A Welfare Criterion for Models with Distorted Beliefs*

Markus K. Brunnermeier† Alp Simsek‡ Wei Xiong§

October 2012

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 an allocation to be *belief-neutral* inefficient if it is inefficient under any convex combination of agents' beliefs. While this criterion gives an incomplete ranking of social allocations, it can identify negative-sum speculation in a broad range of prominent models with distorted beliefs.

*We thank John Geanakoplos, David Hirshleifer, Stephen Morris, Wolfgang Pesendorfer, Chris Sims, and seminar participants at Princeton University, 2012 Behavioral Finance Conference in Amsterdam, 8th Annual Cowles Foundation Conference on General Equilibrium and Its Applications, and 2nd Miami Behavioral Finance Conference for helpful discussion and comments. We also thank Zhenyu Gao and Ji Huang for able research assistance.

†Princeton University and NBER. Email: markus@princeton.edu.

‡Harvard University and NBER. Email: asimsek@fas.harvard.edu.

§Princeton University and NBER. Email: wxiong@princeton.edu.

1 Introduction

The recent financial crisis and subsequent great recession have led to a flurry of new financial regulations. Many measures were undertaken to reduce incentive distortions caused by externalities. While incentive distortions played a prominent role in the build-up of the recent crisis, misconceptions and belief distortions were arguably even more important. Most investors believed a drop in house prices by more than 20% was unrealistic and they underestimated house price correlations across regions.[1] Despite the importance of belief distortions in the formation of bubbles and the subsequent crisis, belief distortions had virtually no impact on financial regulations. This is because economics does not offer a clear welfare criterion.

This paper tries to fill that gap by providing a welfare criterion for models in which individuals hold heterogeneously distorted beliefs. To illustrate the basic idea, we first consider a bet between Joe Stiglitz and Bob Wilson.[2] One day, Joe and Bob argued over the contents of a pillow. Joe maintained that the pillow had a natural down filling, while Bob thought a synthetic filling was more likely. Joe assessed with probability 0.9 that the filling was natural and Bob assessed the probability of 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 only discover the truth by cutting the pillow open, which would destroy it. They agreed to share the cost of buying a new pillow (\$50). 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 \$55 after splitting 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's being destroyed.

To evaluate the social welfare of the bet, it is useful to differentiate two distinct sources of their conflicting beliefs. First, a large body of decision theory literature builds on the personalistic view of probability as advocated by Savage (1954). This literature holds that beliefs reflect personal experience and risk attitude, and are an integrated part of individuals' preferences under uncertainty.[3] According to this view, Joe's and Bob's beliefs might have

[1] See Cheng, Raina, and Xiong (2012) for evidence that finance industry employees showed little concern right before the crisis about the possibility of a housing market crash in their personal home transactions.

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

[3] See Morris (1995) for extensive arguments that rationalize heterogeneous prior beliefs.

reflected their preferences for betting under the particular circumstance rather than their views on the likelihood of the outcome.

Alternatively, psychological biases might have distorted the belief of one (or both) of them. One particularly prominent bias is overconfidence, which is commonly observed among successful individuals in experimental studies. As a result, Joe and Bob held drastically different beliefs, which, in turn, motivated them to take on the value-destroying bet. Under this interpretation, they chose to bet because each believed he would win and the other would lose, instead of simply enjoying the bet. Our welfare criterion emphasizes this concern. In general, the academic literature has widely recognized that people suffer from a range of well-documented behavioral biases that can distort their beliefs and cause them to take actions harmful to the welfare of themselves and others.[4] The presence of distorted beliefs challenges the standard notion of Pareto efficiency, and motivates a benevolent social planner to use the correct belief to evaluate the agents' welfare. Under this paternalistic view, the behavioral economics and finance literature commonly assumes the existence of an objective belief measure and allows the social planner to use it to evaluate welfare issues.

However, a challenge arises because the social planner may not observe the objective belief measure. Commonly, in realistic economic situations, available data does not allow the social planner to discriminate between different beliefs. In fact, such an environment fosters belief distortions among agents in the first place. To cope with this challenge, we acknowledge the relevance of a set of reasonable beliefs and propose to use any reasonable belief as the common belief measure to evaluate all agents' expected utilities despite their different beliefs. The key insight of our welfare criterion is that if a social allocation is (in)efficient under any reasonable belief, then it is *belief-neutral* (in)efficient.

Specifically, we accept any convex combination of agents' beliefs as a reasonable belief and propose to use all of them to extend the two standard welfare analysis approaches—the expected social welfare approach and the Pareto efficiency approach. In implementing our welfare criterion, we strictly interpret agents' beliefs as their views of likelihood of economic outcomes, and incorporate all other aspects such as agents' risk-seeking preferences and preference-driven differences in prior beliefs by appropriate choice of their (state dependent) utility functions.

The expected social welfare approach directly compares two social allocations x and y

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

for a given welfare function. Our welfare criterion posits that x is belief-neutral inferior to y if the expected total welfare from x is lower than that from y under *any* convex combination of the agents' beliefs. Regarding the bet between Joe and Bob, it is difficult for the social planner to tell ex ante who was right. However, it is reasonable to assume that the objective probability that the pillow had natural down lay between their beliefs (i.e., between 0.1 and 0.9). Suppose that Joe and Bob are both risk neutral and that the social planner assigns them equal weights in summing up their utilities in the social welfare function. Then, it is immediately 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 any social welfare function, we can also adopt the Pareto dominance approach. Our criterion asserts that an allocation x is belief-neutral Pareto inefficient if under any reasonable belief, there exists an alternative allocation x' 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 to be destroyed. 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 any convex combination of Joe's and Bob's beliefs, the planner can always find a suitable (belief-measure dependent) transfer without the bet to strictly improve everyone's expected utility. Thus, the bet is belief-neutral inefficient under any welfare function that increases with agents' utilities.

In this example, without taking a stand on which belief is correct, the planner can categorically determine that the bet leads to an inefficient social outcome. The key is that the bet is a negative-sum game. This attribute is also present in many other models with distorted beliefs—agents are willing to speculate against each other, as each believes he will win at the expense of the other parties, even though the game has a negative sum. Our criterion is particularly useful in detecting this type of negative-sum speculation, despite that it requires consistent identification of efficiency of a social allocation under different reasonable beliefs and is therefore necessarily incomplete. Our criterion thus extends the "externality view" to settings with distorted beliefs. Interestingly, the Coase Theorem fails in these settings, as bargaining and market trading cannot lead to an efficient outcome even

if there are no transaction costs and property rights are well defined.

In Section 3, we apply our welfare criterion to a set of examples and show that it provides clear welfare ranking for almost all prominent models with heterogeneously distorted beliefs in the literature. Our first three 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 that in these models, trading costs make trading a negative-sum game just like the bet between Joe and Bob. Indeed, by analyzing a large sample of brokerage accounts held by individual households, Barber and Odean (2000) show that trading costs led to severe under-performance of those who trade most often. Our second example highlights how overinvestment induced by price bubbles makes speculative trading a negative-sum game even in the absence of trading costs (e.g., Bolton, Scheinkman and Xiong (2006), Gilchrist, Himmelberg, and Huberman (2005), and Panageas (2006)). Our third 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 fourth example builds on leverage cycles caused by heterogeneous beliefs (e.g., Geanakoplos (2003, 2009), Fostel and Geanakoplos (2008), Simsek (2010), 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 and thus regulatory implications of such leverage cycles.

The fifth example concerns excessive risk takings induced by agents' heterogeneous beliefs in general equilibrium models of asset markets (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)). By making agents' consumptions more volatile than their endowments, the trading induced by heterogeneous beliefs is a negative-sum game in expected utility terms regardless of the belief the planner uses to evaluate agents' expected utilities. Of course, in more general settings, some trading allows agents with risky en-

dowment streams to share endowment risks. In that case, there is a trade-off between the welfare gain from risk sharing and the welfare loss from speculative trading (e.g., Kubler and Schmedders (2011), Simsek (2011), and Posner and Weyl (2012)). Our criterion provides a tool to analyze such a trade-off, which is particularly relevant in the ongoing debate regarding the roles of financial innovations in facilitating hedging and speculation.

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 leads to unappealing outcomes when agents hold conflicting beliefs. Early general equilibrium literature, e.g., von Weizsäcker (1962), Dreze (1970), Starr (1973), Harris (1978), Hammond (1981), noted that an allocation that is Pareto optimal in the usual sense might feature less-than-perfect risk sharing. This literature made a distinction between ex ante efficiency and various versions of ex post efficiency (with better risk sharing properties), and went on to characterize the properties of ex post efficient allocations. In contrast to this literature, we keep the emphasis on ex ante welfare but propose to use all reasonable beliefs. As a consequence, our criterion is able to identify speculation induced by heterogeneous beliefs as belief-neutral efficient or inefficient in a variety of different environments, some of which do not feature any risk sharing considerations (such as the bet between Joe and Bob).

Independent decision theory literature, e.g., Mongin (1997) and Gilboa, Samet, and Schmeidler (2004), has also pointed out that the standard Pareto principle can be spurious when agents hold conflicting beliefs. To circumvent this problem, Gilboa and Schmeidler (2012) propose the so-called rationalizable Pareto dominance, which augments the standard Pareto condition by requiring the existence of a set of common beliefs to further rationalize a Pareto efficient choice. This extra restriction prevents the bet between Joe and Bob from rationalizably Pareto dominating no betting. It also makes the ranking more incomplete and gives no indication for either betting or no betting being more efficient. In contrast, our criterion identifies the bet having a negative sum and thus being belief-neutral inefficient.

Our paper is part of a growing literature that analyzes welfare in the presence of behav-

ioral biases (see Bernheim and Rangel (2005) for a survey). We focus on heterogeneously distorted beliefs and provide a discerning welfare criterion in this context. As with the other papers that study belief distortions, our criterion replaces individuals' revealed preferences with preferences that would obtain under an "objective" belief. Our contribution is to show that, with heterogeneous distortions, this approach leads to useful welfare comparisons even if the planner does not know the objective belief. Gabaix and Laibson (2006), Weyl (2007), Spinnewijn (2010), and Gennaioli, Shleifer, Vishny (2011) assume that the social planner knows the objective belief measure, while Nielsen (2003) adopts the rational belief in the sense of Kurz (1994).

Recent contributions in behavioral welfare economics, e.g., Bernheim and Rangel (2007, 2009) and Koszegi and Rabin (2008), have (rightly, in our opinion) set a relatively high standard for overruling revealed preference. For example, Bernheim and Rangel (2009) suggest that individuals' choices should be disregarded only if it can be demonstrated, using sound non-choice evidence, that individuals have made those choices under an incomplete understanding of their choice sets. In our context with distorted beliefs, we could provide two types of non-choice evidence for overruling revealed preference. First, a growing psychology literature suggests that individuals have limited capacity for Bayesian inference: That is, they often operate in a cognitive mode that leads to systematic mistakes in estimating probabilities from observed data (see Kahneman, 2011). One could make the case that individuals would behave differently if they were cognitively capable of Bayesian reasoning. Second, we could also check the validity of individuals' choices by considering their reported intentions. In the above example, suppose both Joe and Bob report that they are betting precisely because they expect to make money at the expense of their counterparty. Given that they cannot both make money, one could argue that some of these choices must be based on an incorrect understanding of the environment (although we do not know whose choices). Consequently, reported intentions suggest that individuals' choices in this example are collectively unreasonable.

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 popular models 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 T periods and $T+1$ dates: $t = 0, 1, ..., T$. The evolution of the state of the economy is represented by a binomial-tree process: $\{s_t\}_{t=0}^{T}$. In each period, the state variable can either increase or decrease by a discrete level. The tree is recombining and can take $t+1$ possible values on date t.

There are N agents, indexed by $i \in \{1, 2, ..., N\}$. On each date, each agent holds a belief about the probability of the tree increasing in the following period, which we denote by $\pi_{t,s}^{i}$, where t is the date and s is a state on the date. As this belief can vary across dates and states, we summarize agent i's beliefs by $\Pi^{i} = \left\{\pi_{t,s}^{i}\right\}$. We restrict the agent's belief in each period and each state to be strictly positive: $\pi_{t,s}^{i} > 0$. One can determine the agent's probability assessments of all future states from Π^{i}.

Suppose that the agents consume only on the final date T. A social choice x represents a set of consumption allocations to the agents across the final states s_T: $x = \{x_T^{i}(s_T)\}$. A feasible allocation satisfies the aggregate budget constraint in each final state.

Suppose that agent i has state-dependent utility function $u_i[s_T, x_T^{i}(s_T)]$, which 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 and, as we will discuss later, to accommodate differences in agents' preference-driven prior beliefs. 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.

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 to 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 the welfare of themselves and others. Thus, it is important that a social planner evaluates their 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 can cause rational agents with a common prior belief neither to hold common knowledge heterogeneous beliefs nor to trade with each other. This result motivates us to mostly ignore asymmetric information in our analysis, except in our example considered in Section 3.3.

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 individuals' background and experience. Morris (1995) summarized a series of arguments to advocate the view that rational agents may hold heterogeneous prior beliefs, just like heterogeneous risk preferences.[5]

Wilson (1968) adopts the Savage view of subjective probabilities to analyze social welfare in a setting with heterogeneous priors. As prior beliefs are part of each agent's preferences, the social planner uses each agent's own beliefs to determine his expected future utility, i.e.,

[5]Heterogeneous prior beliefs can also endogenously arise from agents' anticipatory utility (e.g., Brunnermeier and Parker (2005).)

$E_0^i\left[u_i\left[s_T, x_T^i\left(s_T\right)\right]\right]$ for agent i. Suppose the social planner holds a probability measure of his own, which we denote by Π^{SP}. We can rewrite agent i's expected utility as

$$E_0^i\left[u_i\left[s_T, x_T^i\left(s_T\right)\right]\right] = \sum_{s_T} \pi^{SP}\left(s_T\right) \frac{\pi^i\left(s_T\right)}{\pi^{SP}\left(s_T\right)} u_i\left[s_T, x_T^i\left(s_T\right)\right] = E_0^{SP}\left[\frac{\pi^i\left(s_T\right)}{\pi^{SP}\left(s_T\right)} u_i\left[s_T, x_T^i\left(s_T\right)\right]\right],$$

where $\frac{\pi^i(s_T)}{\pi^{SP}(s_T)}$ is the Radon-Nikodym derivative of agent i's probability measure with respect to the social planner's measure. The product $\frac{\pi^i(s_T)}{\pi^{SP}(s_T)} u_i\left[s_T, x_T^i\left(s_T\right)\right]$ acts as the agent's effective utility under the social planner's probability measure. Thus, one can always incorporate agents' preference-driven heterogeneous prior beliefs by their state-dependent utility functions. In the rest of the paper, we treat agents' heterogeneous beliefs as caused by their distorted beliefs, assuming that any heterogeneity in their priors has been already moved into their utility functions.

2.2 Welfare Analysis with Distorted Beliefs

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

Without taking a stand on which agent's beliefs are correct, we allow the objective probability measure to coincide with either the beliefs of one of the agents or a convex combination of their beliefs. In other words, the objective probability measure lies between the agents' beliefs. Denote Π^h to be a convex combination of the agents' beliefs with weight $h = \left\{h^1, ..., h^N\right\}$:

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

The space spanned by $\left\{\Pi^h\right\}$ contains a large set of reasonable belief measures based on the given environment. Because any measure outside this set implies categorical biases in agents' aggregate beliefs, we do not include such measures in our analysis in order to focus on welfare implications of speculation induced by heterogeneous beliefs, as opposed to inefficiencies associated with aggregate biases.[6]

[6]It shall become clear later that one can extend the set of reasonable beliefs to include any measure that

The key insight of our welfare criterion is to analyze the efficiency of a social allocation across all of these reasonable probability measures. Specifically, we propose the following belief-neutral criterion:

Definition 1 *A social allocation x is called* ***belief-neutral inefficient (efficient)*** *if the social planner finds it inefficient (efficient) by using any reasonable probability measure Π^h as the common measure to evaluate all agents' expected utilities.*

We can use two different approaches to implement this welfare criterion, one based on a given social welfare function and the other through the notion of Pareto efficiency. As well known in standard economic theory, in the absence of belief distortions these two approaches are internally consistent, as any Pareto efficient social allocation corresponds an optimal allocation that maximizes the agents' aggregate expected utilities under a set of nonnegative weights.

2.2.1 Expected Social Welfare

The so-called 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) based on 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 so-called utilitarian social welfare function:[7]

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

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

Definition 2 *Consider two social allocations, x and y. If the expected social welfare of allocation x dominates that of allocation y under any reasonable probability measure Π^h,*

$$\begin{aligned} &W\left(E_0^h\left[u_1\left(s_T, x_T^1\left(s_T\right)\right)\right], ..., E_0^h\left[u_N\left(s_T, x_T^N\left(s_T\right)\right)\right]\right) \\ &\quad \geq W\left(E_0^h\left[u_1\left(s_T, y_T^1\left(s_T\right)\right)\right], ..., E_0^h\left[u_N\left(s_T, y_T^N\left(s_T\right)\right)\right]\right) \end{aligned}$$

assigns non-zero probability to all relevant states in endowment-economy settings as agents' aggregate biases do not affect the social welfare. However, such an extension is inappropriate in production-economy settings as agents' aggregate biases can lead to distortions in aggregate investment.

[7] 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.

with the inequality holding strictly under 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 in any convex combination of the agents' beliefs is required. This proposed belief-neutral superiority is a partial ordering of social allocations. In the case of two social allocations x and y, 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 with any fixed 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 being destroyed. The money transfer has no impact on the social welfare regardless of its direction and 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 regardless of any 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 the social welfare function.

2.2.2 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. As we discussed before, in the presence of distorted beliefs, the social planner uses a common probability measure from the set of reasonable measures to evaluate each agent's expected utility. As the social planner cannot identify which measure is correct, he shall use all of them. This logic leads to the following implementation of our criterion:

Definition 3 *Consider a social allocation y. Suppose that under any reasonable probability measure Π^h, there always exists another (measure dependent) allocation y' such that it improves some agents' expected utilities without reducing anyone's, i.e., $\forall i$, $E_0^h\left[u_i\left(s_T, y_T^i\left(s_T\right)\right)\right] \leq E_0^h\left[u_i\left(s_T, y_T'^i\left(s_T\right)\right)\right]$ with the inequality holding strictly for at least one agent. Then, allocation y is **belief-neutral Pareto inefficient**. In contrast, if under any Π^h, there is not such a dominating alternative, then allocation y is **belief-neutral Pareto efficient**.*

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', 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 which keeps the pillow intact and transfers \$80 from Joe to Bob. More generally, under any convex combination of Joe's and Bob's beliefs, there is always 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 being destroyed. The bet is thus belief-neutral Pareto inefficient.

Recall from the standard welfare theory (e.g., Mas-Colell et al, 2005, 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. Then, an allocation, $x \in X$, is **belief-neutral Pareto efficient (inefficient)** if and only if under any reasonable probability measure Π^h, there exists (does not exist) a set of Pareto weights $\{\lambda_i\}$ (with $\lambda_i \geq 0$ for all i and $\sum_i \lambda_i = 1$) such that:*

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

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 according to all welfare functions and all reasonable beliefs; 2) those that are belief-neutral efficient because under any reasonable belief they are superior at least according to some welfare functions; and 3) those that are neither uniformly efficient nor uniformly inefficient across all reasonable beliefs.

2.3 Comments on the Criterion

Our belief-neutral welfare criterion is designed to detect inefficiencies (or efficiencies) associated with negative-sum (or positive-sum) speculation between agents. This criterion is not suited for analyzing efficiencies induced by irrational behavior of an individual agent. Consider an agent who invests a large fraction of her wealth in her own company's stocks. This investment decision may appear inefficient to a conscientious observer who holds a more neutral view of the company's stocks 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.

It is useful to compare our criterion to the rationalizably Pareto dominance criterion of Gilboa and Schmeidler (2012). They propose to extend the Pareto criterion in the presence of heterogeneous subjective beliefs by defining a choice x to rationalizably 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 at least one set of beliefs, under which the expected utility of any agent from x is higher than or equal to that from y. This second condition is extra and requires a set of

common beliefs to rationalize the efficiency of x.

The extra condition is able to prevent the bet between Joe and Bob from being efficient as no common beliefs can rationalize the bet. However, the extra condition also makes the criterion more incomplete than the standard Pareto criterion. As a result, the rationalizably Pareto dominance criterion cannot find the bet as being more inefficient than no betting, or vice versa.

Our criterion builds on the premise that in the presence of distorted beliefs, the planner should ignore agents' expected utilities under their own, possibly distorted, beliefs and instead uses a common probability measure to evaluate their welfare. The use of a common probability measure is analogous to the extra condition imposed by Gilboa and Schmeidler. In contrast, our criterion also requires the planner to vary the common measure across a large set of reasonable measures so that the resulting welfare ranking is belief neutral. Despite this seemingly restrictive belief-neutral requirement, our criterion is able to determine the bet between Joe and Bob, as well as many other examples discussed in the next section, as either efficient or inefficient.

3 Examples

This section provides a series of examples to demonstrate that, despite its incompleteness, the simple welfare criterion we propose can produce surprisingly sharp welfare ranking in a wide range of prominent economic models with heterogeneously distorted beliefs. The key is that such beliefs can lead to negative-sum (or positive-sum) games between agents. In the example of Joe and Bob, the destroyed pillow makes the bet between them a negative-sum game. More generally, trading costs, overinvestments, bankruptcy costs, excessive risk taking, and distorted consumption/savings decisions can make speculative transactions between agents in a broad range of economic models negative-sum games, while the benefits from overcoming market breakdowns induced by adverse selection can also lead to positive-sum games in the presence of asymmetric information. This section uses simplified variants of well-known models to illustrate these different sources of losses and gains, and demonstrates that our welfare criterion provides clear welfare ranking in each case.

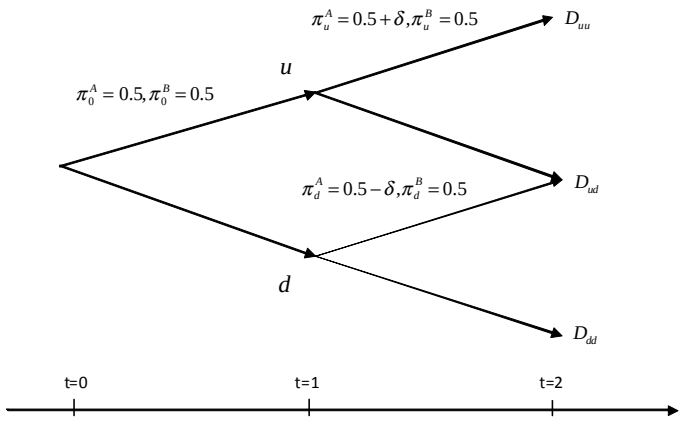


Figure 1: Payoff and belief structure in the bubble model.

3.1 Trading Costs in Bubble Models

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 lead risk-neutral agents not only to trade against each other but also to overvalue assets. Overvaluation does not reduce social welfare by itself as it is simply a welfare transfer across agents. However, in the presence of practical trading frictions such as brokerage fees and bid-ask spreads, excessive trading can reduce the total welfare of all investors. Our criterion illustrates this point.

We focus on a simple binomial setting with three dates (i.e., $t = 0, 1, 2$), two risk-neutral agents (A and B), and a risky asset. The asset's final payoff across the four possible states (uu, ud, du, and dd) on $t = 2$ are D_{uu}, $D_{ud} = D_{du}$, and D_{dd}, respectively. Figure 1 depicts the dynamics of the fundamental state and the two agents' beliefs. We assume that the two agents have time-varying beliefs: they start with the same beliefs on date 0 but hold different beliefs on date 1:

$$\pi_0^A = \pi_0^B = 0.5,\ \pi_u^A = 0.5 + \delta > \pi_u^B = 0.5,\ \text{and}\ \pi_d^A = 0.5 - \delta < \pi_d^B = 0.5. \tag{1}$$

In particular, agent A becomes more optimistic than agent B in state u of date 1 and less optimistic in state d. The parameter $\delta > 0$ determines the two agents' belief dispersion in both states u and d.

To facilitate our analysis, we also impose the following final payoff structure:

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

where $R > 1$ is a constant. Based on this payoff structure, it is straightforward to verify that at $t = 0$ the two agents share the same expectation of the asset's final payoff:

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

Suppose that at $t = 0$ the total supply of the asset is equally divided among the two groups. The fluctuations of the two agents' beliefs at $t = 1$ give each agent 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 the trading price is determined by the buyer's reservation value.[8] It is straightforward to derive the following market price in state u: $p_u = R+1/2+\delta$, which is paid by agent A, and in state d: $p_d = R - 1/2$, which is paid by agent B. By backward induction, both agents on date 0 value the asset by $p_0 = R + \delta/2$. Despite that each agent's expectation of the asset payoff is R, their valuation of the asset is $R + \delta/2$. 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 on date 1. This resale option contributes a non-fundamental component to asset prices in the aforementioned bubble models.

To make the welfare analysis meaningful, we also assume that the seller incurs a cost of κ in selling each unit of the asset. We allow the cost to be modest so that it does not prevent agents from trading: $\kappa < \delta$. Because this cost has to be borne by the seller, it does not affect the market prices derived earlier. However, like the destroyed pillow in the bet between Joe and Bob, the trading cost makes trading a negative-sum game in each agent's belief, and, more generally, in any convex combination of the agents' beliefs. As a result, our welfare criterion can detect the equilibrium being inefficient.

First, consider the welfare-function-based version of the criterion and suppose the planner has the utilitarian welfare function. We will compare the social welfare from the status quo of no trade with that of the market equilibrium, using any convex combination of the two agents' beliefs, $\Pi^h = h\Pi^A + (1-h)\,\Pi^B$, $\forall h \in [0,1]$. In the status quo, agents consume the payoffs from their asset holdings. Since agents are risk-neutral and agree (at date 0) about

[8] It should be clear that as long as the price on date 1 is between the buyer's reservation value and the mid-point of the buyer and seller's reservation values, the asset owner's resale option is valuable.

the asset's expected payoff, the expected utilitarian social welfare is given by:

$$W\left(E_0^h\left[u_A\right], E_0^h\left[u_B\right]\right) = E_0^h\left[\widetilde{R}\right] = R, \quad \forall h \in (0,1).$$

In contrast, in the market equilibrium, agents always trade half of the assets, either from agent B to agent A in state u or vice versa in state d. The trading transfers wealth across agents at a cost, κ /2 (which will be incurred with certainty). Thus, the expected utilitarian welfare is now given by the expected asset payoff net of the trading cost:

$$W\left(E_0^h\left[u_A\right], E_0^h\left[u_B\right]\right) = E_0^h\left[\widetilde{R}\right] - \frac{\kappa}{2} = R - \frac{\kappa}{2}, \ \forall h \in (0,1).$$

It follows that the status quo is belief-neutral superior to the market equilibrium.

This result holds for any welfare function. Using the second version of our welfare criterion, the market equilibrium in this example is also belief-neutral Pareto inefficient. To see this, consider the status quo allocation with an initial transfer of $T \in [-R, R]$ from agent B to agent A. Given this allocation, agent A's expected payoff is given by $R + T$ while agent B's payoff is $R - T$. In contrast, agents' expected payoffs in the equilibrium sum to $R - \frac{\kappa}{2}$ under any convex combination of their beliefs. It follows that, for any reasonable belief, Π^h, the market equilibrium is Pareto dominated by the status quo allocation with some transfer $T \in [-R, R]$.[9]

The key reason for the inefficiency of the market equilibrium in this example is the trading cost, κ. Extensive evidence shows that excessive trading severely undercuts portfolio performance of individual investors in the US, Finland, Taiwan, and China, e.g., Odean (1999), Barber and Odean (2000), Grinblatt and Keloharju (2000), and Barber, Lee, Liu, and Odean (2009). For example, Barber and Odean (2000) analyze performance of a large sample of brokerage accounts held by individual households in the US. They showed that the average household under-performed the market return by 1.5% each year, with trading costs contributing to a majority of the under-performance. For those who traded most actively, trading costs caused under-performance of over 5% each year. One possibility is that individuals are trading for non-speculative reasons, e.g., to rebalance their portfoliios or to meet their liquidity needs. However, it is difficult to reconcile this explanation with the sheer size of the trading volume, e.g., 250% annual turnover rate for the 20% most actively trading investors in the US (Barber and Odean, 2000) and 300% for those in Taiwan (Barber,

[9]Note that in this example the market equilibrium is belief-neutral inefficient even if we extend the set of reasonable beliefs to include any belief that assigns positive probability to both states u and d on date 1.

Lee, Liu, and Odean, 2009). Moreover, individual investors' total annual losses from trading are quite significant, e.g., 2.2% of total GDP in Taiwan. Our criterion creates a presumption that these losses are socially inefficient.[10]

Excessive trading is particularly worrisome during bubble episodes, as emphasized in the above model, because bubbles tend to occur with trading frenzies (e.g., Scheinkman and Xiong (2003), Hong and Stein (2007), Xiong and Yu (2011)). For example, Xiong and Yu (2011) analyze a bubble on deep out-of-the-money Chinese warrants traded between 2005-2008 in an environment with short selling restrictions. Even though these warrants could be reliably valued to have almost zero fundamentals, investors traded them at an average *daily* turnover rate of 200% and at largely inflated prices. The aggregate losses from trading were once again quite significant, as the average daily dollar volume of these warrants was in billions and the investors needed to pay brokerage fees ranging from 0.1-0.3% per side. It is difficult to explain this much trading with non-speculative motives, especially given that the warrants were essentially worthless. In line with this interpretation, an independent experimental literature initiated by Smith, Suchanek, Williams (1988) shows that bubbles and trading frenzies emerge also in fully controlled environments in which there is no non-speculative trading motive by design (see Hussam, Porter, and Smith (2008), Kirchler, Huber, and Stockl (2012), and Andrade, Odean, and Lin (2012) for recent contributions). Taken together, this evidence along with our criterion suggests that trading losses associated with bubble episodes are socially inefficient.

3.2 Overinvestment in Bubble Models

A severe consequence of asset price bubbles is overinvestment. Several recent papers build on the aforementioned bubble models to analyze overinvestment (e.g., Bolton, Scheinkman and Xiong (2006), Gilchrist, Himmelberg, and Huberman (2005), and Panageas (2006)). The key idea is that even in the absence of any governance failure, a firm may choose to overinvest at the expense of its long-run fundamental value in order to maximize its current market value, which contains not only the long-run fundamental value but also the "bubbly" resale option value. Our criterion can highlight welfare losses induced by such overinvestment.

[10] Trading costs represent a social loss even though they are mostly received by other agents in the economy, e.g., brokerage firms and market makers in the form of bid-ask spreads. This is because there is an opportunity cost of employing those agents (and their equipments) to facilitate these transactions as opposed to using them elsewhere in the economy.

We extend the two-period setting from the bubble example to incorporate firm investment. We remove the trading cost by setting κ to be zero, but include firm investment. Suppose that the risky asset is equity issued by a firm. The firm chooses its investment at date 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 across the three states on date 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 on date 2.

Suppose that the firm issues one share of equity for each unit of production capacity. We denote the market price of each share on date 0 by p_0. Following the beliefs specified from the bubble example in equation (1) for the two risk-neutral agents A and B, it is straightforward to derive that at date 1 agent A will acquire all the shares at state u at a price of

$$p_u = (1/2 + \delta)(R + 1 - n) + (1/2 - \delta)(R - n) = R - n + (1/2 + \delta),$$

and agent B will acquire all the shares at state d at a price of

$$p_d = 1/2\,(R - n) + 1/2\,(R - 1 - n) = R - n - 1/2.$$

On date 0 both agents A and B value each share at the same price of $p_0 = R - n + \delta/2$, which is higher than their expectation of the share's final payoff $E^A\left[\widetilde{R}\right] = E^B\left[\widetilde{R}\right] = R - n$. The difference is again due to the value of each asset owner's resale option in anticipation of the fluctuations of the two agents' beliefs on date 1.

Since both A and B agree about the initial share price, the firm chooses its production capacity, n, to maximize its market value given by: $n \cdot p_0 = n \cdot (R - n + \delta/2)$. Thus, the firm's optimal investment level is given by: $n^* = \frac{1}{2}\left(R + \frac{\delta}{2} - 1\right)$, which depends on δ, the magnitude of the two agents' belief dispersion on date 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, $\Pi^h = h\Pi^A + (1 - h)\,\Pi^B$, $\forall h \in (0, 1)$. As in the previous subsection, 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 $n^{**} = \frac{1}{2}(R - 1) < n^*$. 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.[11]

As before, 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.

As in the first example, the driving force behind the inefficient overinvestment is exactly the value of the resale option in the firm's date-0 market valuation. Anticipating the possibility of reselling the share to the other agent at date 1 at a profit, each agent overvalues the share at date 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. Consistent with this 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.

3.3 Benefits of Speculation in Lemons Models

The previous two examples show that excessive trading and overinvestment in heterogeneous-beliefs-induced bubble models lead 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 presents an example to illustrate 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 which attempts to finance a new investment project by selling its legacy asset. However, the firm is asymmetrically informed about the

[11] 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 the agents' aggregate biases is not our focus.

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, 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 that has access to a new project that costs I and generates a payoff of $I+G$. The payoff of the project is not pledgeable (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 which is pledgeable. 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 pledgeable 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 without and with heterogeneous 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 heterogeneous 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 buyer breaks 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 with $\theta < \theta^*$ sells the asset 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, consider (as in Tirole, 2012) the ex-ante utilitarian social welfare function, i.e., the sum of the seller's and buyer's 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 the buyer 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 heterogeneous 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, and there are two groups of risk-neutral buyers for the asset: one believing the probability of $\widetilde{R} = \overline{R}+1$ is 1, while the other group believing the probability is 0. Suppose that no one can short sell the asset and each group has sufficient cash to acquire the asset. Like the previous two examples, buyers in the optimistic group acquire the asset and bid up the asset price to their expectation of the asset payoff. A key feature of the model is that the asset overvaluation induced by agents' heterogeneous beliefs helps to 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, we conjecture that the seller chooses to sell and finance the project regardless of θ. The optimistic buyers break even only if $p^{spec} = \left(\overline{R}+1\right) E\left[\theta\right] = \frac{\overline{R}+1}{2}$. 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, which provides further contrast to the Pareto inefficiency of the earlier equilibrium. 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 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.

3.4 Bankruptcy Costs in Leverage Cycle Models

The literature on leverage cycles based on agents' heterogeneous beliefs (e.g., Geanakoplos (2003, 2009), Fostel and Geanakoplos (2008), Simsek (2010), Cao (2010), and He and Xiong (2012)) is growing. 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 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 2 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

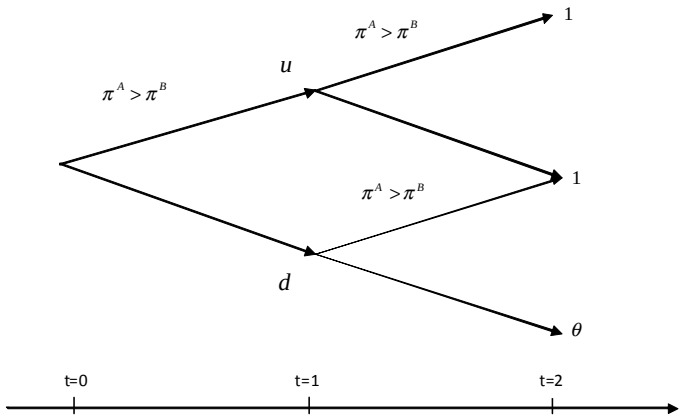


Figure 2: Payoff and belief structure in the leverage cycle model.

about the asset payoff in the lower state d of date 1, which eventually triggers a leverage cycle.[12]

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 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

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

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:

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

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.3 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

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

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.[13] As before, this result holds for any welfare function because the equilibrium is also belief-neutral Pareto inefficient.[14]

[13] 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 $\widetilde{R} = \theta$.

[14] 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 (2010). 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, $\pi_L\left(1-\iota\right)L + \left(1-\pi_L\right)M$, 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.

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.

There is some evidence that speculative motives played a role in the US mortgage market during the recent housing boom-bust cycle. Haughwout et al. (2011) identify "investors" as those individuals that own more than one property financed by a distinct mortgage. Compared to single-home owners, investors are more likely to be driven by speculative trading motives since they are less likely to be holding all of their properties for consumption value. Haughwout et al. (2011) show that investors contributed significantly to the increase in aggregate mortgage debt until 2006, as well as the delinquincies and defaults in the more recent deleveraging phase. In particular, the fraction of new mortgages issued by investors increased from 20% in 2000 to 35% in 2006. The increase was even more dramatic in states that experienced a greater housing bubble, e.g., Arizona, California, Florida, and Nevada. In the bust phase, investors' share in severe delinquincies (i.e., mortgages whose payment is more than 90 days past due) also increased from about 10% in 2004 to over 25% in 2009, with a greater increase in bubble states. Investors' total contribution to delinquincies reached $250 billion in 2009 in non-prime mortgages alone. Our criterion suggests that the deadweight losses (e.g., liquidation or foreclosure costs) associated with these delinquincies might be socially inefficient.

3.5 Excessive Risk Taking in Speculative Trading Models

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. More specifically, a positive shock increases the wealth of optimists more than that of pessimists, as optimists tend to take larger asset

positions. The optimists' greater wealth increases allow them to take even larger positions and thus amplify the impact of the shock on equilibrium asset prices.

Despite the capability of these models to 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. Our simple criterion can potentially fill this gap by offering a useful insight for these types of models. The key point is that trading induced by heterogeneous beliefs makes agents' consumption excessively risky. Each agent takes these risks because she expects to earn high returns in expectation. However, each agent also recognizes that these returns will come from other agents with different beliefs. Thus, when agents are risk averse, trading makes consumption of all agents more volatile and is thus a negative-sum game in *expected utility terms* under any convex combination of their beliefs.

We consider a one-period, endowment economy setting with two agents, A and B, to illustrate the welfare implication (although this static setting is insufficient to highlight the rich asset pricing implications of the aforementioned studies). Each agent is endowed with half dollars and lives from $t = 0$ to $t = 1$. There is neither aggregate nor idiosyncratic endowment risk. Suppose that each agent consumes at $t = 1$ and has an increasing and strictly concave utility function $u(c_i)$. The two agents hold heterogeneous beliefs about a random variable, say $\widetilde{D}$, which can take two possible values, either H or L. One may interpret this random variable as sunspot, which is independent of the agents' endowment risk. Suppose agent A assigns a probability of π^A to state $\widetilde{D} = H$, while agent B assigns π^B. The difference in beliefs causes the agents to engage in speculative trades against each other. We allow them to trade a contract that pays 1 if $\widetilde{D} = H$ and 0 if $\widetilde{D} = L$.

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

$$\max_{k^i} \ \pi^i u\left(0.5 + k^i(1-p)\right) + \left(1-\pi^i\right) u\left(0.5 - k^i p\right).$$

The first order condition gives:

$$(1-p)\,\pi^i u'\left(0.5 + k^i(1-p)\right) = p\left(1-\pi^i\right) u'\left(0.5 - k^i p\right).$$

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.

The market equilibrium in this example is inefficient according to our criterion. To see this, first consider the welfare-function version of the criterion. Suppose the planner has a utilitarian welfare function: $W(u_A, u_B) = u_A + u_B$. We compare the social welfare based on the agents' equilibrium consumption:

$$x = \left\{\left(x_H^i, x_L^i\right)\right\}_{i\in\{A,B\}} = \left\{\left(0.5 + k^i(1-p), 0.5 - k^i p\right)\right\}_{i\in\{A,B\}},$$

with that based on the status quo allocation with no trading: $y = \{(y_H^i, y_L^i) \equiv (0.5, 0.5)\}_{i\in\{A,B\}}$. We allow the planner to use any measure that assigns positive probability to the two possible values of $\widetilde{D}$. The following proposition shows a welfare ranking based on this comparison.

Proposition 2 *If $\pi^A \neq \pi^B$ and the social planner has the utilitarian welfare function, then the social welfare of the status quo allocation dominates that of the market equilibrium allocation according to any measure that assigns positive probability to relevant states of $\widetilde{D}$.*

The mechanism that underlies Proposition 2 is simply that the trade makes their consumption riskier than their endowments. Due to risk aversion, the utilitarian social welfare falls according to each agent (as well as any belief measure that assigns positive probability to the two possible values of $\widetilde{D}$).

The second version of our criterion–belief-neutral Pareto inefficiency–allows for a transfer of endowments in welfare comparison in the same way Pareto inefficiency does. Specifically, we compare the market equilibrium allocation to that from the status quo with a transfer of $T \in [-0.5, 0.5]$ from agent B to agent A:

$$y(T) = \{(0.5 + T, 0.5 + T), (0.5 - T, 0.5 - T)\}.$$

Proposition 3 *If $\pi^A \neq \pi^B$, the market equilibrium is Pareto dominated by the status quo allocation with a certain transfer $T \in [-0.5, 0.5]$ under any measure that assigns positive probability to relevant states of $\widetilde{D}$.*

One should be cautious not to overinterpret this example to mean that trading always reduces social welfare. Richer economic settings often feature a trade-off between welfare-enhancing risk sharing and speculation. To illustrate this trade-off, consider a variant of the earlier example in which there is also a risk sharing motive for trading. In particular, suppose the random variable, $\widetilde{D}$, also affects agents' endowment risks. One may interpret this random variable as corresponding to a relative price shock (e.g., the price of corn) that

leads to a reallocation of wealth between agents. In particular, if $\widetilde{D} = H$, then agent A (e.g., the miller) incurs a loss of e, while agent B (e.g., the farmer) incurs a gain of e. If $\widetilde{D} = L$ then the agents' endowments are the same as before. Thus, the status quo allocation is now given by:

$$y = \left\{ \left(y_H^A, y_L^A \right), \left(y_H^B, y_L^B \right) \right\} = \left\{ (0.5 - e, 0.5), (0.5 + e, 0.5) \right\}.$$

The equilibrium is characterized by the following first order condition for agent A,

$$(1-p)\,\pi^A u'\left(0.5 - e + k^A(1-p)\right) = p\left(1-\pi^A\right) u'\left(0.5 - k^A p\right),$$

a similar condition for agent B, and the market clearing condition $k^A + k^B = 0$.

When agents have common beliefs, $\pi^A = \pi^B = \pi$, the equilibrium is given by $p = \pi$ and $k^{opt} \equiv \left(k^A = e, k^B = -e\right)$. Both agents fully diversify their idiosyncratic risks—agent A (agent B) consumes a constant amount $0.5 - e\pi$ $(0.5 + e\pi)$ regardless of the state.

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, k^{opt}. Our next result illustrates the inefficiency of this equilibrium by comparing it with the common-belief allocation, k^{opt}, with a transfer $T \in [-0.5, 0.5]$:

$$y\left(T, k^{opt}\right) = \left\{ (0.5 - e\pi + T, 0.5 - e\pi + T), (0.5 + e\pi - T, 0.5 + e\pi - T) \right\}.$$

Proposition 4 *If $\pi^A \neq \pi^B$, the market equilibrium in the presence of idiosyncratic endowment risks is Pareto dominated by the optimal risk sharing allocation k^{opt} with a certain (belief-measure dependent) transfer $T \in [-0.5, 0.5]$ under any measure that assigns positive probability to relevant states of $\widetilde{D}$.*

Unlike the setting of Proposition 2, it is now not optimal to prevent trading completely. Rather, the planner would like to prevent *speculative trading*, that is, trading in excess of the optimal risk sharing benchmark. While this example is stylized, the point applies more generally. Consider any economy with complete financial markets in which agents have the same preferences, assumed to be separable over time and states, but potentially distorted beliefs. Suppose there are at least two agents, $\{A, B\}$, and at least two states, $\{H, L\}$. Regardless of the belief measure used by the planner, full risk sharing features the equalization of agents' marginal utilities across states: $\frac{u'\left(x_H^A\right)}{u'\left(x_L^A\right)} = \frac{u'\left(x_H^B\right)}{u'\left(x_L^B\right)}$. In contrast, the equilibrium allocation features the equalization of agents' marginal utilities multiplied

by their distorted beliefs, e.g., $\frac{\pi_H^A}{\pi_L^A}\frac{u'\left(x_H^A\right)}{u'\left(x_L^A\right)} = \frac{\pi_H^B}{\pi_L^B}\frac{u'\left(x_H^B\right)}{u'\left(x_L^B\right)}$.[15] Thus, when agents disagree about the relative probabilities of states H and L, speculative trading causes deviations from the optimal risk sharing, which is inefficient regardless of the planner's belief measure.

In standard asset pricing theories, e.g., the capital asset pricing model, optimal risk sharing requires investors to hold highly diversified portfolios. However, empirical evidence suggests that many individual investors fail to diversify adequately (see Barber and Odean, 2011, for a recent survey). They tend to overinvest in domestic stocks (French and Poterba, 1991), as well as regionally close and familiar stocks (Huberman, 2001, Grinblatt and Keloharju, 2001). In more direct conflict with risk sharing, they also tend to invest considerably in own company stocks (Benartzi, 2001, Poterba, 2003) and professionally close stocks. For example, Doskeland and Hvide (2011) analyze Norwegian individual investors that hold stocks directly, and find that 32% of these holdings correspond to stocks of companies in the same two-digit industry as the investor's employer. Moreover, investors on average earn *negative* risk-adjusted returns from these professionally proximate holdings. Our criterion, along with Proposition 4, creates a presumption that these investments (or similar deparatures from optimal risk sharing) might be socially inefficient.

The analysis so far assumes that the planner is sufficiently informed about the nature of agents' endowment risks. In some contexts, e.g., derivatives trading by firms, the planner might not be as precisely informed as agents about their endowment risks. As a result, it might be difficult for the planner to implement the optimal risk sharing allocation. Instead, the planner's options might be either to allow unrestricted trading in the risky asset or to prevent trading completely. Since the gain from risk sharing increases with the magnitude of the agents' endowment shocks e, our next result shows that the planner would prefer no trading as long as e is sufficiently small.

Proposition 5 *If $\pi^A \neq \pi^B$, then there exists $\overline{e} > 0$ (which depends on π^A and π^B) such that when the agents' endowment risks are below this threshold, $e < \overline{e}$, the market equilibrium is Pareto dominated by the status quo allocation with some transfer $T \in [-0.5, 0.5]$ under any measure that assigns positive probability to relevant states.*

As this proposition illustrates, our welfare criterion can also detect inefficiency in the

[15] The earlier general equilibrium literature, e.g., von Weizsäcker (1962), Dreze (1970), Starr (1973), Harris (1978), Hammond (1981), has noted this deviation of the market equilibrium in the presence of heterogeneous beliefs from the full risk sharing outcome.

presence of a trade-off between risk sharing and speculation. Simsek (2011) and Kubler and Schmedders (2011) analyze richer settings that feature a similar trade-off. Our welfare criterion is useful to analyze the inefficiency of speculative trading in these richer settings.

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 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

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

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:

$$Y = 1.6, R = 1.1, \tau_f = 1.1, \tau_m = 1.65, \alpha = 0.3, \beta = 0.9, A = 1.2, \gamma = 0.5, B_0 = 1.5. \quad (3)$$

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, $\{\pi_a = 0.3, \pi_b = 0.7\}$.

$\{\pi_a, \pi_b\}$	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.

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 ($K = 1.03$).[16]

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 3 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.

[16] In contrast, if $1/\sigma < 1$, then the income effect dominates and agents save less with heterogeneous beliefs relative to the homogeneous-beliefs benchmarks.

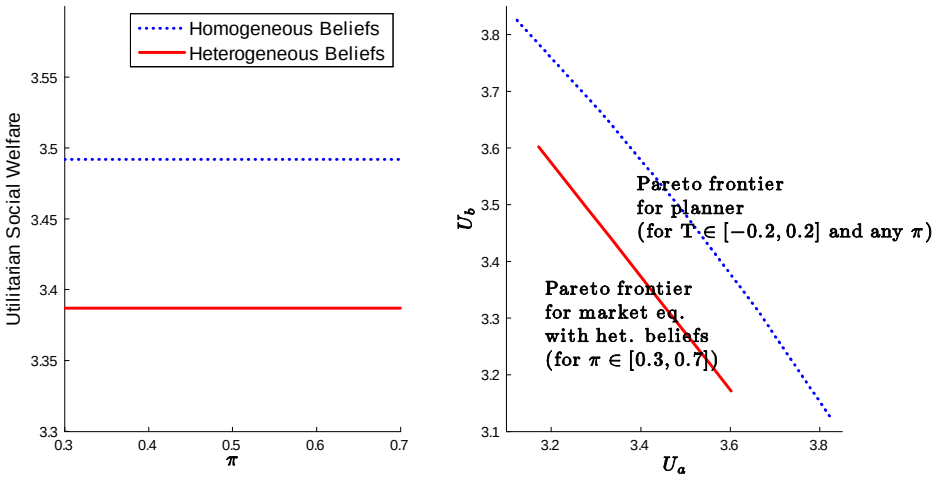


Figure 3: Social welfare and Pareto frontier in homogeneous-beliefs benchmarks and heterogeneous-beliefs equilibrium. The left panel plots the utilitarian social welfare based on any convex combination of the two types of beliefs. The right panel plots the Pareto frontiers respectively for the market equilibrium and a belief-neutral planner.

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 3 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 uses a common probability measure to evaluate the welfare of different agents but cannot differentiate whose beliefs are correct. 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 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 prominent models 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.

Andrade, Eduardo, Shengle Lin, and Terrance Odean (2012), Bubbling with Excitement: An Experiment, working paper.

Barber, Brad and Terrance Odean (2000), Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773-806.

Barber, Brad and Terrance Odean (2011), The behavior of individual investors. In G. Constantinides, M. Harris, and R. Stulz (eds.), *Handbook of the Economics of Finance*.

Barber, Brad, Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean (2009), Just how much do individual investors lose by trading? *Review of Financial Studies* 22, 609-632.

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.

Benartzi, Shlomo (2001), Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock, *Journal of Finance*, 56, p.1747-1764.

Bernheim, Douglas and Antonio Rangel (2005), "Behavioral public economics: Welfare and policy analysis with non-standard decision makers, NBER working paper, No. 11518.

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

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

Brunnermeier, Markus, and Lasse Pedersen (2009), Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201-2238.

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.

Cheng, Ing-haw, Sahil Raina, and Wei Xiong (2012), Wall Street and the housing bubble, Working paper, Princeton University.

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

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, T. and H. Hvide (2011), Do individual investors have asymmetric information based on work experience? *Journal of Finance* 66, p.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.

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

French, Kenneth and James Poterba (1991), International diversification and international equity markets, *American Economic Review*, 81, p.222-226.

Friedman, Milton and L. J. Savage (1948), The utility analysis of choices involving risk, *Journal of Political Economy* 56, 279-304.

Gabaix, Xavier and David Laibson (2006), Shrouded attributes, consumer myopia, and information suppression in competitive markets, *Quarterly Journal of Economics* 121, 505-540.

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 and David Schmeidler (2012), A difficulty with Pareto dominance, Working paper.

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

Grinblatt, Mark and Matti Keloharju (2000), The investment behavior and performance of various investor types: a study of Finland's unique data set, *Journal of Finance* 55, 43-67.

Grinblatt, Mark and Matti Keloharju (2001), How distance, language, and culture influence stock-holdings and trades, *Journal of Finance*, 56, p.1053-1073.

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.

Haughwout, A., D. Lee, J. Tracy, and W. van der Klaauw (2011), Real estate investors, the leverage cycle, and the housing market crisis, *Federal Reserve Bank of New York Staff Reports*, 514.

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 Jeremy Stein (2007), Disagreement and the stock market, *Journal of Economic Perspectives* 21, 109-128.

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

Huberman, G. (2001), Familiarity breeds investment, *Review of Financial Studies*,14, p.659-680.

Hussam, Reshmaan, David Porter, and Vernon Smith (2008), Thar She Blows: Can Bubbles Be Rekindled with Experienced Subjects? *American Economic Review*, 98(3): 924–37.

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

Kahneman, Daniel (2011), *Thinking, fast and slow*, Macmillan, New York.

Kirchler, M., J. Huber and T. Stockl (2012), Thar She Bursts: Reducing Confusion Reduces Bubbles, *American Economic Review*, forthcoming.

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

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

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

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.

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.

Odean, Terrance (1999), Do investors trade too much?, *American Economic Review* 89, 1279-1298.

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

Posner, Eric and Glen Weyl (2012), An FDA for financial innovation: Applying the insurable interest doctrine to 21st century financial markets, *Northwestern University Law Review*, forthcoming.

Poterba, James (2003), Employer Stock and 401(k) Plans, *American Economic Review*, 93, p.398-404.

Savage, L. J. (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.

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

Simsek, Alp (2010), Belief disagreements and collateral constraints, *Econometrica*, forthcoming.

Simsek, Alp (2011), Speculation and risk sharing with new financial assets, Working paper, Harvard University.

Smith, Vernon, Gerry Suchanek, and Arlington Williams (1988), Bubbles, crashes, and endogenous expectations in experimental spot asset markets, *Econometrica*, 56(5), p.1119–51.

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.

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

von Weizsäcker, Christian (1962), General equilibrium and uncertainty, Teaching note, MIT.

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

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.

Xiong, Wei and Jialan Yu (2012), The Chinese Warrants Bubble, *American Economic Review*, 101(6), p.2723-53.

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 A Proofs of Propositions

A.1 Proof of Propositions 2 and 3

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

$$\frac{\pi^A}{1-\pi^A}\frac{u'\left(0.5+k^A\left(1-p\right)\right)}{u'\left(0.5-k^Ap\right)}=\frac{\pi^B}{1-\pi^B}\frac{u'\left(0.5+k^B\left(1-p\right)\right)}{u'\left(0.5-k^Bp\right)}.$$

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.

To prove the propositions, consider any measure that assigns probability $\pi \in (0,1)$ to $\widetilde{D} = H$. First, consider agents' utilitarian social welfare in equilibrium, which is given by:

$$\begin{aligned} U^h &= \pi \left[u\left(0.5 + k^A (1-p)\right) + u\left(0.5 - k^A (1-p)\right) \right] \\ &\quad + (1-\pi) \left[u\left(0.5 - k^A p\right) + u\left(0.5 + k^A p\right) \right]. \end{aligned}$$

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

$$\begin{aligned} u\left(0.5 + k^A (1-p)\right) + u\left(0.5 - k^A (1-p)\right) &< 2u(0.5), \\ u\left(0.5 - k^A p\right) + u\left(0.5 + k^A p\right) &< 2u(0.5). \end{aligned}$$

Thus, $U^h < \pi \cdot 2u(0.5) + (1-\pi) \cdot 2u(0.5) = 2u(0.5)$, which is the utilitarian social welfare under the status quo. This completes the proof of Proposition 2.

Next, consider each agent's certainty-equivalent wealth, $w^{i,eq}$, given by the solution to:

$$u\left(w^{i,eq}\right) = \pi u\left(0.5 + k^i (1-p)\right) + (1-\pi) u\left(0.5 - k^i p\right), \forall i \in \{A, B\}.$$

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

$$u\left(w^{i,eq}\right) < u\left(\pi \left(0.5 + k^i (1-p)\right) + (1-\pi)\left(0.5 - k^i p\right)\right).$$

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

$$w^{i,eq} < \pi \left(0.5 + k^i (1-p)\right) + (1-\pi)\left(0.5 - k^i p\right), \quad \forall i \in \{A, B\}.$$

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, completing the proof of Proposition 3.

A.2 Proof of Propositions 4 and 5

Recall that the optimal risk sharing trade is given by $k^{opt,A} = e$ and $k^{opt,B} = -e$. 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, $k^A \neq k^{opt,A}$. Next fix a belief, Π^h, and consider each agent's certainty-equivalent wealth, $w^{i,h}$, given by the solution to:

$$\begin{aligned} u\left(w^{A,h}(e)\right) &= \pi u\left(0.5 - e + k^A (1-p)\right) + (1-\pi) u\left(0.5 - k^A p\right), \\ u\left(w^{B,h}(e)\right) &= \pi u\left(0.5 + e + k^B (1-p)\right) + (1-\pi) u\left(0.5 - k^B p\right). \end{aligned}$$

Since $k^A \neq k^{opt,A}$, an agent's consumption is not constant across the states. Then, the strict concavity of $u(\cdot)$ implies:

$$\begin{aligned} w^{A,h}(e) &< \pi \left(0.5 - e + k^A (1-p)\right) + (1-\pi)\left(0.5 - k^A p\right), \\ w^{B,h}(e) &< \pi \left(0.5 + e + k^B (1-p)\right) + (1-\pi)\left(0.5 - k^B p\right). \end{aligned}$$

Adding these inequalities and using market clearing, we have:

$$w^{A,h}(e) + w^{B,h}(e) < 1. \tag{4}$$

On the other hand, the status quo allocation combined with the optimal risk sharing trade, k^{opt}, gives each agent a constant consumption of $\frac{1}{2}$. It follows that this allocation combined with an appropriate transfer, T, Pareto dominates the equilibrium allocation, proving Proposition 4.

To prove Proposition 5, consider agents' certainty-equivalent wealth from the status quo allocation with no trade. This is found by solving:

$$\begin{aligned} u\left(w^{status,A,h}(e)\right) &= \pi u(0.5-e) + (1-\pi)\, u(0.5), \\ u\left(w^{status,B,h}(e)\right) &= \pi u(0.5+e) + (1-\pi)\, u(0.5). \end{aligned}$$

A similar analysis to above shows that the sum, $w^{status,A}(e) + w^{status,B}(e)$, is also less than 1. Nonetheless, we claim that there exists $\overline{e} > 0$ such that:

$$w^{A,h}(e) + w^{B,h}(e) < w^{status,A,h}(e) + w^{status,B,h}(e).$$

Once we show this claim, it follows that the status quo allocation with an appropriate transfer, T, dominates the equilibrium allocation under any $\pi \in (0,1)$.

To prove the claim, first note that the left hand-side of equation (4) is a continuous function of e and that its limit $\lim_{e\to 0} w^{A,h}(e) + w^{B,h}(e) < 1$. This is because agents take non-zero speculative positions also in the limit as $e \to 0$ (as discussed in the proof of Proposition 3). In contrast, as $e \to 0$, the status quo approximates riskless consumption, which implies: $\lim_{e\to 0} w^{status,A,h}(e) + w^{status,B,h}(e) = 1$. Combining these two limiting cases leads to the claim, and thus, also Proposition 5.

A.3 Characterization of Equilibrium in Section 3.4

The following proposition summarizes the market equilibrium:

Proposition 6 *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*

$$p_0^* = \frac{\left(2-\pi^A\right)\pi^A[\pi^B + (1-\pi^B)\theta](1-\theta) - [\pi^A(1-\pi^B)(1-\theta) - (1-\pi^A)^2\alpha]\theta}{\left(2-\pi^A\right)\pi^A(1-\theta) - \pi^A(1-\pi^B)(1-\theta) + (1-\pi^A)^2\alpha}. \quad (5)$$

  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 (6) 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:* $p_0 = c + \theta$*.*

- *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 $(\theta,\ K_d)$. 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 (2), $p_0^* > E_0^B[\widetilde{R}]$. 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 6

- 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: $c_2 \leq c < c_3$.

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 (5).

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:

$$(1 - \mu)\frac{c}{p_0^* - \theta} + \mu\frac{c}{p_0^* - K_d} = 1. \tag{6}$$

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 6.