WISHFUL THINKING IN STRATEGIC ENVIRONMENTS

MUHAMET YILDIZ

Abstract. Towards developing a theory of systematic biases about strategies, I analyze strategic implications of a particular bias: wishful thinking about the strategies. Considering canonical state spaces for strategic uncertainty, I identify a player as a wishful thinker at a state if she hopes to enjoy the highest payoff that is consistent with her information about the others' strategies at that state. I develop a straightforward elimination process that characterizes the strategy profiles that are consistent with wishful thinking, mutual knowledge of wishful thinking, and so on. Every pure-strategy Nash equilibrium is consistent with common knowledge of wishful thinking. For generic two-person games, I further show that the pure Nash equilibrium strategies are the only strategies that are consistent with common knowledge of wishful thinking, providing an unusual epistemic characterization for equilibrium strategies. I also investigate the strategic implications of rationality and ex-post optimism, the situation in which a player's expected payoff weakly exceeds her actual payoff. I show that these strategic implications are generically identical to those of wishful thinking whenever each player's payoff is monotone in others' strategies.

Date: July, 2004.

I thank Daron Acemoglu, Haluk Ergin, Casey Rothschild, Marzena Rostek and Jonathan Weinstein for detailed comments.

1. Introduction

Self-serving biases, such as optimism and wishful thinking, are reportedly common among economic agents.[1] There is a burgeoning theoretical literature that investigates the role of such optimistic and heterogeneous beliefs in economic applications, such as financial markets (Harrison and Kreps (1978), Morris (1996)), bargaining (Posner (1972), Landes (1971), Yildiz (2003,2004), Ali (2003)), collective action (Wilson (1968), Banerjee and Somanathan (2001)), lending (Manove and Padilla (1999)), and theory of the firm (van den Steen (2002)). Another line of research has already set out to test these theories empirically (e.g., Farber and Bazerman (1989)). These models all assume that players hold heterogeneous priors about some underlying parameters, and then apply equilibrium analysis. But equilibrium analysis assumes that players correctly guess the other players' strategies[2]—a suspect assumption in above applications when equilibrium is not derived from some reasonable dominance conditions. Indeed, Dekel, Fudenberg, and Levine (2004) demonstrate that it is hard to interpret such equilibria as outcomes of a learning process.

Most observers agree that it is often harder to predict the outcome of strategic interactions than predicting some physical reality. After all, present theories in physical sciences have much sharper predictions than the theories in social sciences, and well-grounded game-theoretical solution concepts have weak predictive power. This suggests that it would be methodologically preferable to model players having heterogenous beliefs about strategic uncertainty whenever they have heterogenous priors about the underlying parameters of the real world. In particular, it seems more likely that players have more substantial

[1] The literature is large. For some examples, see Larwood and Wittaker (1977), Weinstein (1980), and Babcock and Loewenstein (1997). Self-serving biases disappear in some experiments that control for strategic reporting (Bar-Hillel and Budescu (1995), Kaplan and Ruffle (2004)).

[2] Even when there is uncertainty about strategies in the form of mixed strategies, in order to justify equilibrium analysis, one needs to make similar assumptions, such as mutual knowledge of players' conjectures—which implies that players guess other players' conjectures correctly—or a common prior about strategies (Aumann and Brandenburger (1995)).

self-serving biases about the strategies of others instead of about physical reality. Metaphorically, one would expect that an individual who buys a lottery ticket believing in her luck would also drive over the speed limit and get a speeding ticket, believing that a police officer would forgive her. It is unfortunate then that we do not have a theory of strategic interaction that incorporates self-serving biases about the others' strategies. This paper takes a step towards such a theory, by analyzing two benchmark cases, namely: wishful thinking and ex-post optimism. Taking the dictionary definition of wishful thinking (i.e., identification of one's wishes or desires with reality) literally, I will use this term to refer precisely to the extreme case of optimism.

Let me emphasize that non-equilibrium analysis, which allows for heterogeneous priors about the others' strategies, is one of the most developed areas in game theory. We have solution concepts, such as rationalizability (Bernheim (1984), Pearce (1984)). We have many models of strategic uncertainty, and we know how to check whether the players share a common prior about the others' strategies (Aumann (1976,1987), Feinberg (2000)). Nevertheless, we do not have a theory that analyzes *systematic deviations* from the common-prior assumption, such as optimism, about the others' strategies. Such systematic deviations are the subject matter of this paper. The standard models of strategic uncertainty do contain types with optimism, pessimism, or wishful thinking, but no special attention has been paid to these types. I will simply identify these types and investigate their behavior.

I will use Aumann's (1976) canonical partition model for strategic uncertainty to identify whether a player is a wishful thinker. In this model, at any state, there is a unique strategy profile to be played. Players do not necessarily know the state. At each state, each player has an information cell consisting of the states that she cannot rule out at that state. This cell represents the set of correct assumptions that she takes as given, which is sometimes referred to as her outside information. She also has a (conditional) probability distribution on this cell, which is taken to represent her "subjective" beliefs (about other players' strategies, etc.)

Consider an information cell of a player. This player takes the set of strategy profiles played by the other players on this cell as given. She also knows that she can choose any strategy from the set of her own strategies. The product of these two sets is the set of all possible outcomes according to the information cell. I identify a player as *a wishful thinker* at a state if her expected payoff (according to her own probability distribution) at that state coincides with the highest possible expected payoff one can ever expect within the set of these possible outcomes.

This formalization has two noteworthy properties. Firstly, wishful thinking is a property of the information cell. Therefore, whenever a player is to be identified as a wishful thinker, she knows that an outside observer identifies her as such (although she would probably not accept that she is afflicted with the psychological anomaly of wishful thinking). Secondly, a wishful thinker's strategy-belief pair maximizes her expected payoff among all such pairs. In particular, her strategy maximizes her expected payoff given her beliefs. Hence, she must be rational. While wishful thinking is popularly considered to be a form of irrationality involving self deception and misperception of the world, wishful thinking formally *implies* the standard game-theoretical notion of rationality.

In application, we are mostly interested in the strategic implications of wishful thinking. We also want to know how these implications change when wishful thinking is mutually known, when this mutual knowledge is mutually known, and so on. It is especially important to understand the implications of common knowledge of wishful thinking. This is firstly because the strategy profiles that are consistent with common knowledge of wishful thinking[3] are also consistent with mutual knowledge of wishful thinking at arbitrary order, and hence these are the strategy profiles that remain possible as we allow more and more knowledge of wishful thinking. More importantly, it is highly desirable from a methodological point of view to disentangle the implications of heterogeneous priors or self-serving biases from those of asymmetric information. Researchers,

[3]A strategy profile is consistent with (common knowledge of) wishful thinking means that there is a model and a state at which the strategy profile is played and (it is common knowledge that) the players are wishful thinkers.

such as those in the literature discussed above, commonly accomplish this by assuming that there is no asymmetric information.[4] (This also allows the researcher to focus on belief differences without having to deal with asymmetric information.) In keeping with this methodology, it is desirable to examine the implications of wishful thinking when there is no asymmetric information about it—i.e., when wishful thinking is common knowledge.

Unfortunately, it is often too cumbersome to determine such implications through an epistemic model, which consists of a state space, partitions and conditional probability distributions. More problematically, the set of implications often depends on the model, as models often contain many "common knowledge" assumptions. To overcome this, I develop a straightforward iterated elimination process on the strategy profiles that characterizes the strategies that are consistent with wishful thinking, mutual knowledge of wishful thinking, and so on. The application of this procedure allows a researcher to use wishful thinking—the extreme form of optimism—as an alternative benchmark to equilibrium.

To fix the ideas, consider the simple case of the battle of the sexes game:

	l	r
t	**2,1**	**0,0**
b	0,0	**1,2**

The strategy profile (b, l) is inconsistent with wishful thinking. To see this, suppose that there is a model with a state at which player 1 is a wishful thinker and the outcome is (b, l). At this state player 1 must expect the payoff of 1; a wishful thinker always plays a best reply to a pure strategy. But since (b, l) is the outcome at this state, it is possible that player 2 plays l according to the information cell of player 1. Hence, player 1 could have hoped a higher payoff of 2, by believing that player 2 plays l and playing t in response—a contradiction. Hence the strategy profile (b, l) is eliminated. All the remaining strategy profiles are consistent with wishful thinking and in fact with common knowledge of wishful thinking. To see this, consider a model in which the states

[4]See Squintani (2001) for an exception.

are the remaining three strategy profiles, and each player knows only her own strategy. For example, at a state in which player 1 plays t, she finds both l and r possible. As a wishful thinker, she assigns probability 1 to the state at which player 2 plays l. At the state at which she plays b, she finds only r possible. The only belief she can entertain here assigns probability 1 to r. Since she plays a best reply to this belief, she is a wishful thinker at this state, too. Similarly, player 2 is also a wishful thinker at all states. Since both players are wishful thinkers at all states, it is common knowledge that they are wishful thinkers. Hence, the elimination procedure ends here.

Since wishful thinking implies rationality, only rationalizable strategies can be consistent with common knowledge of wishful thinking, and non-rationalizable strategies are eventually eliminated. It turns out that there is a strong relationship between common knowledge of wishful thinking and Nash equilibria in pure strategies. Firstly, every pure-strategy Nash equilibrium is consistent with common knowledge of wishful thinking, simply because it does not leave any room for any strategic uncertainty (and the players are rational). Hence, they survive the elimination process. More interestingly, for generic,[5] two-person games, I show that if a strategy is not played in a pure-strategy Nash equilibrium, then it must be eliminated, eventually. That is, each strategy that is consistent with common knowledge of wishful thinking must be played in some pure-strategy Nash equilibrium. This yields a characterization: a strategy is consistent with common knowledge of wishful thinking if and only if it is played in a pure-strategy Nash equilibrium. This becomes especially surprising when one recognizes that the analysis of wishful thinking is naturally sensitive to strategically irrelevant transformations, such as adding 2 to the payoff of player 1 when player 2 plays r in the battle of the sexes. This characterization illustrates that, unlike most existing iterative elimination processes, such as iterated

[5]Since wishful thinking is an ordinal notion (Remark 2), the precise assumptions are ordinal and much milder than full genericity assumption in normal form. The precise assumptions are (i) each player has a unique best reply for each pure strategy of the other player and (ii) player 1 is never indifferent between two pure strategy profiles at which she plays a best response to the strategy of player 2.

dominance, the elimination process above leads to strong predictions. Notice, however, that the predictions are not so strong as to imply equilibrium outcomes, as the equilibrium strategies played by different players need not match. For example, in the battle of the sexes, the non-equilibrium strategy profile (t, r) is consistent with common knowledge of wishful thinking. At that state, each player plays according to her favorite equilibrium, believing that *her* favorite equilibrium is to be played. It turns out that, in generic two-person games, it is possible to compute the set of strategy profiles that are consistent with common knowledge of wishful thinking by applying the elimination process only once to the product set of equilibrium strategies.

The above characterization suggests that wishful thinking may become common knowledge only in strategic situations in which there is not much room for wishful thinking, when wishful thinking has little impact on players' beliefs. The strategic uncertainty is reduced to the uncertainty about the equilibrium that is to be played. This does not, however, mean that strategic uncertainty needs to vanish. At the states in which the outcome is not an equilibrium (but wishful thinking is common knowledge), the players have substantial uncertainty about which equilibrium is played and exhibit a clear form of wishful thinking.

This characterization also provides partial support for the theoretical literature that uses equilibrium analysis to study the behavior of optimistic players. It shows that if a researcher allows wishful thinking but sticks to the methodology in this literature (for the motivation above), then she can simply focus on equilibrium strategies. The support is partial because (i) this is true only for generic two-person games while the above models are typically non-generic and (ii) outcomes need not be equilibria due to mismatch, and economic implications of these profiles may substantially differ.

It turns out that the analysis of wishful thinking is easily extended to a natural notion of optimism, namely ex-post optimism. A player is said to be ex-post optimistic at a state if her expected payoff at that state weakly exceeds her actual payoff at the state. A player need not know that she is ex-post optimistic. When it is mutually known that players are ex-post optimistic, however, she knows that she is ex-post optimistic. In that case, her expected payoff must be equal to the highest payoff in her information cell, exhibiting a somewhat weaker form of wishful thinking. I slightly modify the elimination procedure for wishful thinking to characterize the set of strategy profiles that are consistent with common knowledge of rationality and ex-post optimism. I further show that when each player's payoff function is monotonic in other players' strategies, generically, these two elimination procedures are equivalent. That is, the strategic implications of rationality and the knowledge of one's own ex-post optimism are the same as those of wishful thinking. In particular, in a generic, two-person, monotonic game, the strategic uncertainty is reduced to the uncertainty about which equilibrium is played whenever it is common knowledge that players are rational and ex-post optimistic.

A number of authors have developed general game theoretical models that incorporate deviations from expected utility maximization and psychological motivations (such as Geanakoplos, Pearce, Stacchetti (1989)). These models led to incorporation of non-traditional motivations, such as fairness (Rabin (1993)), into game theory. Eyster and Rabin (2000) propose an "equilibrium" notion in which players underestimate the correlation between the other players' action and private information while they correctly estimate the distribution of actions. Rostek (2004) proposes some set-valued solution concepts and discusses various decision rules on these sets. One of these decision rules reflects wishful thinking, though with a somewhat different formulation.

I formulate the problem in the next section and investigate the strategic implications of wishful thinking in Section 3. In Section 4, I define ex-post optimism and investigate its strategic implications. Section 5 concludes. The proofs are relegated to a technical appendix.

2. Formulation

Consider a game (N, S, u) where $N = \{1, 2, \ldots, n\}$ is the set of players, $S = S_1 \times \cdots \times S_n$ is the finite set of strategy profiles,[6] and $u_i : S \to \mathbb{R}$ is the utility function of player i for each $i \in N$. In this set up, players well may have asymmetric information or heterogeneous priors about the physical parameters.[7] These aspects are suppressed in the notation, in order to focus on strategic uncertainty. Strategic uncertainty is modeled through a triple (Ω, I, p), which is called a *model*. Here, Ω is a state space that represents the players' uncertainty about the strategies, where a generic member $\omega \in \Omega$ contains all information about players' strategies, including what each player knows. For each $i \in N$, I_i is the information partition of player i; I write $I_i(\omega)$ for the information cell that contains ω. Here, $I_i(\omega)$ is the set of states that are indistinguishable from ω according to player i. An information partition may be interpreted in two ways. First, it may represent all the "objective" information player i has. In this interpretation, at ω, player i *knows* that one of the states in $I_i(\omega)$ occurs but cannot rule out any of these states. Alternatively, one may take an information partition as a representation of possible sets of assumptions a player may take as given. In this interpretation, $I_i(\omega)$ corresponds to states at which a certain set of assumptions holds. Given any event $F \subseteq \Omega$,

$$K_i(F) = \{\omega | I_i(\omega) \subseteq F\}$$

denotes the set of states at which player i knows that F occurs. The mutual knowledge at any order $m \geq 0$ is represented by operator K^m where $K^0(F) = F$ and $K^m(F) = \cap_{i \in N} K_i(K^{m-1}(F))$. The set of states at which F is common knowledge is denoted by $CK(F)$. Write $\sigma_i(\omega)$ for the strategy played by player i at state ω. Each player knows her strategy so that σ_i is constant on each cell $I_i(\omega)$.

[6] I use the notational convention of $x = (x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n$, $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in X_{-i} = \Pi_{j \neq i} X_j$, and $x = (x_i, x_{-i})$.

[7] The strategic-form representation here allows chance moves, about which players' may have non-common priors.

At ω, player i also has a probability distribution $p_{i,\omega}$ on $I_i(\omega)$, representing her beliefs about the state of the world. The expectation operator with respect to $p_{i,\omega}$ is denoted by $E_{i,\omega}$. In this paper, the probability distribution $p_{i,\omega}$ is taken to be a representation of player's "subjective" beliefs. I will use these probability distributions to identify whether a player is a wishful thinker.

I will define a wishful thinker as a player who expects to enjoy the highest payoff that is possible given her information about the other players' strategies and given that she can choose any of her own strategies. A player i is said to be *a wishful thinker at* ω if and only if

$$E_{i,\omega}\left[u_i(\sigma)\right] = \max_{\nu=\delta_{s_i}\times\mu} E_\nu\left[u_i(s)\right], \tag{2.1}$$

where maximization is taken over all beliefs $\nu = \delta_{s_i} \times \mu$ on $S_i \times \sigma_{-i}(I_i(\omega))$ in which the player knows her own strategy s_i. Here, δ_{s_i} is the probability distribution that puts probability 1 on $\{s_i\}$, and $\sigma_{-i}(I_i(\omega))$ is the set of all strategy profiles s_{-i} of other players played at some state in $I_i(\omega)$. Notice that the standard assumption that player knows her own strategy has no bite in this definition. In fact,

$$\max_s u_i(s) = \max_{\delta_{s_i}\times\mu} E_{\delta_{s_i}\times\mu}\left[u_i(s)\right] = \max_\nu E_\nu\left[u_i(s)\right], \tag{2.2}$$

where s is maximized over $S_i \times \sigma_{-i}(I_i(\omega))$, and the second and third maximizations are over the set of beliefs on $S_i \times \sigma_{-i}(I_i(\omega))$ with and without the above assumption, respectively. Clearly, one can rewrite the condition (2.1), so that a player i is *a wishful thinker at* ω if and only if

$$E_{i,\omega}\left[u_i(\sigma)\right] = \max_{\mu,s_i} E_\mu\left[u_i(s_i, s_{-i})\right], \tag{2.3}$$

where maximization is taken over all beliefs μ on $\sigma_{-i}(I_i(\omega))$ and strategies $s_i \in S_i$. That is, the strategy and the beliefs of a wishful thinker are *as if* she chooses her strategies and beliefs in order to make herself feel happy. Wishful thinking can be defined equivalently in terms of extreme optimism. To this end, rewrite (2.3) as

$$E_{i,\omega}\left[u_i(\sigma)\right] = \max_\mu \max_{s_i} E_\mu\left[u_i(s_i, s_{-i})\right].$$

Hence, a wishful thinker can be considered as a person who chooses a belief to maximize her payoff knowing that she will act rationally with respect to her beliefs. Hence, wishful thinking is maximal optimism under rationality.[8] These equivalent definitions are used interchangeably. The set of all states at which a player i is a wishful thinker is denoted by W_i; $W = \cap_{i \in N} W_i$ is the event that everybody is a wishful thinker.

Notation 1. Write $BR_i(\mu)$ for the set of best responses of player i against his belief μ about the other players' strategies. Write $BR_i(s_{-i})$ instead of $BR_i(\mu)$ if μ assigns probability 1 on s_{-i}. Write

$$B_i = \{(\hat{s}_i, s_{-i}) \,|\, \hat{s}_i \in BR_i(s_{-i}), s_{-i} \in S_{-i}\}$$

for the graph of BR_i on pure strategy profiles.

Remark 1. A player always knows whether she is to be identified as a wishful thinker in this paper: if $\omega \in W_i$, then $I_i(\omega) \subseteq W_i$. Therefore,

$$K_i(W_i) = W_i. \tag{2.4}$$

Remark 2. The first equality in (2.2) shows that wishful thinking is an ordinal notion. A player's attitudes towards risk are irrelevant in determining whether she is a wishful thinker at a given state. The strategic implications of wishful thinking are invariant to monotonic transformations of payoff functions. The only non-generic situations one must ever rule out regarding wishful thinking are indifferences between certain pure strategy profiles. Addition of mixed strategies to the game as pure strategies, again, has at most only trivial effect on the strategic implications of wishful thinking.

[8] Extreme pessimism can be defined similarly as the maximal pessimism under rationality. A player is said to be extremely pessimistic at ω if $E_{i,\omega}[u_i(\sigma)] = \min_\mu \max_{s_i} E_\mu [u_i(s_i, s_{-i})]$. That is, she holds the most pessimistic belief under the constraint that she acts rationally. Notice that extreme pessimism is similar to (but distinct from) ambiguity aversion, defined by $E_{i,\omega}[u_i(\sigma)] = \max_{s_i} \min_\mu E_\mu [u_i(s_i, s_{-i})]$.

Remark 3. Use of Aumann's (1976) partition model is justified on two grounds. Firstly, it is considered to be the canonical model of interactive epistemology, as it reflects the most stringent assumptions about knowledge, namely: a person does not know a false statement as truth; a person knows whether she knows; and she can use conjunction to make further inferences. Exploring the strategic implications of wishful thinking within this model allows me to disentangle the effects of wishful thinking from the possible effects that come form dropping some of these standard assumptions of game theory. Nevertheless, all of these assumptions have been challenged in modeling bounded rationality and strategic uncertainty. It will therefore be important to extend the analysis to such more permissive models. Secondly, my formulation of wishful thinking requires a set of assumptions or an objective assessment of the world as a reference in addition to individuals' subjective assessments. Aumann's canonical model perfectly meets this need as it consists of partitions of the state space and conditional probability distributions. This allows a parsimonious model of wishful thinking. In purely subjective models of beliefs, such as the universal type space of Mertens and Zamir (1985) on the strategy space, one needs to add such sets of correct assumptions or reference beliefs to the model. Since these reference beliefs are irrelevant to the players' decision problems, such a model will be less parsimonious and may appear ad hoc.

3. Strategic Implications of Wishful Thinking

In this section, I will characterize the strategies that are consistent with wishful thinking, mutual knowledge of wishful thinking, and so on. The characterization will be given by an iterative elimination process. Firstly, since wishful thinking is stronger than rationality, common knowledge of wishful thinking will lead to a refinement of rationalizability, i.e., all non-rationalizable strategies will eventually be eliminated. Secondly, any Nash equilibrium in pure strategies will survive the iterated elimination process. More surprisingly, I will show that for generic two person games, these are the only strategies that survive. This will yield an unlikely epistemic characterization for pure-Nash-equilibrium strategies.

The next lemma describes a defining property of the strategies played by a wishful thinker. I will use this property to define the elimination procedure.

Lemma 1. *For any $F \subseteq \Omega$, $i \in N$, and any $\hat{s} \in \sigma(K_i(F) \cap W_i)$, there exists $(\hat{s}_i, s_{-i}) \in B_i \cap \sigma(F)$ such that*

$$u_i(\hat{s}_i, s_{-i}) \geq \max_{s_i} u_i(s_i, \hat{s}_{-i}). \tag{3.1}$$

The intuition behind this lemma is simple. A wishful thinker always plays a best reply against some pure strategy profile (by (2.2)). Hence, if a wishful thinker plays a strategy $\hat{s}_i$ knowing that some event F is true, then she must be targeting a best reply against a pure strategy s_{-i} that is consistent with event F. Moreover, the targeted payoff, $u_i(\hat{s}_i, s_{-i})$, cannot be lower than the payoff from playing a best reply to a strategy profile $\hat{s}_{-i}$ of others that is possible in her information set. For otherwise, she would obtain a higher payoff by believing that the other players play $\hat{s}_{-i}$.

Lemma 1 rules out most strategy profiles as strategic outcomes when some of the players are wishful thinkers. For any generic game and any two distinct strategy profiles $s, s' \in B_i$ in which player i plays a best response, either $\left(s_i, s'_{-i}\right)$ or (s'_i, s_{-i}) is eliminated depending on whether $u_i(s) < u_i(s')$ or $u_i(s) > u_i(s')$. This leads to a powerful elimination procedure, which characterizes the strategy profiles that are consistent with wishful thinking, mutual knowledge of wishful thinking, and so on.

Elimination Procedure.

1. Initialization: Set $X^{-1} = S$.
2. Elimination: For any $m \geq 0$, eliminate all the strategy profiles $\hat{s}$ that fail (3.1) for some i and for $\sigma(F) = X^{m-1}$. (That is, eliminate $\hat{s}$ if there exists i for which there does not exists any $(\hat{s}_i, s_{-i}) \in B_i \cap X^{m-1}$ with $u_i(\hat{s}_i, s_{-i}) \geq \max_{s_i} u_i(s_i, \hat{s}_{-i})$.) Call the remaining strategy profile X^m.
3. Iterate step (2).

Note that this is an elimination of *strategy profiles*. Moreover, it contains the iterated elimination of strategies that are not a best reply to a pure strategy. In this elimination, if $\hat{s}_i$ is a best reply only to s_{-i} and the strategy profile $(\hat{s}_i, s_{-i})$ was eliminated at some previous iteration, then the entire strategy $\hat{s}_i$ is to be eliminated—even if some part of $\hat{s}_{-i}$ is still available.

More formally, I define a mapping $\phi : 2^S \to 2^S$ by setting

$$\phi(X) = X \cap \left\{ \hat{s} | \forall i : \max_{s_i} u_i(s_i, \hat{s}_{-i}) \leq \max_{(\hat{s}_i, s_{-i}) \in B_i \cap X} (\hat{s}_i, s_{-i}) \right\} \tag{3.2}$$

at each X where I use the usual convention that the maximum over the empty set yields $-\infty$. The sequence $(X^{-1}, X^0, \ldots)$ is recursively defined by $X^{-1} = S$ and

$$X^m = \phi\left(X^{m-1}\right) \qquad (m \geq 0).$$

The limit of the sequence is

$$X^\infty = \bigcap_{m=0}^{\infty} X^m.$$

Since there are only finitely many strategy profiles, the elimination process stops at some iteration m, and we have $X^\infty = X^m$ for some m. I will now illustrate how the elimination procedure is applied.

Example 1. Consider the following two-person game, where player 1 chooses between the rows, and player 2 chooses between the columns:

	α	β	γ	δ
a	$3^*, 0$	$-1, 0$	0,0	0,2*
b	$0, 0$	$2^*, 1^*$	0,0	0,0
c	0,0	0,0	$1^*, 2^*$	1*,0
d	2,3*	1,0	0,0	0,0

In this table, the asterisk after a player's payoff indicates that the player is playing a best reply to the other player's strategy at that profile. Notice that no strategy is weakly dominated, and hence all strategies are admissible (and rationalizable). Let us apply the above elimination procedure. Take $m = 0$. For Player 1, the strategy d is eliminated, as it is not a best reply to any pure strategy. For player 1 again, the profile (b, α) is eliminated, as $\max_{(b,s_2)\in B_1} u_1(b, s_2) = 2 < 3 = \max_{s_1} u_1(s_1, \alpha)$. Similarly, (c, α) and (c, β) are eliminated. No other profile is eliminated for player 1 at this round. For example, (a, β) is not eliminated because player 1 targets the payoff of 3, while she can get only 2 by responding to β. For player 2, among the strategy profiles that have not been eliminated already, (a, β) is eliminated as $u_2(a, \delta) > u_2(b, \beta)$; no other profile is eliminated. X^0 consists of the strategy profiles in bold:

	α	β	γ	δ
a	$\mathbf{3^*, 0}$	$-1, 0$	**0,0**	**0,2***
b	$0, 0$	$\mathbf{2^*, 1^*}$	**0,0**	**0,0**
c	0,0	0,0	$\mathbf{1^*, 2^*}$	**1*****,0**
d	2,3*	1,0	0,0	0,0

For $m = 1$, strategy α (or simply the strategy profile (a, α)) is eliminated for player 2. This is because α is a best reply only against d, but (d, α) has been eliminated. There are no more eliminations for $m = 1$: $X^1 = X^0 \setminus \{(a, \alpha)\}$. For $m = 2$, strategy a is eliminated as a is not a best reply for player 1 in any remaining profile in the first row. This is the only elimination. Hence, X^2 consists of the boldface strategy profiles in the second and third rows above. For $m = 3$, strategy δ is eliminated. The elimination process stops here. Therefore, $X^3 = X^4 = \cdots = X^\infty = \{(b, \beta), (b, \gamma), (c, \gamma)\}$.

Remark 4. The elimination process is monotonic, i.e., $X \subseteq Y \Rightarrow \phi(X) \subseteq \phi(Y)$. Hence, the limit set X^∞ does not change if one fails to eliminate certain strategies at some step or applies different orders.

The next result states that X^m is precisely the strategies that are consistent with mth-order mutual knowledge of wishful thinking. Therefore, using the elimination procedure above, a researcher can investigate the strategic implications of wishful thinking directly from the strategy profiles—without dealing with abstract, complicated models of strategic uncertainty.

Proposition 1. *For any model* (Ω, I, p), *and any* $m \geq 0$,

$$\sigma\left(K^m\left(W\right)\right) \subseteq X^m;$$

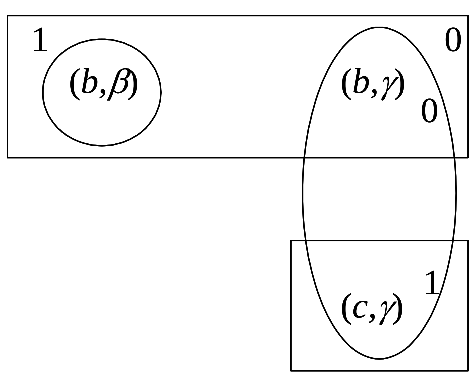
in particular,

$$\sigma\left(CK\left(W\right)\right) \subseteq X^{\infty}.$$

Moreover, there exist models (Ω, I, p) *in which the above inclusions are equalities.*

The first statement is given by inductive applications of Lemma 1. The proof of the second part involves constructing a submodel for each strategy profile in X^m, where the strategy profile is played at a state in which wishful thinking is mth-order mutual knowledge. One constructs such a model using information sets with only one or two states, and at such information cells, Lemma 1 characterizes the wishful thinking behavior. Integrating these models into one model, one obtains the desired model. I will next present a model for the common knowledge case in Example 1.

Example 1 (*continued*). In the following model, wishful thinking is common knowledge, and each strategy profile in X^{∞} is played at some state. Take $\Omega = X^{\infty}$ and $\sigma(\omega) = \omega$ at each ω. Each player knows her own strategy; hence $I_1 = \{\{(b,\beta), (b,\gamma)\}, \{(c,\gamma)\}\}$ and $I_2 = \{\{(b,\beta)\}, \{(b,\gamma), (c,\gamma)\}\}$. Take also $p_{1,(b,\beta)}((b,\beta)) = 1$ and $p_{2,(c,\gamma)}((c,\gamma)) = 1$. This model is described in the following diagram, where the information cells of Player 1 are rectangular:



Clearly, each player is a wishful thinker at each state, and hence $W = K^1(W) = K^2(W) = \cdots = CK(W) = \Omega = X^{\infty}$. Therefore, $\sigma(CK(W)) = X^{\infty}$.

Notice in Example 1 that every strategy that is part of a profile in X^{∞} is also a part of a pure-strategy Nash equilibrium. This is, in fact, not a coincidence. Let

$$NE = \bigcap_{i \in N} B_i$$

be the set of all pure-strategy Nash equilibria; recall that B_i is the set of profiles in which player i plays a best reply to others' strategies. Let also

$$NE_i = \{s_i | \exists s_{-i} : (s_i, s_{-i}) \in NE\}$$

be the set of all strategies of a player i that are played in some pure-strategy Nash equilibrium. Similarly, let

$$X_i^{\infty} = \{s_i | \exists s_{-i} : (s_i, s_{-i}) \in X^{\infty}\}$$

be the set of all strategies of a player i that is consistent with common knowledge of wishful thinking. The next two propositions establish the close relationship between the pure strategy Nash equilibria and common knowledge of wishful thinking.

Proposition 2. *Every pure strategy Nash equilibrium is consistent with common knowledge of wishful thinking:*

$$NE \subseteq X^{\infty}.$$

In a pure-strategy Nash equilibrium there is no strategic uncertainty, and hence players do not have any freedom of entertaining different beliefs. Hence, all players, independent of their level of optimism, hold the same correct beliefs. Formally, any pure-strategy Nash equilibrium is consistent with a model with a single state at which each player plays according to the equilibrium. In such a model, it is common knowledge that each player is a wishful thinker. The next proposition states a more surprising and substantive fact about two-player games. The assumptions of this proposition generically hold.

Proposition 3. *For any two-person game, assume (i) for each s_{-i}, there exists a unique best reply $s_i \in BR_i(s_{-i})$, and (ii) Player 1 is not indifferent between* *any two distinct strategy profiles $s, s' \in B_1$. Then, only pure Nash equilibrium strategies are consistent with common knowledge of wishful thinking:*

$$X_i^\infty = NE_i \qquad (\forall i \in N). \tag{3.3}$$

Moreover,

$$\begin{aligned} X^\infty &= \phi\left(NE_1 \times NE_2\right) \\ &= \{(\hat{s}_1, \hat{s}_2) | \exists (\hat{s}_1, s_2), (s_1, \hat{s}_2) \in NE : u_1(\hat{s}_1, s_2) \geq u_1(s_1, \hat{s}_2), \\ &\qquad\qquad\qquad\qquad\qquad\qquad u_2(s_1, \hat{s}_2) \geq u_2(\hat{s}_1, s_2)\}. \end{aligned} \tag{3.4}$$

The first part of the proposition characterizes common knowledge of wishful thinking in terms of strategies. It states that in a generic two-person game, the only strategies that are consistent with common knowledge of wishful thinking are the pure equilibrium strategies. The second part gives a practical characterization for the strategy profiles that are consistent with common knowledge of wishful thinking. It states that this set can be computed by simply applying the elimination procedure only once to the set $NE_1 \times NE_2$.

The proof can be summarized as follows. First, under assumption (i), one can show that the restrictions of the best response functions to X^∞ (which are defined from X_j^∞ to X_i^∞) are well-defined, one-to-one, and onto. Together with assumption (ii), this allows one to rearrange the strategies so that $B_1 \cap X^\infty$ is equal to the diagonal of $X_1^\infty \times X_2^\infty$ and the payoff of player 1 is decreasing along this diagonal—as in Example 1. As in Example 1, this implies that all the strategy profiles that are under the diagonal and in $X_1^\infty \times X_2^\infty$ must have been eliminated. Using arguments similar to the one used to eliminate strategy profile (a, α), one then shows that if Player 2 does not give a best reply in a strategy profile on the diagonal, then the strategy of player 2 must have been eliminated. Therefore, the diagonal of $X_1^\infty \times X_2^\infty$ is equal to NE. This proves (3.3). Since the elimination process is monotonic, (3.3) implies that the result would not change if one started elimination from the set $NE_1 \times NE_2$. The latter elimination stops at the first step, yielding (3.4).

This provides an unusual epistemic characterization for pure-Nash-equilibrium strategies in terms of common knowledge of wishful thinking. More importantly, it suggests that there is little room left for optimism or pessimism when the wishful thinking is common knowledge. In fact, (3.4) establishes that the strategic uncertainty is reduced to uncertainty about the equilibrium played. Nevertheless, common knowledge of wishful thinking is characterized by equilibrium strategies—*not by equilibria*. As stated in (3.4), two players may play equilibrium strategies that correspond to two different equilibria at some state in which wishful thinking is common knowledge. In such a state there is still substantial strategic uncertainty remaining, and players exhibit a clear form of wishful thinking. For example, at state (b, γ) in Example 1, each player incorrectly believes that they will play *her* favorite equilibrium. This is a general fact. By (3.4), if the outcome $\hat{s} = (\hat{s}_1, \hat{s}_2) \in X^{\infty}$ is not an equilibrium already, then there are two equilibria $(\hat{s}_1, s_2)$ and $(s_1, \hat{s}_2) \in NE$ that the players consider as possible and rank in diagonally opposing orders: $u_1(\hat{s}_1, s_2) > u_1(s_1, \hat{s}_2)$ and $u_2(s_1, \hat{s}_2) \geq u_2(\hat{s}_1, s_2)$. Each player plays according to her own favorite equilibrium, believing that her own favorite equilibrium is to be played.

If the equilibria are strictly Pareto-ranked, then the players cannot have such opposing rankings. In that case, the outcome is necessarily an equilibrium.

Corollary 1. *Under the (generic) assumptions of Proposition 3, if the equilibria are Pareto-ranked with strict inequalities, then*

$$X^{\infty} = NE.$$

Proposition 3 has established already that for generic two-person games, when wishful thinking is common knowledge, strategic uncertainty is reduced to uncertainty about which equilibrium strategies are played. As in Example 1, one can indeed construct a model in which wishful thinking is common knowledge, all strategy profiles that are consistent with common knowledge of wishful thinking are played, and at each state each player assigns probability 1 to an equilibrium. That is, the players are in agreement that an equilibrium is played, but they may disagree about which equilibrium is played. This is stated by the following corollary, which is suggested by Haluk Ergin.

Corollary 2. *Under the assumptions of Proposition 3, there exists a model in which* $\sigma(CK(W)) = X^{\infty}$ *and for each* $\omega \in CK(W)$ *and* i, $p_{i,\omega}\left(\sigma^{-1}(NE) \cap I_i(\omega)\right) = 1$*, i.e., each player is certain that an equilibrium is played.*

Remark 5. Common knowledge of wishful thinking differs from usual epistemic foundations for equilibrium, such as the sufficient conditions of Aumann and Brandenburger (1995). For example, at state (b, γ) in Example 1, rationality is common knowledge, but the players do not know each other's conjectures, and these conjectures do not form a Nash equilibrium. Notice that each player is certain about the other player's conjecture, but she is wrong.

Proposition 3 immediately implies that, for generic two-person games, existence of a strategy profile that is consistent with common knowledge of wishful thinking is equivalent to existence of a Nash equilibrium in pure strategies.

Corollary 3. *Under the assumptions of Proposition 3,*

$$X^{\infty} \neq \varnothing \iff NE \neq \varnothing.$$

The next two examples show that the assumptions in Proposition 3 are not superfluous. The first one shows that the characterizations need not be true when there are three or more players.

Example 2. Consider the following three-player game where player 3 chooses the matrices (λ and ρ denote the matrices on the left and right, respectively). Below $\varepsilon_i : S \to (0, 1)$, $i \in N$, are arbitrary functions; arguments are suppressed.

	l	r
t	$2, -\varepsilon_2, -\varepsilon_3$	$-\varepsilon_1, 0, -\varepsilon_3$
b	$-\varepsilon_1, -\varepsilon_2, -\varepsilon_3$	$1,1,1$

	l	r
t	$-\varepsilon_1, \varepsilon_2, \varepsilon_3$	$-\varepsilon_1, -\varepsilon_2, \varepsilon_3$
b	$\varepsilon_1, \varepsilon_2, 3$	$\varepsilon_1, -\varepsilon_2, -\varepsilon_3$

Check that $NE = \{(b, r, \lambda), (b, l, \rho)\}$, but $X^{\infty} = S \setminus \{(b, l, \lambda)\}$, so that the non-equilibrium strategy t is consistent with common knowledge of wishful thinking.

Example 3. In the matching-penny game, $X^\infty = S$, while $NE = \varnothing$, showing that assumption (i) is not superfluous in Proposition 3. In the game

	l	r
t	2^*,2^*	1^*,0
b	2^*,0	0,1^*

$NE = \{(t,l)\}$, but $X^\infty = S \backslash \{(t,r)\}$, containing non-equilibrium strategies b and r. Hence, assumption (ii) is not superfluous, either.

The next example illustrates that, in a generic two-person game, $X^\infty = \varnothing$ whenever $NE = \varnothing$.

Example 4. Consider the following game with no pure-strategy Nash equilibrium.

	l	r
t	0,2^*	3^*,1
b	2^*,0	1,3^*

Now, $X^0 = \{(t,l), (t,r)\}$ as (b,l) and (b,r) are deleted at the first iteration for players 2 and 1, respectively. Since $B_2 = \{(b,r), (t,l)\}$, $X^0 \cap B_2 = \{(t,l)\}$, and hence (t,r) is eliminated. No strategy is eliminated for player 1, and $X^1 = \{(t,l)\}$. But, $B_1 = \{(t,r), (b,l)\}$, and hence $X^1 \cap B_1 = \varnothing$. Therefore, $X^2 = \varnothing$.

4. Ex-post optimism and its strategic implications

Wishful thinking serves well as an alternative benchmark to equilibrium. Nevertheless it is clearly an extreme case. A general theory must then consider a more general notion of optimism. Optimism is a straightforward notion when there is no private information. In the presence of private information, there are many different notions of optimism, and the very task of finding a suitable notion of optimism seems to be a challenge. Since strategic uncertainty must involve some asymmetric information whenever it is present (as each player knows her own strategy), a theory of optimism about strategies must face this challenge. In this section, I will consider an apparently natural notion of optimism, namely ex-post optimism. I will show that its strategic implications will be similar to that of wishful thinking. It will be clear, however, that this is also a relatively extreme notion, and thus the challenging task of a suitable non-extreme notion of optimism is left to further research.

4.1. **Definitions.** A player i is said to be *ex-post optimistic at* ω if her expected payoff at ω (according to her own expectations) is at least as high as her actual payoff at ω, i.e.,

$$E_{i,\omega}\left[u_i\left(\sigma\right)\right] \geq u_i\left(\sigma\left(\omega\right)\right). \tag{4.1}$$

Notice that, since each player knows her own strategy, the actual outcome at any state is a distributed information among the players. That is, one can figure it out by pooling each players' private information. Hence, this is the only notion of optimism with respect to a set of reference beliefs that contains the distributed information among players. I will write O_i for the set of states at which player i is ex-post optimistic; $O = \cap_{i \in N} O_i$ denotes the set of states at which every player is ex-post optimistic.

An ex-post optimistic player need not know that she is ex-post optimistic. Indeed it will be an extreme case of optimism when one assumes that a player knows that she is ex-post optimistic. This assumption corresponds to the event

$$\begin{aligned} K_i\left(O_i\right) &= \left\{\omega | \forall \omega' \in I_i\left(\omega\right): E_{i,\omega}\left[u_i\left(\sigma\right)\right] \geq u_i\left(\sigma\left(\omega'\right)\right)\right\} \\ &= \left\{\omega | E_{i,\omega}\left[u_i\left(\sigma\right)\right] = \max_{\omega' \in I_i(\omega)} u_i\left(\sigma\left(\omega'\right)\right)\right\}. \end{aligned}$$

Here, player i expects to enjoy the highest payoff available at $I_i\left(\omega\right)$. This corresponds to a somewhat weaker form of wishful thinking. It defines a situation in which a player takes her own strategy as predetermined (as she knows it) and chooses a belief in order to make herself feel happy. Hence I will call such a player a *fatalistically wishful thinker* (at ω). The set of all such states is denoted by $\tilde{W}_i \equiv K_i\left(O_i\right)$. One can easily check that wishful thinking implies fatalistic wishful thinking and hence ex-post optimism:

$$W_i \subseteq \tilde{W}_i \equiv K_i\left(O_i\right) \subseteq O_i.$$

Optimism alone is not related to rationality, and hence I will make separate knowledge assumptions on rationality.

Notation 2. Write $\mu_{i,\omega} \equiv p_{i,\omega} \circ \sigma_{-i}^{-1}$ for the belief of player i about the other players' strategies at ω.

A player i is said to be *rational* at ω if $\sigma_i(\omega) \in BR_i\left(\mu_{i,\omega}\right)$. The set of all states at which player i is rational is denoted by R_i; $R = \cap_{i \in N} R_i$ denotes the event that every player is rational. Recall that rationality is implied by wishful thinking:

$$W_i \subseteq R_i. \tag{4.2}$$

The next example shows that all of the above inclusions can be strict.

Example 5. Consider the following game in which the column player can be taken to be Nature:

	l	r
t	0,0	3,0
b	2,0	1,0

Consider $I_1(\omega) = \{\omega, \omega'\}$ with $p_{1,\omega}(\omega) = 1$, $\sigma_1(\omega) = \sigma_1(\omega') = b$, $\sigma_2(\omega) = l$, and $\sigma_2(\omega') = r$. Clearly, Player 1 is rational at ω as b is the unique best reply to her belief that action l is played with probability 1. She is a fatalistically wishful thinker (and thus ex-post optimistic) because she expects to enjoy the highest possible payoff at $I_1(\omega)$: $E_{1,\omega}(u_1(\sigma)) = 2 = u_1(\sigma(\omega)) > u_1(\sigma(\omega')) = 1$. But she is not a wishful thinker, for she could have expected even a higher payoff of 3, by being certain that player 2 plays r and playing t in response.

4.2. **Strategic Implications.** I will now investigate the strategic implications of ex-post optimism and rationality. In analyzing the strategic implications of wishful thinking it was very helpful to note that a wishful thinker always plays a best reply to a pure strategy. Although this will also be generically true for rational, fatalistically wishful thinkers, it will not be the case in general.

Example 6. In Example 5, consider the game

	l	r
t	0,0	3*,0
m	3*,0	0,0
b	2,0	2,0

instead, and let $p_{1,\omega}(\omega) = p_{1,\omega}(\omega') = 1/2$, leaving the rest of the example unchanged. Check that player 1 is rational and a fatalistically wishful thinker, but she does not play a best reply to any pure strategy.

I will not make any genericity assumption for the general analysis. In general, if a player i is a fatalistic wishful thinker at ω, then her payoff is constant on the support of her probability distribution $\mu_{i,\omega}$. If she is rational, she plays a best reply against such a belief. Let B_i^X be the set of such strategy profiles when i knows that a strategy profile in $X \subseteq S$ is played, i.e.,

$$B_i^X = \{s | s_i \in BR_i(\mu_i),\ s_{-i} \in \operatorname{supp}(\mu_i), \forall s'_{-i} \in \operatorname{supp}(\mu_i) : \\ u_i\left(s_i, s'_{-i}\right) = u_i(s_i, s_{-i}),\ \left(s_i, s'_{-i}\right) \in X\}.$$

Clearly,

$$B_i \cap X \subseteq B_i^X. \tag{4.3}$$

Now, if a fatalistic wishful thinker plays a strategy, knowing that a strategy profile in X is played, then she must be targeting a payoff level that is obtained in a strategy profile in B_i^X. Her payoff from the actual outcome cannot be higher than this targeted payoff. This is stated in the following lemma.

Lemma 2. *For any $F \subseteq \Omega$, $i \in N$, and any $\hat{s} \in \sigma\left(K_i(F) \cap \tilde{W}_i \cap R_i\right)$, there exists $(\hat{s}_i, s_{-i}) \in B_i^{\sigma(F)}$ such that*

$$u_i(\hat{s}_i, s_{-i}) \geq u_i(\hat{s}). \tag{4.4}$$

As before, this summarizes the restrictions imposed by rationality and fatalistic wishful thinking on strategies. Towards, characterizing the strategic implications of these assumptions, I define a mapping $\psi : 2^S \to 2^S$ by setting

$$\psi(X) = X \cap \left\{ \hat{s} | \forall i : u_i(\hat{s}) \leq \max_{(\hat{s}_i, s_{-i}) \in B_i^X} (\hat{s}_i, s_{-i}) \right\} \tag{4.5}$$

at each X where I use the usual convention that the maximum over the empty set yields $-\infty$. I iteratively define a decreasing family $(Y^{-1}, Y^0, Y^1, \ldots)$ of sets by $Y^{-1} = S$ and

$$Y^m = \psi\left(Y^{m-1}\right) \qquad (m \geq 0).$$

Let Y^∞ be the limit set. This sequence is the outcome of the following elimination process:

Elimination Procedure for Ex-post Optimism.

(1) Initialization: Set $Y^{-1} = S$.
(2) Elimination: For any $m \geq 0$, eliminate all the strategy profiles $\hat{s}$ that fail (4.4) for some i and for $\sigma(F) = Y^{m-1}$. Call the remaining strategy profile Y^m.
(3) Iterate step (2).

Example 1 (*continued*). Now, apply this elimination procedure to the original game

	α	β	γ	δ
a	$3^*, 0$	$-1, 0$	0,0	0,2*
b	$0, 0$	$2^*, 1^*$	0,0	0,0
c	0,0	0,0	$1^*, 2^*$	1^*,0
d	2,3*	1,0	0,0	0,0

where B_i is indicated by a single asterisk, and $B_i^S = B_i$. For $m = 0$, strategy d of player 1 is eliminated. There are no more eliminations for $m = 0$. For $m = 1$, strategy α of player 2 is eliminated. For $m = 2$, strategy a of player 1 is eliminated. Finally, for $m = 3$, strategy δ of player 2 is eliminated, and the elimination process stops there. Y^{∞} consists of the strategy profiles in bold:

	α	β	γ	δ
a	$3^{*},0$	$-1,0$	0,0	0,2^{*}
b	$0,0$	$\mathbf{2^{*},1^{*}}$	**0,0**	0,0
c	0,0	**0,0**	$\mathbf{1^{*},2^{*}}$	1^{*},0
d	2,3^{*}	1,0	0,0	0,0

The next proposition establishes that, using this elimination process, one can indeed characterize the set of strategy profiles that are consistent with rationality and knowledge of one's own ex-post optimism, strategy profiles that are consistent with mutual knowledge of this, and so on. In the limit, one characterizes the set of strategy profiles that are consistent with common knowledge of rationality and ex-post optimism.

Proposition 4. *For any model* (Ω, I, p)*, and any* $m \geq 0$*,*

$$\sigma\left(K^{m}\left(\cap_{i\in N} K_{i}\left(O_{i}\right)\cap R\right)\right)\subseteq Y^{m};$$

in particular,

$$\sigma\left(CK\left(R\cap O\right)\right)\subseteq Y^{\infty}.$$

Moreover, there exist models (Ω, I, p) *in which the above inclusions are equalities.*

Since wishful thinking implies both rationality and knowledge of one's own ex-post optimism (i.e., $W_i \subseteq R_i \cap K_i(O_i)$), each set Y^m must contain X^m. Indeed, by (4.3), $\phi(X) \subseteq \psi(X)$ for each X, and hence $X^m \subseteq Y^m$ for each m. Of course, Y^{∞} is a subset of rationalizable strategy profiles.

The strategic implications of rationality and ex-post optimism may be significantly weaker in general. I will now provide conditions under which the gap closes significantly. Firstly, under the generic assumption that

$$s_{-i}\neq s'_{-i}\Rightarrow u_{i}\left(s_{i},s_{-i}\right)\neq u_{i}\left(s_{i},s'_{-i}\right)\qquad\left(\forall i,s_{i}\right), \tag{4.6}$$

$B_i^X \equiv B_i \cap X$, and hence a rational and fatalistically wishful thinker plays a best reply to a pure strategy. Hence, the elimination procedure for ex-post optimism requires elimination of strategies that are not a best reply to a pure strategy. As shown in application to Example 1, this may lead to strong predictions. I will now describe an important class of games, which includes many classical models, in which the strategic implications of rationality and ex-post optimism are the same as those of wishful thinking.

Monotonic Games. I will say that a game is *monotonic* if and only if for all $s_i, s_i', s_{-i}, s_{-i}'$

$$u_i(s_i, s_{-i}) > u_i\left(s_i, s_{-i}'\right) \iff u_i(s_i', s_{-i}) > u_i\left(s_i', s_{-i}'\right).$$

This condition is clearly satisfied in many classical applications, such as Cournot competition, Bertrand competition, partnership games, and provision of public goods.

Lemma 3. *For any monotonic game that satisfies the generic condition (4.6),*

$$\phi = \psi.$$

The proof is based on two observations. Under condition (4.6), $B_i^X = B_i \cap X$. In that case, for a monotonic game (4.4) implies (3.1). Lemma 3 immediately implies that strategic implications of knowledge of ex-post optimism and rationality are the same as those of wishful thinking, as stated in the next result.

Proposition 5. *For any monotonic game that satisfies the generic condition (4.6),*

$$X^m = Y^m \qquad (\forall m \geq 0).$$

In particular, generically, when it is common knowledge that players are ex-post optimistic and rational, then strategic uncertainty is reduced to uncertainty about which equilibrium is played.

Proposition 6. *Consider any monotonic two-person game that satisfies (4.6) and the conditions of Proposition 3, which are all generically satisfied. Pure*

Nash equilibrium strategies are the only strategies that are consistent with common knowledge of rationality and ex-post optimism:

$$Y_i^\infty = NE_i \qquad (\forall i \in N). \tag{4.7}$$

Moreover,

$$Y^\infty = \phi\left(NE_1 \times NE_2\right).$$

Remark 6. Although ex-post optimism itself is a weak condition, the strategic implications of common knowledge of ex-post optimism and rationality are highly strong. This is because knowledge of ex-post optimism rules out equally common case of ex-post pessimism. This is necessitated by the modeling assumption that players cannot know a false statement as truth, namely the Truth Axiom of the standard model of interactive epistemology, which is also used here. Clearly, it is highly desirable to develop a more permissive model of optimism by considering some other notions of optimism and also considering a more permissive model of knowledge, such as a Bayesian model with certainty as the knowledge operator, in which one can assume "common knowledge" of optimism, without assuming that a player knows that she is optimistic.

The following table summarizes the findings of this paper; S^∞ denotes the set of rationalizable strategy profiles.

Case	Relationship
General	$NE \subseteq X^\infty \subseteq Y^\infty \subseteq S^\infty$
Two-person, generic	$NE \subseteq X^\infty = \phi\left(NE_1 \times NE_2\right) \subseteq NE_1 \times NE_2$
Monotonic, generic	$X^\infty = Y^\infty$
Two-person, monotonic, generic	$NE \subseteq Y^\infty = \phi\left(NE_1 \times NE_2\right) \subseteq NE_1 \times NE_2$

5. Conclusion

Self-serving biases are reportedly common. Moreover, given the elusive nature of strategic uncertainty, one would expect to have more substantial self-serving biases about other players' strategies. It is surprising then such systematic biases about the other players' strategies are typically assumed away by modelers, even in the literature on optimism. What is worse, there is no game theoretical framework that incorporates such systematic deviations from the common-prior assumption—although the use of heterogenous priors about strategies is becoming mainstream in game theory. In this paper, I take a first step towards a theory of such deviations, focusing on the extreme form of optimism, namely wishful thinking. I develop a framework for analyzing wishful thinking about strategies. I use the canonical model for strategic uncertainty to identify whether a player is a wishful thinker and develop a straightforward elimination process directly on strategy profiles to characterize the set of strategy profiles that are consistent with wishful thinking, mutual knowledge of wishful thinking, and so on. I further show that in generic two-person games, pure Nash-equilibrium strategies are the only strategies that are consistent with common knowledge of wishful thinking. In such games, wishful thinking can be common knowledge only in cases in which strategic uncertainty is reduced to uncertainty about the equilibrium that is played, and if a researcher assumes away asymmetric information about whether a player is a wishful thinker in order to disentangle the effects of wishful thinking from that of informational differences, then he may as well focus on equilibrium strategies.

With a slight modification, I extend my analysis of wishful thinking to ex-post optimism. For generic, monotonic games, the analyses of the two notions are identical. Of course, ex-post optimism also becomes an extreme notion in the standard epistemic model, once one considers a case in which players' optimism is known. Therefore, it is highly desirable to analyze other natural notions of optimism—possibly in other subjective models of knowledge in which one can assume common knowledge of optimism without assuming that a player knows that she is optimistic. This rather challenging task is left to future research. My analysis in this paper is elementary, raising the hope that one may be able to develop a tractable framework to incorporate self-serving (or other) biases into game theory.

Appendix A. Technical Appendix—Proofs

A.1. **Strategic Implications of Wishful Thinking.** I will now prove Lemma 1 and Proposition 1.

Proof of Lemma 1. Let $\hat{s} = \sigma(\omega)$ for some $\omega \in K_i(F) \cap W_i$. Since $\sigma_i(I_i(\omega)) = \{\hat{s}_i\}$ and the expectation of a random variable cannot be strictly higher than the variable everywhere, there exists $(\hat{s}_i, s_{-i}) \in \sigma(I_i(\omega))$ such that

$$u_i(\hat{s}_i, s_{-i}) \geq E_{i,\omega}[u_i(\sigma)] = \max_{s \in S_i \times \sigma_{-i}(I_i(\omega))} u_i(s), \tag{A.1}$$

where the equality is by (2.2) and the fact that $\omega \in W_i$. Now, since $s_{-i} \in \sigma_{-i}(I_i(\omega))$, (A.1) implies that $u_i(\hat{s}_i, s_{-i}) \geq \max_{s_i} u_i(s_i, s_{-i})$, showing that $\hat{s}_i \in BR_i(s_{-i})$, and hence $(\hat{s}_i, s_{-i}) \in B_i$. Moreover, since $\omega \in K_i(F)$, $I_i(\omega) \subseteq F$, and hence $(\hat{s}_i, s_{-i}) \in \sigma(I_i(\omega)) \subseteq \sigma(F)$. Thus, $(\hat{s}_i, s_{-i}) \in B_i \cap \sigma(F)$. But, since $\hat{s}_{-i} = \sigma_{-i}(\omega) \in \sigma_{-i}(I_i(\omega))$, (A.1) implies that

$$u_i(\hat{s}_i, s_{-i}) \geq \max_{s_i} u_i(s_i, \hat{s}_{-i}).$$

□

Proof of Proposition 1. For $m = 0$, the statement, $\sigma(W) \subseteq X^0$, is immediately implied by Lemma 1. For any m, assume that $\sigma(K^{m-1}(W)) \subseteq X^{m-1}$. Take any $\hat{s} = \sigma(\omega)$ for some $\omega \in K^m(W)$. Firstly, since $K^m(W) \subseteq K^{m-1}(W)$, $\hat{s} \in X^{m-1}$. Moreover, for any i, $\omega \in K_i(K^{m-1}(W)) \cap W_i$. Hence, by Lemma 1, there exists $(\hat{s}_i, s_{-i}) \in B_i \cap \sigma(K^{m-1}(W)) \subseteq B_i \cap X^{m-1}$ such that

$$\max_{s_i} u_i(s_i, \hat{s}_{-i}) \leq u_i(\hat{s}_i, s_{-i}) \leq \max_{\left(\hat{s}_i, s'_{-i}\right) \in B_i \cap X^{m-1}} u_i\left(\hat{s}_i, s'_{-i}\right).$$

(The inclusion above is by the induction hypothesis.) Thus, $\hat{s} \in \phi(X^{m-1}) = X^m$. Therefore, $\sigma(K^m(W)) \subseteq X^m$.

For any given m, I will now construct a model (Ω, I, p) in which $\sigma(K^m(W)) = X^m$. Fix any $\hat{s} \in X^m$. I will construct a model $\left(\Omega^{\hat{s}}, I^{\hat{s}}, p\right)$ in which $\hat{s} \in \sigma(K^m(W))$. Then, the model with $\Omega = \cup_{\hat{s} \in X^m} \Omega^{\hat{s}}$ and $I = \cup_{\hat{s} \in X^m} I^{\hat{s}}$ satisfies the bill. Take some ω_0 as the first member of $\Omega^{\hat{s}}$ and set $\sigma(\omega_0) = \hat{s}$. Set $F^m = \{\omega_0\}$. (I will define a sequence $F^0, \ldots, F^m$ such that $F^k \subseteq K^k(W)$ for each k.) For any player i, if $\hat{s} \in B_i$, then set $I_i^{\hat{s}}(\omega_0) = \{\omega_0\}$. Clearly, $\omega_0 \in W_i$ for such a player i. If $\hat{s} \notin B_i$, since $\hat{s} \in X^m$, there exists $s^{i,m-1} \equiv \left(\hat{s}_i, s_{-i}^{i,m-1}\right) \in B_i \cap X^{m-1}$ such that $u_i\left(s^{i,m-1}\right) \geq \max_{s_i} u_i\left(s_i, \hat{s}_{-i}\right)$. For each such i, consider a new state $\omega_{i,m-1}$ and set $\sigma\left(\omega_{i,m-1}\right) = s^{i,m-1}$, $I_i^{\hat{s}}(\omega_0) = \{\omega_0, \omega_{i,m-1}\}$, and $p_{i,\omega_0}\left(\omega_{i,m-1}\right) = 1$. By construction, $\omega_0 \in W_i$ for each such i. Let F^{m-1} be the set of states that are defined so far. Recall that, for each $\omega_{i,m-1}$, $I_i^{\hat{s}}\left(\omega_{i,m-1}\right) = I_i^{\hat{s}}(\omega_0)$ and $\sigma_i\left(\omega_{i,m-1}\right) = \hat{s}_i$ have been defined already. Now for each $j \neq i$, conduct the last operation again: if $s^{i,m-1} \in B_j$, then set $I_j^{\hat{s}}\left(\omega_{i,m-1}\right) = \{\omega_{i,m-1}\}$, yielding $\omega_{i,m-1} \in W_j$. If $s^{i,m-1} \notin B_j$ and $m-1 \geq 0$, then there exists $s^{j,m-2} = \left(s_j^{i,m-1}, s_{-j}^{j,m-2}\right) \in B_j \cap X^{m-2}$ such that $u_j\left(s^{j,m-2}\right) \geq \max_{s_j} u_j\left(s_j, s_{-j}^{j,m-2}\right)$. For each such j, consider a new member $\omega_{j,m-2}$ and set $\sigma\left(\omega_{j,m-2}\right) = s^{j,m-2}$, $I_j^{\hat{s}}\left(\omega_{i,m-1}\right) = \{\omega_{i,m-1}, \omega_{j,m-2}\}$, and $p_{j,\omega_{i,m-1}}\left(\omega_{j,m-2}\right) = 1$. Once again $\omega_{i,m-1} \in W_j$. Conduct this for each $\omega_{i,m-1}$, and let F^{m-2} be the set of states that are defined so far. Clearly, one can define such a sequence of sets $F^m, F^{m-1}, \ldots, F^0, F^{-1}$ following the above procedure. Set $\Omega^{\hat{s}} = F^{-1}$. For the states $\omega_{i,-1} \in F^{-1} \backslash F^0$, for which $\sigma\left(\omega_{i,-1}\right) = s^{i,-1} \in X^{-1} = S$, $I_j^{\hat{s}}$ remains to be defined for $j \neq i$; set $I_j^{\hat{s}}\left(\omega_{i,-1}\right) = \{\omega_{i,-1}\}$. Clearly, such j need not be a wishful thinker or rational at $\omega_{i,-1}$. But by construction each player is a wishful thinker at each state in F^0. Therefore, $W \supseteq F^0$. Hence, for each i, $K_i(W) \supseteq K_i\left(F^0\right) \supseteq F^1$, so that $K^1(W) \supseteq F^1$. Similarly, $K^k(W) \supseteq F^k$ for each $k \leq m$. In particular, $\omega_0 \in F^m \subseteq K^m(W)$, showing that $\hat{s} = \sigma(\omega_0) \in \sigma\left(K^m(W)\right)$. Finally, for the case of X^∞, such a sequence of increasing sets could be defined indefinitely without ever going out of X^∞, and each player will be a wishful thinker at each state, so that wishful thinking is common knowledge, and at the initial state the fixed profile $\hat{s} \in X^\infty$ is played. □

A.2. **Wishful Thinking and Nash Equilibrium.** Here, I will explore the relationship between Nash equilibrium and common knowledge of wishful thinking and prove Propositions 2 and 3. The next lemma states some straightforward but very useful facts about the elimination process and its relationship to equilibrium.

Lemma 4. *The following are true.*

(1) *ϕ is monotonic (i.e., $X \subseteq Y \Rightarrow \phi(X) \subseteq \phi(Y)$);*
(2) *NE is a fixed point of ϕ (i.e., $\phi(NE) = NE$);*
(3) *X^∞ is a fixed point of ϕ (i.e., $\phi(X^\infty) = X^\infty$).*

Lemma 4 immediately implies Proposition 2.

Proof of Proposition 2. By Lemmas 4.2, 4.1, and the definition of X^∞,

$$NE = \phi^\infty(NE) \subseteq \phi^\infty(S) = X^\infty.$$

□

Recall that

$$X_i^\infty = \{s_i | \exists s_{-i} : (s_i, s_{-i}) \in X^\infty\}$$

is the set of strategies for player i that are consistent with common knowledge of wishful thinking. Since $\phi_a(X^\infty) = X^\infty$, each such strategy must be a best reply to a surviving strategy:

Lemma 5. *For each i and $s_i \in X_i^\infty$, there exists s_{-i} such that $(s_i, s_{-i}) \in B_i \cap X^\infty$.*

The next lemma establishes some useful facts for two-player games with unique best replies. It states that X^∞ is closed under best reply and that the restriction of the best-response function to X_i^∞ is a bijection. Most notably, part 3 states that, when applied to $NE_1 \times NE_2$, the elimination process stops at the first iteration.

Lemma 6. *For any two-player game assume that, for each i, BR_i is singleton-valued. Then, the following are true.*

(1) $|X_1^\infty| = |X_2^\infty|$;
(2) *For each i, there exists a one-to-one and onto mapping $\rho_i : X_j^\infty \to X_i^\infty$ such that $BR_i(s_j) = \{\rho_i(s_j)\}$ for each s_j;*
(3) $\phi^\infty(NE_1 \times NE_2) = \phi(NE_1 \times NE_2)$.

Proof. By Lemma 5, for each $s_i \in X_i^\infty$, there exists a $\rho_i^{-1}(s_i) \in X_j^\infty$ such that $s_i \in BR_i\left(\rho_i^{-1}(s_i)\right)$. Since BR_i is singleton-valued, ρ_i^{-1} is one-to-one. Hence, $|X_i^\infty| \leq \left|X_j^\infty\right|$. Since i is arbitrary, this yields (1). But (1) in turn implies that the one-to-one function ρ_i^{-1} is also onto. Thus, $\rho_i : X_j^\infty \to X_i$ is well-defined. Since ρ_i^{-1} is a bijection, so is ρ_i, yielding (2). To show (3), check that, when BR_i is singleton-valued, for each $\hat{s}_i \in NE_i$, there exists a unique $(\hat{s}_i, s_{-i}) \in B_i \cap (NE_1 \times NE_2)$. In that case, $\phi(\phi(NE_1 \times NE_2)) = \phi(NE_1 \times NE_2)$, and hence $\phi^k(NE_1 \times NE_2) = \phi(NE_1 \times NE_2)$ for each k. □

Proof of Proposition 3. To prove (3.3), first note that if $X^\infty = \varnothing$, then by Proposition 2, $NE = \varnothing$, and thus $X^\infty = NE$. Assume that $X^\infty \neq \varnothing$. Using Lemma 6 and assumption (ii), one can rename the strategies as $X_1^\infty = \left\{s_1^1, \ldots, s_1^k\right\}$ and $X_2^\infty = \left\{s_2^1, \ldots, s_2^k\right\}$ for some $k \geq 1$ so that $B_1 \cap X^\infty$ is the diagonal of $X_1^\infty \times X_2^\infty$, i.e.,

$$B_1 \cap X^\infty = \left\{\left(s_1^l, s_2^l\right) | 1 \leq l \leq k\right\} \subseteq X_1^\infty \times X_2^\infty, \tag{A.2}$$

and

$$u_1\left(s_1^l, s_2^l\right) \text{ is strictly decreasing in } l. \tag{A.3}$$

Now, for any $l > m$, (A.2) and (A.3) imply that

$$\max_{s_1} u_1(s_1, s_2^m) = u_1(s_1^m, s_2^m) > u_1\left(s_1^l, s_2^l\right) = \max_{\left(s_1^l, s_2\right) \in X^\infty \cap B_1} u_1\left(s_1^l, s_2\right),$$

which in turn implies that $\left(s_1^l, s_2^m\right) \notin \phi(X^\infty) = X^\infty$. Therefore,

$$X^\infty \subseteq \left\{\left(s_1^l, s_2^m\right) | 1 \leq l \leq m \leq k\right\}. \tag{A.4}$$

Now, I will use mathematical induction (on l) to show that $NE = B_1 \cap X^\infty$. Together with (A.2), this implies (3.3). For $l = 1$, by Lemma 5, there exists s_1 such that $(s_1, s_2^1) \in B_2 \cap X^\infty$. But (A.4) states that s_1^1 is the only strategy that can satisfy this. Hence, $(s_1^1, s_2^1) \in B_2$. Together with (A.2), this shows that $(s_1^1, s_2^1) \in NE$. Assume that $(s_1^1, s_2^1), \ldots, \left(s_1^{l-1}, s_2^{l-1}\right) \in NE$ for some $l > 1$. Since BR_2 is singleton-valued (by assumption (i)), $s_2^l \notin BR_2(s_1)$ for any $s_1 \in \left\{s_1^1, \ldots, s_1^{l-1}\right\}$. Hence, $s_2^l \in BR_2(s_1)$ for some $s_1 \in \left\{s_1^l, \ldots, s_1^k\right\}$; recall from Lemma 6.2 that ρ_2 is onto. But since $\left\{\left(s_1^{l+1}, s_2^l\right), \ldots, \left(s_1^k, s_2^l\right)\right\} \cap X^\infty = \varnothing$

(by (A.4)), it must be that $s_2^l \in BR_2\left(s_1^l\right)$, showing that $\left(s_1^l, s_2^l\right) \in B_1 \cap B_2 = NE$. Therefore, $B_1 \cap X^\infty \subseteq NE$. ($NE \subseteq B_1 \cap X^\infty$ by the first part of this proposition.)

To prove (3.4), write

$$X^\infty = \phi\left(X^\infty\right) \subseteq \phi\left(NE_1 \times NE_2\right) = \phi^\infty\left(NE_1 \times NE_2\right) \subseteq \phi^\infty\left(S\right) = X^\infty,$$

where the first equality is by Lemma 4.3, the next inclusion is by Lemma 4.1 and the first part of the proposition (i.e., $X^\infty \subseteq NE_1 \times NE_2$), the next equality is by Lemma 6.3, the next inclusion is again by Lemma 4.1, and the last equality is by definition. □

Proof of Corollary 2. Let $\Omega = X^\infty$, and let σ be the identity mapping. Set $I_i\left(\omega\right) = \left\{s \in X^\infty | s_i = \sigma_i\left(\omega\right)\right\}$ for each i and ω. By (3.3), for each i and ω, there exists a (unique) Nash equilibrium $\hat{s}$ with $\hat{s} \in I_i\left(\omega\right)$. Set $p_{i,\omega} = \delta_{\hat{s}}$. Now, by (3.4), for any $s \in I_i\left(\omega\right)$, there exists $\left(\tilde{s}_i, s_{-i}\right) \in NE$ such that $u_i\left(\hat{s}\right) \geq u_i\left(\tilde{s}_i, s_{-i}\right)$. Since $u_i\left(\tilde{s}_i, s_{-i}\right) = \max_{s_i'} u_i\left(s_i', s_{-i}\right)$ (by definition of NE) for each s_{-i}, this shows that $u_i\left(\hat{s}\right) = \max_{s_i, s_{-i} \in \sigma_{-i}\left(I_i\left(\omega\right)\right)} u_i\left(s\right)$, showing that $\omega \in W_i$. □

A.3. **Strategic Implications of Ex-post Optimism.**

Proof of Lemma 2. Take any $\hat{s} = \sigma\left(\hat{\omega}\right)$ for some $\hat{\omega} \in K_i\left(F\right) \cap \tilde{W}_i \cap R_i$. Let $\tilde{\omega} \in \arg\max_{\omega \in I_i\left(\hat{\omega}\right)} u_i\left(\hat{s}_i, \sigma_{-i}\left(\omega\right)\right) \subseteq K_i\left(F\right) \cap \tilde{W}_i \cap R_i$. Since $\tilde{\omega} \in R_i$, $\hat{s}_i = \sigma_i\left(\tilde{\omega}\right) \in BR_i\left(\mu_{i,\tilde{\omega}}\right)$. Since $\tilde{\omega} \in K_i\left(F\right)$, $\mathrm{supp}\left(\mu_{i,\tilde{\omega}}\right) \subseteq \sigma\left(F\right)$, and since $\tilde{\omega} \in \tilde{W}_i$, $u_i\left(\hat{s}_i, s_{-i}\right) = u_i\left(\hat{s}_i, \sigma_{-i}\left(\tilde{\omega}\right)\right)$ for each $s_{-i} \in \mathrm{supp}\left(\mu_{i,\tilde{\omega}}\right)$. Hence, $\left(\hat{s}_i, \sigma_{-i}\left(\tilde{\omega}\right)\right) \in B_i^{\sigma(F)}$. Moreover, since $\hat{\omega} \in I_i\left(\hat{\omega}\right)$, by definition of $\tilde{\omega}$, $u_i\left(\hat{s}_i, \sigma_{-i}\left(\tilde{\omega}\right)\right) \geq u_i\left(\sigma\left(\hat{\omega}\right)\right) = u_i\left(\hat{s}\right)$. □

Proof of Proposition 4. The first inclusion is given by Lemma 2 as in the proof of Proposition 1. To see the second inclusion, write $CK\left(R \cap O\right) = CK\left(R\right) \cap CK\left(O\right) = CK\left(R\right) \cap CK\left(\cap_{i \in N} K_i\left(O_i\right)\right) = CK\left(\cap_{i \in N} K_i\left(O_i\right) \cap R\right)$. Then, the inclusion is obtained by taking intersections on both sides of the first inclusion. One can construct a model with equality as in the proof of Proposition 1. I will illustrate the main step. Fix any $\hat{s} \in Y^m$. Towards a model $\left(\Omega^{\hat{s}}, I^{\hat{s}}, p\right)$ in which $\hat{s} \in \sigma\left(K^m\left(\cap_{i \in N} K_i\left(O_i\right) \cap R\right)\right)$, take some ω_0 as the first member of $\Omega^{\hat{s}}$ and set $\sigma(\omega_0) = \hat{s}$. Towards defining a $F^0, \ldots, F^m$ with $F^k \subseteq K^k(\cap_{i \in N} K_i(O_i) \cap R)$ for each k, set $F^m = \{\omega_0\}$. For any player i, if $\hat{s} \in B_i^{Y^{m-1}}$, then $\hat{s}_i \in BR_i(\mu_i)$ for some μ_i with support $\{\hat{s}_{-i}, s_{-i,1}, \ldots, s_{-i,l}\}$ as in (**??**). Consider new members $\omega_{i,m-1,1}, \omega_{i,m-1,2}, \ldots, \omega_{i,m-1,l}$, set $\sigma(\omega_{i,m-1,1}) = (\hat{s}_i, s_{-i,1}), \ldots, \sigma(\omega_{i,m-1,l}) = (\hat{s}_i, s_{-i,l})$. Set $I_i^{\hat{s}}(\omega_0) = \{\omega_0, \omega_{i,m-1,1}, \ldots, \omega_{i,m-1,l}\}$. Set $p_{i,\omega_0}(\omega) = \mu_i(\sigma(\omega))$ for each $\omega \in I_i^{\hat{s}}(\omega_0)$. Clearly, $\omega_0 \in K_i(O_i) \cap R_i$ for such a player i. If $\hat{s} \notin B_i^{Y^{m-1}}$, since $\hat{s} \in Y^m$, there exists $s^{i,m-1} \equiv \left(\hat{s}_i, s_{-i}^{i,m-1}\right) \in B_i^{Y^{m-1}}$ such that $u_i(s^{i,m-1}) \geq u_i(s_i, \hat{s}_{-i})$. Now, $\hat{s}_i \in BR_i(\mu_i)$ for some μ_i with support $\left\{s_{-i}^{i,m-1}, s_{-i,2}, \ldots, s_{-i,l}\right\}$. Consider, again, new members $\omega_{i,m-1,1}, \omega_{i,m-1,2}, \ldots, \omega_{i,m-1,l}$, but set $\sigma(\omega_{i,m-1,1}) = \left(\hat{s}_i, s_{-i}^{i,m-1}\right), \sigma(\omega_{i,m-1,2}) = (\hat{s}_i, s_{-i,2}), \ldots, \sigma(\omega_{i,m-1,l}) = (\hat{s}_i, s_{-i,l})$. $I_i^{\hat{s}}(\omega_0)$ and p_{i,ω_0} are defined as above. □

Proof of Proposition 3. Take any $X \subseteq S$. Suppose $\hat{s} \in X \backslash \phi(X)$. Then, there exists i such that

$$u_i(\tilde{s}_i, \hat{s}_{-i}) \equiv \max_{s_i} u_i(s_i, \hat{s}_{-i}) > \max_{(\hat{s}_i, s_{-i}) \in B_i \cap X} u_i(\hat{s}_i, s_{-i}) \equiv u_i(\hat{s}_i, \tilde{s}_{-i}) \geq u_i(\tilde{s}_i, \tilde{s}_{-i}),$$

where the last inequality is due to the fact that $\hat{s}_i$ is a best reply to $\tilde{s}_{-i}$ by definition of B_i. Since the game is monotonic, this implies that

$$u_i(\hat{s}) > u_i(\hat{s}_i, \tilde{s}_{-i}).$$

But by (4.6), $B_i^X = B_i \cap X$, and hence $u_i(\hat{s}_i, \tilde{s}_{-i}) = \max_{(\hat{s}_i, s_{-i}) \in B_i^X} u_i(\hat{s}_i, s_{-i})$. Thus,

$$u_i(\hat{s}) > \max_{(\hat{s}_i, s_{-i}) \in B_i^X} u_i(\hat{s}_i, s_{-i}),$$

showing that $\hat{s} \notin \psi(X)$. Therefore, $\psi(X) \subseteq \phi(X)$. Since the reverse inequality is always true, this proves that $\psi(X) = \phi(X)$. □

References

[1] Ali, S. Nageeb M. (2003): "Multilateral Bargaining with Subjective Biases: Waiting To Settle," Stanford University, mimeo, 2003.

[2] Aumann, Robert (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics* 1, 67-96.

[3] Aumann, Robert (1976): "Agreeing to disagree," *Annals of Statistics* 4, 1236-1239.

[4] Aumann, Robert (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1-18.

[5] Aumann, R. and A. Brandenburger (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161-1180.

[6] Babcock, Linda and George Loewenstein (1997): "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *Journal of Economic Perspectives* 11, 109-126.

[7] Bar-Hillel, Maya and David Budescu (1995): "The Elusive Wishful Thinking Effect," *Thinking and Reasoning*, 1, 71-103.

[8] Banerjee, Abhijit and Rohini Somanathan (2001): "A Simple Model of Voice," *Quarterly Journal of Economics*, 116, 189-227.

[9] Bernheim Douglas (1984): "Rationalizable Strategic Behavior," *Econometrica* 52, 1007-1028.

[10] Brandenburger, Adam and Eddie Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1391-1402.

[11] Dekel, Eddie, Drew Fudenberg, and David Levine (2004) "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282-303.

[12] Eyster, Eric and Mathew Rabin (2000): "Cursed Equilibrium," mimeo.

[13] Farber, Henry S. and Max H. Bazerman (1989): "Divergent expectations as a cause of disagreement in bargaining: Evidence from a comparison of arbitration schemes," *Quarterly Journal of Economics* 104, 99-120.

[14] Feinberg, Yossi (2000): "Characterizing Common Priors in the Form of Posteriors," *Journal of Economic Theory* 91, 127-179.

[15] Geanakoplos, John, David Pearce, and Ennio Stacchetti (1989): "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1, 60-79.

[16] Harrison, Michael and David Kreps (1978): "Speculative Investor Behavior in a Stock Market with Heterogenous Expectations" *Quarterly Journal of Economics* 92, 323-36.

[17] Kaplan, Todd and Bradley Ruffle (2004): "The Self-serving Bias and Beliefs about Rationality," *Economic Inquiry* 42, 237-246.

[18] Landes, William M. (1971): "An economic analysis of the courts," *Journal of Law and Economics* 14, 61-107.

[19] Larwood, Laurie and William Whittaker (1977): "Managerial Myopia: Self-serving Biases in Organizational Planning," *Journal of Applied Psychology*, 194-198.

[20] Manove, Michael and Jorge Padilla (1999): "Banking (Conservatively) with Optimists," *Rand Journal of Economics*, 30, 324-350.

[21] Morris, Stephen (1995): "The Common Prior Assumption in Economic Theory," *Economics and Philosophy* 11, 227-53.

[22] Morris, Stephen (1996): "Speculative Investor Behavior and Learning" *Quarterly Journal of Economics* 111, 1111-33.

[23] Pearce, David (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica* 52, 1029-1050.

[24] Posner, Richard (1972): "The Behavior of Administrative Agencies," *Journal of Legal Studies* 1, 305.

[25] Rabin, Matthew (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 5, 1281-1302.

[26] Squintani, Francesco (2001): "Mistaken self-perception and Equilibrium," mimeo.

[27] Rostek, Marzena (2004): "p-dominant sets", mimeo.

[28] Van den Steen, Eric (2001): "Organizational Beliefs and Managerial Vision," forthcoming in *Journal of Law, Economics, and Organization.*

[29] Weinstein, Neil (1980): "Unrealistic Optimism about Future Life Events," *Journal of Personality and Social Psychology*, 860-20.

[30] Wilson, Robert (1968): "The Theory of Syndicates," *Econometrica*, 36, 119-132.

[31] Yildiz, Muhamet (2003): "Bargaining without a Common Prior — An Agreement Theorem," *Econometrica* 71, 793-811.

[32] Yildiz, Muhamet (2004): "Waiting to Persuade," *Quarterly Journal of Economics* 119, 223-248.

MIT

URL: `http://econ-www.mit.edu/faculty/myildiz/index.htm`