A Welfare Criterion For Models with Distorted Beliefs

Markus Brunnermeier, Alp Simsek & Wei Xiong

Welfare Analysis for Behavioral Models

- Vast evidence on people holding wrong beliefs and making inefficient decisions.
  - e.g., reviews of Hirshleifer (2001), Barberis and Thaler (2003), Della Vigna (2009)
- For normative analysis, a welfare criterion is needed.
- BF literature commonly assumes a true belief and the planner knows the true belief
  - e.g., Gabaix and Laibson (2006), Weyl (2007), Spinnewijn (2010), and Gennaioli, Shleifer, Vishny (2011)
- Whose belief is wrong? And which belief should a planner use?
  - One cannot liberally overturn revealed beliefs.
- This question becomes even more serious for models with heterogeneous beliefs
  - Harrison and Kreps (1978), Detemple and Murthy (1994), Scheinkman and Xiong (2003), Geanakoplos (2009), and others

A Belief-Neutral Criterion

- This paper provides a belief-neutral welfare criterion.
  - The planner is aware of presence of distorted beliefs but not sure of the true belief.

- Negative or positive sums often appear in models with heterogeneously distorted beliefs.
  - One can evaluate welfare without taking a stand on whose belief is right or wrong.

- Negative-sum speculation in macro and finance models
  1. Over-investment in bubble models
  2. Bankruptcy costs in leverage cycle models
  3. Excessive risk taking in speculative trading models
  4. Consumption-savings distortions in macro models
- Positive-sum speculation
  5. Overcoming market breakdown in Lemons models

An Example

Joe Stiglitz:

With 90% chance it is made of cotton

Bob Wilson:

With 90% chance it is made of polyester

Joe and Bob took a bet:

- If it is made of cotton, Bob pays Joe $100; otherwise, Joe pays Bob $100.

They had to cut the pillow open to find out its content

- It cost $50 to replace the pillow, which is paid by the winner.

An Example

Expected return from the bet:

$$90\% \times (\$100 - \$50) - 10\% \times \$100 = \$35$$

Expected return from the bet:

$$90\% \times (\$100 - \$50) - 10\% \times \$100 = \$35$$

Both Joe and Bob found the bet desirable

- The bet is Pareto efficient!

An Example

$100

- The bet induces a wealth transfer between them, but a perfect pillow is destroyed!

Negative Externality

- The bet is a negative-sum game!
  - Socially inefficient regardless of whose belief is right or wrong.

- Externality driven by conflicting beliefs
  - Joe believes that he will win and thus the cost of replacing the pillow goes to Bob
  - Bob believes that he will win and thus the cost of replacing the pillow goes to Joe
  - The presence of the negative externality holds in both Joe and Bob's beliefs.

A Belief-Neutral Welfare Criterion

We propose a belief-neutral welfare criterion.

- A set of reasonable beliefs, spanned by convex combinations of agents' beliefs.
  - The objective measure lies between agents' beliefs.
  - Can be further expanded in some settings.

- A choice x is efficient (inefficient) if the planner finds it efficient (inefficient) by using every reasonable belief as the common measure to evaluate all agents' welfare.

- Two ways to implement
  - Social welfare function
  - Pareto efficiency

Welfare Analysis with Conflicting Beliefs

- Divergence between ex-ante and ex-post efficiencies, e.g., von Weizsäcker (1962), Dreze (1970), Starr (1973), Harris (1978), Hammond (1981).
- Spurious unanimity problem of Pareto criterion, e.g., Mongin (1997) and Gilboa, Samuelson, and Schmeidler (2012)

Sources of heterogeneous beliefs

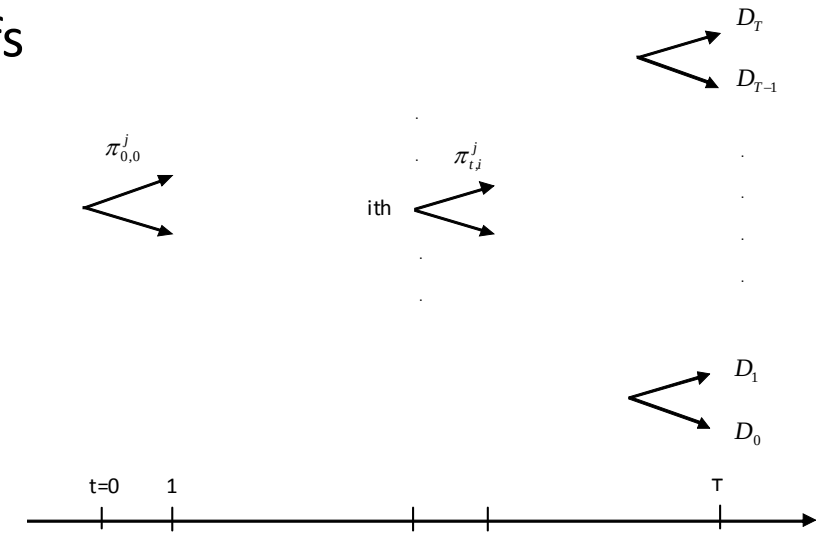
- Subjective beliefs
  - Savage's view: beliefs are part of their preferences under uncertainty.
  - Beliefs may reflect state-dependent preferences.
  - The bet did not help Bob and Joe hedge their state-dependent risk, rather each believed he would win and the other would lose.

- Distorted beliefs
  - Mounting evidence that biases, like overconfidence, representativeness, etc., can distort people's beliefs.
  - Then, social planner needs to use a common, objective measure to evaluate agents' welfare on their behalves.
  - Our framework allows state-dependent utility functions for subjective beliefs.
  - Our criterion requires only presence of belief distortion but not precise identification of whose beliefs are distorted, complementary to Bernheim and Rangel (2009).

Model Setting

- Consider a generic setting with T periods: $t = 0,1, \dots, T$.
  - The state follows a binomial tree.

- N agents holding different beliefs
  - $\Pi^i = \{\pi^i_{t,s}\}, i \in \{1, \dots, N\}$
  - $\pi^i_{t,s} > 0$



- A social choice:
  - $x = \{x^i_T(s_T)\}$

- State-dependent utility
  - $u_i[s_T, x^i_T(s_T)]$
  - Capturing state-dependent preferences and subjective priors

- Set of reasonable beliefs:
  - any convex combination of agents' beliefs: $\Pi^h = \sum_i h^i \Pi^i$, where $h^i \geq 0$ and $\sum_i h^i = 1$.
  - Includes all extreme beliefs that are present in the system.

Implementation by a Social Welfare Function

- For a given social welfare function
  - Bergsonian social welfare function: $W(u^1, u^2, \dots, u_N) = \sum_i \lambda_i u_i$, where $\{\lambda_i\}$ are non-negative weights
    - Varying the weights gives Pareto frontier
  - Utilitarian social welfare function: $W(u^1, u^2, \dots, u_N) = \sum_i u_i$

- Allocation x is belief-neutral superior to y if $\forall \Pi^h$, the planner finds that

$$W(E_0^h[u^1(s_T, x_T^1(s_T))], \dots, E_0^h[u_N(s_T, x_T^N(s_T))]) \geq W(E_0^h[u^1(s_T, y_T^1(s_T))], \dots, E_0^h[u_N(s_T, y_T^N(s_T))])$$

Implementation by a Social Welfare Function

- Back to the bet between Joe and Bob.
- Suppose that both of them are risk neutral and that the planner uses a utilitarian welfare function:

$$W(u_{Joe}, u_{Bob}) = u_{Joe} + u_{Bob} = w_{Joe} + w_{Bob}$$

  - Social welfare is equivalent to expected social wealth.

- The bet generates a wealth transfer and a pillow being destroyed.
  - The destroyed pillow leads to a negative sum, which is independent of the beliefs used by the planner.

- What if Bob and Joe have unequal weights?
  - The bet can transfer wealth from the low-weight person to the other.

Implementation by Pareto Dominance

- An allocation x is called belief-neutral Pareto efficient if under any measure Π^h there does not exist another allocation x' such that it improves some agents' expected utilities without reducing anyone's, i.e.,
$\forall i,\ E_0^h[u_i(s_T, x_T^i(s_T))] \leq E_0^h[u_i(s_T, x'^i_T(s_T))]$.

  - Different from standard Pareto dominance, the planner uses a common measure to evaluate all agents' welfare, instead of their own.
  - The standard economic theory: for a given, common belief measure, each allocation on the Pareto frontier maximizes a linear social welfare function with a certain set of Pareto weights.

Implementation by Pareto Dominance

- Back to the bet between Joe and Bob.
- Suppose that the planner uses Joe's beliefs.
  - The bet leads to an expected gain of $35 to Joe and an expected loss of $85 to Bob.
  - An alternative by transferring $35 to Joe from Bob makes Joe indifferent and improves Bob's welfare by $50.
- Suppose that the planner uses any convex combination of their beliefs, say with weight $h \in (0,1)$ to Joe.
  - A higher h means a larger expected gain to Joe from the bet under the measure.
  - Still, an appropriate transfer from Bob to Joe can make Joe indifferent and save Bob some money.
- Thus, the bet is belief-neutral Pareto inefficient.
  - The belief-neutral inefficiency of the bet does not rely on any particular welfare function.
  - The bet is belief-neutral inefficient, even though which allocation dominates the bet may depend on the belief measure or welfare function.

Generalize the Bet

State-dependent replacement cost:

- It costs \$50 if it is made of cotton but \$20 if it is made of polyester.

- The externality is still belief neutral negative
- Under Joe's belief:
  - His expected return is $90\% \cdot (\$100 - \$50) - 10\% \cdot \$100 = \$35$;
  - while expected return to Bob is $-90\% \cdot \$100 + 10\% \cdot (\$100 - \$20) = -\$82.$

- Under Bob's belief:
  - His expected return is $90\% \cdot (\$100 - \$20) - 10\% \cdot \$100 = \$62$;
  - while expected return to Joe is $-90\% \cdot \$100 + 10\% \cdot (\$100 - \$50) = -\$85$

Generalize the Bet

What if there is **benefit** from knowing the pillow's content?

- The pillow is either made of cotton or a poisonous material.
  - In the first case, the winner pays \$50 to replace the pillow;
  - in the latter, the winner gets another reward of \$100 from turning in the poisonous pillow.

- The externality is not belief-neutral negative

- Under Joe's belief:
  - Expected return to himself is $90\% \cdot (\$100 - \$50) - 10\% \cdot \$100 = \$35$;
  - Expected return to Bob is $-90\% \cdot \$100 + 10\% \cdot (\$100 + \$100) = -\$70$.
  - A transfer of \$35 from Bob to Joe dominates the bet.

- Under Bob's belief:
  - Expected return to himself is $90\% \cdot (\$100 + \$100) - 10\% \cdot \$100 = \$170$;
  - Expected return to Joe is $-90\% \cdot \$100 + 10\% \cdot (\$100 - \$50) = -\$85$.
  - No transfer can improve one's welfare without hurting the other.

Applications

- The belief-neutral criterion is incomplete
- Nevertheless, useful for spotting negative-sum & positive-sum speculation induced by heterogeneously distorted beliefs.
- These applications are simplified versions of prominent economic models.
  - We aim to demonstrate the relevance of the belief-neutral criterion rather than to advocate for any specific policy recommendation.

Application 1: Over-investment in Bubble Models

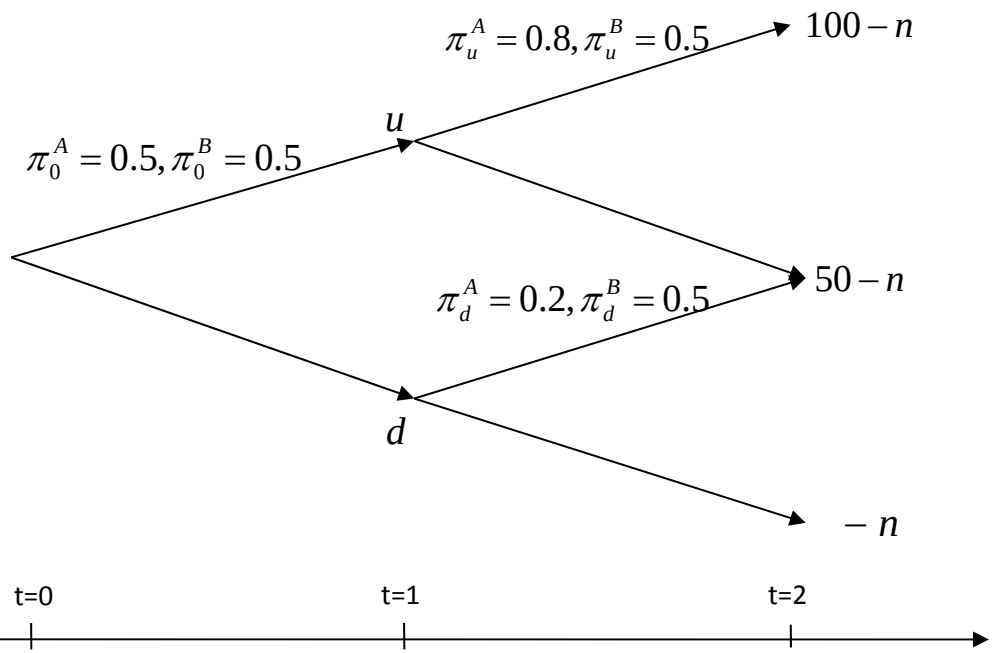
- Bubble models with heterogeneous beliefs and short-sales constrains, e.g., Harrison and Kreps (1978), Morris (1996), Scheinkman and Xiong (2003)
- Price bubbles drive over-investment, e.g., Bolton, Scheinkman and Xiong (2006), Gilchrist, Himmelberg, and Huberman (2005).
- Decreasing return to scale and invest n units at $t = 0$.
  - Firm objective: $\max_n \; n \cdot p_0$

Market setting:
$p_0 = 57.5 - n$
$\max_n \; n \cdot (57.5 - n) \;\Rightarrow n_* = \frac{57.5}{2}.$

If the planner adopts A's beliefs:
$p_0 = E_0^A[\tilde{R}] = 50 - n$
$\max_n \; n \cdot (50 - n) \;\Rightarrow n_* = 25.$

If the planner adopts B's beliefs:
$p_0 = E_0^B[\tilde{R}] = 50 - n$
$\max_n \; n \cdot (50 - n) \;\Rightarrow n_* = 25.$

Over-investment from both A and B's beliefs!

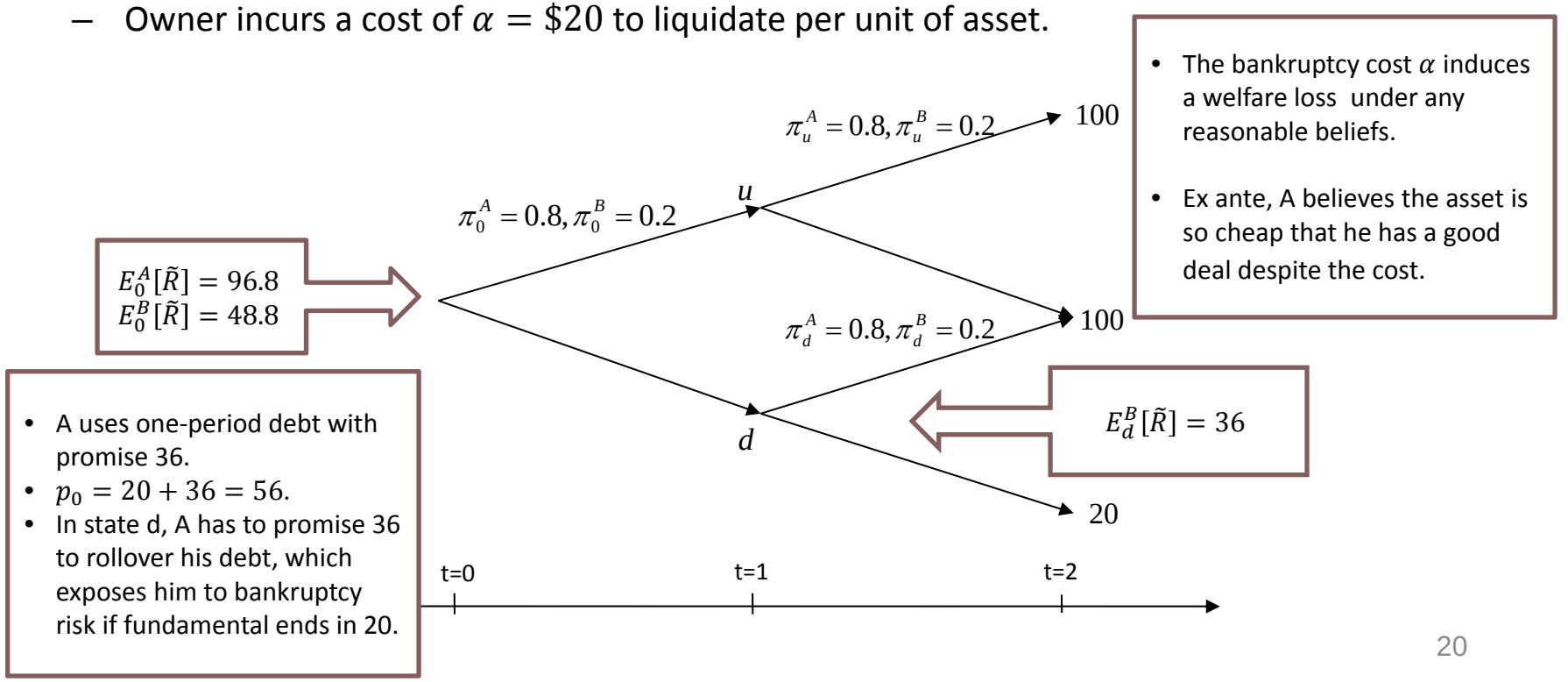


Application 2: Benefits of Speculation in Lemons Model

- Speculation caused by heterogeneous beliefs can be beneficial in lemons model, a la Akerlof (1970).
- We adopt a simple version of Tirole (2012):
  - A seller needs to liquidate a legacy asset to finance a profitable investment opportunity.
  - A lemons problem arises as the seller knows more about the quality of the legacy asset than potential buyers.
  - Speculation induced by the heterogeneously distorted beliefs of potential buyers can lead to positive externality as it overcomes the adverse selection problem.
  - Belief-neutral efficient outcome!

Application 3: Bankruptcy Costs in Leverage Cycle Models

- Cash-constrained optimists tend to use collateralized debt to finance their investments, which fuels initial price boom and later price bust.
  - Geanakoplos (2003, 2009), Fostel and Geanakoplos (2008), Simsek (2010), and He and Xiong (2012).
- A is always more optimistic than B, both risk neutral.
  - At $t = 0$, A is endowed with \$20 and no asset.
  - Owner incurs a cost of $\alpha = \$20$ to liquidate per unit of asset.



- The bankruptcy cost α induces a welfare loss under any reasonable beliefs.
- Ex ante, A believes the asset is so cheap that he has a good deal despite the cost.

- A uses one-period debt with promise 36.
- $p_0 = 20 + 36 = 56$.
- In state d, A has to promise 36 to rollover his debt, which exposes him to bankruptcy risk if fundamental ends in 20.

Application 4: Excessive Risk Taking in Speculative Trading Models

- Many general equilibrium models with heterogeneous beliefs:
  - Speculation between optimists and pessimists lead to endogenous risk and amplified price volatility
  - 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)

- When agents are risk averse, trading makes each agent's consumption more volatile without changing the aggregate wealth.
  - negative-sum game in utility terms if speculation is induced by belief distortions rather than differences in preferences.

Application 5: Consumption-Savings Distortions in Macro Models

- In macro models with investment, speculation induced by belief disagreements can also distort savings and thus investments.
  - Speculation not only makes their consumption more volatile but also distorts the aggregate consumption:

- Sims (2008)
  - Two types of agents disagree about future inflation.
  - Inflation optimists prefer to borrow nominal from pessimists.
    - Substitution effect: speculation motivates both types to save
    - Wealth effect: expectations of speculation gains induce both to consume more
    - Depending on their rate of relative risk aversion, substitution effect dominates wealth effect or vice versa, and thus leads to over- or under-investment.
- Our criterion can also detect belief-neutral inefficiency of such distortions.

Comparing to Gilboa, Samuelson, and Schmeidler (2012)

- Gilboa, Samuelson, and Schmeidler (2012) also recognize that the standard Pareto criterion can be spurious in the presence of conflicting beliefs.
- They propose to weaken the criterion: Allocation f no-betting Pareto dominates g (i.e., $f \succ_{NBP} g$) if
  1. $\forall i, f \succcurlyeq_i g; \exists j, f \succcurlyeq_j g$
  2. There exists at least one measure p_0 such that, for all i,
     $$\int u_i(f(s))dp_0 > \int u_i(g(s))dp_0$$

- The non-betting Pareto criterion rules the bet between Bob and Joe as neither efficient nor inefficient.
  - The additional requirement makes the criterion more restrictive and thus more incomplete than the standard Pareto criterion.

- Our criterion let the planner use a common belief to evaluate each agent's welfare but let the common belief to vary across a large set.
  - The key premise is that the planner is sure of the presence of distorted beliefs.
  - Our criterion gives more clear-cut ranking.

Conclusion

- A belief-neutral welfare criterion for behavioral models
- Opens normative analysis for financial regulation
  - Avoid negative-sum speculation and facilitate positive-sum one
- Separate "preferences" from "belief distortions"
  - Only require presence of belief distortions
  - Don't need to know the truth
- Negative externality
  - Over-investment in bubble models
  - Bankruptcy costs in leverage cycle models
  - Excessive risk taking in speculative trading models
  - Consumption-savings distortions in macro models
- Positive externality
  - Overcoming market breakdown in Lemons models