NBER WORKING PAPER SERIES

A WELFARE CRITERION FOR MODELS WITH DISTORTED BELIEFS

Markus K. Brunnermeier
Alp Simsek
Wei Xiong

Working Paper 20691
http://www.nber.org/papers/w20691

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2014

A Welfare Criterion for Models with Distorted Beliefs
Markus K. Brunnermeier, Alp Simsek, and Wei Xiong
NBER Working Paper No. 20691
November 2014, Revised February 2015
JEL No. D61,D62,D84,G12,G14

ABSTRACT

This paper proposes a welfare criterion for economies in which agents have heterogeneously distorted beliefs. Instead of taking a stand on whose belief is correct, our criterion asserts that an allocation is belief-neutral efficient (inefficient) if it is efficient (inefficient) under any convex combination of agents' beliefs. While this criterion gives an incomplete ranking of social allocations, it can identify positive- and negative-sum speculation driven by conflicting beliefs in a broad range of economic environments.

Markus K. Brunnermeier
Princeton University
Department of Economics
Bendheim Center for Finance
Princeton, NJ 08540
and NBER
markus@princeton.edu

Wei Xiong
Princeton University
Department of Economics
Bendheim Center for Finance
Princeton, NJ 08450
and NBER
wxiong@princeton.edu

Alp Simsek
Department of Economics, E17-244
MIT
77 Massachusetts Avenue
Cambridge, Ma. 02139
and NBER
asimsek@mit.edu

1   Introduction

The burgeoning behavioral finance and economics literature has identified a set of psychological biases that distort people's beliefs and decisions in various economic situations.[1] The presence of belief distortions stimulates normative analysis of welfare consequences of belief distortions. A standard approach taken by the literature is to assume that the social planner knows the objective belief measure and uses the objective measure to evaluate agents' welfare.[2] This approach, however, faces a major challenge in implementation—which belief should the planner use? In many realistic situations, the planner does not observe the objective belief and faces the same difficulty as individuals in discriminating different beliefs based on available data. Perhaps due to this challenge, many studies in the behavioral literature shy away from making any normative statement.

This paper proposes a belief-neutral welfare criterion, which requires the planner to be sure of the presence of belief distortions by some agents but without having to precisely identify the objective belief. To illustrate the basic idea, we first consider a bet between Joe Stiglitz and Bob Wilson.[3] One day, Joe and Bob argued over the contents of a pillow. Joe maintained that the pillow had a natural filling, while Bob thought a polyester filling was more likely. Joe assessed with probability 0.9 that the pillow had natural down and Bob assessed the probability at 0.1. They decided to construct a bet as follows: If the pillow had natural down, Bob would pay Joe $100, but if it had artificial down, Joe would pay Bob $100. They could discover the truth only by cutting the pillow open, which would destroy it. They agreed that the winner would replace the pillow at a cost of $50.[4] It is clear that both Joe and Bob preferred the bet relative to no betting at all, as each expected to make a net profit of $35 after deducting the cost of replacing the pillow. This bet was desirable from each individual's perspective, and thus it Pareto dominated no betting under the standard Pareto principle. However, the outcome of the bet was worrisome—it led to a wealth transfer between Joe and Bob and a perfectly good pillow being destroyed.

Joe and Bob might have taken the bet for its entertainment value, which could have

---

[1]See Hirshleifer (2001), Barberis and Thaler (2003), and Della-Vigna (2009) for extensive reviews of the literature.

[2]For example, see Weyl (2007), Sandroni and Squintani (2007), Spinnewijn (2010), Hassan and Mertens (2011), Gennaioli, Shleifer, and Vishny (2011), and Bianchi, Boz, and Mendoza (2012).

[3]See Kreps (2012, p.193) for more details of the story.

[4]We can also make the example more realistic by making the replacement cost of the pillow state dependent, i.e., the cost being $50 if it had natural down and $20 if it had artificial down. Our welfare analysis of the bet is robust to such a state-dependent replacement cost.

justified the cost of destroying the pillow. Another possible motive was that each bettor believed he would win and the other would lose. A planner could simply verify the reason by asking them. If the bet was motivated by a belief in winning, then at least one of them was overconfident, even though it was still difficult to tell who was overconfident. In this case, it is immediately obvious that the bet was a negative-sum game *regardless* of whose belief the planner uses to evaluate the social welfare. The resulting social loss is exactly the destroyed pillow.

In fact, the conflicting beliefs of Joe and Bob induced a form of externality. Despite knowing the bet would lead to the pillow being destroyed, each believed that he would win and the other party would lose. In this setting with conflicting beliefs, the meaning of "externality" needs to be broadened.[5] From Bob's perspective, his action causes an externality on Joe, even though Joe does not see it this way. From Joe's perspective, there is an analogous externality. The standard libertarian view does not restrict the choice of any individual if it does not cause a negative externality on others. We modify/extend this libertarian viewpoint to a setting with heterogeneous beliefs. In this setting, externality has to be evaluated under the belief of the individual whose choice causes the externality rather than under the belief of the individual who is exposed to the externality. In our pillow example, under the belief of either Joe or Bob, the negative externality on the other even exceeds his own gain. This negative sum serves as the basis for our welfare criterion.

To generalize the key insight of this example, we acknowledge the relevance of a set of reasonable beliefs and require efficiency to be robust across all of the reasonable beliefs. Our welfare criterion asserts a social choice to be *belief-neutral* (in)efficient if and only if it is (in)efficient under every reasonable belief. A key presumption of this criterion is that the planner is sure that some agents' beliefs are distorted.[6] Specifically, we accept all convex combinations of agents' beliefs as reasonable beliefs, as long as they are consistent with the commonly agreed upon aggregate statistics. We propose to use all of them to extend the two standard welfare analysis approaches—the expected social welfare approach and the Pareto efficiency approach.

---

[5] One might disagree about this broadened use of the term externality—in fact, one of us does so. Alternatively, one could coin a new term for a situation in which one benefits at the expense of others under one's own beliefs but not under the other's. In related work, Nielsen (2009b) have also used the word "externality" to describe a similar phenomenon.

[6] One can use sound non-choice evidence to rule the presence of belief distortions in observed choices. See the recent contributions of Bernheim and Rangel (2007, 2009) and Koszegi and Rabin (2007) along this line.

The expected social welfare approach directly compares two social allocations x and y for a given welfare function. Our welfare criterion posits that y is belief-neutral inferior to x if the expected total welfare from y is lower than that from x under *every* convex combination of the agents' beliefs. Let's go back to the bet between Joe and Bob. Suppose that the planner is sure that the bet was induced by belief distortions and that Joe and Bob were both risk neutral. If the social planner assigns Joe and Bob equal weight in summing up their utilities in the social welfare function, it is clear that the bet is belief-neutral inferior to the status quo (no betting). This is because, regardless of which reasonable belief the social planner adopts to evaluate Joe's and Bob's expected utilities, the transfer of $100 between them has no impact on the expected social welfare, but destroying the pillow leads to a sure social loss of $50.

Without relying on a particular social welfare function, we can also adopt the Pareto dominance approach. Our criterion asserts that an allocation y is belief-neutral Pareto inefficient if, under every reasonable belief, there exists an alternative allocation y' that improves the expected utilities of some agents without hurting anyone else. Returning to the example, suppose that the planner adopts Joe's belief. Under this belief, the bet leads to an expected wealth transfer of $80 from Bob to Joe and the pillow's destruction. Alternatively, a direct transfer of $80 from Bob to Joe without the bet improves everyone's expected utility by saving them the cost of replacing the pillow. Similarly, under every convex combination of Joe's and Bob's beliefs, the planner can find a suitable (belief-measure dependent) transfer without the bet to strictly improve everyone's expected utility. Thus, the bet is belief-neutral inefficient with respect to any social welfare function that increases with agents' utilities.

In summary, without taking a stand on which belief was correct, the planner can categorically determine that the bet leads to an inefficient social outcome. The key is that the externality induced by the conflicting beliefs of Joe and Bob is uniformly negative under every reasonable belief. Of course, this feature may not always hold in a more general situation. For illustration, let us extend the bet. Suppose that Bob believed the pillow contained poisonous materials (instead of polyester) with 90% probability and that there is a social gain of $100 from removing a poisonous pillow from the public (instead of the $50 cost of replacing the pillow). The bet had a positive sum under Bob's belief but still had a negative sum under Joe's belief. Thus, it is neither belief-neutral efficient nor belief-neutral inefficient.

Despite its incompleteness, our belief-neutral criterion is able to identify negative and

positive externalities from belief distortions in a variety of settings. Section 3 illustrates the potential applications using a series of examples, some of which are simplified versions of prominent models in the literature.

The first example extends and generalizes the pillow example in three ways. First, we make agents risk averse. Trading induced by heterogeneous beliefs makes agents' consumption more volatile than their endowments and results in a negative-sum game in expected utility terms regardless of the planner's belief. Second, we add a hedging motive to the speculative motive, by making agents' endowments negatively correlated. In this case, a trade-off arises between the welfare gain from risk sharing and the welfare loss from speculative trading (see, e.g., Kubler and Schmedders (2012), Simsek (2013a), and Posner and Weyl (2013)). Third, we also allow for endogenous information acquisition along the lines of Grossman and Stiglitz (1980). In a fully rational model with homogenous prior beliefs, agents' incentive to collect information is subdued since trading reveals part of the information to other traders for free. This positive externality leads to an under-provision of information that enhances socially beneficial hedging. Heterogeneous beliefs can provide a counterforce to this externality as the speculative motive provides additional incentive to acquire information, as informally discussed in Black (1986). Our criterion enables us to compare the hedging and information benefits with the costs of speculation due to belief heterogeneity.

Our second example investigates whether agents adequately *self-insure*, e.g., whether motorists wear seat belts as a precaution for traffic accidents or banks retain capital as a precaution for financial losses. Suppose agents are optimistic about their own idiosyncratic risks although they know and agree on the aggregate risks, e.g., each motorist knows the average accident probability but believes the accidents will happen to other motorists. In this setting, each agent chooses not to self-insure, even though she recognizes the benefit of self-insurance for the average agent. While this example does not involve any negative externality, the restriction that agents agree on aggregate risks nevertheless makes it possible for our criterion to evaluate mandatory self-insurance policies such as seat-belt laws or bank capital requirements.

The third and fourth examples involve speculative bubbles. A number of recent studies (e.g., Harrison and Kreps (1978), Morris (1996), Scheinkman and Xiong (2003), Wu and Guo (2004), Hong, Scheinkman, and Xiong (2006), and Hong and Sraer (2011)) emphasize that

the option to resell assets to future optimists can induce bubbles in asset prices. Our first example highlights how overinvestment induced by price bubbles makes speculative trading a negative-sum game just like the bet between Joe and Bob (e.g., Bolton, Scheinkman, and Xiong (2006), Gilchrist, Himmelberg, and Huberman (2005), and Panageas (2006)). Our next example highlights that bubbles caused by heterogenous beliefs can help overcome market breakdowns induced by the adverse-selection problem in lemons models (as in Akerlof (1970)) and thus lead to a positive-sum game. Our criterion can also identify the consequent belief-neutral welfare gains.

Our fifth example builds on leverage cycles caused by heterogeneous beliefs (e.g., Geanakoplos (2003, 2009), Fostel and Geanakoplos (2008), Simsek (2013b), Cao (2010), and He and Xiong (2012)). In these models, binding collateral constraints force optimistic asset owners to liquidate positions. The liquidation costs associated with forced selling make the initial leveraged asset acquisition a negative-sum game. Our criterion provides a tool to analyze welfare implications of such leverage cycles.

The last example illustrates consumption/savings distortions induced by heterogeneous beliefs in macroeconomic models, e.g., Sims (2008). In production-economy settings, trading between them not only makes their consumption excessively volatile, but also induces them to save either too much or too little relative to homogeneous-economy benchmarks. The consequent distortion in aggregate investment again leads to a negative-sum game, which our criterion can identify.

Economists have long recognized that the standard Pareto criterion can lead to unappealing outcomes when agents hold conflicting beliefs. Early general equilibrium literature, e.g., von Weizsäcker (1969), Dreze (1970), Starr (1973), Harris (1978), and Hammond (1981), noted that an allocation that is Pareto optimal in the usual sense might feature less-than-perfect risk sharing. In particular, it made a distinction between ex ante efficiency and various versions of ex post efficiency (with better risk-sharing properties). In recent work, Nielsen (2003, 2009a-b) have utilized the ex post welfare criterion to investigate optimal exchange rate and social security policies in environments with belief disagreements. Other than using a different welfare criterion, these papers assume that agents agree on holding rational beliefs in the sense of Kurz (1994). In contrast, in most of our analysis we do not place rationality restrictions on agents' beliefs, except for the self-insurance example in Section 3.2, where we assume agents know and agree on the aggregate risks.

5

The independent decision theory literature, e.g., Mongin (1997), Gilboa, Samet, and Schmeidler (2004), has also pointed out that the standard Pareto principle can be spurious when agents hold conflicting beliefs. Our contribution to these studies is to propose a belief-neutral criterion, which circumvents the spurious unanimity problem under the premise that the planner is aware of the presence of belief distortions but unaware of the objective belief. In parallel work, Gilboa, Samuelson, and Schmeidler (2012) have proposed an alternative criterion, which we compare with our criterion in Section 2.3.3.

Another strand of the literature, e.g., Stiglitz (1989), Summers and Summers (1989), and, more recently, Davila (2014), has emphasized the negative-sum nature of speculation in financial markets to make a case for a financial transaction tax. Our criterion and examples capture the costs as well as the benefits of speculation, therefore providing a framework for analyzing how speculative activities in financial markets should be regulated.

The paper is organized as follows: Section 2 describes the welfare criterion in a generic setting. Section 3 provides a series of examples to demonstrate the capability of the criterion to generate clear welfare ranking in a variety of environments with distorted beliefs. We conclude in Section 4. The technical proofs are provided in the appendix.

2　The Welfare Criterion

We introduce the welfare criterion in a generic setting with N agents, indexed by $i \in \{1, 2, ..., N\}$ and with $T + 1$ dates: $t = 0, 1, ..., T$. At time t, let s_t be the state of the economy, which summarizes the history of the economy up to t. Agent i's conditional transition probability at time t from state s_t to state s'_{t+1} at $t+1$ is $\pi_t^i \left(s_t, s'_{t+1} \right) \geq 0$. The subscript t indicates that beliefs can be time-varying, and superscript i indicates that beliefs are potentially heterogeneous across agents. We summarize agent i's beliefs by $\Pi^i = \left\{ \pi_t^i \left(s_t, s'_{t+1} \right) \right\}_{t=0}^{T-1}$. We assume that agents consume only at the final time T. A social choice x represents a set of consumption allocations to all agents along the path of s_T: $x = \{ x^i \left(s_T \right) \}$. Note that $x^i \left(s_T \right)$ is a vector of consumption to agent i. A feasible allocation satisfies the aggregate budget constraint at each final state.

Suppose that agent i has state-dependent utility function $u_i \left[s_T, x^i \left(s_T \right) \right]$ over the consumption stream $x^i \left(s_T \right)$. This function is strictly increasing and locally concave with respect to consumption. This utility specification is sufficiently general to capture the standard utility functions used in most economic models. Based on the utility specification and the agent's

beliefs, his expected utility at time 0 is $E_0^i[u_i[s_T, x_T^i(s_T)]]$, where the superscript i denotes the expectation under agent i's beliefs. By building on expected utilities, our framework ignores preferences that feature ambiguity aversion.[7]

2.1   Heterogeneous Beliefs

We let agents hold different beliefs (i.e., $\Pi^i \neq \Pi^{i'}$ if $i \neq i'$) and assume the beliefs are common knowledge among the agents. Before we dive into welfare analysis, it is useful to sort out different sources of heterogeneous beliefs. Throughout our later analysis, we treat agents' beliefs as given. It is straightforward to think of the beliefs as outcomes of the agents' learning processes. Suppose that an unobservable variable π determines the probability of the tree moving up each period. Each agent has a prior belief about the distribution of π, observes some information about π in each period, and uses Bayes' rule to update his belief about π. Through this learning process, three sources may lead to heterogeneous beliefs among agents: 1) distortions in updating, 2) different information, and 3) different prior beliefs.

We emphasize distortions in updating as a key source of heterogeneous beliefs. A large branch of the academic literature highlights that people suffer from a range of well-established psychological biases, such as overconfidence, limited attention, representativeness bias, and conservatism in making financial decisions. See Hirshleifer (2001), Barberis and Thaler (2003), and Della-Vigna (2009) for extensive reviews of the literature. These biases cause agents to react differently to information. In particular, overconfidence causes agents to exaggerate the precision of certain noisy signals and thus overreact to the signals. When agents overreact to different signals, they may end up with substantially different beliefs and, as a result, may speculate against each other.

The presence of belief distortions prompts welfare concerns. Some agents may be unaware of their belief distortions and, as a result, take actions that hurt their own and others' welfare. Thus, it is important that a social planner evaluates each agent's welfare by using the objective probability measure, which serves as the premise of our welfare criterion.

A second source of belief differences is asymmetric information. The well-known no-trade theorem (e.g., Aumann (1976), Milgrom and Stokey (1982) and Sebenius and Geanakoplos (1983)) shows that asymmetric information cannot cause rational agents with a common

---

[7]Our later examples all use Markovian state structures and consumption at the final date T.

prior belief to hold common knowledge of heterogeneous beliefs or to trade with each other. This result motivates us to mostly ignore asymmetric information in our analysis, except in our example considered in Section 3.4.

A third source of belief differences is heterogeneous prior beliefs. The decision theory literature that builds on Savage's (1954) notion of subjective probability treats beliefs separately for individual agents. As economics does not offer much guidance on how individuals form their prior beliefs, economists tend to agree that prior beliefs probably depend on an individual's background and experience. Morris (1995) summarizes a series of arguments to advocate the view that rational agents may hold heterogeneous prior beliefs, just like heterogeneous risk preferences. In Brunnermeier and Parker (2005) and Brunnermeier, Gollier, and Parker (2007), heterogeneous prior beliefs arise endogenously from agents' anticipatory utility. In our analysis, we abstract away from agents' heterogeneous priors and instead focus on heterogeneous beliefs caused by distortions in updating.

2.2   Welfare Analysis with Distorted Beliefs

In the presence of distorted beliefs, it is important that the social planner uses an objective probability measure to evaluate agents' expected utilities in the welfare analysis. The challenge here is that the social planner may not observe the probability that drives economic uncertainty. Given the agents' different belief measures, whose measure is appropriate for welfare analysis? Is there an even more appropriate one outside of those used by the agents? We now introduce a belief-neutral welfare criterion.

Without taking a stand on which agent's belief is correct, we allow the planner to consider every belief from a set of reasonable beliefs. In our baseline analysis, this set contains all convex combinations of the agents' beliefs. Denote Π^h to be a convex combination of the agents' beliefs with weight $h = \{h^1, ..., h^N\}$:

$$\Pi^h = \sum_i h^i \Pi^i, \text{ where } h^i \geq 0 \text{ and } \sum h^i = 1.$$

In settings with multiple events, we further restrict the set of reasonable beliefs to satisfy commonly agreed upon aggregate statistics.[8] As we discuss in Section 2.3.3, our criterion can be generalized to use alternative specifications for the set of reasonable beliefs.

The key contribution of our welfare criterion is that it allows for analysis of the efficiency of a social allocation according to all reasonable beliefs. Specifically, we propose to identify

---

[8] The example in Section 3.2 clarifies this aggregate statistics restriction on reasonable beliefs.

an allocation as inefficient (or efficient) if the social planner finds it inefficient (or efficient) under every reasonable probability measure Π^h that is commonly used to evaluate all agents' expected utilities. We can use two different approaches to implement our proposal, one based on a given social welfare function and the other through the notion of Pareto efficiency. As is well known from standard economic theory, in the absence of belief distortions these two approaches are internally consistent. In particular, any Pareto-efficient social allocation corresponds to an optimal allocation that maximizes the agents' aggregate expected utilities under a set of nonnegative weights.

2.2.1 Belief-neutral Social Welfare Criterion

The Bergsonian social welfare function is a sum of agents' expected utilities $\left\{ E_0^h \left[u_i \right] \right\}$ (calculated according to a common measure Π^h) using a set of nonnegative weights $\{\lambda_i\}$:

$$W \left(E_0^h \left[u_1 \right], E_0^h \left[u_2 \right], ..., E_0^h \left[u_N \right] \right) = \sum_{i=1}^{N} \lambda_i E_0^h \left[u_i \right] = E_0^h \left[\sum_{i=1}^{N} \lambda_i u_i \right].$$

If the weights are all equal, it becomes the utilitarian social welfare function:[9]

$$W \left(E_0^h \left[u_1 \right], E_0^h \left[u_2 \right], ..., E_0^h \left[u_N \right] \right) = \sum_{i=1}^{N} E_0^h \left[u_i \right] = E_0^h \left[\sum_{i=1}^{N} u_i \right].$$

Based on a given welfare function, we can implement our criterion as follows.

Definition 1 *Consider two social allocations, x and y. If the expected social welfare of allocation x dominates that of allocation y for every reasonable probability measure Π^h,*

$$W \left(E_0^h \left[u_1 \left(s_T, x_T^1 \left(s_T \right) \right) \right], ..., E_0^h \left[u_N \left(s_T, x_T^N \left(s_T \right) \right) \right] \right)$$
$$\geq W \left(E_0^h \left[u_1 \left(s_T, y_T^1 \left(s_T \right) \right) \right], ..., E_0^h \left[u_N \left(s_T, y_T^N \left(s_T \right) \right) \right] \right)$$

with the inequality holding strictly for at least one reasonable measure, then allocation x is **belief-neutral superior** *to allocation y.*

To establish the superiority of one social allocation relative to another, a higher expected social welfare according to every convex combination of the agents' beliefs is required. This

---

[9]Given that these social welfare functions are linear and that the social planner uses the same probability measure to evaluate the expected utilities of all agents, the expected social welfare is independent of the order of aggregating welfare and computing expectations. In our analysis, we find it more convenient to first aggregate agents' welfare in each of the final states and then compare the expected social welfare under different probability measures.

proposed belief-neutral superiority is a partial ordering of social allocations. Given two social allocations x and y, the allocation x might dominate y in one measure and y might dominate x in another measure. In such cases, we would say that x and y are incomparable.

Despite its incompleteness, this criterion is nevertheless useful in detecting negative-sum speculation driven by distorted beliefs. We now apply this criterion to analyze the bet between Joe and Bob described in the introduction. Suppose that both Joe and Bob are risk neutral, $u_{Joe}(w) = w$ and $u_{Bob}(w) = w$, and that the social planner uses the utilitarian social welfare function for a reasonable belief:

$$W\left(E_0^h\left[u_{Joe}\right], E_0^h\left[u_{Bob}\right]\right) = E_0^h\left[w_{Joe} + w_{Bob}\right] = w_{Joe} + w_{Bob}.$$

It is obvious that without any betting, regardless of the probability measure the social planner adopts, the social welfare is simply the sum of Joe's and Bob's initial wealth. The bet causes a transfer of $100 between them and the pillow's destruction. The money transfer has no impact on the social welfare regardless of its direction or the probability measure the social planner adopts to evaluate the welfare. However, replacing the pillow incurs a sure cost of $50 and therefore makes the bet a negative-sum game for every reasonable, common probability measure used to evaluate Joe's and Bob's expected utilities. Thus, the status quo allocation is belief-neutral superior to the bet.

The utilitarian social welfare function assigns equal weights to all agents. If the social welfare function puts a sufficiently high weight on one agent, say, Joe, then we cannot directly compare the two allocations x and y. This is because, under Joe's belief, the bet increases his own expected utility and thus the social welfare relative to the status quo allocation. However, this may not be the case under Bob's belief. The second version of our criterion addresses this concern by generalizing the notion of Pareto efficiency and establishes that the bet is belief-neutral inefficient regardless of the choice of social welfare function.

2.2.2 Belief-neutral Pareto Efficiency

The essence of Pareto efficiency is to determine whether there exists an alternative feasible allocation that improves the welfare (i.e., expected utility) of some agents without hurting any other agent. If such an alternative exists, the allocation under evaluation is Pareto inefficient. We next generalize this logic to environments with distorted beliefs to obtain a second implementation of our criterion.

10

Definition 2 *Consider a social allocation y. Suppose that, for every reasonable probability measure Π^h, there exists another (measure dependent) allocation y^h such that it improves some agents' expected utilities without reducing anyone's, i.e., $\forall i$, $E_0^h\left[u_i\left(s_T, y_T^i\left(s_T\right)\right)\right] \leq E_0^h\left[u_i\left(s_T, y_T^{hi}\left(s_T\right)\right)\right]$ with the inequality holding strictly for at least one agent. In that case, allocation y is **belief-neutral Pareto inefficient**. In contrast, if, for every Π^h, there does not exist a dominating alternative, then allocation y is **belief-neutral Pareto efficient**.*

If agents have common beliefs, i.e., if Π^i is the same for each i, then the belief-neutral Pareto criterion coincides with the usual Pareto criterion. In the presence of distorted beliefs, as we discussed before, the social planner uses a common probability measure from the set of reasonable measures to evaluate each agent's expected utility. The belief-neutral criterion then identifies an allocation as inefficient (or efficient) if it is Pareto inefficient (or efficient) under every reasonable measure.

Returning again to the bet between Joe and Bob, we can show that the betting allocation, denoted by y, is belief-neutral Pareto inefficient. Under Joe's belief, y is dominated by an alternative allocation, y^h, which keeps the pillow intact and simply transfers $80 from Bob to Joe. This allocation improves both Joe's and Bob's expected utilities, under Joe's belief. Similarly, under Bob's belief, y is dominated by an alternative allocation that keeps the pillow intact and transfers $80 from Joe to Bob. More generally, under every convex combination of Joe's and Bob's beliefs, there exists an appropriate direct transfer that improves the expected utilities of both Joe and Bob. The gain from such a transfer is due to saving the pillow from destruction. The bet is thus belief-neutral Pareto inefficient. It is also easy to see that the status quo allocation is belief-neutral efficient, as for every reasonable belief the planner cannot find a transfer to improve Joe's or Bob's welfare without hurting the other's. Taken together, the status quo allocation is on the belief-neutral Pareto-efficient frontier while the betting allocation is in the belief-neutral inefficient set.

Recall from the standard welfare theory (e.g., Mas-Colell et al., 1995, Proposition 16.E.2) that each allocation on the Pareto frontier maximizes a linear social welfare function corresponding to some Pareto weights. This observation leads to the following result, which states belief-neutral Pareto inefficiency in terms of social welfare maximization.

Proposition 1 *Let X denote the set of all feasible allocations. Accordingly, an allocation, $x \in X$, is **belief-neutral Pareto efficient (inefficient)** if and only if, for every reasonable*

probability measure Π^h, *there exists (does not exist) a set of Pareto weights* $\left\{\lambda_i^h\right\}$ *(with* $\lambda_i^h > 0$ *for all* i *and* $\sum_i \lambda_i^h = 1$*) such that*

$$x \in \arg\max_{\hat{x} \in X} \sum_{i=1}^{N} \lambda_i^h E_0^h \left[u_i \left(s_T, \hat{x}_T^i \left(s_T\right)\right)\right].$$

Proposition 1 illustrates the relationship between the two versions of our criterion. Both versions consider *all reasonable beliefs* (i.e., convex combinations of agents' beliefs), which is the key characteristic of our approach. However, the welfare-function-based criterion fixes a particular social welfare function (e.g., a particular set of Pareto weights). By doing so, it enables us to compare allocations directly, e.g., to say that the status quo allocation, x, is belief-neutral superior to the betting allocation, y. In contrast, the Pareto-efficiency version is more general because it considers not only all reasonable beliefs, but also all social welfare functions (e.g., all possible Pareto weights). The cost of this generality is that the criterion does not provide direct comparisons between two allocations. Rather, it *categorizes allocations* into three sets: 1) those that are belief-neutral inefficient because they are inferior under every reasonable belief and every welfare function, 2) those that are belief-neutral efficient because under every reasonable belief they are superior at least according to one welfare function, and 3) those that are neither uniformly efficient nor uniformly inefficient across all reasonable beliefs.

2.3 Comments on the Criterion

2.3.1 Incompleteness

Our belief-neutral criterion requires the externality induced by agents' conflicting beliefs to be uniformly positive or negative across the set of reasonable beliefs. This requirement is demanding and may lead to an incomplete ranking in some situations. To illustrate this incompleteness, we will extend the bet between Joe and Bob. Suppose that Joe believed that the pillow was made of cotton with 90% probability as before, while Bob believed that the pillow contained poisonous materials with 90% probability. Again, they had to cut open the pillow to find out its content. If Joe was right, he would win \$100 from Bob and pay \$50 to replace the pillow. If Bob was right, he would win \$100 from Joe. In addition, by removing the poisonous pillow from the public, there is a social benefit of \$100, which we assume goes to Bob. For the sake of our discussion, we implement our criterion by using only a utilitarian social welfare function. It is easy to see that if the planner uses Joe's belief

12

to evaluate the social welfare, the bet induces a negative sum. However, if the planner uses Bob's belief, the bet induces a positive sum due to the reward for discovering a poisonous pillow. Taken together, the bet is neither belief-neutral superior nor belief-neutral inferior relative to the status quo. This incomplete ranking reflects the belief-dependent cost and benefit of the bet in this case.

Furthermore, in this extended example, the sets of belief-neutral efficient and belief-neutral inefficient allocations are both empty. Given that these sets can be empty, our criterion instructs the planner to 1) choose a belief-neutral efficient allocation if it exists, 2) avoid a belief-neutral inefficient allocation if there is any, and 3) otherwise avoid any market intervention.

2.3.2   Collective and Cautious Paternalism

As our criterion ignores agents' preferences under their own beliefs, it is naturally paternalistic. That said, the criterion features a specific and disciplined form of paternalism—which might be called collective paternalism—as it is designed to detect inefficiencies (or efficiencies) based on disagreements between agents in a group setting.

More specifically, our criterion is not designed to analyze the inefficiencies induced by irrational behavior of an individual agent. Consider an agent who invests a large fraction of her wealth in a particular company's stock. This investment decision may appear inefficient to a conscientious observer who holds a more neutral view of the company's stock than the agent and who thus believes the agent should diversify her investment away from the company. However, the decision is optimal under the agent's beliefs. Without taking a stand on the beliefs of the agent and the observer, our criterion cannot identify the agent's investment decision as efficient or inefficient in isolation.

On the other hand, in an equilibrium context, when one group of agents holds different beliefs than another group due to belief distortions, the trading between the two groups can make their consumption excessively volatile (as we will discuss in Section 3.1). Our criterion can identify the negative sum in expected utilities induced by trading without ruling a particular group's choice as inefficient.[10]

---

[10]In this context, it is also useful to contrast our criterion with that offered by Bewley. In Bewley (2002), an individual decision maker holds several belief distributions and overcomes inertia only if the new choice dominates the status quo under all belief distributions. Bewley's theory shares our feature of belief neutrality, but analyzes a single agent's decision problem rather than evaluating the welfare of many decision makers, each with a different (but single) belief distribution.

Even in group settings, our criterion is further disciplined by the requirement that the planner knows agents' beliefs are distorted. In practice, the planner might rule the presence of belief distortions using evidence from psychology or neuroscience (along the lines of Bernheim and Rangel (2007, 2009) and Koszegi and Rabin (2008)). Alternatively, the planner can account for other reasons for trade and obtain belief distortions as a reasonable residual. For example, consider the evidence suggesting that individuals invest considerably in their own companies or in professionally close stocks and that they tend to lose money on these investments (see Doskeland and Hvide (2011)). In this context, risk-sharing can be ruled out since it would require investment in the opposite direction. Trade based on individual investors' informational advantage can also be ruled out since it would result in realized gains on average as opposed to losses. Hence, using this type of reasoning, a planner can conclude observed trades in this context are likely to be based on belief distortions.

2.3.3 Set of Reasonable Beliefs

In our baseline analysis, we take the set of reasonable beliefs as those that correspond to convex combinations of agents' beliefs, subject to the aggregate statistics commonly agreed by the agents. We view this as a reasonable benchmark for two reasons. First, this set is sufficiently large to include all of the extreme beliefs held by any agent in a given environment. As illustrated by the examples in the next section, this set includes the beliefs of the optimists who bid up asset prices and who take highly leveraged positions, as well as the beliefs of pessimists who are constrained by short-sales restrictions from directly participating in asset markets. Second, the set of convex combinations is also appropriate given our focus on detecting the inefficiencies due to belief disagreements as opposed to belief mistakes shared by all agents. If the objective belief is outside our reasonable set, then there might be some welfare losses—due to an irrational behavior common to all agents—that go undetected by our criterion. But our criterion will be useful even in these scenarios to detect part of the welfare losses that stem from negative-sum speculation.

That said, as Definitions 1 and 2 illustrate, our welfare criterion can also be used with more flexible specifications for the set of reasonable beliefs. We envision that, depending on the application, the set of reasonable beliefs can be taken to be larger—or perhaps smaller— than our baseline specification. For instance, if the planner has a priori knowledge that the objective belief is likely to be in a particular set, then reasonable beliefs can be extended

to include that set.[11] Conversely, if the planner has a priori knowledge that certain beliefs do not correspond to the objective belief, then those beliefs could be excluded from the reasonable set even if they are held by some agents.

2.3.4 Improving an Inefficient Allocation

In welfare analysis, the planner is often concerned not only with whether a social allocation is efficient or not, but also with how to improve upon an inefficient allocation. If the planner follows a specific welfare function, the first version of our belief-neutral criterion can be applied to address both of these issues because it can directly rank an allocation y against any alternative y'.

If the planner cannot rely on a specific welfare function, one might wonder whether the second version of our criterion, the belief-neutral Pareto criterion, can address the issue of how to improve upon an inefficient allocation, say, y. This concern arises because, according to Definition 2, an alternative allocation y^h that dominates y may depend on the belief measure Π^h that the planner uses to evaluate y. In other words, the planner has to specify a particular belief Π^h in order to implement a belief-dependent alternative allocation. However, as we will show in several of our examples, the belief-neutral Pareto-efficient frontier is often non-empty. When this is the case, we can indirectly rank y as inferior to any allocation on the belief-neutral Pareto-efficient frontier. This is also the case in our initial example of the bet between Bob and Joe, in which the status quo allocation is on the belief-neutral Pareto efficient frontier while the betting allocation is in the belief-neutral inefficient set. Thus, without relying on a particular welfare function, the belief-neutral Pareto criterion would nevertheless suggest that the status quo is preferred to the betting allocation.

2.3.5 An Alternative Criterion

It is useful to compare our welfare criterion to the no-betting Pareto-dominance criterion of Gilboa, Samuelson, and Schmeidler (GSS, 2012). They propose to extend the Pareto criterion in the presence of heterogeneous subjective beliefs by defining a choice x to dominate another choice y based on two conditions. First, each agent's expected utility under her own beliefs from x is higher than or equal to that from y, which is the standard Pareto condition. Second, there exists one belief, under which the expected utility of each agent from x is higher than

---

[11] As shall become clear, in endowment-economy settings, our criterion becomes discerning even if the reasonable beliefs are extended to include any belief that assigns nonzero probability to all relevant states.

or equal to that from y. This second condition is additional and requires the existence of a common belief to rationalize the efficiency of x.

In the example of Joe and Bob, there is no common belief that can rationalize the betting allocation. Hence, GSS's second condition prevents the bet from no-betting Pareto dominating the status quo. However, their first condition also prevents the status quo from no-betting Pareto dominating the bet—since the bet is desirable according to traders' own beliefs. As a result, the betting and the status quo allocations cannot be compared according to their no-betting Pareto dominance criterion. In contrast, our belief-neutral criterion identifies the betting allocation as being belief-neutral inferior.

The differences between the two criteria can be understood by considering the main premises behind our approach. First, we envision scenarios in which the planner knows that agents' beliefs are distorted. Consequently, unlike GSS, we ignore agents' expected utilities under their own, possibly distorted beliefs. Our second premise is that the planner does not know the objective belief. Consequently, we require the planner to vary the common-belief measure across a large set of reasonable beliefs so that the resulting welfare ranking is robust. In contrast, GSS require the existence of a single common belief to rationalize the efficiency of an allocation. Despite the seemingly restrictive robustness requirement, our criterion is able to identify positive and negative externalities in the bet between Joe and Bob, as well as in many other examples discussed in the next section.

3   Examples

This section provides a series of examples to demonstrate that the simple welfare criterion we propose, despite its incompleteness, can produce a surprisingly sharp welfare ranking in a wide range of economic environments with heterogeneously distorted beliefs. The key is that the externality induced by conflicting beliefs in these models is often uniformly positive or negative across different beliefs. In the example of the bet between Joe and Bob, the negative externality is reflected in the destroyed pillow. More generally, similar negative externalities can emerge through excessive risk taking, overinvestments, bankruptcy costs, and distorted consumption/saving decisions. Meanwhile, positive externalities can arise from alleviating the free-rider problem in information acquisition and overcoming market breakdowns induced by adverse selection. This section uses simple models to illustrate these different sources of externalities and demonstrates that our welfare criterion provides

a clear welfare ranking in each case. The second example on self-insurance shows that our criterion can provide a welfare ranking even without any externality, provided that there are certain restrictions on agents' belief disagreements—as for example a commonly agreed upon aggregate statistic. Our analysis offers important policy implications. However, the purpose of our simple examples is to isolate particular externalities rather than cover all relevant features necessary for making specific policy recommendations.

3.1 Speculation, Hedging, and Informational Efficiency

Our first example focuses on excessive trading induced by distorted beliefs, an issue that often arises in policy discussions. We present this example in three stages to explore several closely related conceptual issues. In the first stage, agents have constant endowments, but trading induced by their belief disagreements makes their consumption risky. As a result, trading is a negative-sum game in *expected utility terms* under any convex combination of the agents' beliefs. As in the pillow example, the welfare cost here stems from excessive risk taking. In the second stage, we introduce hedging motives by allowing agents' endowments to be risky and perfectly negatively correlated. Nevertheless, belief disagreements make the agents trade beyond simply offsetting their endowment risks. Our criterion can again identify the welfare loss induced by distorted beliefs. In the third stage, we introduce informational frictions regarding the state of the economy, which governs the agents' endowment risks. The presence of informational frictions makes costly information acquisition socially desirable. However, without heterogeneous beliefs, agents cannot fully appropriate the social value of their information acquisition. Distorted beliefs in this case can help overcome this free-rider problem and restore informational efficiency.

3.1.1 Speculative Motive

We start with a one-period endowment economy setting with two agents, A and B. Each agent invests at time $t = 0$ and consumes only at $t = 1$. Each agent has a constant endowment at $t = 1$ denoted by w. That is, there is neither aggregate nor idiosyncratic endowment risk. Each agent has an increasing and strictly concave utility function $u(c^i)$. The two agents hold heterogeneous beliefs about the state of the world, which takes two values, a and b. For now, since the endowment is constant, one may interpret the state as a sunspot. There is a single risky asset with payoff $V(a) = 1$ in state a and $V(b) = -1$ in state b. The asset is in

zero net supply, like derivative contracts or bets. Agent A assigns a probability of $\pi^A \in (0,1)$ to state a, while agent B assigns $\pi^B \in (0,1)$. The difference in beliefs causes the agents to engage in speculative trades against each other.[12]

Suppose that the asset is traded at a price of p at $t = 0$. Agent i $(i \in \{A, B\})$ chooses k^i, the number of contracts, to maximize his strictly concave expected utility:

$$\max_{k^i} \ \pi^i u \left(w + k^i \left(1 - p\right)\right) + \left(1 - \pi^i\right) u \left(w - k^i \left(1 + p\right)\right).$$

In the appendix, we formally derive each agent's optimal trading strategy and the resulting asset price p. It is intuitive that, in the presence of belief disagreements, each agent takes a nonzero position in the asset. The trading turns the agents' constant consumption allocation in the status quo into a risky one. The resulting market equilibrium is inefficient by either version of our welfare criterion. Suppose the planner has a utilitarian welfare function, $W(u_A, u_B) = u_A + u_B$. For any probability measure that lies between agents' belief measures, it is socially optimal to maintain deterministic consumption.[13] If $\pi^A \neq \pi^B$, the status quo allocation is belief-neutral superior to the market equilibrium allocation. We also apply the second version of our criterion. Under any belief measure $\pi \in (0,1)$, the status quo with an appropriate (measure dependent) transfer Pareto dominates the equilibrium allocation (again, because the latter is more volatile). Thus, the equilibrium allocation is belief-neutral inefficient.

3.1.2 Speculative and Hedging Motives

To make the example more realistic, we now introduce risky endowments. Agent A's endowment is $w - z$ in state a and $w + z$ in state b, while agent B's endowment is $w + z$ in state a and $w - z$ in state b. Without loss of generality, we assume $z \in (0, w)$. The negatively

---

[12]A large class of economic models analyzes trading between agents who hold heterogeneous beliefs regarding economic fundamentals and the impact of their trading on equilibrium asset price dynamics (e.g., Detemple and Murthy (1994), Kurz (1996), Zapatero (1998), Basak (2000), Buraschi and Jiltsov (2006), Jouini and Napp (2007), David (2008), Dumas, Kurshev, and Uppal (2009), Xiong and Yan (2010), and Dumas, Lewis, and Osambela (2011)). A key insight of these models is that trading induced by heterogeneous beliefs can lead to endogenous fluctuations in agents' wealth distribution, which, in turn, amplifies asset price volatility and induces time-varying risk premia. While these models can capture important dynamics of asset prices and risk premia, researchers tend to avoid making any welfare statement due to the lack of a well-specified welfare criterion. This simple example serves to illustrate that our criterion can potentially fill this gap, albeit without delivering any implication for asset price dynamics.

[13]In fact, for this example, the set of reasonable beliefs can be extended to include any measure that assigns a probability $\pi \in (0,1)$ to state a.

correlated endowments motivate the two agents to trade for hedging purposes, in addition to acting on the speculative motive discussed in the previous version.

Suppose that agent i ($i \in \{A, B\}$) takes on a position of k^i in the risky asset. If agents have common beliefs, $\pi^A = \pi^B = \pi$, they will trade the asset to fully hedge their endowment risk. As a result, each agent consumes a constant amount w regardless of the state. However, when agents have different beliefs, $\pi^A \neq \pi^B$, their equilibrium consumption is risky. This is because their pursuit of speculative gains causes them to deviate from the optimal risk-sharing allocation. Like in the previous version, the market equilibrium is again belief-neutral Pareto inefficient.[14] Specifically, we formally prove in the appendix that, for any measure with $\pi \in (0, 1)$, the equilibrium allocation is dominated by the optimal risk-sharing allocation with a certain transfer (which depends on the belief measure).

However, in this case, the status quo allocation with endowment risks is also belief-neutral inefficient. The welfare ranking between the status quo and the equilibrium allocation depends on the relative magnitude of the agents' endowment risk and their belief disagreement. For example, if the endowment risk is large and the belief disagreement is small, then the equilibrium allocation belief-neutral dominates the status quo. This sheds light on the ongoing debate over financial innovation, see, e.g., Posner and Weyl (2013). Introducing new tradable securities on the one hand allows agents to hedge their risk and on the other hand opens room for welfare-reducing speculation.

3.1.3 Social Value of Information

Besides speculation and hedging, trading can also occur for informational reasons. Traders can collect information and make trades based on it. The trading reveals (part of) the information to all market participants. This information is socially desirable if it improves investment efficiency in physical projects or enhances risk sharing. However, each agent has an incentive to free-ride on the costly information acquisition of others (e.g., Grossman and Stiglitz (1980)). When agents are fully rational, the level of information acquisition may be suboptimally low. Distorted beliefs can help mitigate this inefficiency. We now extend the example to show that our criterion permits a welfare analysis of the social value of information in the presence of distorted beliefs.

---

[14]Kubler and Schmedders (2012) and Simsek (2013a) analyze richer settings that feature a similar trade-off between hedging and risk sharing. Our welfare criterion is also useful for analyzing the inefficiency of speculative trading in these richer settings.

Consider a setting in which agents A and B live in an economy with two possible regimes that are characterized in the previous two stages. In regime 1, both agents have constant endowments w, as described in Section 3.1.1. In regime 2, they have perfectly negatively correlated endowments $y = \left\{ \left(y_a^A, y_b^A \right), \left(y_a^B, y_b^B \right) \right\} = \left\{ (w - z, w + z), (w + z, w - z) \right\}$, as described in Section 3.1.2. Let us assume that both agents agree that the two regimes are equally likely. At $t = 0$, one of the agents, say, agent A, can acquire information, which perfectly reveals the regime the economy is in, at a personal acquisition cost $c \geq 0$.[15] This acquisition cost is in utility terms and is deducted from the agent's utility from wealth.

As before, agents can trade a risky asset with payoff $V(a) = 1$ in state a and $V(b) = -1$ in state b. The agents have (possibly) heterogenous beliefs about probabilities of states a and b. For simplicity, we impose symmetry, i.e., $\pi^A = 1 - \pi^B$. If $\pi^A = 1/2$, beliefs are homogenous. For explicit derivation, we assume that both agents have a logarithmic utility function over wealth, $u(W^i) = \ln(W^i)$.

Let us first consider the case with common beliefs $\pi^A = \pi^B$. In this case, agents trade only to hedge the risk embedded in their endowments. If they know the regime the economy is in, they can mutually diversify away their endowment risk in regime 2 by trading the risky asset. Otherwise, any trading position on the risky asset yields a risky consumption stream in either regime 1 or regime 2, or in both. Acquiring information then plays a critical role in resolving this situation. While the costly signal is observed only by the information acquirer, i.e., agent A, trading perfectly reveals the private information to agent B as well. Specifically, if the signal reveals to agent A that the economy is in regime 2, agent A will initiate a trade with agent B; on the other hand, if the signal reveals to agent A that the economy is in regime 1, agent A is indifferent between trading or not trading. Thus, in the absence of any strategic behavior (which we assume), the trade initiated by agent A reveals his private signal to agent B.

Despite being socially desirable, the costly information may be under-provided in a competitive equilibrium. This follows from the public-good nature of information: both traders benefit from knowing the regime, but a single trader, agent A, bears the cost of acquiring the information. As in Grossman and Stiglitz (1980), information revealed through trading conveys a positive externality to the uninformed agent that cannot be fully captured by

---

[15]Strictly speaking, there is a continuum of each type of agent. In this example, we abstract from the free-rider problem among type-A agents in acquiring the information. One can think of a monopolistic information seller, who sells the information only to an investment club of type-A agents at a fixed cost higher than any single agent can afford.

the information acquirer. This externality may lead to an (inefficient) under-provision of information—specifically, when the information acquisition cost c exceeds agent A's individual gain but not the social gain.

The particular depiction of the information acquisition problem changes once traders are endowed with distorted beliefs. Under distorted beliefs, agents trade not only for risk-sharing purposes but also for speculative reasons. Speculation may be welfare enhancing, as it could give agent A additional incentives to acquire the costly information. Specifically, knowing which regime the economy is in allows agent A to fully hedge his endowment risk in regime 2, which, in turn, frees up his risk-bearing capacity to take a greater speculative position against agent B based on their heterogeneous beliefs about states a and b.[16] This speculative motive for acquiring information thus mitigates the externality in acquiring information and may lead to a belief-neutral efficient private provision of information, provided that the distortions of beliefs are not sufficiently strong.

In the appendix, we formally derive the subgame perfect equilibrium of the model by first computing both agents' trading strategies at $t = 0$ while taking as given agent A's information acquisition decision, and then solving agent A's optimal information acquisition policy. It is intuitive that agent A chooses to acquire information if and only if the acquisition cost c is lower than a threshold $c^{\text{eq}}(\pi^A)$, his private value of information. His value of information depends on the two agents' beliefs, π^A and $\pi^B = 1 - \pi^A$. By letting the planner use a utilitarian welfare function, $W(u_A, u_B) = u_A + u_B$, we also define two other threshold levels, $\overline{c}^{\text{eff}}(\pi^A)$ and $\underline{c}^{\text{inef}}(\pi^A)$, $\forall \pi^A \in [0, 1]$. If $c \leq \overline{c}^{\text{eff}}(\pi^A)$, acquiring information is belief-neutral efficient, while if $c \geq \underline{c}^{\text{inef}}(\pi^A)$, acquiring information is belief-neutral inefficient. These two thresholds measure the social value of information. It turns out that, due to the symmetry in the two agents' final wealth in this example, these two thresholds coincide: $\overline{c}^{\text{eff}}(\pi^A) = \underline{c}^{\text{inef}}(\pi^A)$, which we simply denote by c^{eff}.

Figure 1 plots the two thresholds for the information acquisition cost against agent A's belief π^A, based on the following parameter values: $w = 10$ and $z = 5$. As π^A deviates from 0.5, there is greater belief disagreement between agents A and B, since $\pi^B = 1 - \pi^A$. The solid line plots the private value of information $c^{\text{eq}}(\pi^A)$, below which agent A chooses to acquire information in the market equilibrium. The dotted line depicts $c^{\text{eff}}(\pi^A)$, below which acquiring information is belief-neutral efficient. The shapes of these lines are symmetric

---

[16]This feature is reminiscent of Simsek (2013a), in that by helping agents to better hedge their endowment risks, financial innovations allow agents to speculate more based on their heterogeneous beliefs.
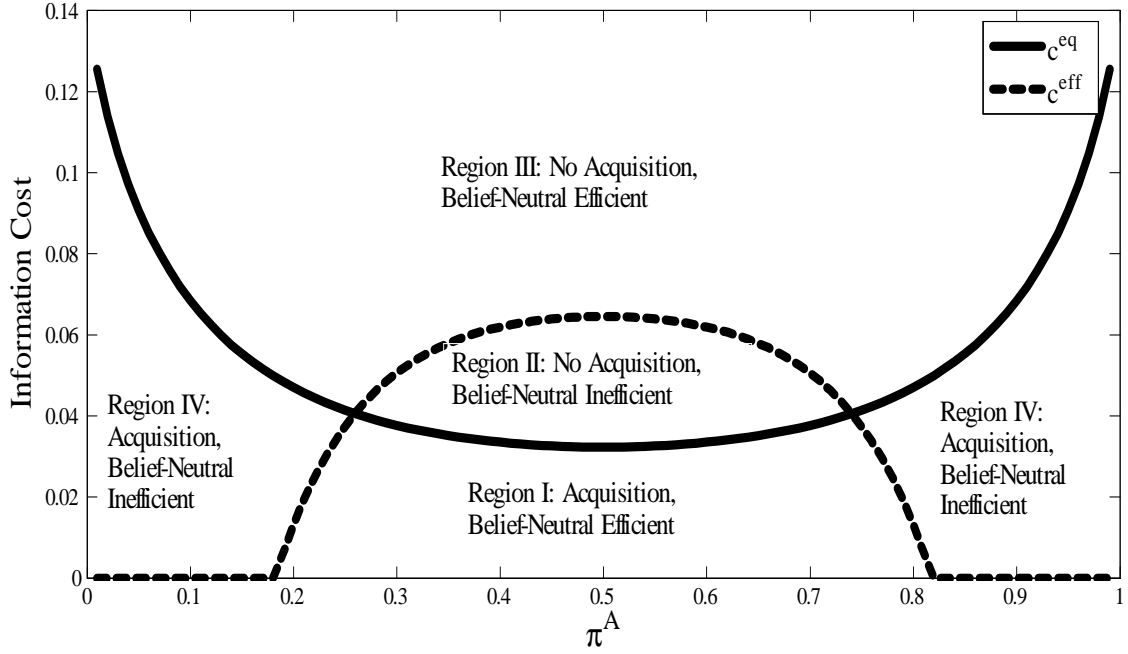
Figure 1: Threshold levels for information acquisition cost. c^{eq} denotes the threshold for agent A to acquire information in equilibrium. c^{eff} refers to the threshold for acquiring information to be belief-neutral efficient. This plot uses the following parameter values: $w = 10$ and $z = 5$.

around $\pi^A = 0.5$ due to the symmetric structure in the two agents' beliefs.

At the benchmark level $\pi^A = 0.5$, agents A and B have the same (correct) beliefs about the probabilities of the two states. In this case, $c^{\mathrm{eff}}(0.5) > c^{\mathrm{eq}}(0.5)$, which reflects the under-provision of information in the equilibrium. This is because agent A alone pays for the information acquisition, while both agents A and B benefit from the information. In fact, $c^{\mathrm{eff}}(0.5)$ is exactly double $c^{\mathrm{eq}}(0.5)$ due to the symmetric structure of this example.

As π^A deviates from 0.5, $c^{\mathrm{eq}}(\pi^A)$ rises. This is because, as the two agents' belief disagreement increases, agent A perceives a greater profit from trading against agent B. This opportunity for increased speculation motivates agent A to acquire information at a larger cost. Thus, by raising c^{eq}, belief distortions mitigate the under-provision of information in the market equilibrium.

As π^A deviates from 0.5, $c^{\mathrm{eff}}(\pi^A)$ drops. That is, as the two agents' belief disagreement increases, the social value of information decreases. This is because the acquired information not only improves the sharing of endowment risks between the two agents, but also allows them to speculate more based on their disagreements. The latter effect makes both agents' final consumption more volatile and thus reduces the social value of information.

Based on the way $c^{\text{eq}}\left(\pi^A\right)$ and $c^{\text{eff}}\left(\pi^A\right)$ intersect each other, Figure 1 illustrates four different regions. In region I, the information acquisition cost c is lower than both $c^{\text{eq}}\left(\pi^A\right)$ and $c^{\text{eff}}\left(\pi^A\right)$ and, as a result, agent A acquires information in the equilibrium and the information acquisition is belief-neutral efficient. In region II, c is higher than $c^{\text{eq}}\left(\pi^A\right)$ but lower than $c^{\text{eff}}\left(\pi^A\right)$ and, as a result, there is no information acquisition and the lack of information acquisition is belief-neutral inefficient. As discussed earlier, a key insight of our analysis is that as the two agents' belief disagreement rises (i.e., π^A deviates further away from 0.5), this region narrows. In region III, c is higher than both $c^{\text{eq}}\left(\pi^A\right)$ and $c^{\text{eff}}\left(\pi^A\right)$. In this case, agent A does not acquire information and the lack of information acquisition is belief-neutral efficient. Finally, in region IV, c is higher than $c^{\text{eff}}\left(\pi^A\right)$ but lower than $c^{\text{eq}}\left(\pi^A\right)$. In this case, agent A acquires information and the information acquisition is belief-neutral inefficient.

The finance literature, e.g., Black (1986), has long recognized the presence of noise trading induced by potential belief distortions of certain market participants as the key to resolving the free-rider problem in information acquisition. However, there is little formal analysis of this issue due to the challenge in performing welfare analysis with the presence of distorted beliefs. This example shows that our criterion can help fill this gap in the literature.

3.2 Self-insurance with Optimism

It is well known that insurance markets might malfunction because of information asymmetries. However, these asymmetries cannot explain the failure of self-insurance arrangements that do not require market interaction. For instance, survey evidence suggests that less than 15% of motorists in the US would wear seat belts voluntarily (see Williams and Lund (1986)). This means that seat belt laws in the US represent mandatory self-insurance. Regulations of this type are common also in other contexts. For example, financial regulation typically imposes on banks capital requirements that serve as insurance against potential losses.[17] We next present a model of the failure of self-insurance arrangements based on optimism. In this model, our criterion can identify mandatory insurance allocations as belief-neutral superior to laissez-faire allocations. This example also highlights that our criterion can lead to clear welfare ranking even in the absence of any externality, as long as agents' belief disagreements satisfy certain restrictions.

---

[17]These requirements are typically justified by moral hazard or fire-sale externalities. However, given that they are conceptually similar to seat-belt laws, there might be room for an additional justification.

We develop the model in the context of seat-belt laws. There are a large number of motorists denoted by $i \in I = \{1, .., N\}$. Each motorist i takes a precautionary action, $b_i \in \{0, 1\}$, where $b_i = 1$ corresponds to wearing a seat belt and $b_i = 0$ corresponds to not wearing one. After this decision, the motorist can be in one of two states denoted by $s_i \in \{0, 1\}$, where $s_i = 1$ corresponds to an accident and $s_i = 0$ corresponds to no accident. In case of an accident, the motorist suffers physical damage, which we capture with a monetary equivalent loss denoted by $d > 0$. For simplicity, suppose wearing a seat belt enables the motorist to completely avoid the damage. But wearing a seat belt is also inconvenient, which we capture with a monetary equivalent cost $c \in (0, d)$. We also assume the motorist is risk-neutral, so that the state utility function can be written as

$$u_i = -s_i (1 - b_i) d - b_i c.$$

The aggregate state of the economy is given by $s = (s_i)_{i \in I} \in S$. Unlike our other examples, here the state involves multiple events. For simplicity, the economy features no aggregate uncertainty in the sense that exactly a fraction $\mu \in (0, 1)$ of motorists will have an accident, where μN is also an integer, so that

$$\sum_{i=1}^{N} s_i = \mu N \text{ for each } s \in S. \tag{1}$$

Importantly, all agents agree on this aggregate statistic, although each agent also believes her own state satisfies $s_i = 0$ with certainty. Put differently, motorists know and agree on the average accident probability (perhaps because they observe the historical accident statistics). Nonetheless, each motorist is optimistic in the sense that she believes these accidents will happen to other motorists.

Absent any policy requirement, each motorist chooses not to wear a seat belt so that the laissez-faire allocation features $b_i = 0$ for each $i \in I$. To evaluate welfare, consider the utilitarian social welfare function under any convex combination of motorists' beliefs. Since each motorist's belief satisfies the aggregate restriction in (1), so does any convex combination of their beliefs, which implies the utilitarian welfare,

$$E_0^h \left[\sum_{i=1}^{N} u_i \right] = -E_0^h \left[\sum_{i=1}^{N} s_i \right] d = -\mu N d.$$

Intuitively, the fraction μ of motorists will have an accident, which will lead to a social

loss, $\mu N d$. Importantly, the inefficiency is detected by our belief-neutral criterion because motorists agree on the average accident probability.

Next consider a policy that makes it mandatory for all motorists to wear a seat belt so that $b_i = 1$ for each $i \in I$. The corresponding belief-neutral utilitarian welfare is

$$E_0^h \left[\sum_{i=1}^N u_i \right] = -Nc.$$

This policy features a different type of inefficiency. In particular, according to each motorist, at most $N - 1$ motorists—namely, motorists excluding himself—can have positive accident probability. Thus, each motorist believes that motorists in total are forced to wear more seat belts than is strictly necessary to reduce accident damage. The inefficiency is again detected by our belief-neutral criterion.

Comparing the two cases, the mandatory seat belt allocation is belief-neutral superior to the laissez-faire allocation if and only if

$$c < \mu d. \tag{2}$$

Intuitively, a seat belt is socially beneficial in this model as long as its inconvenience cost, c, is smaller than its average damage reduction, μd.[18] While there is not any externality in this example, the assumption that agents agree on the average accident probability—that restricts agents' belief disagreements—enables us to provide a clear welfare ranking.

In the context of capital requirements for banks, the counterpart of an accident $s_i = 1$ can be suffering financial losses. The counterpart of wearing a seat belt $b_i = 1$ can be retaining sufficient equity capital (as opposed to paying out dividends) to absorb potential losses. The counterpart of the optimism assumption can be that banks know and agree on the average probability of suffering losses, but each bank believes these losses will be borne by other banks. Our analysis then suggests banks in a laissez-faire allocation will not retain sufficient capital to absorb losses. Moreover, as long as retaining capital is relatively cheap, mandatory bank capital requirements will generate a belief-neutral welfare improvement over the laissez-faire allocation. More broadly, our criterion would identify mandatory self-insurance policies

---

[18]The same point can be made using the Pareto version of our criterion, although the argument is more subtle than that derived from the utilitarian social welfare function. In this model, there are no belief-neutral Pareto-efficient allocations. This is because efficiency requires a subset of motorists to wear a seat belt, but which subset of motorists does so depends on the belief used for welfare analysis. That said, when condition (2) is satisfied, the laissez-faire allocation is belief-neutral Pareto dominated by the mandatory seat-belt allocation combined with appropriate ex ante transfers. In this sense, the Pareto version of our criterion also favors the mandatory seat-belt allocation over the laissez-faire allocation.

$\pi_u^A = 0.5 + \delta, \pi_u^B = 0.5$ D_{uu}

u

$\pi_0^A = 0.5, \pi_0^B = 0.5$

$\pi_d^A = 0.5 - \delta, \pi_d^B = 0.5$ D_{ud}

d
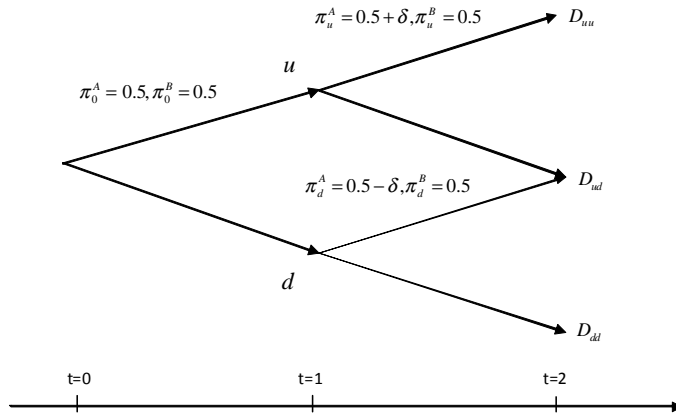
D_{dd}

t=0   t=1   t=2

Figure 2: Payoff and belief structure in the bubble model.

as belief-neutral superior under the assumptions that agents agree on the aggregate risks, as in (1), and the cost of self-insurance is smaller than its average benefit, as in (2).

3.3   Bubble Models of Overinvestment

A segment of the literature emphasizes that when short sales are constrained, heterogeneous beliefs can lead to price bubbles as asset owners anticipate reselling their assets to other, more optimistic agents in the future (e.g., Harrison and Kreps (1978), Morris (1996), Scheinkman and Xiong (2003), Wu and Guo (2004), Hong, Scheinkman, and Xiong (2006), and Hong and Sraer (2011)). In these models, heterogeneous beliefs induce risk-neutral agents not only to trade against each other but also to overvalue assets. Overvaluation does not reduce social welfare by itself, because it is simply a welfare transfer across agents. However, overvaluation of equity can lead to firms' overinvestments (e.g., Bolton, Scheinkman, and Xiong (2006), Gilchrist, Himmelberg, and Huberman (2005), and Panageas (2006)), which reduce the total welfare of all investors. Our criterion can identify overinvestments independent of the belief used to evaluate firms' fundamental values.

We focus on a simple binomial setting with three dates (i.e., $t = 0, 1, 2$) and two risk-neutral agents (A and B). These agents trade the equity issued by a firm. The firm chooses its investment at time 0. Suppose that the firm's investment is cost-free but the investment return has a decreasing return to scale. If the firm chooses to establish a production capacity of n units, the dollar return to per unit of capacity is determined by a binomial tree depicted in Figure 2. There are three possible states (uu, ud, and dd) on $t = 2$. Suppose that the

26

return per unit across the states at time 2 is

$$\widetilde{R} = \{D_{uu}, D_{ud}, D_{dd}\} = \{R + 1 - n, R - n, R - 1 - n\},$$

where $R > 1$ is a constant. Due to the firm's decreasing return to scale, a larger investment scale n reduces the per-unit return by n across all states at time 2. Suppose that the firm issues one share of equity for each unit of production capacity. The shares are equally distributed to A and B.

Before analyzing the firm's investment decision, we first examine the market price of each share of equity. Figure 2 depicts the dynamics of the two agents' beliefs. We assume that the two agents have time-varying beliefs: They start with the same beliefs at time 0 but hold different beliefs at time 1:

$$\pi_0^A = \pi_0^B = 0.5, \ \pi_u^A = 0.5 + \delta > \pi_u^B = 0.5, \ \text{and} \ \pi_d^A = 0.5 - \delta < \pi_d^B = 0.5. \tag{3}$$

In particular, agent A becomes more optimistic than agent B in state u at time 1 and less optimistic in state d. The parameter $\delta > 0$ determines the two agents' belief dispersion in both states u and d. It is straightforward to verify that, at $t = 0$, the two agents share the same expectation of the asset's final payoff:

$$E_0^A\left[\widetilde{R}\right] = E_0^B\left[\widetilde{R}\right] = R - n.$$

Following this literature, we assume that short-sale of the equity is not allowed. Accordingly, the fluctuations of the two agents' beliefs at $t = 1$ give an asset owner, who can be either A or B, an option to resell his holding to the other agent: more specifically, for agent A to sell to agent B in state d and for agent B to sell to agent A in state u. To obtain a bubble, we assume that each agent has sufficient cash to acquire the asset so that the competitive price is determined by the buyer's reservation value. It is straightforward to derive the following market price in state u: $p_u = R + 1/2 + \delta - n$, which is paid by agent A, and in state d: $p_d = R - 1/2 - n$, which is paid by agent B. By backward induction, both agents at time 0 value the asset by $p_0 = R + \delta/2 - n$. Even though each agent's expectation of the asset payoff is $R - n$, their valuation of the asset is $R + \delta/2 - n$. The difference is driven by the resale option, i.e., the speculative motive to resell the asset to the other agent at a price higher than his own valuation at time 1. This resale option contributes a non-fundamental component to asset prices in the aforementioned bubble models.

We now analyze the firm's investment decision. Suppose the firm chooses its production capacity, n, to maximize its market value, given by $n \cdot p_0 = n \cdot (R - n + \delta/2)$. This is the appropriate objective function since the owners, agents A and B, agree on the firm's valuation at time 0. The firm's optimal investment level is then given by $n^* = \frac{1}{2}\left(R + \frac{\delta}{2}\right)$, which depends on δ, the magnitude of the two agents' belief dispersion at time 1.

Is this investment decision socially efficient? Suppose the planner uses the utilitarian social welfare function along with a convex combination of the two agents' beliefs, $\Pi^h = h\Pi^A + (1 - h)\Pi^B$, $\forall h \in (0, 1)$. Since both A and B are risk neutral, the expected utilitarian social welfare is equal to the firm's expected final payoff, given by $n \cdot E^h\left[\widetilde{R}\right] = n(R - n)$. This expression is maximized by choosing $n^{**} = \frac{1}{2}R < n^*$. This implies that the firm overinvests in the market equilibrium relative to the level that maximizes the expected utilitarian social welfare (or the firm's long-run fundamental value) under any convex combination of the agents' beliefs.[19]

This result does not need to rely on any social welfare function because the market equilibrium is, in fact, belief-neutral Pareto inefficient. In particular, it can be checked that for any reasonable belief, Π^h, the market equilibrium with investment level n^* is Pareto dominated by an alternative allocation with investment $n^{**} < n^*$ combined with some initial transfer, $T \in [-n^*(R - n^*), n^*(R - n^*)]$, from agent B to agent A.

The driving force behind the inefficient overinvestment is exactly the value of the resale option in the firm's time-0 market valuation. Anticipating the possibility of reselling the share to the other agent at time 1 at a profit, each agent overvalues the share at time 0 relative to his own expectation of the share's long-run fundamental value. This, in turn, induces the firm to overinvest. Note that each agent recognizes that this level of investment reduces the firm's long-run value. However, each agent also thinks that these losses will be borne by the other agent. A negative externality emerges just like in the bet between Joe and Bob. Consistent with this overinvestment example, Gilchrist, Himmelberg, and Huberman (2005) provide evidence that firms tend to increase investment in response to increased heterogeneous beliefs proxied by dispersion in analysts' earnings forecasts.

---

[19]Given the presence of the firm's investment decision, it is important to restrict the set of reasonable beliefs to the convex combinations of agents' beliefs. This is because a measure outside the convex combinations of agents' beliefs would imply that the agents' aggregate belief is biased and thus rule the firm's investment decision in the equilibrium as inefficient, even in the absence of any belief dispersion between the two agents. As stated previously, analyzing inefficiencies associated with agents' aggregate biases is not our focus.

3.4 Benefits of Speculation in Lemons Models

The previous example shows that overinvestment in heterogeneous-beliefs-induced bubble models leads to belief-neutral welfare losses. However, speculation and bubbles induced by heterogeneously distorted beliefs can also be beneficial. Among other things, bubbles help overcome market breakdown in "lemons" models caused by adverse selection (as in Akerlof (1970)). This subsection illustrates this point by introducing heterogeneous beliefs into a recent model of Tirole (2012). Also see Morris (1994) for a model in which heterogeneous beliefs help break the no-trade theorem and Zhuk (2012) for a model in which bubbles induced by heterogeneous beliefs help overcome the information externalities among firms.

The model of Tirole (2012) considers a firm that attempts to finance a new investment project by selling its legacy asset. However, the firm is asymmetrically informed about the payoff from the legacy asset, which creates a lemons problem. As in Akerlof (1970), the equilibrium features a low price and reduced trade and, in some extreme cases, a complete market breakdown. We show that bubbles induced by heterogeneous beliefs mitigate the lemons problem by allowing the firm to sell its asset and invest in the new project even if the quality of its legacy asset is relatively high. Our criterion can detect the consequent welfare gain.

Consider a seller who has access to a new project that costs I and generates a payoff of $I + G$. The payoff of the project is not pledgable (that is, it accrues to the seller but cannot be promised to others). Thus, the seller needs to finance the project by selling a legacy asset that is pledgable. This asset returns $\overline{R}$ with probability θ, and 0 otherwise. The probability, θ, itself is uniformly distributed over $[0, 1]$. The prior value of the pledgable asset exceeds the investment cost, $p^{prior} \equiv \frac{\overline{R}}{2} > I$, so that the project is always financed in a constrained efficient allocation.

The key friction is that the seller is asymmetrically informed about the success probability of the legacy asset. In particular, the seller receives a signal and fully learns θ, while potential buyers continue to believe that θ is distributed according to the uniform prior. The rest of the section analyzes the effect of this friction on the efficiency of the equilibrium allocation with and without heterogeneously distorted beliefs. Suppose also that $G < \frac{\overline{R}}{2}$, which rules out the extreme case in which the seller is always able to finance the project despite having asymmetric information.

First, consider the benchmark without distorted beliefs among potential buyers. Let p^*

denote the equilibrium asset price. If $p^* < I$, then there is no trade because the seller is unable to finance the new project by selling the legacy asset. If $p^* > I$, then a trade is possible. In particular, the seller will sell the asset only if $\theta \overline{R} \leq p^* + G$. In a competitive equilibrium, the buyers break even, which implies $p^* = \overline{R} E \left[\theta | \theta \leq \frac{p^*+G}{R} \right]$. Solving this further gives the equilibrium price $p^* = G$. It follows that there is no trade when $G < G^* = I$. When $G \geq G^*$, the seller will sell the asset only if $\theta < \theta^*$, and she will do so at a price p^*, where

$$\theta^* = \frac{2G}{\overline{R}} < 1 \text{ and } p^* = G < p^{prior}.$$

In particular, the adverse selection induced by the asymmetric information between buyers and the seller reduces the level of asset trading and the asset price. Intuitively, sellers with low-quality assets ("lemons") exert a negative externality on sellers with higher-quality assets. In some cases (i.e., $G < I$), there is a complete market breakdown.

To formally discuss social welfare, we consider (as in Tirole (2012)) the ex ante utilitarian social welfare function, i.e., the sum of the seller's and buyers' expected utilities under the prior distribution for θ. Since the trading profits represent a pure transfer between the seller and buyers, the ex ante social welfare is simply

$$E \left[\widetilde{R}\theta + I_{\{G>G^*, \theta<\theta^*\}} G \right] < \frac{\overline{R}}{2} + G.$$

Here, $I_{\{G>G^*, \theta<\theta^*\}}$ is an indicator function for whether the seller manages to invest in the project, and the inequality follows since there is investment with probability strictly less than 1. In contrast, an alternative (feasible) allocation that always transfers the asset from the seller to buyers at price p^{prior} ensures that the project is always financed and the social welfare is $\frac{\overline{R}}{2} + G$. Hence, the competitive equilibrium is (constrained) Pareto inefficient.

Next, we consider the case of buyers holding heterogeneously distorted beliefs regarding the asset return. Suppose that the asset return in the event of success is random and independent of the asset's success. We denote it by $\widetilde{R}$ and assume that it can take two possible values, $\overline{R} + 1$ and $\overline{R} - 1$. The seller believes the probability of $\widetilde{R} = \overline{R} + 1$ is 0.5. There are two groups of risk-neutral buyers for the asset. One group believes the probability of $\widetilde{R} = \overline{R} + 1$ is 1, while the other group believes the probability is 0. Suppose that no one can short-sell the asset and each group has sufficient cash to acquire the asset. As in the previous example, buyers in the optimistic group acquire the asset and bid up its price to their expectation of the asset payoff. A key feature of the model is that the asset overvaluation induced by agents' heterogeneous beliefs (as in Miller (1977)) helps overcome

the lemons problem. To see this most starkly, suppose $G > \frac{\overline{R}-1}{2}$ (while continuing to assume $G < \frac{\overline{R}}{2}$). Under this assumption, it can be seen that the seller chooses to sell and finance the project regardless of θ. The optimistic buyers break even only if $p = \left(\overline{R}+1\right) E\left[\theta\right] = \frac{\overline{R}+1}{2}$. At this price, the seller in turn finds it optimal to sell because $\theta\overline{R} \leq \overline{R} < p + G$, where the last inequality follows since $G > \frac{\overline{R}-1}{2}$. Consequently, unlike the earlier case (for the same parameters), the competitive equilibrium with belief heterogeneity features trade and investment with probability 1.

We can apply our welfare criterion to show that the equilibrium with belief heterogeneity is, in fact, belief-neutral efficient. To see this, let Π^h denote a probability measure, which assigns probability $h \in [0, 1]$ to $\widetilde{R} = \overline{R}+1$ and which is a convex combination of all buyers' beliefs. The ex ante social welfare under this belief can be written as

$$E^h\left[\widetilde{R}\theta + G\right] = E^h\left[\overline{R}\right]\frac{1}{2} + G,$$

since the project is invested with probability 1. As this expression illustrates, regardless of the probability measure, the ex ante welfare is at its highest possible level. This is because there is no disagreement about G, the gains from undertaking the project. This in turn implies that the equilibrium is belief-neutral efficient. Thus, speculation induced by heterogeneous beliefs mitigates the lemons problem and leads to belief-neutral welfare gains.

3.5   Bankruptcy Costs in Leverage Cycle Models

A growing literature builds on agents' heterogeneous beliefs to analyze leverage cycles (e.g., Geanakoplos (2003, 2009), Fostel and Geanakoplos (2008), Simsek (2013b), Cao (2010), Shen, Yan, and Zhang (2011), and He and Xiong (2012)). The key feature of those models is that optimism can motivate cash-constrained optimists to use collateralized short-term debt to finance their asset acquisition. The leverage initially fuels the price boom but later forces the optimists to deleverage after bad shocks, resulting in a leverage cycle. This framework nicely integrates the optimists' leverage cycle with the asset price cycle. Both cycles are important for understanding various historical episodes of financial crises, including the recent one. To use this framework to analyze relevant policy issues (such as regulation over financial institutions' leverage), it is important to discuss welfare implications. Our criterion can generate clear welfare ranking in this framework. The key insight is that over-optimism causes optimists to use excessive leverage in asset acquisition despite the possibility
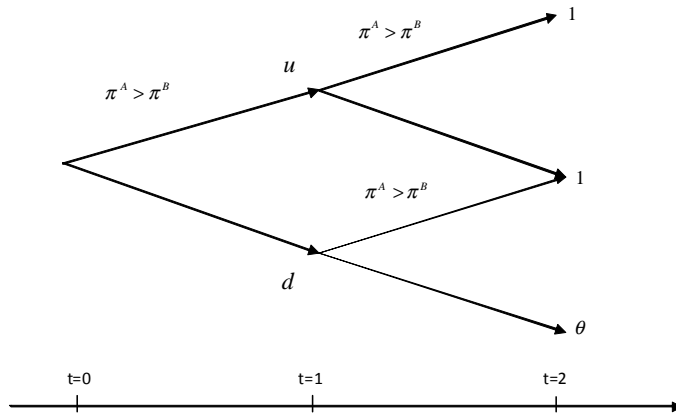
Figure 3: Payoff and belief structure in the leverage cycle model.

of incurring bankruptcy costs in the future. Bankruptcy costs make the excessive leverage a negative-sum game between optimistic buyers and pessimistic creditors.

Consider a setting with 3 dates, i.e., $t = 0, 1, 2$, and two types of risk-neutral agents (A and B). Figure 3 depicts the asset payoff and the beliefs of the two types. Suppose that the final payoff of a risky asset across the three final states at date 2 is $\widetilde{R} = \{1, 1, \theta\}$, where $\theta \in (0, 1)$. The asset gives a low payoff of θ after two negative fundamental moves and gives 1 in other final states. We normalize the net supply of the asset to one unit and the risk-free interest rate to zero. Each type holds a constant belief about the probability of the fundamental state rising on the tree in the following period. We denote the two groups' beliefs by $\pi^A \in (0, 1)$ and $\pi^B \in (0, 1)$ with $\pi^A > \pi^B$. A key feature of this setting is that the specified payoff and belief structures lead to an increased divergence in the agents' fundamental expectations about the asset payoff in the lower state d of date 1, which eventually triggers a leverage cycle.[20]

Suppose that the pessimists (type-B agents) initially own all of the asset at $t = 0$. It is desirable for the optimists (type-A agents) to acquire all of the assets. However, they face a practical problem in that they may not have sufficient cash endowments to make the purchases. To highlight this problem, we assume that there is one unit of optimists, each with an initial cash endowment of $c > 0$. They can use asset holdings as collateral to raise debt financing. If a borrower is unable to make the promised debt payment, the creditor can seize the collateral. This in turn makes the availability and cost of the borrower's debt

---

[20]In Brunnermeier and Pedersen (2009), bad fundamental news leads to higher fundamental volatility, which in turn triggers an increase in margin.

financing dependent on the future value of the collateral. On the other hand, the availability of debt financing directly determines how much the optimists can bid up the asset price beyond the pessimists' asset valuation.

In deciding how much to borrow, type-A agents face two sources of costs. First, as the creditors (likely type-B agents) are more concerned about the potential default risk than the borrowers, higher leverage tends to be more costly. Second, if a type-A agent defaults on the debt and is forced to sell his asset on either date 1 or 2, he faces a personal liquidation cost, α. One can interpret this cost as the inconvenience cost of vacating a house, which is incurred by the borrower. At the end of this subsection, we also describe a version of the model in which costs are incurred by the creditor when the borrower defaults. These two versions have similar welfare implications.

Our setting maintains several key features used by Geanakoplos (2009), including the same binomial payoff structure and the same collateralized debt contract. We add liquidation costs, which is a realistic feature, and one that was especially relevant during the recent subprime mortgage crisis. Since this feature complicates the analysis, we allow for only two types of beliefs rather than a continuum. The model derivation follows He and Xiong (2012), who analyze equilibrium debt financing in a setting with two types of agents whose beliefs vary over time, but without liquidation costs.

There are two relevant debt contracts in equilibrium. One contract promises a payment of θ at date 1 collateralized by one unit of the asset. Because the asset's fundamental value in the worst state of date 2 is able to cover θ, this debt contract is riskless throughout and can thus give the borrower an initial credit of θ. The second contract promises a payment on date 1 equal to type-B agents' (the creditors') asset valuation in state d of date 1:

$$K_d \equiv E_d^B[\widetilde{R}] = \pi^B + \left(1 - \pi^B\right)\theta > \theta.$$

As the creditors value the collateral for at least K_d on date 1, this debt is also riskless and allows a borrower to borrow at the risk-free interest rate for the initial period. However, to refinance this debt in state d of date 1, the borrower has to make a greater promise of paying 1 at date 2. This new promise allows him to raise K_d from type-B agents to pay off his initial debt, but exposes him to the risk of defaulting and being forced to liquidate the asset if the asset's fundamental value eventually turns out to be θ on date 2. Relative to the first contract, the second one gives higher leverage at the expense of a higher refinancing cost in state d of date 1 as well as the possibility of incurring the liquidation cost on date 2.

We prove in the appendix that these two debt choices dominate the other alternatives.

We assume that the liquidation cost, α, is modest so that in some scenarios the type-A agents will choose the higher leverage (i.e., the contract with promise K_d) and thus face the liquidation risk:

$$\alpha < \frac{\pi^A \pi^B (\pi^A - \pi^B)}{(1 - \pi^A)^2 [1 - (1 - \pi^B)^2]}(1 - \pi^B)(1 - \theta). \tag{4}$$

Under this assumption, the analysis in the appendix shows that there is a price threshold $p_0^* \in \left(E_0^B \left[\tilde{R} \right], E_0^A \left[\tilde{R} \right] \right)$, such that type-A agents choose the debt with promise K_d if and only if $p_0 \leq p_0^*$. Intuitively, when the price is low, type-A agents see a bargain and are willing take the high-leverage debt despite the refinancing and liquidation costs it entails.

Appendix A.4 characterizes the equilibrium in five different cases based on type-A agents' initial cash c. We are particularly interested in three cases, in which c is sufficiently low so that at least some of type-A agents choose to finance their asset purchases by using the high-leverage debt with promise K_d. This debt financing exposes them to the liquidation cost on date 2. They make this choice purely for speculative reasons—because they perceive the asset to be significantly underpriced, $p_0 \leq p_0^* < E_0^A \left[\tilde{R} \right]$.

We next apply our welfare criterion to illustrate that this equilibrium is indeed inefficient. To see this, first suppose the planner has the utilitarian welfare function. We use a convex combination of the two types' beliefs, $\Pi^h = h\Pi^A + (1 - h) \Pi^B$, $\forall h \in (0, 1)$, to calculate welfare. The risk neutrality of both types of agents implies that the social welfare is given by the asset's expected fundamental value plus optimists' cash, c, and minus the expected liquidation costs, which amount to

$$W \left(E_0^h \left[u_A \right], E_0^h \left[u_B \right] \right) = c + E_0^h \left[\tilde{R} - \alpha \mu I_{\tilde{R}=\theta} \right],$$

where μ is the fraction of type-A agents using high-leverage K_d debt contract and $I_{\tilde{R}=\theta}$ denotes the indicator function for the realization of the state $\tilde{R} = \theta$. Since both type-A and type-B agents assign a positive probability to this state, the social welfare is lower than that of the status quo allocation with no asset trading:

$$W \left(E_0^h \left[u_A \right], E_0^h \left[u_B \right] \right) < c + E_0^h \left[\tilde{R} \right].$$

Thus, our criterion identifies, regardless of the beliefs, a strict welfare loss in these cases due to the liquidation costs incurred by the borrowers.[21] As before, this result holds for any

---

[21] The welfare loss is present even if the planner adopts a belief measure outside the convex combinations of the two agents' beliefs, as long as the measure assigns a positive probability to the state $\tilde{R} = \theta$.

welfare function because the equilibrium is also belief-neutral Pareto inefficient.[22]

In more general settings, agents acquire assets not just for speculative purposes but also for consumption. For example, people buy houses not only because they expect housing prices to appreciate but also because they enjoy living in their house. It is important to incorporate both speculative incentives and consumption values in evaluating the welfare consequences of leverage cycles. Our criterion provides a useful tool for such an evaluation.

3.6 Consumption/Savings Distortions in Macro Models

In macroeconomic models, belief disagreements can also distort aggregate investment through individuals' consumption/savings decision, e.g., Sims (2008). Belief disagreements cause individuals to perceive greater expected returns from their investments. This affects their savings decision in the same way an increase in the real interest rate does. It creates not only a substitution effect, which tends to increase savings, but also an income effect, which tends to increase current consumption and thus reduce savings. Depending on which effect dominates, individuals might save too much or too little relative to a homogeneous-beliefs benchmark. The net saving in turn leads to over- or under- investment. Our criterion can help detect these types of inefficiencies.

As the setting used by Sims is simple enough, we adopt it in full. The setting has two dates and two types of agents. We normalize the size of the population to one. Each agent starts with an endowment of B_0 dollars of nominal bonds issued by the government and an endowment of Y units of goods. At the initial date, he can consume part of the goods endowment and invest the rest either in the nominal bonds or in a real asset.

There are two possible states of the world on the second date $s \in \{f, m\}$. In state s, the

---

[22] As an alternative, we briefly describe a setting in which bankruptcy costs are borne by creditors instead of borrowers. This alternative setting follows that of Simsek (2013b). Suppose there are only two dates, $t \in \{0, 1\}$, but three states, $\{H, M, L\}$, in which the asset price will be either high, medium, or low. The agents agree about the probability of the low payoff state, π_L, but disagree about the probabilities of the remaining states. In particular, type-A agents are more optimistic about the high state, i.e., $\pi_H^A > \pi_H^B$ (and thus, $\pi_M^A < \pi_M^B$). As before, type-A agents borrow from type-B agents using collateralized debt contracts. Suppose a fraction, $\iota \in (0, 1)$, of the value of the asset is lost in a foreclosure, which is the main difference from the earlier setting. In this case, it can be seen that type-A agents face a trade-off between choosing a safe debt contract with face value L, and a risky debt contract with face value M. The risky debt enables them to borrow a larger amount, $\pi_L (1 - \iota) L + (1 - \pi_L) M$, but is also expensive (i.e., it has a high yield). This is because it leads to bankruptcy costs in some states. As before, under appropriate conditions, the speculative motive induces type-A agents to finance their purchases with the risky debt. This arrangement generates expected bankruptcy costs according to any reasonable belief measure, and is thus belief-neutral inefficient.

government fixes the state-dependent lump-sum tax on each agent to be τ_s and the gross nominal interest rate to R. In state f, the tax backing for bonds is low and hence prices are high, while in state m, taxes are high and prices are therefore lower. Thus, the government's second date budget constraints determine the bond price: $P_{2s} = \frac{RB_0}{\tau_s}$, where $s = f, m$.

The economy has a representative firm, which produces at the second date according to a decreasing return to scale production function: $g(K) = AK^{1-\alpha}$, where K is the capital input and A is a constant. The firm has to rent capital from individual agents at a market rental rate of ρ. We normalize the firm's ownership to one share, which is equally divided among the agents. Thus, the firm's profit per unit of ownership is $\Psi = AK^{1-\alpha} - \rho K$. The firm's profit optimization requires that $\rho = A(1-\alpha)K^{-\alpha}$.

There are two types of agents: $i \in \{a, b\}$. Type i agents believe that the probability of state f is $\pi_i \in (0, 1)$. Each type contributes to half of the population. Each agent maximizes his aggregate utility across the two dates:

$$\max \ U(C_{i1}) + \beta \left[\pi_i U(C_{if}) + (1 - \pi_i) U(C_{im}) \right]$$

where C_{i1}, C_{if}, and C_{im} are a type i agent's consumption on date 1 and in states f and m of date 2, and β is the agent's time discount rate. On the first date, the agent can allocate his initial good endowment Y to consumption C_{i1}, renting capital to the firm K_i, and buying more nominal bonds $B_i - B_0$ at a nominal price of P_1:

$$C_{i1} + K_i + \frac{B_i - B_0}{P_1} = Y.$$

Note that the agent can take a short position in the capital, which is equivalent to borrowing in real terms at a rate of ρ. He can also take a short position in the nominal bonds, which is equivalent to borrowing in nominal terms at a rate of R. His consumption in state s of the second date is given by

$$C_{is} = \rho K_i + \frac{RB_i}{P_{2s}} - \tau_s + \frac{\Psi}{2}$$

where P_{2s} is the nominal bond price in the state. Suppose that both types of agents have a power utility function: $U(C) = \frac{C^{1-\gamma}}{1-\gamma}$ with γ as the rate of relative risk aversion.

The first order condition for the agent with respect to K_i gives

$$C_{i1}^{-\gamma} = \beta\rho \left[\pi_i C_{if}^{-\gamma} + (1 - \pi_i) C_{im}^{-\gamma} \right], \quad i \in \{a, b\}$$

and with respect to B_i gives

$$\frac{1}{P_1} C_{i1}^{-\gamma} = \beta R \left[\frac{\pi_i C_{if}^{-\gamma}}{P_{2f}} + \frac{(1 - \pi_i) C_{im}^{-\gamma}}{P_{2m}} \right], \quad i \in \{a, b\}.$$

$\{\pi_a, \pi_b\}$	K_a	K_b	K	B_a	B_b	P_1	C_{a1}	C_{af}	C_{am}	C_{b1}	C_{bf}	C_{bm}
$\{0.3, 0.3\}$	0.51	0.51	1.03	1.50	1.50	0.84	1.09	0.61	0.61	1.09	0.61	0.61
$\{0.7, 0.7\}$	0.51	0.51	1.03	1.50	1.50	0.98	1.09	0.61	0.61	1.09	0.61	0.61
$\{0.3, 0.7\}$	-2.19	3.30	1.12	3.94	-0.94	0.89	1.04	0.20	1.09	1.04	1.09	0.20

Table I: Equilibrium under homogeneous and heterogeneous beliefs.

The market clearing condition for the capital gives $K = K_a + K_b$ and for the nominal bonds gives $B_0 = B_a + B_b$. These conditions allow us to determine a unique equilibrium represented by $\{K_a, K_b, B_a, B_b, P_1\}$.

While analytical solution of the equilibrium is not available, it is numerically tractable. We adopt the same parameter values used by Sims to illustrate the equilibrium:

$$Y = 1.6, R = 1.1, \tau_f = 1.1, \tau_m = 1.65, \alpha = 0.3, \beta = 0.9, A = 1.2, \gamma = 0.5, B_0 = 1.5. \quad (5)$$

We compare the equilibrium outcomes under three sets of beliefs: two homogeneous-beliefs benchmarks, $\{\pi_a = 0.3, \pi_b = 0.3\}$ and $\{\pi_a = 0.7, \pi_b = 0.7\}$, and a heterogeneous-beliefs economy in which each agent believes in one of the benchmarks, $\{\pi_a = 0.3, \pi_b = 0.7\}$.

Table I lists the equilibrium quantities in the three settings. First note that the two homogeneous-beliefs equilibria have some common (belief-neutral) properties. In particular, while beliefs about inflation affect the nominal bond price, P_1, they have no effect on real allocations such as investment and consumption. In contrast, the equilibrium with heterogeneous beliefs has two main differences in terms of real allocations. First, with heterogeneous beliefs, agents have more volatile consumption across the two states of the second date. Like the last example, this increased variability is due to the speculation between the agents about the nominal price inflation. The type a agents (the inflation pessimists) invest more in nominal bonds and at the same time short-sell the capital (i.e., borrow in real terms). Second, with heterogeneous beliefs, agents also save more (and consume less) on date 1. Intuitively, belief disagreements induce agents to perceive a greater expected return from their investments, which creates both substitution and income effects. Given the elasticity of intertemporal substitution, $1/\gamma = 2 > 1$, the substitution effect dominates. Thus, in this case agents save more to engage in more speculation. This leads to a greater aggregate investment ($K = 1.12$) than in homogeneous-beliefs benchmarks ($K = 1.03$).[23]

---

[23]In contrast, if $1/\sigma < 1$, then the income effect dominates and agents save less with heterogeneous beliefs relative to the homogeneous-beliefs benchmarks.
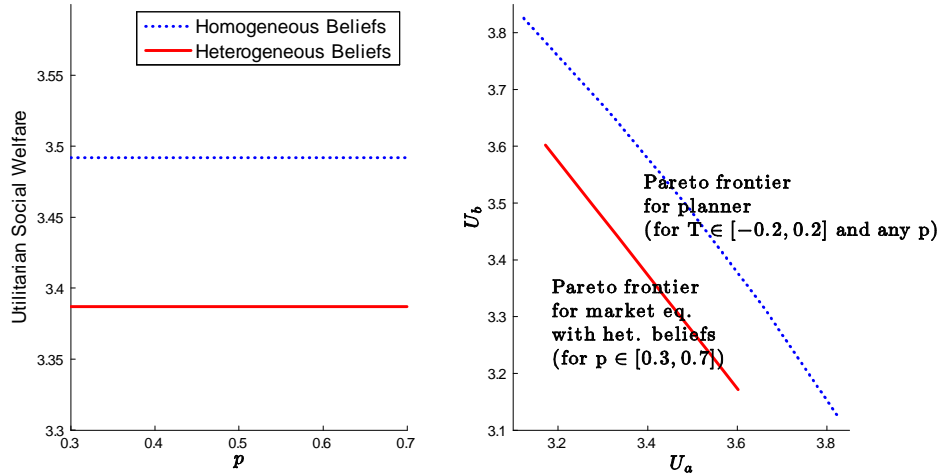
Figure 4: Social welfare and Pareto frontier in homogeneous-beliefs benchmarks and heterogeneous-beliefs equilibrium. The left panel plots the utilitarian social welfare based on any convex combination of the two types of beliefs. The right panel plots the Pareto frontiers respectively for the market equilibrium and a belief-neutral planner.

Taken together, this setting with heterogeneous beliefs exhibits two types of inefficiency: more volatile consumption and distorted savings (and investment). To discuss welfare implications of heterogeneous beliefs, we start by considering the utilitarian social welfare function. Instead of taking a stance on whose beliefs are superior, the planner evaluates the social welfare using any convex combination of the two types of beliefs: $\pi \in [\pi_a, \pi_b]$. The left panel of Figure 4 depicts the social welfare based on the equilibrium consumption of the two types of agents in the heterogeneous-beliefs and homogeneous-beliefs settings as π varies between π_a and π_b. Heterogeneous beliefs reduce the expected social welfare regardless of the belief measure one uses to evaluate the agents' expected utilities.

As before, this result holds for any welfare function because the market equilibrium is in fact belief-neutral Pareto inefficient. To illustrate this point, define $y(T)$ as an allocation in which a fraction, T, of all of agent B's endowments (bonds, goods, and shares of the representative firm) are transferred to agent A. For each T, consider the common-beliefs equilibrium starting with this initial allocation $y(T)$, which is a feasible allocation available to the planner. The case $T = 0$ corresponds to the homogeneous-beliefs benchmarks displayed in Table I. The second panel of Figure 4 plots the slightly curved Pareto frontier corresponding to this allocation as the transfer, T, varies. The same panel also plots the Pareto frontier for the equilibrium with heterogeneous beliefs as π, the belief the planner

uses to evaluate agents' expected utilities, varies between π_a and π_b. The figure shows that, for any belief $\pi \in [\pi_a, \pi_b]$, the equilibrium with heterogeneous beliefs is Pareto dominated. The intuition is the same as in the earlier sections: In this economy, more volatile consumption and distorted savings is sub-optimal *according to any reasonable belief*. A planner who corrects these inefficiencies can redistribute wealth to improve over the market equilibrium. This example demonstrates that our criterion is able to give clear welfare ranking in a macro setting with distorted consumption/savings decisions induced by heterogeneous beliefs.

4 Conclusion

This paper proposes a belief-neutral welfare criterion for models in which agents have heterogeneously distorted beliefs. The criterion builds on the premise that a planner is aware of belief distortions by some agents but cannot differentiate whose beliefs are distorted. The criterion rules that an allocation is belief-neutral efficient (inefficient) if it is efficient (inefficient) under any convex combination of the agents' beliefs. We can implement this criterion either through a given social welfare function or the notion of Pareto efficiency. While this criterion gives an incomplete welfare ranking, it is nevertheless useful in identifying negative-sum or positive-sum speculation. Through a series of examples, we show that this criterion is capable of identifying welfare gains/losses in a wide range of economic environments with heterogeneously distorted beliefs.

References

Akerlof, George (1970), The market for lemons, qualitative uncertainty and the market mechanism, *Quarterly Journal of Economics* 84, 488-500.

Aumann, Robert (1976), Agreeing to disagree, *The Annals of Statistics* 4, 1236-1239.

Barberis, Nicholas and Richard Thaler (2003), A survey of behavioral finance, in Constantinides, G., Harris, M., Stulz, R. (eds.), *Handbook of the Economics of Finance*, North-Holland, Amsterdam.

Basak, Suleyman (2000), A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk, *Journal of Economic Dynamics and Control* 24, 63-95.

Bernheim, Douglas and Antonio Rangel (2007), Toward choice-theoretic foundations for behavioral welfare economics, *American Economic Review Papers and Proceedings* 97, 464-470.

Bernheim, Douglas and Antonio Rangel (2009), Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics, *Quarterly Journal of Economics* 124, 51-104.

Bewley, Truman (2002), Knightian decision theory, Part I, *Decisions in Economics and Finance* 25, 79-110.

Bianchi, Javier, Emine Boz, and Enrique Mendoza (2012), Macro-prudential policy in a Fisherian model of financial innovation, working paper, University of Maryland.

Black, Fischer (1986), Noise, *Journal of Finance* 41, 529-543.

Bolton, Patrick, Jose Scheinkman, and Wei Xiong (2006), Executive compensation and short-termist behavior in speculative markets, *Review of Economic Studies* 73, 577-610.

Brunnermeier, Markus, Christian Gollier, and Jonathan A. Parker (2007), Optimal beliefs, asset prices, and the preference for skewed returns, *American Economic Review Papers and Proceedings* 97, 159-165.

Brunnermeier, Markus and Jonathan A. Parker (2005), Optimal expectations, *American Economic Review* 95, 1092-1118.

Buraschi, Andrea and Alexei Jiltsov (2006), Model uncertainty and option markets with heterogeneous beliefs, *Journal of Finance* 61, 2841-97.

Cao, Dan (2010), Collateral shortages, asset price and investment volatility with heterogeneous beliefs, working paper, MIT.

David, Alexander (2008), Heterogeneous beliefs, speculation, and the equity premium, *Journal of Finance* 63, 41-83.

Davila, Eduardo (2014), Optimal financial transaction taxes, working paper, Harvard University.

DellaVigna, Stefano (2009), Psychology and economics: Evidence from the field, *Journal of Economic Literature* 47, 315-372.

Detemple, Jerome and Shashidhar Murthy (1994), Intertemporal asset pricing with heterogeneous beliefs, *Journal of Economic Theory* 62, 294-320.

Doskeland, Trond and Hans Hvide (2011), Do individual investors have asymmetric information based on work experience?" *Journal of Finance* 66, 1011-1041.

Dreze, Jacques (1970), Market allocations under uncertainty, *European Economic Review* 2, 133-165.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal (2009), Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility, *Journal of Finance* 64, 579-629.

Dumas, Bernard, Karen Lewis, and Emilio Osambela (2011), Differences of opinion and international equity markets, working paper, Wharton School, University of Pennsylvania.

Fostel, Ana and John Geanakoplos (2008), Leverage cycles and the anxious economy, *American Economic Review* 98, 1211-1244.

Geanakoplos, John (2003), Liquidity, default, and crashes: endogenous contracts in general equilibrium, *Advances in Economics and Econometrics: Theory and Applications, Eighth World Conference*, Volume II, 170-205.

Geanakoplos, John (2009), The leverage cycle, *NBER Macroeconomics Annual* 24, 1-65.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny (2011), Neglected risks, financial innovation, financial fragility, *Journal of Financial Economics*, forthcoming.

Gilboa, Itzhak, Dov Samet, and David Schmeidler (2004), Utilitarian aggregation of beliefs and tastes, *Journal of Political Economy* 112, 932-938.

Gilboa, Itzhak, Larry Samuelson, and David Schmeidler (2012), No-betting Pareto dominance, *Econometrica*, forthcoming.

Gilchrist, Simon, Charles Himmelberg, and Gur Huberman (2005), Do stock price bubbles influence corporate investment? *Journal of Monetary Economics* 52, 805-827.

Grossman, Sanford and Joseph Stiglitz (1980), On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hammond, Peter (1981), Ex-ante and ex-post welfare optimality under uncertainty, *Economica* 48, 235-250.

Harris, Richard (1978), Ex-post efficiency and resource allocation under uncertainty, *Review of Economic Studies* 45, 427-436.

Harrison, Michael and David Kreps (1978), Speculative investor behavior in a stock market with heterogeneous expectations, *Quarterly Journal of Economics* 92, 323–336.

Hassan, Tarek and Thomas Mertens (2011), The social cost of near-rational investment, working paper, University of Chicago.

He, Zhiguo and Wei Xiong (2012), Debt financing in asset markets, *American Economic Review Papers and Proceedings*, forthcoming.

Hirshleifer, David (2001), Investor psychology and asset pricing, *Journal of Finance* 56, 1533-1597.

Hong, Harrison, Jose Scheinkman, and Wei Xiong (2006), Asset float and speculative bubbles, *Journal of Finance* 61, 1073-1117.

Hong, Harrison and David Sraer (2011), Quiet bubbles, working paper, Princeton University.

Jouini, Elyes and Clotilde Napp (2007), Consensus consumer and intertemporal asset pricing with heterogeneous beliefs, *Review of Economic Studies* 74, 1149-1174.

Koszegi, Botond and Matthew Rabin (2007), Mistakes in choice-based welfare analysis, *American Economic Review* 97, 477-481.

Kubler, F. and K. Schmedders (2012), Financial innovation and asset price volatility, *American Economic Review: Papers and Proceedings* 102, 147–151.

Kreps, David (2012), *Choice and Competitive Markets*, Princeton University Press, forthcoming.

Kurz, Mordecai (1994), On the structure and diversity of rational beliefs, *Economic Theory* 4, p.877-900.

Kurz, Mordecai (1996), Rational beliefs and endogenous uncertainty, *Economic Theory* 8, 383-397.

Mas-Colell Andreu, Michael Whinston, and Jerry Green (1995), *Microeconomic Theory*, Oxford University Press.

Milgrom, Paul and Nancy Stokey (1982), Information, trade and common knowledge, *Journal of Economic Theory* 26, 17-27.

Miller, E. (1977), Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Mongin, Philippe (1997), Spurious unanimity and the Pareto principle, working paper, Universite Cergy-Pontoise, France.

Morris, Stephen (1994), Trade with heterogeneous prior beliefs and asymmetric information, *Econometrica* 62, 1327-1347.

Morris, Stephen (1995), The common prior assumption in economic theory, *Economics and Philosophy* 11, 227-253.

Morris, Stephen (1996), Speculative investor behavior and learning, *Quarterly Journal of Economics* 110, 1111-1133.

Nielsen, Carsten Krabbe (2003), Floating exchange rates versus a monetary union under rational beliefs: the role of endogenous uncertainty, *Economic Theory*, 21, p.293-315.

Nielsen, Carsten Krabbe (2009a), Optimal economic institutions under rational overconfidence, with applications to the choice of exchange rate regime, *International Journal of Economic Theory*, 5, p.375-407.

Nielsen, Carsten Krabbe (2009b), Rational overconfidence and social security, discussion paper no. 916, The Institute of Economic Research, Korea University.

Panageas, Stavros (2006), The neoclassical q theory of investment in speculative markets, working paper, University of Chicago.

Posner, Eric and Glen Weyl (2013), An FDA for financial innovation: Applying the insurable interest doctrine to 21st century financial markets, *Northwestern University Law Review* 107, 1307-1485.

Sandroni, Alvaro and Francesco Squintani (2007), Overconfidence, insurance, and paternalism, *American Economic Review* 97, 1994-2004.

Savage, Leonard (1954), *The Foundation of Statistics*, John Wiley and Sons.

Scheinkman, Jose and Wei Xiong (2003), Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1219.

Sebenius, James and John Geanakoplos (1983), Don't bet on it: Contingent agreements with asymmetric information, *Journal of the American Statistical Association* 78, 424-426.

Shen, Ji, Hongjun Yan, and Jinfan Zhang (2011), Collateral-motivated financial innovation, working paper, Yale University.

Sims, Christopher (2008), Inflation expectations, uncertainty, and monetary policy, working paper, Princeton University.

Simsek, Alp (2013a), Speculation and risk sharing with new financial assets, *Quarterly Journal of Economics* 128, 1365-1396.

Simsek, Alp (2013b), Belief disagreements and collateral constraints, *Econometrica* 81, 1-53.

Spinnewijn, Johannes (2010), Unemployed but optimistic: Optimal insurance design with biased beliefs, working paper, London School of Economics.

Starr, Ross (1973), Optimal production and allocation under uncertainty, *Quarterly Journal of Economics* 87, 81-93.

Stiglitz, Joseph (1989), Using tax policy to curb speculative trading, *Journal of Financial Services Research* 3, 101-115.

Summers, Lawrence and Victoria Summers (1989), When financial markets work too well: A cautious case for a securities transaction tax, *Journal of Financial Services Research* 3, p.261-86.

Tirole, Jean (2012), Overcoming adverse selection: How public intervention can restore market functioning, *American Economic Review* 102, 29-59.

von Weizsäcker, Christian (1969), General equilibrium and uncertainty, teaching note, MIT.

Weyl, Glen (2007), Is arbitrage socially beneficial?, working paper, Princeton University.

Williams, Allan and Adrian Lund (1986), Seat belt use laws and occupant crash protection in the United States, *American Journal of Public Health* 76, 1438-42.

Wu, Ho-Mou and Wen-Chung Guo (2004), Asset price volatility and trading volume with rational beliefs, *Economic Theory* 23, 795-829.

Xiong, Wei and Hongjun Yan (2010), Heterogeneous expectations and bond markets, *Review of Financial Studies* 23, 1433-1466.

Zapatero, Fernando (1998), Effects of financial innovations on market volatility when beliefs are heterogeneous, *Journal of Economic Dynamics and Control* 22, 597-626.

Zhuk, Sergei (2012), Bubbles and technological innovations, working paper, Princeton University.

Appendix: Appendix

A.1   Technical Derivation for Section 3.1.1

The first-order condition to agent i's utility maximization implies

$$(1-p)\,\pi^i u'\left(w + k^i\,(1-p)\right) = p\left(1-\pi^i\right) u'\left(w - k^i\,(1+p)\right).$$

The market-clearing condition requires that $k^A + k^B = 0$. The standard results hold that there is a market equilibrium allocation, $\left\{k^A, k^B, p\right\}$, which solves each agent's optimality condition and the market-clearing condition.

First, we establish that, if $\pi^A \neq \pi^B$, the two agents will take a nonzero position in the contract. The first-order condition implies that

$$\frac{\pi^A}{1-\pi^A} \frac{u'\left(w + k^A\left(1-p\right)\right)}{u'\left(w - k^A p\right)} = \frac{\pi^B}{1-\pi^B} \frac{u'\left(w + k^B\left(1-p\right)\right)}{u'\left(w - k^B p\right)}.$$

Suppose that $k^A = k^B = 0$. Then, we must have $\frac{\pi^A}{1-\pi^A} = \frac{\pi^B}{1-\pi^B}$, which contradicts $\pi^A \neq \pi^B$. Thus, k^A and k^B cannot both be zero, which in turn implies that both are nonzero.

We now prove that if the social planner has the utilitarian welfare function $W\left(u_A, u_B\right) = u_A + u_B$, then the status quo allocation $y = \left\{\left(y_a^i, y_b^i\right) \equiv \left(w, w\right)\right\}_{i \in \{A,B\}}$ is belief-neutral superior to the market equilibrium allocation:

$$x = \left\{\left(x_a^i, x_b^i\right)\right\}_{i \in \{A,B\}} = \left\{\left(w + k^i\left(1-p\right), w - k^i\left(1+p\right)\right)\right\}_{i \in \{A,B\}}.$$

Consider any measure with $\pi \in \left[\pi^B, \pi^A\right]$. The agents' utilitarian social welfare in the market equilibrium is given by

$$\begin{aligned}
U^h &= \pi\left[u\left(w + k^A\left(1-p\right)\right) + u\left(w - k^A\left(1-p\right)\right)\right] \\
&\quad + \left(1-\pi\right)\left[u\left(w - k^A p\right) + u\left(w + k^A p\right)\right].
\end{aligned}$$

The strict concavity of $u\left(\cdot\right)$ implies that

$$\begin{aligned}
u\left(w + k^A\left(1-p\right)\right) + u\left(w - k^A\left(1-p\right)\right) &< 2u\left(w\right), \\
u\left(w - k^A p\right) + u\left(w + k^A p\right) &< 2u\left(w\right).
\end{aligned}$$

Thus, $U^h < \pi \cdot 2u\left(w\right) + \left(1-\pi\right) \cdot 2u\left(w\right) = 2u\left(w\right)$, which is the utilitarian social welfare under the status quo. This proves that the status quo allocation is belief-neutral superior to the market equilibrium allocation.

Next, we show that for any measure with $\pi \in \left(0, 1\right)$, the equilibrium allocation is Pareto dominated by the status quo allocation with a certain transfer $T^\pi \in \left[-w, w\right]$, which leads to the following allocation: $y\left(T^\pi\right) = \left\{\left(w + T^\pi, w + T^\pi\right), \left(w - T^\pi, w - T^\pi\right)\right\}$. Consider each agent's certainty-equivalent wealth, $w^{i,eq}$, given by the solution to

$$u\left(w^{i,eq}\right) = \pi u\left(w + k^i\left(1-p\right)\right) + \left(1-\pi\right) u\left(w - k^i p\right), \forall i \in \{A, B\}.$$

The strict concavity of $u\left(\cdot\right)$ (along with the fact that $k^i \neq 0$) implies that

$$u\left(w^{i,eq}\right) < u\left(\pi\left(w + k^i\left(1-p\right)\right) + \left(1-\pi\right)\left(w - k^i p\right)\right).$$

Since $u\left(\cdot\right)$ is strictly increasing, this further implies

$$w^{i,eq} < \pi\left(w + k^i\left(1-p\right)\right) + \left(1-\pi\right)\left(w - k^i p\right), \quad \forall i \in \{A, B\}.$$

Adding these inequalities and using market clearing, $k^A + k^B = 0$, we have $w^{A,eq} + w^{B,eq} < 1$. It follows that the status quo with an appropriate transfer Pareto dominates the equilibrium.

A.2 Technical Derivation for Section 3.1.2

Suppose that agent i $(i \in \{A, B\})$ takes on a position of k^i in the risky asset. The position is characterized by the following first-order condition:

$$(1 - p) \pi^A u' \left(w - z + k^A (1 - p)\right) = p \left(1 - \pi^A\right) u' \left(w + z - k^A (1 + p)\right).$$

This is also a similar condition for agent B. In equilibrium, the market-clearing condition is $k^A + k^B = 0$.

Recall that the optimal risk-sharing trade is given by $k^{*A} = z$ and $k^{*B} = -z$. An analysis similar to the previous proof shows that, when $\pi^A \neq \pi^B$, agents deviate from the optimal risk-sharing trade, that is, $k^A \neq k^{*A}$. Next, fix a belief, π, and consider each agent's certainty-equivalent wealth under this belief given by the solution to

$$
\begin{aligned}
u \left(w^{A,eq} (e, \pi)\right) &= \pi u \left(w - z + k^A (1 - p)\right) + (1 - \pi) u \left(w + z - k^A (1 + p)\right), \\
u \left(w^{B,eq} (e, \pi)\right) &= \pi u \left(w + z + k^B (1 - p)\right) + (1 - \pi) u \left(w - z - k^B (1 + p)\right).
\end{aligned}
$$

Since $k^A \neq k^{*A}$, an agent's consumption is not constant across the states. Using the strict concavity of $u(\cdot)$, we obtain

$$w^{A,eq} (e, \pi) + w^{B,eq} (e, \pi) < 2w.$$

On the other hand, the status quo allocation, combined with the optimal risk-sharing trade, $\left(k^{*A}, k^{*B}\right)$, gives each agent a constant consumption of w. It follows that, under any belief π, this allocation, combined with an appropriate transfer T^π, Pareto dominates the equilibrium allocation.

A.3 Technical Derivation for Section 3.1.3

To solve the subgame perfect equilibrium of the model, we first compute both agents' trading strategies at $t = 0$, while taking as given agent A's information acquisition decision, and then solve agent A's optimal information acquisition policy.

Consider the case in which agent A acquires the information at $t = 0$. Recall that agent A's trading position perfectly reveals the informational content of his private signal to agent B. Thus, both agents have the same information set at the time of choosing their optimal trading strategies. Let $W^i (S, s)$ denote the endowment of agent i in regime S and state s, where $S \in \{1, 2\}$ and $s \in \{a, b\}$. In view of log utility, the agent's problem can be written as

$$U_i^e (S) = \max_{k^i(S) \in \mathbb{R}} \left\{ \pi^i \ln \left[W^i (S, a) + k^i (S)\right] + \left(1 - \pi^i\right) \ln \left[W^i (S, b) - k^i (S)\right] \right\}. \quad (6)$$

The first-order condition for $k^i(S)$ is given by

$$\pi^i \frac{1}{W^i(S,a) + k^i(S)} - (1 - \pi^i) \frac{1}{W^i(S,b) - k^i(S)} = 0.$$

Note that, due to symmetry, the market-clearing condition $k^A(S) + k^B(S) = 0$ is automatically satisfied.

Consider now the case in which agent A does not acquire information at $t = 0$. As before, the agents' problems are symmetric. Specifically, agent i's problem is given by

$$U_i^e(N) = \max_{k^i(N) \in \mathbb{R}} \frac{1}{2} \left[\pi^i \ln \left[W^i(1,a) + k^i \right] + (1 - \pi^i) \ln \left[W^i(1,b) - k^i \right] \right. \tag{7}$$
$$\left. + \pi^i \ln \left[W^i(2,a) + k^i \right] + (1 - \pi^i) \ln \left[W^i(2,b) - k^i \right] \right].$$

The first-order condition for $k^i(N)$ is given by

$$\pi^i \frac{1}{W^i(1,a) + k^i} - (1 - \pi^i) \frac{1}{W^i(1,b) - k^i}$$
$$+ \pi^i \frac{1}{W^i(2,a) + k^i} - (1 - \pi^i) \frac{1}{W^i(1,a) - k^i} = 0.$$

As in the previous case, note that the market-clearing condition $k^A(N) + k^B(N) = 0$ is automatically satisfied.

Agent A acquires information if and only if

$$\frac{1}{2} U_A^e(1) + \frac{1}{2} U_A^e(2) - c \geq U_A^e(N),$$

where $U_A^e(N)$, given in (7), is agent A's expected utility by not acquiring the information and $U_A^e(S)$, given in (6), is his expected utility conditional on acquiring the signal and the signal reveals regime S. This condition is equivalent to

$$c \leq c^{eq}(\pi^A) \equiv \frac{1}{2} U_A^e(1) + \frac{1}{2} U_A^e(2) - U_A^e(N). \tag{8}$$

Here, $c^{eq}(\pi^A)$ denotes the cost threshold below which agent A chooses to acquire information, characterized by equations (6) and (7).

Having solved for agent A's information acquisition policy, we now apply our welfare criterion to determine whether the private provision or lack of provision of information is either 1) belief-neutral efficient, 2) belief-neutral inefficient, or 3) neither belief-neutral efficient nor inefficient.

The set of reasonable beliefs is given by

$$B_R(\pi^A) = \left[\min\left\{ \pi^A, 1 - \pi^A \right\}, \max\left\{ \pi^A, 1 - \pi^A \right\} \right].$$

For any reasonable belief $\Pr(a) \in B_R$, the social welfare obtained when acquiring information, $W(I)$, is given by

$$W(I) = \sum_{S=1}^{2} \sum_{s=a}^{b} \sum_{i=A}^{B} \frac{1}{2} \Pr(s) \left\{ \ln \left[W^i(S,s) + V(s) k^i(S) \right] - c I_{(i=A)} \right\},$$

while that obtained when no information is acquired, $W(N)$, is given by

$$W(N) = \sum_{S=1}^{2} \sum_{s=a}^{b} \sum_{i=A}^{B} \frac{1}{2} \Pr(s) \ln \left[W^i(S,s) + V(s) k^i \right].$$

Define the belief-neutral-efficient cost threshold $\bar{c}^{\text{eff}}\left(\pi^A\right)$ as

$$\bar{c}^{\text{eff}}\left(\pi^A\right) \equiv \sup \left\{ c \geq 0 : W(I) \geq W(N), \forall \Pr(a) \in B_R\left(\pi^A\right) \right\}$$

and the belief-neutral-inefficient cost threshold $\underline{c}^{\text{inef}}\left(\pi^A\right)$ as

$$\underline{c}^{\text{inef}}\left(\pi^A\right) \equiv \inf \left\{ c \geq 0 : W(I) \leq W(N), \forall \Pr(a) \in B_R\left(\pi^A\right) \right\}.$$

We now characterize $\bar{c}^{\text{eff}}\left(\pi^A\right)$ and $\underline{c}^{\text{inef}}\left(\pi^A\right)$, and prove that $\bar{c}^{\text{eff}}\left(\pi^A\right) = \underline{c}^{\text{inef}}\left(\pi^A\right), \forall \pi^A \in [0,1]$. Define $x^i(S,s)$ and $y^i(S,s)$ as trader i's final wealth in state (S,s) when trader A acquires and does not acquire information, respectively:

$$x^i(S,s) \equiv W^i(S,s) + V(s) k^i(S)$$

and

$$y^i(S,s) \equiv W^i(S,s) + V(s) k^i(N).$$

From the market-clearing condition, it follows that

	$S = 1$		$S = 2$	
	$s = a$	$s = b$	$s = a$	$s = b$
$x^A(S,s)$	$w + k^A(1)$	$w - k^A(1)$	$w - z + k^A(2)$	$w + z - k^A(2)$
$x^B(S,s)$	$w - k^A(1)$	$w + k^A(1)$	$w + z - k^A(2)$	$w - z + k^A(2)$
$y^A(S,s)$	$w + k^A(N)$	$w - k^A(N)$	$w - z + k^A(N)$	$w + z - k^A(N)$
$y^B(S,s)$	$w - k^A(N)$	$w + k^A(N)$	$w + z - k^A(N)$	$w - z + k^A(N)$

It follows directly from the symmetry in the agents' payoffs that both $W(I)$ and $W(N)$ are independent of the belief that the planner uses to evaluate the social welfare:

$$W(I) = \sum_{S=1}^{2} \frac{1}{2} \left\{ \ln \left[x^A(S,a) \right] + \ln \left[x^B(S,a) \right] \right\} - c$$

and

$$W(N) = \sum_{S=1}^{2} \frac{1}{2} \left\{ \ln \left[y^A(S,a) \right] + \ln \left[y^B(S,a) \right] \right\}.$$

Accordingly, we have

$$\bar{c}^{\text{eff}}\left(\pi^A\right) = \underline{c}^{\text{inef}}\left(\pi^A\right) = \sum_{S=1}^{2} \frac{1}{2} \left\{ \ln \left[x^A(S,a) \right] + \ln \left[x^B(S,a) \right] - \ln \left[y^A(S,a) \right] - \ln \left[y^B(S,a) \right] \right\},$$

$$(9)$$

completing the characterization. The main text compares the cost thresholds characterized in (8) and (9) to assess the belief-neutral efficiency of agent A's private information acquisition decision.

A.4 Characterization of Equilibrium in Section 3.5

The following proposition summarizes the market equilibrium:

Proposition 2 *Depending on type-A agents' cash endowment c, the following five cases can emerge in equilibrium.*

- *Case 1: $c < c_1$, where $c_1 = E_0^B[\widetilde{R}] - K_d$. In this case, type-A agents acquire the asset at $t = 0$ by using a one-period debt contract with a promise of K_d. However, their purchasing capacity is insufficient to lift the asset price, p_0, above type-B agents' expectation of the asset's fundamental value. Consequently, $p_0 = E_0^B[\widetilde{R}]$.*

- *Case 2: $c \in [c_1, c_2)$, where $c_2 = p_0^* - K_d$ and*

$$p_0^* = \frac{\left(2 - \pi^A\right)\pi^A[\pi^B + (1 - \pi^B)\theta](1 - \theta) - [\pi^A(1 - \pi^B)(1 - \theta) - (1 - \pi^A)^2\alpha]\theta}{(2 - \pi^A)\pi^A(1 - \theta) - \pi^A(1 - \pi^B)(1 - \theta) + (1 - \pi^A)^2\alpha}. \quad (10)$$

  In this case, type-A agents acquire the asset at $t = 0$ by using one-period debt contract with a promise of K_d. The asset price p_0 is given by type-A agents' aggregate purchasing capacity: $p_0 = c + K_d$.

- *Case 3: $c \in [c_2, c_3)$, where $c_3 = p_0^* - \theta$. In this case, type-A agents acquire the asset at $t = 0$ and are indifferent to using debt contracts with promises of θ and K_d. The asset price p_0 remains at a constant level $p_0 = p_0^*$. The fraction of borrowers who choose to use debt face value K_d is given by equation (11) below.*

- *Case 4: $c \in [c_3, c_4)$, where $c_4 = E_0^A[\widetilde{R}] - \theta$. In this case, type-A agents acquire the asset by using riskless debt with a promise of θ. The asset price p_0 is determined by their aggregate purchasing capacity: $p_0 = c + \theta$.*

48

- *Case 5: $c \geq c_4$, where $c_4 = E_0^A[\widetilde{R}] - \theta$. In this case, type-A agents have ample cash endowments to support their asset acquisition at a price equal to their expectation of the asset's fundamental value, $p_0 = E_0^A[\widetilde{R}]$, by using debt with a promised payment less than θ.*

We prove this proposition in two steps. First, we characterize type-A agents' optimal debt contract. We show that the relevant debt contracts are short-term debt with face value θ and K_d, and we characterize the choice between these two contracts. Second, we consider market clearing and characterize the equilibrium price for cases 1-5. In each case, we also show that (unlike in Geanakoplos, 2009) type-A agents do not have an incentive to hold cash to buy assets in state d of date 1. In particular, type-A agents use all of their purchasing power to buy the assets at date 0.

Step 1. First consider type-A agents' debt contract choice. We start with short-term debt with maturity at $t = 1$. It can be seen that the face value of short-term debt should lie in the range of $[\theta, 1]$, i.e., between the two possible payoffs of the collateral. If the agent chooses to borrow short-term debt at $t = 0$, he has to roll over his debt at $t = 1$. If he fails to obtain refinancing, he will default and incur a personal liquidation cost of α. In state u, the subsequent asset payoff is surely 1; thus there is no problem rolling over the debt. In state d, the maximum debt financing the borrower can obtain from the pessimistic creditors is

$$K_d = E_d^B[\widetilde{R}] = \pi^B + \left(1 - \pi^B\right)\theta.$$

Thus, the borrower is able to structure a new debt contract with creditors if his initial debt promise is not higher than K_d. By making a new promise of F_d, he can obtain the following credit to repay his initial debt:

$$C(F_d) = \begin{cases} F_d & \text{if } F_d \leq \theta, \\ \pi^B F_d + (1 - \pi^B)\theta & \text{if } \theta < F_d \leq 1. \end{cases}$$

Note that the new debt is risk-free if $F_d \leq \theta$ or risky if $\theta < F_d \leq 1$. In the latter case, the lender will be paid with F_d in the good du state but receive the asset in the bad dd state. Thus, if the borrower's initial debt promiseF_0 is lower than or equal to K_d, he can obtain refinancing even in the lower state d at $t = 1$; and if F_0 is higher than K_d, he will have to default in the lower state d.

We now discuss the borrower's debt promise choice in using short-term debt. First consider the range, $[\theta, K_d]$. If the borrower promises $F_0 = \theta$, he can obtain an initial credit of θ, which allows him to establish an initial position of $c/(p_0 - \theta)$ units of asset. The

expected return on his cash is

$$R_0^\theta = \frac{(2 - \pi^A)\pi^A(1 - \theta)}{p_0 - \theta}.$$

If he chooses a promise $F_0 \in (\theta, K_d]$, he can obtain an initial credit of F_0. The expected return on his cash after accounting for the possible liquidation cost α is

$$
\begin{aligned}
R_0^S &= \frac{\pi^A(1 - F_0) + (1 - \pi^A)\pi^A(1 - F_d) + (1 - \pi^A)^2(-\alpha)}{p_0 - F_0} \\
&= \frac{\pi^A(1 - F_0) + (1 - \pi^A)\frac{\pi^A}{\pi^B}[\pi^B + (1 - \pi^B)\theta - F_0] + (1 - \pi^A)^2(-\alpha)}{p_0 - F_0}.
\end{aligned}
$$

Note that while he can refinance his initial debt in state d on date 1, he will eventually default in state dd on date 2. It is straightforward to verify that $\frac{dR_0^S}{dF_0} < 0$ if and only if

$$p_0 > \widetilde{p}_0^* \equiv \frac{\pi^A + (1 - \pi^A)\frac{\pi^A}{\pi^B}[\pi^B + (1 - \pi^B)\theta] - (1 - \pi^A)^2\alpha}{\pi^A + (1 - \pi^A)\frac{\pi^A}{\pi^B}}.$$

Thus, if $p_0 > \widetilde{p}_0^*$, $F_0 = \theta$ is the optimal choice. If $p_0 = \widetilde{p}_0^*$, any $F_0 \in (\theta, K_d]$ would yield the same expected return. If $p_0 < \widetilde{p}_0^*$, $F_0 = K_d$ is superior to any promise in (θ, K_d). But we still need to compare this choice with $F_0 = \theta$ debt. Suppose that at a critical level p_0^*, the expected returns from $F_0 = \theta$ and K_d are equal:

$$\frac{\pi^A(1 - K_d) + (1 - \pi^A)^2(-\alpha)}{p_0^* - K_d} = \frac{(2 - \pi^A)\pi^A(1 - \theta)}{p_0^* - \theta}$$

which gives

$$p_0^* = \frac{[1 - (1 - \pi^A)^2][\pi^B + (1 - \pi^B)\theta](1 - \theta) - [\pi^A(1 - \pi^B)(1 - \theta) - (1 - \pi^A)^2\alpha]\theta}{[1 - (1 - \pi^A)^2](1 - \theta) - \pi^A(1 - \pi^B)(1 - \theta) + (1 - \pi^A)^2\alpha} < \widetilde{p}_0^*.$$

Therefore, if $p_0 < p_0^*$, $F_0 = K_d$ is the optimal face value; if $p_0 > p_0^*$, $F_0 = \theta$ dominates; when $p_0 = p_0^*$, the borrower is indifferent between $F_0 = K_d$ and θ.

We now consider short-term debt with promise higher than K_d. For such a choice, the debt is no longer riskless as the borrower cannot refinance it in state d on date 1 and has to turn over the asset to the creditor. Anticipating this possibility, the creditor is willing to grant the following credit on date 0:

$$C_0(F_0) = \pi^B F_0 + (1 - \pi^B)[\pi^B + (1 - \pi^B)\theta].$$

Then, the expected return to the borrower is

$$R_0^S = \frac{\pi^A(1 - F_0) + (1 - \pi^A)(-\alpha)}{p_0 - \pi^B F_0 - (1 - \pi^B)[\pi^B + (1 - \pi^B)\theta]}.$$

It is straightforward to verify that $\frac{dR_0^S}{dF_0} < 0$ iff

$$p_0 > p_0^* \equiv 1 - (1 - \pi^B)^2 + (1 - \pi^B)^2\theta - \frac{\pi^B}{\pi^A}(1 - \pi^A)\alpha.$$

Note that the asset price p_0 is bounded from below by the asset valuation of pessimists

$$E_0^B[\widetilde{R}] \equiv 1 - (1 - \pi^B)^2 + (1 - \pi^B)^2\theta.$$

As $E_0^B[\widetilde{R}] > p_0^*$, it is not optimal for the borrower to choose a debt promise above K_d.

It is also straightforward to verify that under condition (4), $p_0^* > E_0^B[\widetilde{R}]$. Therefore, the borrower's optimal short-term debt promise at $t = 0$ is

$$F_0 = \begin{cases} K_d, & \text{if } p_0 \in [E_0^B[\widetilde{R}], p_0^*); \\ \theta \text{ or } K_d, & \text{if } p_0 = p_0^*; \\ \theta, & \text{if } p_0 \in (p_0^*, E_0^A[\widetilde{R}]]. \end{cases}$$

Step 2. We now discuss different cases based on group-A agents' cash endowment c from high to low, in reverse order from those cases listed in Proposition 2

- Case 5: $c \geq c_4$.

In this case, the asset price is determined by type-A agents' beliefs at each date. Moreover, at these prices, type-A agents are able to finance their asset acquisition by using debt with promise less than θ. In fact, each type-A agent is indifferent between acquiring or not acquiring the asset. To ensure this case holds true, c has to satisfy

$$c \geq c_4 \equiv E_0^A\left[\widetilde{R}\right] - \theta.$$

- Case 4: $c_3 \leq c < c_4$.

In this case, type-A agents use debt with promise θ to finance their asset acquisition. However, their aggregate purchasing power is unable to sustain the price at their asset valuation. Instead, at $t = 0$, the price is determined by their purchasing power:

$$p_0 = c + \theta.$$

Going forward, in state d of date 1, type-A agents can still refinance their debt and thus keep the asset price at their valuation, i.e., $p_d = E_d^A\left[\widetilde{R}\right]$. To ensure that optimists' debt contract choice is optimal, we need to ensure that $p_0 > p_0^*$, which is equivalent to

$$c > c_3 \equiv p_0^* - \theta.$$

We next check type-A agents' incentive to save cash to date 1 in this case. First consider their return from buying at date 0 (and holding until date 2), which is given by:

$$\frac{[\pi^A + (1 - \pi^A)\pi^A](1 - \theta)}{p_0 - \theta} > 1,$$

where the inequality follows since $p_0 \in [p_0^*, E_0^A[\widetilde{R}])$. If instead they save cash to date 1, they will have to buy the asset from other type-A agents (since these agents hold all the assets in the conjectured equilibrium). In view of liquidation costs, other type-A agents would sell at a price $E_d^A[\widetilde{R}] + \alpha$. Thus, the return from saving cash is given by:

$$\pi^A + (1 - \pi^A)\frac{\pi^A(1 - \theta)}{E_d^A[\widetilde{R}] + \alpha - \theta} < 1.$$

Thus, type-A agents have no incentive to save cash.

- Case 3: $c_2 \le c < c_3$.

In this case, type-A agents are indifferent to using debt with promises of θ and K_d to purchase asset at price p_0^*. The expected return is

$$
\begin{aligned}
& \frac{\left[\pi^A + (1 - \pi^A)\pi^A\right](1 - \theta)}{p_0^* - \theta}. \\
= {}& \frac{[1 - (1 - \pi^A)^2](1 - \theta) - \pi^A(1 - \pi^B)(1 - \theta) + (1 - \pi^A)^2\alpha}{\pi^B(1 - \theta)} \\
> {}& \frac{[1 - (1 - \pi^A)^2](1 - \theta) - \pi^A(1 - \pi^B)(1 - \theta)}{\pi^B(1 - \theta)} = \pi^A + (1 - \pi^A)\frac{\pi^A}{\pi^B},
\end{aligned}
$$

where the equality follows from the definition of p_0^* in (10).

Next consider a type-A agent, which we refer to as an arbitrageur, and consider his incentive to save cash to date 1. If the state goes to u at $t = 1$, the arbitrageur cannot profit from his cash. If the state goes to d, he can potentially profit. He has three options. First, he could buy the asset from type-A agents who initially purchased with a debt contract with face value θ. To buy from these agents, the arbitrageurs would have to pay $p_d^{liq} = \alpha + E_d^A[\tilde{\theta}]$, which exceeds her valuation. Second, he could buy from type-A agents who initially purchased with a debt contract with face value K_d. These agents are distressed in the sense that they have collateralized all of their asset in exchange for K_d. At the same time, they incur a liquidation cost, α, from selling the asset at date 1. If instead they wait until date 2, then they incur the liquidation cost only if state dd is realized. Thus, they would be willing to sell the asset to the arbitrageur at a price:

$$p_d^{liq} = K_d - (1 - \pi^A)\alpha + \alpha.$$

Third, instead of buying the asset, the arbitrageurs could also refinance the debt contract of other optimists. This gives a payoff of K_d. The expected return to holding cash at date $t = 0$ is:

$$\pi^A + (1 - \pi^A)\frac{\pi^A(1 - \theta)}{K_d - \theta} = \pi^A + (1 - \pi^A)\frac{\pi^A}{\pi^B}.$$

This shows that taking an asset position at $t = 0$ dominates saving cash.

Next consider the fraction of optimists, μ, that uses debt with promise K_d. By market clearing, μ is determined as the solution to:

$$(1 - \mu)\frac{c}{p_0^* - \theta} + \mu\frac{c}{p_0^* - K_d} = 1. \tag{11}$$

At the lower end of the region c_2, $\mu = 0$, i.e., all optimists use short-term debt with promise K_d. Thus,

$$c_2 = p_0^* - K_d.$$

- Case 2: $c_1 \leq c < c_2$.

In this case, each optimist uses debt with promise K_d to finance his asset acquisition at $t = 0$, and the asset price is determined by the aggregate purchasing power of the optimists:

$$p_0 = c + K_d < p_0^*.$$

As the asset price is even lower than the previous case, the expected return to an optimist from taking a levered position with debt promise K_d is at least $\pi^A + (1 - \pi^A)\frac{\pi^A}{\pi^B}$. However, the expected return from saving cash is at most $\pi^A + (1 - \pi^A)\frac{\pi^A}{\pi^B}$. Thus, there is no incentive for any optimist to save cash at $t = 0$.

Once the optimists' cash endowment drops to a critical level c_1, the asset price becomes the pessimists' asset valuation: $E_0^B[\widetilde{R}]$. This determines c_1:

$$c_1 = E_0^B[\widetilde{R}] - K_d.$$

- Case 1: $c < c_1$.

In this case, each optimist acquires the asset by using debt with promise K_d, but his aggregate purchasing power is insufficient to maintain a level above the pessimists' valuation. The low price implies a high expected return, which makes it undesirable for any optimist to save cash at $t = 0$. This completes the proof of Proposition 2.