^3 is, however, better than CE^3 because unlike the price-taking intermediaries, the planner understands how his contracts affect the private market interest rate. In short, the allocations of CE^3 suffer due to a pecuniary externality. The key theorem of Farhi et al. (2009) explains how simple liquidity controls on intermediaries drive the economy from CE^3 to SP^3. They show that if the impatient agents have, all else equal, more (less) ex ante utility weight, then CE^3 underprovides (overprovides) liquidity. However, to sign the link between initial investment and ex ante utility in any case, the authors require either preference homotheticity or a bound on the liquidity shock variance. Thus, our analysis is distinct both in that we study the case without intermediaries and require just nonincreasing absolute risk aversion.

Bhattacharya and Gale (1987) study a related setting in which intermediaries insure agents who want to consume either early or late. However, at the time of initial investment, each intermediary is uncertain about the fraction of "early diers" he will face and wants to insure against having too many of them. Consequently, the model of Bhattacharya and Gale (1987) is mathematically similar to the general Diamond–Dybvig setting but has intermediaries instead of individual investors as its agents. In the parlance of Farhi et al. (2009), the focus of Bhattacharya and Gale (1987) is a

characterization of SP^2. In particular, they show how different assumptions regarding preferences affect the sign of the optimal wedge between the technological return and the effective interest rate. Also, they observe that Walrasian interbank markets exhibit "free-rider" problems, but they do not formalize this argument.

Finally, using the original Diamond–Dybvig preferences, Yared (2013) also considers a case without intermediaries. In his model, agents cannot commit to honoring debt contracts, but the government can issue bonds for trade (financed by taxes). When the government issues a sufficient number of bonds, the economy is equivalent to one with unconstrained, uncontingent trade between agents. In this case, the equilibrium allocations are inefficient. However, by restricting the supply of bonds, the government effectively imposes a $t = 1$ borrowing constraint and under certain conditions can implement the first best. The borrowing constraint helps because it forces agents to store more initially. So, Yared (2013) also derives an overinvestment result. Our analysis is distinct from his because (i) he uses the original Diamond–Dybvig preference/shock assumptions, (ii) his policy intervention involves using fiscal policy to set a borrowing constraint in the middle period of the model, and (iii) he does not allow for aggregate risk.[3]

3 Model

Consider an economy consisting of three time periods, $t \in \{0, 1, 2\}$, and a unit mass of ex ante identical investors. There is a single consumption good in each time period, and each agent is endowed with $e > 0$ units of the good at $t = 0$. At $t = 0$, each agent allocates his endowment between two investment technologies. The first is a one period saving technology that yields a gross return of 1 at $t = 1$. The second is a long-term illiquid investment that gives a gross return of $R_s > 0$ at $t = 2$ in aggregate state $s \in S$, where $S \geq 1$. Let x and y denote the respective allocations to the short- and long-term assets. Let $e_{s1}(x)$ and $e_{s2}(x)$ denote the $t = 1$ and $t = 2$ endowments conditional on liquid investment x in state s. By definition, $e_{s1}(x) = x$ and $e_{s2}(x) = R_s(e - x) = R_s y$.[4] At $t = 1$, the agents trade one period, riskless bonds in competitive, anonymous markets.

While investors do not consume at $t = 0$, they derive utility from consumption at $t = 1$ and $t = 2$ according to the twice continuously differentiable Bernoulli utility function u, which satisfies $u' > 0$, $u'' < 0$, and $\lim_{x \downarrow 0} u'(x) = \infty$. At $t = 0$, the agents face two forms of uncertainty. First, agents randomly have different tastes for consumption at $t = 1$ and $t = 2$. That is, they have different degrees of impatience, and each investor learns his type at $t = 1$. Second, there is aggregate risk concerning

---

[3] See Allen and Gale (2004), Fostel and Geanakoplos (2008), and Grochulski and Zhang (2016) for additional treatments of constrained inefficiency in Diamond–Dybvig-like models. See Villamil (1991), Lin (1996), Hazlett (1997), Ennis and Keister (2016), Kashyap et al. (2014), and many others for additional applications of the Diamond–Dybvig model to banking, bank runs, and contracting.

[4] Since e is arbitrary and completely invested in the two assets (there is no $t = 0$ consumption), it is without loss of generality to fix the short-term return at 1. If the short-term return was $R_0 > 0$, then the $t = 0$ trade-off would be between a $t = 1$ return of R_0 or a $t = 2$ return of R_s. This is equivalent to letting the initial endowment be $\widetilde{e} = R_0 e$ and setting the $t = 1$ and $t = 2$ returns to 1 and R_s/R_0, respectively.

the probability distribution of types and the $t = 2$ return. As the $t = 1$ interest rate depends on the distribution of types, this uncertainty introduces interest rate risk.

Suppose there are $I > 1$ impatience types indexed by i. Type i is distinguished by the pair $(w^i, \beta^i) \gg 0$. We suppose the state $s, t = 1$ present value utility of impatience type i is

$$w^i \left[u \left(c^i_{s1} \right) + \beta^i u \left(c^i_{s2} \right) \right],$$

where c^i_{st} is the consumption of type i at $t \in \{1, 2\}$ in state $s \in S$. Parameter β^i represents the patience of type i. If $\beta^i < \beta^j$, then i is more impatient than j. Parameter w^i weights the utility of different types. If in addition to $\beta^i < \beta^j$ we also have $w^i \beta^i \geq w^j \beta^j$, then ex ante type i is valued more than type j, because the present value weights are higher for i in period 1 and at least as high in period 2. We suppose that impatience increases because of emergencies, which increase the utility of contemporaneous consumption. The case with $w^i \beta^i = w^j \beta^j$ corresponds to the one described in footnote 5 of independent and equally likely emergencies each period.[5]

There are S possible type distributions: $\Pi_1, \ldots, \Pi_S$, where $\Pi_s = \left(\pi^1_s, \ldots, \pi^I_s \right)$ and π^i_s is the probability of becoming type i conditional on distribution s. Let $P_s > 0$ be the $t = 0$ probability of realizing distribution s at $t = 1$. Assume, as usual, that π^i_s is both the probability of becoming type i as well as the fraction of agents that realize type i. By relabeling the impatience types as necessary, it is without loss of generality to order the agents as follows: $\beta^1 < \cdots < \beta^I$.

Let q_s be the $t = 1$ price of a bond paying 1 at $t = 2$. As the agents are the same ex ante, they choose the same x at $t = 0$ and thus have identical endowments at $t = 1$ and $t = 2$. However, because they differ in impatience, they trade in bonds at $t = 1$. The type distribution determines the gains from trade and thus the market clearing

---

[5] Our analysis would be the same even if emergencies were equally likely to come in period 1 or period 2. Suppose that every agent has probability p of facing an emergency in the middle period and the same probability p of facing an emergency in the last period. These risks are independent across periods and across agents. Suppose that in case of an emergency, the Bernoulli utility of consumption is multiplied by $\lambda > 1$. Let $\bar{\lambda} = p\lambda + (1 - p)1$. Let c_t be consumption in period t if there is no emergency in period t, and let c_{tE} be consumption in period t if there is an emergency in period t. Since there are no insurance markets, conditional on what happens in the middle period, agents will consume the same amount $c_2 = c_{2E}$ whether or not they face an emergency in the last period. Conditional on being in an emergency in period 1, an agent will be maximizing

$$\lambda u(c_{1E}) + \beta \bar{\lambda} u(c_2).$$

Conditional on not being in an emergency in period 1, an agent will be maximizing

$$u(c_1) + \beta \bar{\lambda} u(\hat{c}_2)$$

The upshot is that an agent who faces an emergency in period 1 acts as if he is more impatient than an agent who does not face an emergency in period 1. Just as importantly, assuming that ex ante agents maximize the expectation of these utilities, one can see that more weight is put on the impatient self than the patient self in the sense that the period 1 utility of the impatient (emergency) self is weighted more than the period 1 utility of the patient (nonemergency) self, and the expected period 2 utility (of the period 1 emergency self) is weighted at least as much as the expected period 2 utility of the nonemergency period 1 self.

interest rate. Hence, the bond price has an s subscript. Define $\mathbf{q} \equiv (q_1, \ldots, q_S)$. Let b_s^i denote the bond holdings of type i conditional on state s. The $t = 0$ budget set of an agent is thus

$$
\widetilde{B}(\mathbf{q}, e) = \left\{
\begin{array}{l}
\left((x, y), \left(c_{s1}^i, c_{s2}^i, b_s^i \right)_{s \in S}^{i \in I} \right) \in \mathbb{R}_+^2 \times \mathbb{R}^{3SI} \mid \forall i \in I, s \in S \\
x + y \leq e, \\
c_{s1}^i + q_s b_s^i \leq x, \; c_{s2}^i \leq R_s y + b_s^i, \\
c_{s1}^i, c_{s2}^i \geq 0
\end{array}
\right\}.
$$

Note that because there are no bond market frictions, it is without loss of generality to assume that the agents simultaneously trade $t = 1$ and $t = 2$ consumption goods at respective prices 1 and q_s. Conditional on (x, y) and aggregate state s, the setting becomes a classic I agent, two good endowment economy. We can ignore bond holdings and write the budget set as

$$
B(\mathbf{q}, e) = \left\{
\begin{array}{l}
\left((x, y), \left(c_{s1}^i, c_{s2}^i \right)_{s \in S}^{i \in I} \right) \in \mathbb{R}_+^2 \times \mathbb{R}_+^{2SI} \mid \forall i \in I, s \in S \\
x + y \leq e, \\
c_{s1}^i + q_s c_{s2}^i \leq x + q_s R_s y
\end{array}
\right\}.
$$

The $t = 0$ problem of each investor is thus

$$
\max_{(x,y), \left(c_{s1}^i, c_{s2}^i \right)_{s \in S}^{i \in I}} \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i \left[u \left(c_{s1}^i \right) + \beta^i u \left(c_{s2}^i \right) \right] \text{ such that}
$$

$$
\left((x, y), \left(c_{s1}^i, c_{s2}^i \right)_{s \in S}^{i \in I} \right) \in B(\mathbf{q}, e). \tag{1}
$$

Given e, define $c_{st}^i(x, q_s)$ to be time t consumption demand at bond price q_s and endowment $(x, R_s(e - x))$. We now define competitive equilibrium.

Definition 1 (*Competitive equilibrium*) Competitive equilibrium consists of prices q_s^* ($\forall s \in S$), a short-term asset investment x^* (liquidity), and consumption choices $c_{st}^{i*} = c_{st}^i(x^*, q_s^*)$ ($\forall i \in I, s \in S, t \in \{1, 2\}$) such that:

1. Given prices, $\left((x^*, e - x^*), \left(c_{s1}^{i*}, c_{s2}^{i*} \right)_{i \in I}^{s \in S} \right)$ solves the $t = 0$ investor problem (1).
2. Markets clear for all $s \in S$:

$$
\sum_{i=1}^{I} \pi_s^i c_{s1}^{i*} = x^*
$$

$$
\sum_{i=1}^{I} \pi_s^i c_{s2}^{i*} = R_s \left(e - x^* \right).
$$

From the assumptions on u, competitive equilibrium exists, and in equilibrium there is an interior solution for liquidity:

Proposition 1 (Existence) *A competitive equilibrium exists and $x^* \in (0, e)$.*

Proposition 1 stems from the following observation: as x approaches 0, the interest rate diverges to infinity in all states, encouraging high investment in x. As x approaches e, the gross interest rate goes to 0, encouraging high investment in y. By continuity, there is an intermediate x that constitutes an equilibrium.[6]

Note that Diamond and Dybvig (1983) assume there are two types and that patient and impatient utility are, respectively, $\rho u (c_1 + c_2)$ and $u (c_1)$, where $\rho < 1$, $\rho R > 1$, and relative risk aversion is greater than or equal to 1. While our specification is not technically a generalization of the classical Diamond–Dybvig one, our general setting exhibits the same tensions and trade-offs as those of the classical model. Moreover, our version is the standard representation of time variation in "taste" for consumption. In the classical setting, patient and impatient agents have qualitatively different preferences: while for the former c_1 and c_2 are *perfect* substitutes, for the latter $t = 2$ goods are completely useless. In contrast, we follow Jacklin (1987), Bhattacharya and Gale (1987), and Farhi et al. (2009) and analyze the case in which $t = 1$ and $t = 2$ goods are imperfect substitutes for both types.

4 Welfare analysis

As there are not complete $t = 0$ insurance markets for emergency risk, competitive equilibrium is not ex ante Pareto efficient. In this section, we show that equilibrium is also constrained inefficient (1): decreasing absolute risk aversion implies investors systematically underinvest in liquidity.

Once agents have chosen their initial investment x, the economy reduces to a simple collection of S two period economies. Observe that after each agent chooses x, his budget set in each state s if he is of impatience type i reduces to

$$\mathcal{B}(q_s, e, x, s) = \left\{ \left(c_{s1}^i, c_{s2}^i \right) \in \mathbb{R}_+^2 \mid c_{s1}^i + q_s c_{s2}^i \leq x + q_s R_s (e - x) \right\}.$$

We now define competitive equilibrium conditional on x:

Definition 2 (*Competitive equilibrium conditional on x*) Competitive equilibrium conditional on x consists of prices $q_s (x)$ ($\forall s \in S$) and consumption choices $c_{st}^i (x) = c_{st}^i (x, q_s (x))$ ($\forall i \in I, s \in S, t \in \{1, 2\}$) such that:

1. Given $q_s (x)$, we have $(c_{s1}^i (x), c_{s2}^i (x)) \in \mathcal{B}(q_s (x), e, x, s)$, and $(c_1, c_2) \in \mathcal{B}(q_s (x), e, x, s)$ implies

$$w^i \left[u \left(c_{s1}^i (x) \right) + \beta^i u \left(c_{s2}^i (x) \right) \right] \geq w^i \left[u (c_1) + \beta^i u \left(c_2^i \right) \right].$$

---

[6] Given $x \in (0, e)$, the model reduces to a collection of two good, I agent economies in which existence and continuity follow from standard arguments.

2. Markets clear for all $s \in S$:

$$\sum_{i=1}^{I} \pi_s^i c_{s1}^i (x) = x$$

$$\sum_{i=1}^{I} \pi_s^i c_{s2}^i (x) = R_s (e - x).$$

Given arbitrary bond prices **q** and liquidity x, we can define ex ante utility $V(x, \mathbf{q})$:

$$V(x, \mathbf{q}) = \max_{(c_{s1}^i, c_{s2}^i)_{s \in S}^{i \in I}} \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i \left[u \left(c_{s1}^i \right) + \beta^i u \left(c_{s2}^i \right) \right] \text{ such that}$$

$$\left(c_{s1}^i, c_{s2}^i \right) \in \mathcal{B} (q_s, e, x, s). \tag{2}$$

Therefore, conditional on x equilibrium ex ante utility is $V(x, \mathbf{q}(x))$, and competitive equilibrium ex ante utility is $V(x^*, \mathbf{q}(x^*))$. Consider now the $t = 0$ problem of a benevolent planner who is able to force all of the investors into a particular short-term savings level x but then must stand idly by while markets clear. The planner can anticipate that each agent will wind up with expected utility $V(x, \mathbf{q}(x))$. As his policy affects all agents, the planner, unlike the atomistic individuals, internalizes how x impacts q_s across different realizations of Π_s. Geanakoplos and Polemarchakis (1986) described a method for proving that generically in these types of situations with incomplete markets, the planner can improve utility by choosing x that differs from competitive equilibrium. Here we sharpen that conclusion by proving that the planner can always improve utility by increasing x beyond competitive equilibrium x^*.

Define $a(\cdot) \equiv -u''(\cdot) / u'(\cdot)$ to be the Arrow–Pratt measure of absolute risk aversion. u satisfies declining absolute risk aversion (DARA) if $a(c) \geq a(c')$ whenever $c \leq c'$. We now state our main result.

Theorem 1 *Suppose (i) u satisfies DARA and (ii) if $\beta^i \leq \beta^j$ then $w^i \beta^i \geq w^j \beta^j$. Then every competitive equilibrium is ex ante constrained inefficient. In particular, all investors could be made better off by collectively allocating more to the short-term asset: the function $V : (0, e) \times \mathbb{R}_{++}^S \to \mathbb{R}$ is differentiable, and at every equilibrium $x^* \in (0, e)$ and*

$$\left. \frac{d}{dx} V(x, \mathbf{q}(x)) \right|_{x=x^*} > 0.$$

Proof See Appendix A.3.   □

Condition (ii) says, essentially, that agents put more ex ante weight on their impatient selves than on their patient selves, which causes them to choose higher x than if they put equal weight on their future selves. The theorem says that, despite this,

they would all be better off if they all chose still higher x (and thus lower y). In short, condition (ii) and DARA imply overinvestment in the long-term, illiquid asset. Furthermore, one can show that interest rates are highest exactly when the fraction of impatient agents is highest. Hence, investors realize that they are most likely to become impatient exactly when borrowing rates will be highest. These high rates in impatient times incline them to allocate more to the short investment than they would have otherwise. But, as Theorem 1 tells us, the investors do not allocate enough to the short-term asset: no one agent affects the interest rate, which is the tool by which all are made better off collectively.

Intuitively, increasing liquidity above x^* has two effects. First, by stability the interest rate falls (the price of $t = 2$ goods rises) as the relative supply of $t = 1$ goods rises. For all realizations of s, the falling rate redistributes from the lenders (the patient agents) to the borrowers (the impatient agents). As there is more utility weight on the impatient agents, ex ante utility increases (as does ex post utility, on average). The second effect is that each agent's $t = 1$ ($t = 2$) endowment rises (falls). In states where the interest rate is high, this effect actually benefits all agents, holding constant the price. When the interest rate is low, everyone is hurt. The question is, what is the overall balance of these endowment effects? Can they overwhelm the beneficial redistribution effect? The answer is no. As we see in the Date 0 Lemma (Lemma 1), this balancing is exactly the margin on which each agent is choosing x to begin with. At x^*, these endowment effects completely wash out. Only the redistribution effect remains, which means increasing liquidity (local to x^*) unambiguously improves utility ex ante.

A key step in proving Theorem 1 is establishing stability of equilibrium, which follows from the results of Geanakoplos and Walsh (2016). They show that utility separability, DARA, common endowments, and common Bernoulli utilities are sufficient for uniqueness and stability in two good, I agent economies, which are what emerge in the present analysis after x is chosen and uncertainty is resolved. The key step in proving stability is in establishing that total demand for $t = 2$ goods is downward sloping in q for any x and realization of s. Why is DARA sufficient for downward sloping demand? Proving sufficiency amounts to showing that the income effects of relatively impatient agents are not too strong. Patient agents, who are savers, necessarily have downward sloping demand: The substitution effect in the Slutsky equation is negative by separability and concavity, and the income effect term is negative because they are buyers of $t = 2$ goods. For very impatient agents, however, the income effect may be positive because the most impatient agents are borrowing (they are sellers of $t = 2$ goods). DARA is sufficient for showing that the total income effect is negative. Specifically, this assumption ensures that the patient agents have the largest (absolute value) income effects. Why? The most patient agents consume the most $t = 2$ goods. If $a(c) \geq a(c')$ whenever $c \leq c'$, these agents have the least sensitive marginal utility of consumption for these goods. If you give an investor a splash of income, he allocates between $t = 1$ and $t = 2$ goods, keeping constant the ratio of marginal utilities. With highly insensitive marginal utility for $t = 2$ goods, a patient agent will change consumption of $t = 2$ goods a lot just to the maintain the ratio. See Appendix A.1 and Geanakoplos and Walsh (2016) for more details.

Curtailing short-term investment is not always the right intervention. If instead of condition (ii) we assume $w^1 \leq w^2 \leq \cdots \leq w^N$, then there is always *underinvestment*

in the long-term technology, which we prove as Theorem 3 in Appendix. If the w's satisfy Eq. 8 in Appendix, then equilibrium is constrained efficient. Is Eq. 8 a knife-edge case? When $S = 1$, we can prove the answer is yes: Theorem 4 in Appendix considers the classical case of no aggregate risk about the fraction of the population that turns out to be impatient. We prove that equilibrium is unique, and for almost all ex ante weights $(w^1, \ldots, w^I)$ individuals place on their future patient and impatient selves, equilibrium is constrained inefficient: everybody could be made better off if everybody shifted investment one way or the other.[7]

In comparison with the existing literature, our assumption of DARA is weaker than those employed in previous papers, which impose some combination of CRRA u, homothetic u, numerical bounds on relative risk aversion, and restrictions on R_s. Some variation of condition (ii) has been needed in previous papers to establish underinvestment in liquidity. A special case is $w^i = 1/\beta^i$, which implies $w^1\beta^1 = \cdots = w^I\beta^I = 1$. $w^i = 1/\beta^i$ corresponds to the "liquidity shock" case from Farhi et al. (2009) and is similar to the original Diamond–Dybvig specification. Finally, while most other papers emphasize the role of banks or intermediaries in creating inefficiencies, we derive and sign constrained inefficiency in a simple setting consisting of just anonymous trade in competitive asset markets.

5 Fire sales

We now change the model so that agents are able to liquidate the long-term investment at a fire sale rate at $t = 1$. In particular, after uncertainty is resolved, an agent may liquidate an amount L_s^i, where $0 \leq L_s^i \leq y$. Doing so yields a fire sale return of $r_s L_s^i$ at $t = 1$, where $r_s > 0$, and means that the agent's $t = 2$ endowment falls to $R_s\left(y - L_s^i\right)$. The $t = 0$ problem of each investor thus becomes

$$
\max_{(x,y),\left(c_{s1}^i, c_{s2}^i, L_s^i\right)_{s\in S}^{i\in I}} \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i \left[u\left(c_{s1}^i\right) + \beta^i u\left(c_{s2}^i\right) \right] \text{ such that}
$$

$$
\left((x, y), \left(c_{s1}^i, c_{s2}^i, L_s^i\right)_{s\in S}^{i\in I} \right) \in \mathcal{B}_L\left(\mathbf{q}, e\right)
$$

$$
\mathcal{B}_L\left(\mathbf{q}, e\right) = \left\{ \begin{array}{l} \left((x, y), \left(c_{s1}^i, c_{s2}^i, L_s^i\right)_{s\in S}^{i\in I} \right) \in \mathbb{R}_+^2 \times \mathbb{R}_+^{3SI} \mid \forall i \in I, s \in S \\ x + y \leq e \\ 0 \leq L_s^i \leq y \\ c_{s1}^i + q_s c_{s2}^i \leq \left(x + r_s L_s^i\right) + q_s R_s\left(y - L_s^i\right) \end{array} \right\}.
$$

---

[7] With $S > 1$, we conjecture Theorem 4 would still hold, but the proof would be more difficult. Theorem 3 does not follow from Geanakoplos and Polemarchakis (1986) because in the present paper we have fixed individual endowments at 0 (all consumption comes from production). Geanakoplos and Polemarchakis (1986) prove constrained inefficiency for almost all endowments, not for any fixed set of individual endowments.

Let $L_s = \sum_{i=1}^{I} \pi_s^i L_s^i$ be the total fire sale liquidation. Competitive equilibrium is as in Definition 1, except the market clearing conditions become

$$\sum_{i=1}^{I} \pi_s^i c_{s1}^i = x + r_s L_s$$

$$\sum_{i=1}^{I} \pi_s^i c_{s2}^i = R_s \left(e - x - L_s \right). \tag{3}$$

In contrast with the $r_s = 0$ version in which liquidation is impossible, here a high interest rate ($1/q_s$) induces $t = 1$ liquidation, resulting in a decline in $t = 2$ production. Indeed, if $1/q_s > R_s/r_s$ in state s, agents would fully liquidate the long-term technology, consuming at $t = 2$ only via the bond market. This complete fire sale, however, could never be an equilibrium because it would violate $t = 2$ market clearing (3). Therefore, in state s of the fire sale equilibrium either $1/q_s = R_s/r_s$ and agents are indifferent to liquidation or $1/q_s < R_s/r_s$ and $L_s = 0$. Because agents are indifferent in fire sale states, it is straightforward to extend the proof of Theorem 1 to the fire sale economy, yielding Theorem 2.

Theorem 2 *Suppose (i) u satisfies DARA, (ii) if $\beta^i \leq \beta^j$ then $w^i \beta^i \geq w^j \beta^j$, and (iii) there is at least one state without fire sales (e.g., $r_s = 0$ for some $s \in S$). Then every competitive equilibrium with fire sales is ex ante constrained inefficient. In particular,*
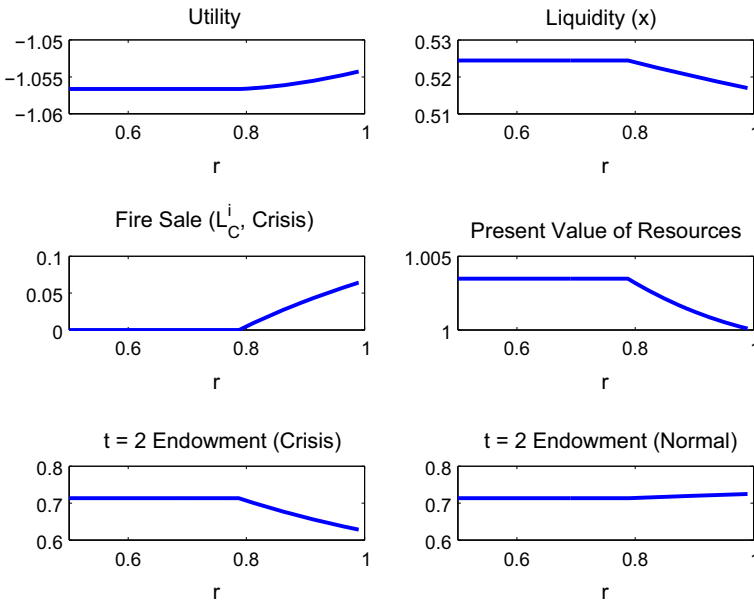

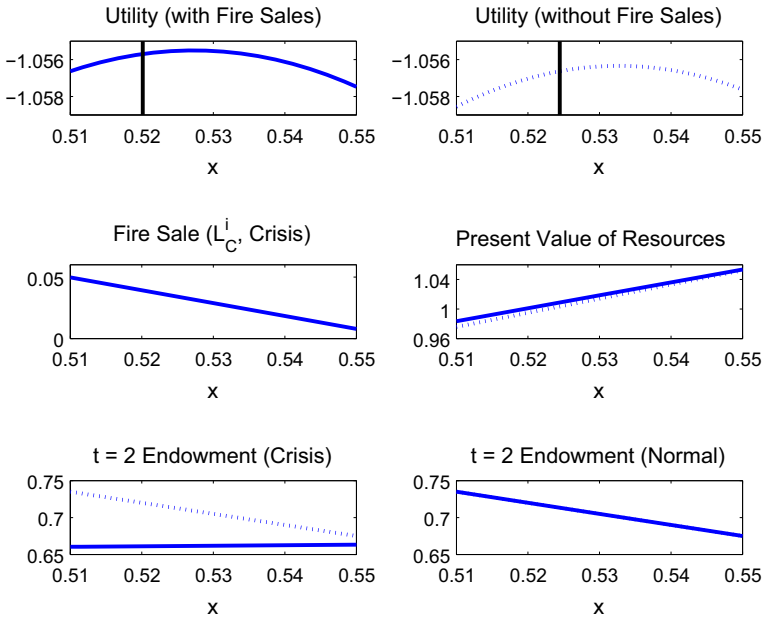**Fig. 1** Varying the liquidation value (r) in the fire sales model

Fig. 2 Varying the initial level of liquidity (x). The *solid lines* correspond to the case with $r = .9$. The *dotted lines* correspond to the case without fire sales. The *vertical lines* correspond to the equilibrium levels of liquidity

all investors could be made strictly better off by collectively allocating more to the short-term asset.

Proof See Appendix A.4. □

5.1 Numerical fire sale example

Consider the special case where there are two types, impatient and patient ($\rho^I > \rho^P$, where $\rho^i = 1/\beta^i$), and two aggregate states, normal and crisis ($\pi_N^I < \pi_C^I$). R and r are constant across states. Let $u(c) = \log(c)$, $(\rho^I, \rho^P) = (2, 1)$, $R = 1.5$, $(\pi_C^I, \pi_N^I) = (.5, .1)$, $(P_C, P_N) = (.1, .9)$, and $e = 1$. In Fig. 1, we see that as liquidation recovery r increases past roughly .8, the agents liquidate part of the long-term technology in the crisis state. And, in anticipation of this option, all else equal, initial liquid short-term investment x declines. The decline in short-term liquid investment, coupled with the fire sale option, generates endogenous volatility in $t = 2$ production. Note, however, that the possibility of liquidation actually increases ex ante utility, as we see in Fig. 1: The fire sale caps the interest rate at R/r and effectively improves insurance.

Figure 2 shows what happens with the imposition of various initial liquidity levels. Increasing x decreases the interest rate in normal times (Fig. 3) and improves utility, up to a point. Most interestingly, we see that while increasing x mechanically decreases the $t = 2$ endowment in normal times, the additional liquidity actually *increases* production in crisis times. The reason is that fire sales decline faster than the fall in
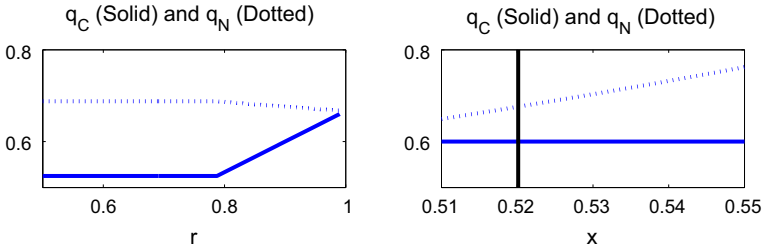
Fig. 3 Bond prices, varying the liquidation value (r) and the initial level of liquidity (x). The *vertical line* corresponds to the equilibrium level of liquidity. $r = .9$ in the *right pane*

long-term investment. Therefore, a liquidity floor not only improves welfare for the investors but also increases output in the worst states of the world.

6 Conclusion

In our generalized version of the seminal Diamond–Dybvig model, incomplete but anonymous and competitive markets yield constrained inefficient allocations. Specifically, investors underprovide the market with liquidity: If they were to collectively invest less in long-term illiquid assets, they would all achieve better insurance and higher ex ante utility. We allow for uncertainty in the liquidity shock distribution and the long-term return, which creates aggregate interest rate risk. Our key assumptions are just that (i) all else equal, investors experiencing liquidity shocks have higher marginal utility of consumption and (ii) absolute risk aversion is nonincreasing in consumption. An intuitive and common argument is that financial institutions and investors overinvest in high-yielding but illiquid assets. We have shown that this over-investment holds across a number of important dimensions of generality.

A Appendix

A.1 Geanakoplos and Walsh (2016) results for the two good, *I* agent economy

This section rehashes the results of Geanakoplos and Walsh (2016), which we use in proving our theorems. Consider the economy that exists after liquidity x has been chosen and all uncertainty has resolved. To simplify notation, drop reference to x and s. This economy consists of the I impatience types, where π^i is the fraction of i types, and two goods, $t = 1$ consumption (c_1^i) and $t = 2$ consumption (c_2^i). The agents have the identical endowment (e_1, e_2) $\gg 0$. Stars denote competitive equilibrium quantities. The utility function of impatience type $i \in I$ is $u\left(c_1^i\right) + \beta^i u\left(c_2^i\right)$, where u is twice continuously differentiable, $u' > 0, u'' < 0$, and $\lim_{x \downarrow 0} u'(x) = \infty$. u satisfies

DARA if $a(c) \geq a(c')$ whenever $c \leq c'$, where $a(\cdot) \equiv -u''(\cdot)/u'(\cdot)$. The agents are ordered by patience: $\beta^1 < \cdots < \beta^I$. The budget constraint is $c_1^i + qc_2^i \leq e_1 + qe_2$. Define $D(q) \equiv \sum_{i=1}^{I} \pi^i c_2^i ((e_1, e_2), q)$ to be market demand for $t = 2$ goods, where $c_2^i((e_1, e_2), q)$ is type i's demand at price q. From Geanakoplos and Walsh (2016), we have the following result:

Proposition 2 *Suppose u satisfies DARA. Then equilibrium is unique and the following hold:*

1. $c_1^{1*} > \cdots > c_1^{I*}$.
2. $c_2^{1*} < \cdots < c_2^{I*}$.
3. $\rho^1 u'\left(c_1^{1*}\right) > \cdots > \rho^I u'\left(c_1^{I*}\right)$, *where* $\rho^i = 1/\beta^i$.
4. *Demand is downward sloping:* $D'(q^*) < 0$.
5. *Equilibrium is stable:* $\partial q^*/\partial e_1 > 0$ *and* $\partial q^*/\partial e_2 < 0$.
6. q^* *is continuously differentiable in* (e_1, e_2)*, and* c_t^i *is continuously differentiable in* (e_1, e_2) *and* q.

Proof Parts 1, 2, 4, and 5 are proved in Geanakoplos and Walsh (2016). Part 3 follows immediately from Part 1 and the Euler equations, $q = \beta^i u'\left(c_2^i\right)/u'\left(c_1^i\right)$, which hold at the optimum by the assumptions on u. The properties in Part 6 are standard, following from the twice continuous differentiability of u, the Euler equation, market clearing conditions, and the implicit function theorem. □

Intuitively, when absolute risk aversion is decreasing, then the derivative of consumption with respect to wealth is increasing in consumption: $\partial c_2^i/\partial \omega$ rises monotonically in c_2^i, meaning buyers have the strongest income effects. It follows that demand is downward sloping, which implies stability.

A.2 Statement and Proof of the Date 0 Lemma (Lemma 1)

Lemma 1 (Date 0 Lemma) *At competitive equilibrium, the partial derivative of ex ante utility with respect to liquidity is* 0*:*

$$\frac{\partial}{\partial x} V(x, \mathbf{q}(x))\bigg|_{x=x^*} = 0$$

Proof Plugging the budget constraint into Eq. 2, ex ante utility is

$$V(x, \mathbf{q}(x)) = \max_{c_{s2}^i \geq 0} \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i \left(u\left(x + q_s(x) R_s(e-x) - q_s(x) c_{s2}^i \right) + \beta^i u\left(c_{s2}^i\right) \right),$$

so by the envelope theorem

$$\frac{\partial V}{\partial x} = \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i u'\left(c_{s1}^i(x, q_s(x)) \right) (1 - q_s(x) R_s). \tag{4}$$

Similarly, plugging in the budget constraint, for any **q** the agent problem (1) becomes

$$\max_{x\in[0,e],c_{s2}^i\geq 0} \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i \left(u\left(x + q_s R_s\left(e - x\right) - q_s c_{s2}^i\right) + \beta^i u\left(c_{s2}^i\right)\right),$$

and the x FOC is

$$0 = \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i u'\left(c_{s1}^i\left(x, q_s\right)\right)\left(1 - q_s R_s\right). \tag{5}$$

By the concavity of the problem and $x^* \in (0, e)$ (see Proposition 1), the FOC holds in equilibrium. Since the right-hand sides of (5) and (4) coincide at $x = x^*$ in equilibrium, we have

$$\frac{\partial}{\partial x} V\left(x, \mathbf{q}\left(x\right)\right)\bigg|_{x=x^*} = 0.$$

$\square$

A.3 Proof of Theorem 1

Plugging the budget constraint into Eq. 2 and optimizing over c_{s2}^i, ex ante utility is

$$V\left(x, \mathbf{q}\left(x\right)\right)$$
$$= \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i \left(\begin{matrix} u\left(x + q_s\left(x\right) R_s\left(e - x\right) - q_s\left(x\right) c_{s2}^i\left(x, q_s\left(x\right)\right)\right) \\ + \beta^i u\left(c_{s2}^i\left(x, q_s\left(x\right)\right)\right) \end{matrix}\right). \tag{6}$$

By Part 6 of Proposition 2, $V : (0, e) \times \mathbb{R}_{++}^S \to \mathbb{R}$ is well defined and differentiable. By the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}x} V = \frac{\partial V}{\partial x} + \sum_{s=1}^{S} q_s'\left(x\right) \frac{\partial V}{\partial q_s}.$$

By the Date 0 Lemma (Lemma 1), at $x = x^*$ the first term on the right-hand side is 0, so

$$\frac{\mathrm{d}}{\mathrm{d}x} V\left(x, \mathbf{q}\left(x\right)\right)\bigg|_{x=x^*} = \sum_{s=1}^{S} q_s'\left(x^*\right) \frac{\partial V\left(x^*, \mathbf{q}\left(x^*\right)\right)}{\partial q_s}.$$

Since $x^* \in (0, e)$ by Proposition 1 and since DARA holds, $q_s'(x^*) > 0$ by stability (Part 5 of Proposition 2). Thus, the theorem follows provided for all $s \in S$

$$\left. \frac{\partial V(x, \mathbf{q}(x))}{\partial q_s} \right|_{x=x^*} > 0.$$

From Eq. 6, we have

$$\left. \frac{\partial V(x, \mathbf{q}(x))}{\partial q_s} \right|_{x=x^*} = P_s \sum_{i=1}^{I} \pi_s^i w^i \beta^i \rho^i u'\left(c_{s1}^{i*}\right) \left(R_s\left(e - x^*\right) - c_{s2}^{i*}\right), \quad (7)$$

where $\rho^i = 1/\beta^i$. By premise (ii) of the theorem and Proposition 2 (Parts 2 and 3), $\left(w^i \beta^i \rho^i u'\left(c_{s1}^{i*}\right)\right)^{i \in I}$ and $\left(R_s\left(e - x^*\right) - c_{s2}^{i*}\right)^{i \in I}$ are both strictly decreasing sequences. By interpreting type $i \in I$ as a random variable drawn with probability π_s^i, these sequences are strictly decreasing functions of the random variable (call them $f(i)$ and $g(i)$, respectively), and we can write the weighted summation in Eq. 7 as expectation with respect to Π_s:

$$E_{\Pi_s}\left[w^i \beta^i \rho^i u'\left(c_{s1}^{i*}\right)\left(R_s\left(e - x^*\right) - c_{s2}^{i*}\right)\right] = E_{\Pi_s}\left[f(i) g(i)\right].$$

Since the functions are strictly decreasing, the Chebyshev Sum Inequality[8] gives us

$$E_{\Pi_s}\left[f(i) g(i)\right] > E_{\Pi_s}\left[f(i)\right] E_{\Pi_s}\left[g(i)\right].$$

So, the theorem follows from market clearing:

$$E_{\Pi_s}\left[g(i)\right] = \sum_{i=1}^{I} \pi_s^i \left(R_s\left(e - x^*\right) - c_{s2}^{i*}\right) = 0.$$

Intuitively, since $w^i \beta^i \rho^i u'\left(c_{s1}^{i*}\right)$ and $R_s\left(e - x^*\right) - c_{s2}^{i*}$ have the same order, they have positive covariance. Since the expectation of $R_s\left(e - x^*\right) - c_{s2}^{i*}$ is 0 by market clearing, positive covariance implies the expectation of the product is positive:

$$\begin{aligned} 0 &< cov_{\Pi_s}\left[f(i), g(i)\right] = E_{\Pi_s}\left[f(i) g(i)\right] - E_{\Pi_s}\left[f(i)\right] E_{\Pi_s}\left[g(i)\right] \\ &= E_{\Pi_s}\left[f(i) g(i)\right]. \end{aligned}$$

$\square$

---

[8] To prove the Chebyshev Sum Inequality, let i be a random variable, and let f and g be either both increasing or decreasing. If i' is an i.i.d. copy of i, then we must always have $\left(f(i) - f(i')\right)\left(g(i) - g(i')\right) \geq 0$ for any realization of i, i'. Taking expectation, we get $E\left[f(i) g(i)\right] - E\left[f(i)\right] E\left[g(i')\right] - E\left[f(i')\right] E\left[g(i)\right] + E\left[f(i')g(i')\right] \geq 0$, which yields the weak version of the inequality. When f and g are strictly decreasing and i is non-constant (as in our setting), the inequality is strict.

A.4 Proof of Theorem 2

Given two observations, the proof of Theorem 1 can be easily adapted to prove Theorem 2. First, the Date 0 Lemma (Lemma 1) still holds with fire sales. To see this, let $q_s(x)$ denote the conditional on x equilibrium price without fire sales, and let $\widetilde{q}_s(x)$ denote the equilibrium price with fire sales. If $1/q_s > R_s/r_s$, then once fire sales are allowed agents will choose $L_s^i = e - x$. In this case, markets will not clear in $t = 2$. So, if $1/q_s(x) \geq R_s/r_s$, then $\widetilde{q}_s(x) = r_s/R_s$ and agents are indifferent to the level of L_s^i. If $1/q_s < R_s/r_s$, then income is maximized by choosing $L_s^i = 0$, and $q_s(x) = \widetilde{q}_s(x)$. From the perspective of the price-taking agents, in equilibrium the economy is as if there were no fire sales. Therefore, since there is one state without fire sales, we will have $x^* \in (0, e)$ (by the argument for Proposition 1) and the x FOC (5) will hold in equilibrium.

Second, even though $\widetilde{q}_s(x)$ is non-differentiable at x if $1/q_s(x) = R_s/r_s$, $\widetilde{q}_s(x)$ is always right differentiable in equilibrium. In states where $1/q_s < R_s/r_s$, $L_s^i = 0$ binds and DARA ensures the interest rate falls as x rises, which only makes $L_s^i = 0$ bind more. If $1/q_s > R_s/r_s$, then the interest rate $1/\widetilde{q}_s(x)$ is stuck at R_s/r_s, even with a little more liquidity (since $q_s(x)$ is continuous). If $1/q_s = R_s/r_s$ and agents are exactly indifferent to using fires or not, then increasing x pushes down the interest rate and makes them strictly prefer to not liquidate any of the long-term asset.

We can now sign the ex ante utility impact of a small increase in x (at x^*) in the fire sales economy. Holding prices constant, the state-by-state reallocation of endowments has no impact on utility by the Date 0 Lemma (Lemma 1). In the strict fire sale states, the interest rate is constant and there is no price effect on utility. In the strict and indifferent $L_s^i = 0$ states (there is at least one by assumption), it is as if there were no fire sales, and the proof of Theorem 1 goes through.   □

A.5 Statement and Proof of Theorem 3

Theorem 3 *Suppose (i) u satisfies DARA and (ii) if $\beta^i \leq \beta^j$ then $w^i \leq w^j$. Then competitive equilibrium is ex ante constrained inefficient. In particular, all investors could be made better off by collectively allocating more to the long-term asset:*

$$\frac{d}{dx} V(x, \mathbf{q}(x)) \bigg|_{x=x^*} < 0.$$

Proof Using $\beta^i \rho^i = 1$ and the fact that $\left(w^i u'\left(c_{s1}^{i*} \right) \right)^{i \in I}$ and $\left(R_s(e - x^*) - c_{s2}^{i*} \right)^{i \in I}$ are, respectively, strictly increasing and decreasing [Part 1 of Proposition 2 and premise (ii)], Theorem 3 follows from the proof of Theorem 1.   □

A.6 Statement and Proof of Theorem 4

From the proofs of Theorems 1 and 3, we can now see that there is either over or underinvestment unless $(w^1, \ldots, w^I)$ is such that

$$\sum_{s=1}^{S} P_s q_s' \left(x^*\right) \sum_{i=1}^{I} \pi_s^i w^i u' \left(c_{s1}^{i*}\right) \left(R_s \left(e - x^*\right) - c_{s2}^{i*}\right) = 0, \tag{8}$$

which seems like a knife-edge case. This insight gives rise to the following theorem, which takes up the classical case of no aggregate uncertainty about the fraction of agents who are impatient. In this case, there is a unique equilibrium, which is generically constrained inefficient (though we do not say whether because of overinvestment or because of underinvestment).

Theorem 4 *Suppose there is no aggregate uncertainty, $S = 1$. Suppose u satisfies DARA and is three times continuously differentiable. Then there is a unique equilibrium. Fix arbitrary discounts $0 < \beta^1 < \cdots < \beta^I$, productivity $R > 0$, and the probabilities π_s^i. Then for almost every choice of $(w^1, \ldots, w^I) \in \mathbb{R}_{++}^I$, the unique competitive equilibrium is ex ante constrained inefficient*

Proof By the assumption on u, $q_s(x)$ and consumptions $c_{st}^i(x)$ are twice continuously differentiable functions of $x \in (0, e)$. Hence, we can define the continuously differentiable function $F : (0, e) \times \mathbb{R}_{++}^I \to \mathbb{R}^2$ by

$$F(x, w) = (F_1(x, w), F_2(x, w)),$$

where

$$F_1(x, w) = \sum_{s=1}^{S} P_s \sum_{i=1}^{I} \pi_s^i w^i u' \left(c_{s1}^i (x)\right) (1 - Rq_s(x))$$

$$F_2(x, w) = \sum_{s=1}^{S} P_s \left(\frac{\partial}{\partial x} q_s (x)\right) \sum_{i=1}^{I} \pi_s^i w^i u' \left(c_{s1}^i (x)\right) [R(e - x) - c_{s2}^i (x)],$$

and $w = (w^1, \ldots, w^I)$. Equilibrium occurs when $F_1(x, w) = 0$ and constrained inefficiency when $F_2(x, w) \neq 0$ also.

First, we show that whenever $F_1(x, w) = 0$, the derivative $\partial F_1(x, w)/\partial x < 0$. The reason is that $\partial q_s(x)/\partial x > 0$ (by Proposition 2), and with no aggregate uncertainty, $1 - Rq_s(x) = 0$. It follows that there must be a unique equilibrium, since any function like $F_1(x|w)$ that always crosses the x-axis with negative slope can only cross once.

We now demonstrate that whenever the two-dimensional function $F(x, w) = 0$, the derivative matrix $D_{(x,w)}F$ has column rank 2. Note first that the most impatient agent 1 must always be a seller of good 2, hence $[R(e - x) - c_{s2}^i (x)] > 0$ for all s. It follows that $\partial F_2(x, w)/\partial w^1 > 0$. Moreover, when $1 - Rq_s(x) = 0$ for all s (as must happen when there is one future aggregate state), $\partial F_1(x, w)/\partial w^1 = 0$. We already saw that $\partial F_1(x, w)/\partial x < 0$. Hence,

$$D_{x,w^1} F = \begin{bmatrix} - & 0 \\ ? & + \end{bmatrix}$$

has full rank.

This rank argument shows that the two-dimensional function F is transverse to 0, $F \pitchfork 0$, that is whenever $F(x, w) = 0$, the derivative matrix $D_{(x,w)}F$ has column rank 2. Then since the functions F are continuously differentiable, by the transversality theorem, for almost all $w \in \mathbb{R}_{++}^I$, we have $F_w \pitchfork 0$, where $F_w(x) \equiv F(x, w)$. But $D_x F_w(x)$ is a 2×1 matrix and thus can never have rank 2. Hence, for almost all w, $F_w(x)$ is never 0, which means that whenever the top expression F_1 is 0, the bottom expression F_2 is nonzero. It then follows that for almost all w, every equilibrium is constrained inefficient. $\qquad\square$

References

Allen, F., Gale, D.: Financial intermediaries and markets. Econometrica **72**(4), 1023–1061 (2004)

Arrow, K.J: Aspects of the theory of risk-bearing. In: Yrjö Jahnsson Lectures, Yrjö Jahnssonin Säätiö, Helsinki (1965)

Bhattacharya, S., Gale, D.: Preference shocks, liquidity and central bank policy. In: Barnett, W.A., Singleton, K.J. (eds.) New Approaches to Monetary Economics. Cambridge University Press, New York (1987)

Diamond, D.W., Dybvig, P.H.: Bank runs, deposit insurance, and liquidity. J. Political Econ. **91**(3), 401–419 (1983)

Ennis, H.M., Keister, T.: Optimal banking contracts and financial fragility. Econ. Theory **61**, 335–363 (2016)

Fairchild, F.R., Furniss, E.S., Buck, N.S.: Economics. The Macmillan Company, London (1937)

Farhi, E., Golosov, M., Tsyvinski, A.: A theory of liquidity and regulation of financial intermediation. Rev. Econ. Stud. **76**, 973–992 (2009)

Fostel, A., Geanakoplos, J.: Collateral restrictions and liquidity under-supply: a simple model. Econ. Theory **35**, 441–467 (2008)

Geanakoplos, J., Polemarchakis, H.M.: Existence, regularity, and constrained suboptimality of competitive allocations when markets are incomplete. In: Heller, W.P., Starr, R.M., Starrett, D.A. (eds.) Uncertainty, Information, and Communication, Essays in Honor of Kenneth Arrow, Chap 3, vol. 3, pp. 65–95. Cambridge University Press, Cambridge (1986)

Geanakoplos, J., Walsh, K.J.: Uniqueness and stability of equilibrium in economies with two goods, Cowles Foundation Discussion Paper No. 2050 (2016)

Grochulski, B., Zhang, Y.: Optimal liquidity regulation with shadow banking, Federal Reserve Bank of Richmond Working Paper WP 15-12R (2016)

Hazlett, D.: Deposit insurance and regulation in a Diamond–Dybvig banking model with a risky technology. Econ. Theory **9**, 453–470 (1997)

Jacklin, C.J.: Demand deposits, trading restrictions, and risk sharing. In: Prescott, E.C., Wallace, N. (eds.) Contractual Arrangements for Intertemporal Trade. University of Minnesota Press, Minneapolis (1987)

Kashyap, A.K., Tsomocos, D.P., Vardoulakis, A.P.: How does macroprudential regulation change bank credit supply? NBER Working Paper No. 20165 (2014)

Lin, P.: Banking, incentive constraints, and demand deposit contracts with nonlinear returns. Econ. Theory **8**, 27–39 (1996)

Villamil, A.P.: Demand deposit contracts, suspension of convertibility, and optimal financial intermediation. Econ. Theory **1**, 277–288 (1991)

Yared, P.: Public debt under limited private credit. J. Eur. Econ. Assoc. **11**, 229–245 (2013)