View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Distribution-Valued Solution Concepts

David H. Wolpert
NASA Ames Research Center
MailStop 269-1
Moffett Field, CA 94035-1000
david.h.wolpert@nasa.gov

James W. Bono
Department of Economics
American University
Washington, D.C. 20016
bono@american.edu

June 4, 2010

Abstract

Under its conventional positive interpretation, game theory predicts that the mixed strategy profile of players in a noncooperative game will satisfy some set-valued solution concept. Relative probabilities of profiles in that set are unspecified, and all profiles not satisfying it are implicitly assigned probability zero. However the axioms underlying Bayesian rationality say that we should reason about player behavior using a probability density over all mixed strategy profiles, not using a subset of all such profiles. Such a density over profiles can be viewed as a solution concept that is distribution-valued rather than set-valued. A distribution-valued concept provides a best single prediction for any noncooperative game, i.e., a universal refinement. In addition, regulators can use a distribution-valued solution concept to make Bayes optimal choices of a mechanism, as required by Savage's axioms. In particular, they can do this in strategic situations where conventional mechanism design cannot provide advice. We illustrate all of this on a Cournot duopoly game.

Keywords: Quantal Response Equilibrium, Bayesian Statistics, Entropic prior, Maximum entropy
JEL Classification codes: C02, C11, C70, C72

We would like to thank Professors George Judge, Julian Jamison, Alan Isaac and audience members at the American University Economics Seminar and the *2010 International Conference on Social Computing, Behavioral Modeling and Prediction*, as well as attendees of the *From Game Theory to Game Engineering* workshop. We would like to extend a special thank you to the NASA-American-Stanford working group.

1 Introduction

It is common practice in game theory to define "rational" decision-makers as Bayesian decision makers. This means that they make their decision(s) so as to maximize their posterior expected utility, conditioned on their information about the world external to them, and on their prior beliefs.

Game theoreticians themselves, in conducting their daily lives, are decision makers. In particular, one decision that game-theoreticians often have to make is what strategy profile q to predict as the outcome of a given N-player non-cooperative game Γ. To make this decision they have information specifying Γ, and they also have prior beliefs concerning the players of Γ and the strategy profile of those players, q.

In this paper we use the Bayesian definition of rational decision-making to analyze this prediction decision that game theoreticians often have to make. More precisely, by identifying rational decision making with Bayesian decision making, we analyze how a "rational" Predictor agent, external to the game Γ, would combine her information concerning that game with her prior beliefs concerning the players of the game and their profiles q, to predict what profile q is jointly adopted by the players of Γ.

Being Bayesian, such a rational Predictor would use a posterior distribution over strategy profiles q, conditioned on her information about the the game Γ and its players. This posterior distribution is essentially a solution concept. However in contrast to set-valued solution concepts over q's, like the Nash Equilibrium (NE), Quantal Response Equilibrium (QRE), etc., this solution concept is not a set of q's. Rather it is a distribution over all q's, providing their relative probabilities.[1] So in particular, this posterior provides the relative probabilities of all the NE profiles, rather than just providing the set of such profiles.

The usual axiomatic justifications for Bayesian reasoning mandate that Predictor use such a distribution-valued solution concept rather than a set-valued one. However there are also numerous practical advantages to using such a distribution-valued solution concept. One advantage is that with distribution-valued solution concepts, there are no difficulties in choosing a single "best prediction" for q, as there often are with set-valued solution concepts. If Predictor must decide on a single "best" q as her prediction, Bayesian decision theory tells Predictor exactly how she must do this: she must combine her posterior distribution over q's, together with her loss function for predicting q when the actual profile is q', to give the (unique) single best prediction for q. In this, Bayesian decision theory, combined with a distribution-valued solution concept, provides a universally applicable refinement.

Another advantage of a distribution-valued solution concept is that it is needed for Predictor to implement a Bayesian alternative to mechanism design. More precisely, say that Predictor is choosing the value of a parameter in the game to be played. Each game parameter Predictor might choose will induce a different associated posterior over q. Say that Predictor is concerned with the value of a function $f : q \rightarrow \mathbb{R}$, e.g., the social welfare

[1] In this paper, to simplify the exposition, whenever we don't need to be formal we loosely use the term "distribution", even if we mean a probability density function, properly speaking.

of q. Since different values of the game parameter induce different distributions over q, they induce different expected values of f. Accordingly, to strictly adhere with Savage's axioms, Predictor should choose the parameter with the maximal associated posterior expected value of f. This cannot be done without a distribution-valued solution concept.

In the rest of this introduction we elaborate these practical advantages of using distribution-valued solution concepts rather than set-valued solution concepts. More general discussion of the relationship between distribution-valued solution concepts and set-valued solution concepts can be found in the appendix. We end the introduction with a brief discussion of the relation of PGT to other work in the literature.

In Sec. 2 we introduce a prior over profiles q, and then a likelihood over q's. Taken together, these give us a distribution-valued solution concept in the form of a posterior distribution over q's. In its homogeneous formulation (all players have the same parameter values), this model has two parameters.

Illustrations of our solution concept for simple games with small pure strategy spaces can be found in Wolpert and Bono (2008). In Sec. 3 we illustrate the solution concept on a more computationally challenging domain, the Cournot game.

We emphasize that we do *not* claim that the prior and likelihood considered in this paper are "the correct" prior and likelihood. Rather we use them to illustrate distribution-valued concepts in general. Just like one can have different set-valued solution concepts, one can have different distribution-valued concepts. Ultimately field and laboratory experiments should be used to determine what such concept to use.[2]

We refer to distribution-valued solution concepts in general as Predictive Game Theory (PGT). PGT is the application of statistical inference *to* games, in contrast to the use of statistical inference by some players *within* a game. As such, it can be used to analyze any type of game.[3]

1.1 Posterior distributions in PGT

Often the information available to Predictor, $\mathscr{I}$, includes the precise game structure, set of strategy profiles and player utility functions. This is not always the case though. Sometimes Predictor will have some uncertainty about such quantities. Or it may be that $\mathscr{I}$ contains non-conventional information that is relevant for predicting behavior, e..g., information about player rationality, focal points and demographic data.

Any Bayesian quantifies whatever information she has about the state of the world in terms of distributions over the set of all possible states of the world. So to be Bayesian, Predictor must quantify $\mathscr{I}$ in terms of such distributions. In noncooperative game theory, the set of "all possible states of the world" is usually the set of mixed strategy profiles available to the players of the game. So to be Bayesian, Predictor must quantify $\mathscr{I}$ in terms of distributions over the set of strategy profiles.

[2] Indeed, to best match experimental data, it may end up being easiest to construct the posterior directly, rather than construct it by first constructing a prior and likelihood that then get combined.

[3] An earlier version of this paper can be found at Wolpert and Bono (2008), and a brief high-level summary of PGT was published in Wolpert and Bono (2010).

To make this more precise, let X_i be the set of player i's pure strategies, $X = \times_{i=1}^N X_i$ be the set of pure strategy profiles, $\Delta(X_i)$ be the set of player i's mixed strategies and $\Delta_{\mathcal{X}} = \times_{i=1}^N$ be the set of mixed strategy profiles. We refer to a generic element of X as $x = (x_1, ..., x_n)$, and we refer to a generic element of $\Delta_{\mathcal{X}}$ as $q = (q_1, ..., q_N)$. For simplicity, we restrict attention to finite pure strategy sets. So a "probability distribution" over mixed strategy profiles is a density function mapping from $\Delta_{\mathcal{X}} \rightarrow \mathbb{R}$.

As a point of notation, we will often use integrals with the measure implicit. So in particular, if the integration variable is finite, we implicitly are using a point-mass measure, and the integral is a sum. Similarly we generically use "P" to indicate either a probability density function or a probability distribution, as appropriate.

The states of the world are all of $\Delta_{\mathcal{X}}$, and the Bayesian posterior over states of the world conditioned on $\mathscr{I}$ is

$$P(q \mid \mathscr{I}) \propto P(q)\mathscr{L}(\mathscr{I} \mid q) \tag{1}$$

where $P(q)$ represents Predictor's prior beliefs about the probability of q being chosen and $\mathscr{L}(\mathscr{I} \mid q)$ is the likelihood of the information in $\mathscr{I}$ given that q is chosen by the players.

As mentioned above, in general $\mathscr{I}$ may not fully specify the game the players are engaged in, i.e., Predictor may have uncertainty about the game. This uncertainty is reflected in $P(q \mid \mathscr{I})$. As an example, say Predictor is uncertain about the utility functions, so that $\mathscr{I}$ is a distribution over possible utility functions. (Recall that allowing Predictor to be uncertain about the utilities does not mean the players themselves are.) Then $P(q \mid \mathscr{I})$ is given by averaging over that distribution over utility functions. To illustrate this, suppose that $\mathscr{I}$ specifies that that the players' utility functions are $\mathscr{U}'$ with probability m, and that they are instead $\mathscr{U}''$ with probability $1 - m$. Let $\mathscr{I}'$ be all information in $\mathscr{I}$ *other* than this information about the players utility functions. Then

$$P(q \mid \mathscr{I}) = mP(q \mid \mathscr{U}', \mathscr{I}') + (1 - m)P(q \mid \mathscr{U}'', \mathscr{I}'). \tag{2}$$

However arrived at, the posterior over mixed strategy profiles induces posterior expected values of any function $f(q)$ of the mixed strategy profiles,

$$\begin{aligned} \mathbb{E}(f \mid \mathscr{I}) &= \int_{\Delta_{\mathcal{X}}} f(q)P(q \mid \mathscr{I})dq \\ &= \frac{\int_{\Delta_{\mathcal{X}}} f(q)P(\mathscr{I} \mid q)P(q)dq}{\int_{\Delta_{\mathcal{X}}} P(\mathscr{I} \mid q)P(q)dq} \end{aligned} \tag{3}$$

Often Predictor is most interested in such posterior expected values. For example, for appropriate choices of f, the associated posterior expectation gives posterior expected posterior expected profits of the two firms, posterior expected social welfare, posterior expected covariance between player pure strategies, etc.

As an example, by choosing $f(q) = q(x)$ for any given x, we get the posterior over pure strategy profiles:

$$\begin{aligned} P(x \mid \mathscr{I}) &= \int P(x \mid q, \mathscr{I}) P(q \mid \mathscr{I}) dq \\ &= \int q(x) P(q \mid \mathscr{I}) dq \\ &= \mathbb{E}[q]. \end{aligned} \tag{4}$$

Typically under $P(x \mid \mathscr{I})$ the pure strategies of the players are not statistically independent. This is true even though the support of the posterior density function $P(q \mid \mathscr{I})$ is restricted to mixed strategy profiles q under which the players *are* statistically independent. Formally, the distribution over pure strategy profiles is given by

$$\begin{aligned} P(x_i, x_{-i}) &= \int dq \; P(x_i, x_{-i} \mid q, \mathscr{I}) P(q \mid \mathscr{I}) \\ &= \int dq \; q(x_i, x_{-i}) P(q \mid \mathscr{I}) \\ &= \int dq_i dq_{-i} \; q_i(x_i) q_{-i}(x_{-i}) P(q_i, q_{-i} \mid \mathscr{I}) \end{aligned} \tag{5}$$

and in general this differs from the product of the distributions over pure strategies,

$$P(x_i)P(x_{-i}) = \left[\int dq_i \; q_i(x_i) P(q_i \mid \mathscr{I})\right]\left[\int dq_{-i} \; q_{-i}(x_{-i}) P(q_{-i} \mid \mathscr{I})\right]. \tag{6}$$

Note that this coupling of pure strategies is different from the correlation that arises in a correlated equilibrium [see Aumann (1974)]. The coupling between x_A and x_B that arises in $P(x \mid \mathscr{I})$ is *from Predictor's perspective*, arising from her averaging over all q's. However, there is no coupling between the player's pure strategies *from the players' perspective* because they choose their strategies independently.

To illustrate this, say we had a game with multiple NE, and we knew that the players of the game were fully rational. Then by observing the pure strategy of one of the players, we would gain statistical information about what equilibrium is most likely being played. That information would in turn tell us something about the move of the other players.

Note there are two different kinds of probability arising in PGT. $q(x)$ refers to the probability that the players choose pure strategy profile x. In contrast, $P(q \mid \mathscr{I})$ refers to the probability that the external modeler of the system, Predictor, assigns to the event that the players jointly decide to choose x with probability q.

As an illustration, even if the players are involved in a complete information game, and so know one another's utilities exactly, this does not mean that Predictor knows those utilities exactly. Indeed, if the game being played were an incomplete information game, there would be yet a third kind of probability, giving the beliefs the players have concerning the types of one another. (Similarly, there are other kinds of probability for games involving imperfect information, states of Nature, etc.)

1.2 Point Prediction in PGT

A Bayesian Predictor's posterior distribution over strategy profiles contains far more information than a single "best" prediction of a strategy profile. However, if Predictor needs to select a single strategy profile as a best prediction, then decision theory tells her how to. First, she must specify her loss function $L(q, q')$, which quantifies the loss experienced by Predictor when she makes prediction q' while the actual profile turns out to be q.

Example 1: A game theoretician working for a company may be required to make a single prediction for how the company's employees will behave under a new organization structure. That game theoretician will have her salary tied, implicitly or explicitly, to how closely her prediction q' for the behavior of the employees matches their actual behavior q. This connection between the accuracy of her prediction and her salary provides the game theoretician's loss function for making prediction q' when the actual profile of the employees turns out to be q.

Example 2: An academic game theoretician publishes a prediction for what profile will be adopted by the subjects of an upcoming behavioral game theory experiment. Her professional prestige is tied to how closely her prediction q' for the outcome of the experiment matches the actual profile adopted by the experimental subjects, q. This connection between the accuracy of her prediction and her prestige provides the game theoretician's loss function for making prediction q' when the actual profile of the subjects turns out to be q.

Given her loss function, the Bayesian Predictor combines it with her posterior distribution over profiles, by choosing as her "best prediction" the strategy profile that minimizes the associated expected loss. If Predictor's loss function is $L(q, q')$, where q' is the predicted profile and q is the realized profile., then that Bayes-optimal prediction is

$$q^* = \operatorname*{argmin}_{q'} \int L(q, q') P(q \mid \mathscr{I}) dq.$$

Since this optimal predicted profile is almost always unique, it can be viewed as a unique "refinement".[4]

Often under q^* no player's strategy is a best response to the strategies of the other players. Assuming there is more than one NE of the game, this is true even if the players are all fully rational, i.e., if the support of the density over strategy profiles is restricted to the NE. In this sense, "predictive" bounded rationality is automatic under PGT.

It is possible to use the posterior $P(q \mid \mathscr{I})$ to specify a unique mixed strategy profile without using a loss function. In particular, the expected posterior mixed strategy profile,

[4]Note that this refinement depends on the loss function of Predictor, *which is not part of the specification of the game.* As such, it varies from one Predictor to another. An alternative that does not have such variability is to predict the posterior expected profile; see below.

$P(x \mid \mathscr{I})$, provides such a unique profile.[5] In yet other circumstances, a researcher may be interested in the mode of the posterior, $\text{argmax}_q \, P(q \mid \mathscr{I})$. Both quantities always exist and are typically unique.

Alternatively, if Predictor wishes to choose a single best *pure strategy* profile, the procedure is again straightforward. She first specifies her loss function, $L(x, x')$, where x' is the predicted pure action profile and x is the realized pure action profile. Predictor then chooses

$$x^* = \underset{x'}{\text{argmin}} \int L(x, x') q(x) P(q \mid \mathscr{I}) dx dq. \tag{7}$$

As an example, if all x_i are real-valued, and Predictor's loss function is given by $L(x, x') = ||x - x'||^2$, then the optimal choice is the expected value of x computed under $P(x \mid \mathscr{I})$, given in Eq. (4). Note that in general, this optimal choice may not be an element of X. If it is required that the prediction lie in X, then Predictor must instead do a search over the (convex) function mapping every x' to $\int L(x, x') q(x) P(q \mid \mathscr{I}) dx dq$.

As another example, say Predictor's loss function is a zero-one loss function, where $L(x, x') = 1$ for $x \neq x'$ and $L(x, x') = 0$ for $x = x'$. Then the optimal choice is the mode of $P(x \mid \mathscr{I})$, $\text{argmax}_x \, P(x \mid \mathscr{I})$.

1.3 PGT Alternative to Mechanism Design

Say Predictor (the regulator) can set a parameter λ specifying some aspect of a game to be played by a set of players who will choose a mixed strategy profile q. As an example, λ might specify the form of an auction, or any similar choice of a mechanism in a mechanism design problem. More generally, λ can be any choice that someone external to the N-player game can make that will modify that game before it is played. Although it is not required mathematically, to ground intuition we can assume that λ is perfectly observed by all N players.

Let $W(q, \lambda)$ be Predictor's utility function, and indicate the game specified by λ as Γ_λ. Let $\mathscr{I}$ be some other information that Predictor has concerning the game and/or player behavior, in addition to the value λ that she will choose. Then the standard approach of optimal control (i.e., Bayesian decision theory) says that Predictor should set λ to

$$\text{argmax}_\lambda \Big[\mathbb{E}(W \mid \mathscr{I}, \lambda) \Big] \;=\; \text{argmax}_\lambda \Big[\int W(q, \lambda) P(q \mid \mathscr{I}, \lambda) dq \Big] \tag{8}$$

So for example, if Predictor's utility function only depends on the pure strategy profile of the players, we can write $W(q, \lambda) = \int dx \; q(x) w(x)$ for some function w. In this case Predictor should set λ to

$$\text{argmax}_\lambda \Big[\int W(q, \lambda) P(q \mid \mathscr{I}, \lambda) dq \Big] \;=\; \text{argmax}_\lambda \Big[\int w(x) q(x) P(q \mid \mathscr{I}, \lambda) dq dx \Big] \tag{9}$$

[5]Note that $P(x \mid \mathscr{I})$ is also the Bayes optimal prediction, if the loss function is quadratic, $L(q, q') = \int dx [q(x) - q'(x)]^2$.

There are many ways to extend the foregoing. As an illustration, consider the case where Predictor's utility is not $W(q, \lambda)$, but rather $W(\theta, \lambda)$ where $\theta \in \Theta$ is set stochastically by $P(\theta \mid q, \lambda)$. In other words, Predictor does not directly care about the mixed strategy profile, but rather about the ramifications of that profile on the state of some other system with state space Θ. As an example, say the player pure strategy profile stochastically sets the state $\theta \in \Theta$ of a system, and Predictor only cares about that value θ. Then using the PGT alternative to mechanism design, Predictor should set λ to maximize

$$\int P(\theta \mid x) q(x) P(q \mid \mathscr{I}, \lambda) W(\theta) dq dx d\theta.$$

In this way Predictor makes her choice of λ in strict accordance with Savage's axioms. That is, Predictor chooses λ by answering the question, "what value of λ (the mechanism) maximizes Predictor's associated expected utility?".

In contrast, for mechanism design based on set-valued solution concepts, this question cannot be properly posed. Perhaps the closest analog of this question would be something like "what value of λ maximizes Predictor's expected utility under a set-valued solution concept?" However, consider the case where the N-player game has multiple equilibria for every value of λ. Let W^j_λ be Predictor's expected utility under the j'th equilibrium for value λ. In particular, there are pairs $\lambda, \lambda' \neq \lambda$ such that the intervals $[\min_j(W^j_\lambda)$, $\max_j(W^j_\lambda)]$ and $[\min_j(W^j_{\lambda'})$, $\max_j(W^j_{\lambda'})]$ overlap. Mechanism design with set-valued solution concepts can provide no advice on whether the controller should choose λ or λ' in this situation (in contrast to the PGT alternative to mechanism design).

By using a distribution-valued solution concept, Predictor can compare choices of λ (i.e., choices of "mechanism") based on other considerations beside the associated values of expected welfare. In particular, Predictor can use the posterior to answer many of the questions that real-world stakeholders often ask concerning the possible policy choices of a regulator, such as:

- "Which policy produces the lowest variance in welfare?"
- "What is the probability that policy λ_A produces greater welfare than policy λ_B?"
- "Which policy minimizes the probability that welfare is below some threshold value?"
- "Which of the policies maximizes welfare subject to the condition that the expected profits of firms are positive with probability greater than some threshold value?"

Answering any of these questions requires a distribution over mixed strategy profiles. This means that set-valued solution concepts cannot answer these questions, since they do not provide such a distribution. In section 4.4 below, we will elaborate in detail the general procedure for using PGT to answer all the questions listed above. We then show how to use that procedure to answer the questions for a Cournot duopoly with negative externalities.

In many situations, a regulator changes their policy choice very infrequently on the timescale of the decision-making of the players subject to that policy. (In fact firms often prefer that regulators change policy infrequently, to "allow them to plan for the long term".) In such situations, often the players effectively view the regulator's choice of policy as the value of a state of Nature. In other words, even if the players know the regulator's utility function, in such situations they don't account for that utility function when making their decisions. We deal with this setting in our Cournot duopoly example below.

1.4 Relation to other work

In this paper we are interested in *ex ante* predictions, not *ex post* fitting of a model to data. However *ex ante* predictions are closely related to Bayesian *ex post* fitting of a model to data. *Ex post* maximum likelihood fitting of the parameters of a model to data can be done with a set-valued solution concept, if one simply treats which of the equilibria under that concept actually arose as another parameter to be chosen by the maximum likelihood. However if one were to use Bayesian fitting, one would need to have a prior over *all* parameters. That includes a prior over the parameter of which equilibrium arises. This prior is nothing other than the *ex ante* relative probabilities of the profiles allowed by the set-valued solution concept. Similarly, since we are interested in making *ex ante* predictions, we need a way to provide *ex ante* relative probabilities of the profiles allowed by any set-valued solution concept.[6]

This need for a distribution over equilibria when fitting set-valued concept models to data is not integrated into most work in the experimental game theory literature. In contrast, it is integrated into some prominent work in the econometric literature on structural modeling (e.g., Bajari et al. (2010); Aguirregabiria and Mira (2009) and references therein). However most structural modeling requires player trembles, incomplete information, observational noise by the researcher or some such source of noise. In contrast to the field experiments studied in econometrics, many of the laboratory experiments studied in game theory are designed so that those types of noise are vanishingly small.[7] Accordingly, structural modeling is not well-suited to analysis of some game theory experiments.

Finally, it should be noted that subsequent to the appearance of early versions of this paper, the "Heterogeneous Logit Quantal Response Equilibrium" was introduced (see Rogers et al. (2009)).[8] For games in which the HQRE results in a unique choice of q_i

[6]Note that in general a prior over which of the profiles q allowed by the set-valued solution concept arises would vary with the game. In other words, it would take the form $P(q \mid \mathscr{I})$. So this "prior" is identical to what in our analysis here PGT serves as the "posterior", since in this paper we do not extend the PGT conditioning argument to include experimental data.

[7]Perhaps the most problematic issue in game theory experiments is ensuring the common knowledge assumption that players know one another's utility functions, not just one another's payoffs. This issue can be minimized by having payoffs be small (so non-concavity of the utility functions is irrelevant), games be anonymous (so there are no reputation effects), and players told they are playing computers (so there are no other-regarding preferences). See Starmer (2000) for more on the issue of designing game theory experiments to match the assumptions of game theory.

[8] That work posits a distribution over exponents λ_i in the logit response functions of the players. It then treats the λ_i's as types in a Bayesian game. (Note that in the logit QRE, it is only the product

by any player i having response function exponent λ_i, a population distribution $P(\lambda)$ over such exponents results in a distribution over q_i's. In this sense, in this special situation, the HQRE is a particular type of PGT.

Note though that even in this special situation, for any given $P(\lambda)$, the HQRE assigns zero probability density to almost all q_i's.[9] Moreover, of course, if some λ do not result in a unique q, we cannot interpret the HQRE as a distribution over q's of any sort, even one whose support has measure zero.

However even in the special situation where each λ results in a unique q, there remains a subtle but important distinction between PGT and the HQRE. In the HQRE, the map from the type of a player i to q_i is single-valued. So if only Predictor knew every player's type, Predictor would know q exactly, with no uncertainty. In contrast, the PGT paradigm is founded on the fact in the real world, even if Predictor knows the type of every player, she is still (very) uncertain about their mixed strategies.

2 Posteriors over Mixed Strategy Profiles

In this section we describe a Bayesian formulation of Predictor's distribution-valued solution concept, the posterior distribution. We write this as

$$P(q \mid \mathscr{I}) \propto P(q)\mathscr{L}(\mathscr{I} \mid q),$$

where $\mathscr{L}(\mathscr{I} \mid q)$ is the likelihood of $\mathscr{I}$ given q, $P(q)$ is the prior distribution over mixed strategy profiles, and as before, $\mathscr{I}$ is Predictor's relevant information.

In this paper we focus on cases where the information $\mathscr{I}$ is an exact and complete specification of the number of players, their associated (finite) strategy spaces and utility functions, for a particular noncooperative game. So formally, any particular $\mathscr{I}$ is an element of the set of all possible such game specifications. Note that this set of game specifications has the same cardinality as $\mathbb{R}$, and so density functions across it across it are well-defined.

2.1 The Likelihood

The likelihood function $\mathscr{L}(\mathscr{I} \mid q)$ is a distribution over game specifications $\mathscr{I}$, conditioned on the mixed strategy profile actually being q. For a fixed game specification $\mathscr{I}$, the likelihood will assign greater weight to q's that better "coincide" with $\mathscr{I}$, as determined by some external criterion. Choosing that criterion is a core component of how Predictor chooses to model human behavior. As such, it is analogous to the choice of what precise set-valued solution concept to adopt when pursuing a set-valued analysis.

$\lambda_i u_i$ that arises in player i's response function; uncertainty in exponents λ_i is on the same footing as uncertainty about utilities u_i.) So a player i's strategy in the Bayesian game is the map from their type λ_i to q_i. The logit QRE of this Bayesian game is the HQRE.

[9] To see this, note that for any fixed u_i and q_{-i}, almost all distributions $q_i(x_i)$ are not proportional to $exp[\lambda_i \mathbb{E}_{q_{-i}}(u_i \mid x_i)]$ for any value of λ_i.

In this paper, we focus on a likelihood that involves a quantification of bounded rationality, i.e., we consider cases where Predictor models human behavior in terms of such a quantification. For any given q_{-i} and $\mathscr{I}$, we will quantify the rationality of player i as the exponent of a logit response by player i that best fits i's actual mixed strategy q_i. For cases of a single player playing against Nature, quantifying rationality in terms of logit responses has a long history [see Train (2003)]. In the context of multi-player games, it goes back at least to the seminal work of McKelvey and Palfrey (1995) on quantal response equilibria. Accordingly, we call the likelihood associated with this rationality quantification the *QR-rationality likelihood*.

To simplify the exposition we introduce some more notation. Let $U^i_{q_{-i}}$ be the vector of expected utilities that player i gets from playing each of his possible pure strategies against the mixture q_{-i}. We call this player i's *environment*. The logit mixed strategy distribution for player i facing environment $U^i_{q_{-i}}$ is

$$\mathbb{L}_{U^i_{q_{-i}},b_i}(x_i) \propto e^{b_i \mathbb{E}_q(u_i|x_{i,j})}$$

where $\mathbb{E}_q(u_i \mid x_{i,j})$ is player i's expected utility from playing her j'th pure strategy against the mixture q_{-i}. Note that as b_i increases, the mixed strategy $\mathbb{L}$ assigns greater probability to those pure strategies of i with greater expected utility. Moreover, as $b_i \to \infty$, the logit mixed strategy is a best response to q_{-i} [see McKelvey and Palfrey (1995)]. Accordingly, in the experimental literature, the constant b_i is commonly interpreted as a measure of i's rationality.

In both laboratory and field experiments, players do not adopt *exactly* logit responses. To be a proper Bayesian, Predictor's likelihood function must reflect this. (Otherwise, strictly speaking, any experimental data would invalidate Predictor's Bayesian model.) Therefore her measure of rationality must be well-defined even for arbitrary, non-logit responses. So we need a way to define rationality for each player i for an arbitrary profile q.

One method of doing so is to define rationality as the value of b_i such that the logit response distribution specified by b_i best fits the actual q_i. This in turn requires choosing how to measure how well one distribution fits another.

Here we use the Kullback-Leibler (KL) divergence to measure that fit. KL divergence is an extremely common measure of how well one distribution fits another, with its origins in information theory [see Kullback and Leibler (1951); Kullback (1951, 1987); Cover and Thomas (1991a); Mackay (2003b)]. The KL divergence from q_i to the logit response distribution parameterized by b_i is

$$\begin{aligned} KL\left(q_i(x_i), \mathbb{L}_{U^i_{q_{-i}},b_i}(x_i)\right) &= \sum_{x_{i,j}\in X_i} q(x_{i,j}) \ln\left(\frac{q(x_{i,j})}{\mathbb{L}_{U^i_{q_{-i}},b_i}(x_i)}\right) \\ &= \sum_{x_{i,j}\in X_i} q(x_{i,j}) \ln\left(\frac{q(x_{i,j})\sum_{x_{i,l}\in X_i} e^{b_i\mathbb{E}_q(u_i|x_{i,l})}}{e^{b_i\mathbb{E}_q(u_i|x_{i,j})}}\right). \end{aligned} \tag{10}$$

We define the rationality of player i for arbitrary q to be the (unique) minimizer over b_i of $KL\left(q_i(x_i), \mathbb{L}_{U^i_{q_{-i}},b_i}(x_i)\right)$. This gives us the following characterization of rationality.

Definition 2.1. The *QR-rationality* of q_i against q_{-i}, written $\beta_i(q, U_i)$, is the value of b_i that minimizes the KL distance from q_i to $\mathbb{L}_{U^i_{q_{-i}}, b_i}(x_i)$, equation 10. The vector of all players' rationalities for the mixed strategy profile q and set of utility functions U is $\beta(q, U)$.

We will often simplify notation and just write "$\beta_i(q, U)$". If U is implicitly held fixed, we will often write "$\beta_i(q)$", or even "β_i".

In the special case where $U^i_{q_{-i}}(x_i)$ is independent of x_i, the QR-rationality parameter β_i can be any real number. (The same theoretical pathology applies when fitting logit-QRE's to experimental data.) Usually the set of q exhibiting this pathology is of zero measure in $\Delta_{\mathcal{X}}$ however, and therefore can be ignored. However, for completeness we define $\beta_i = \infty$ when $U^i_{q_{-i}}(x_i)$ is independent of x_i.

The QR-rationality $\beta_i(q, U_i)$ is invariant under shifts of the player utility functions. However, just like the logit-QRE profile, β_i is not invariant to positive rescalings of utility functions. So QR-rationality depends on utility units. This dependence is simply linear however: Multiplying the scale of the utility function of player i by some constant A is equivalent to using the original utility function but multiplying player i's rationality by A.

$\beta(q)$ is a single-valued vector-valued function of q. In addition, it is bounded function almost everywhere in $\Delta_{\mathcal{X}}$. Moreover, the expected utility of player i is the same under q_i as under the logit mixed strategy of player i for a rationality value $\beta_i(q)$. Formally, $\mathbb{E}_q(u_i) = \mathbb{E}_{q_i^{\beta_i(q)}, q_{-i}}(u_i)$ where $q_i^{\beta_i(q)}$ is the logit distribution given by $\mathbb{L}_{U^i_{q_{-i}}, b_i}$. These and other properties of QR rationality are established inWolpert and Bono (2008).

Here we presume that the likelihood $\mathscr{L}(\mathscr{I} \mid q)$ is a product over each player i of the value of a real-valued function F_i of the QR rationality of player i for mixed strategy profile q, for the game specified in $\mathscr{I}$:

$$\mathscr{L}(\mathscr{I} \mid q) \;\; \propto \;\; \prod_i F_i(\beta_i(q)) \tag{11}$$

We also presume that each F_i is a monotonically increasing function, i.e., we presume that everything else being equal, Predictor expects each player to be more likely to have high rationality rather than lower rationality.

One way to justify this presumption is to assume that a particular human's rationality is a feature of their behavior that is invariant across games. A weaker assumption that justifies our presumption is that the population average of the rationality of humans is invariant across games. Formally though, our likelihood doesn't require the existence of a population of players; it is simply a quantification of Predictor's uncertainty about player behavior.

Given our presumed form of the likelihood, our remaining task is to specify the precise functions F_i. To do so, we proceed by computational expediency.[10] To make calculations

[10]In this, we are inspired by random utility models, where computational expediency is sometimes used to justify modeling player utility uncertainty in a way that makes the resultant decision distributions be functions like the logit and probit.

in PGT typically requires use of Monte Carlo procedures. Accordingly, computational expediency lead us to choose an f so that the resultant Monte Carlo estimators have good convergence behavior, i.e., have low variance. In particular, note that if $\mathscr{L}(\mathscr{I} \mid q)$ were unbounded as $\beta_i(q)$ grows, Monte Carlo estimates of (functions of) the posterior could have infinite variance. Accordingly, we choose an f that is bounded:

$$\mathscr{L}(\mathscr{I} \mid q) \propto \prod_i [\tanh(\beta_i(q) - .5) + 1]^{\alpha_i} \tag{12}$$

This choice for F_i is monotonically increasing and bounded, as required. In addition, using this F_i means that any q under which there is a player with zero QR-rationality will receive zero weight in the posterior. The separate F_i may differ in their values of the parameter α_i. Since an increase in α_i increases the likelihood that i chooses strategies with a high β_i, α_i quantifies Predictor's information/beliefs about how much more likely player i is to be rational rather than irrational. This provides a way for Predictor to use the likelihood to model heterogeneity in player abilities.

It should be emphasized that equation 12 is not the only reasonable choice for a QR-rationality likelihood.[11] Such a multiplicity of modeling choices is ubiquitous in all types of statistics, not just PGT. (In particular, it is ubiquitous in econometrics). When making a prediction about a system, ultimately a Bayesian statistician must choose how to quantify their insight into how the system's state is related to what information they have concerning it, in terms of a likelihood. This is just as true when the system being predicted is a set of players as when it is a more conventional "inanimate" system.

In this regard, in the appendix we discuss two likelihood that do not involve QR rationality. The first of these involves *N-rationality*, and says that the likelihood of a player choosing a specific q_i when the other players choose q_{-i} depends on how close the associated expected utility, $\mathbb{E}_{q_i,q_{-i}}(u_i)$, is to the best response expected utility, $\max_{q'_i}\mathbb{E}_{q'_i,q_{-i}}(u_i)$. This likelihood is closely related to the epsilon equilibrium concept of Radner (1980), which also uses the best response expected utility as a target. The second likelihood involves *intelligence*, and says that the likelihood of a particular q_i given q_{-i} depends on the proportion of strategies q'_i that yield a lower expected utility than q_i, given q_{-i}. The appendix contains a simple example to illustrate the differences among the QR-rationality, N-rationality and intelligence likelihoods.

2.2 The Prior

The role of the prior distribution, $P(q)$, is to quantify Predictor's subjective beliefs about the relative probabilities of mixed strategy profiles without regard to the game-specific

[11] For example, one could use

$$\mathscr{L}(\mathscr{I} \mid q) \propto \prod_i g_i(\beta_i(q)) \tag{13}$$

where

$$g_i(\beta_i(q)) = \begin{cases} \alpha_i \ln(\beta_i(q) + 1) + 1 & \text{if } \beta_i(q) \geq 0 \\ e^{\alpha_i \beta_i(q)} & \text{otherwise.} \end{cases}$$

information $\mathscr{I}$ (which instead goes into the likelihood function). Properly speaking, this should reflect the relative probability Predictor assigns to a population average, over all game-playing situations she might encounter, of the profile arising in those games. To avoid pre-judging the results of such an average, here we instead use a prior over distributions with roots in information theory.

Based on any of several separate sets of simple desiderata, there is a unique real-valued quantification of the amount of syntactic information in a distribution $q(x)$ [see Shannon (1948), Mackay (2003a), Cover and Thomas (1991b)]. That quantification, the Shannon entropy of a density q, is written as $S(q) = -\sum_x q(x)\ln(q(x))$. The entropic prior density is

$$P(q) \propto \exp(\delta S(q))$$

for a real-valued parameter δ. This prior has proven extremely powerful in many branches of Bayesian statistics [see Mackay (2003a), Gull (1988), Strauss et al. (1994)]. Loosely speaking, for $\delta > 0$ it says that *everything else being equal*, mixed strategy profiles are more likely the flatter they are. This means it favors those profiles that have least effect on the posterior over pure strategy profiles,

$$P(x \mid \mathscr{I}) = \int q(x)P(q \mid \mathscr{I})dq. \tag{14}$$

Similarly, the prior over pure strategy profiles, $P(x) = \int q(x)P(q)dq$ is flat under the entropic prior $P(q)$. (By symmetry, this is true regardless of δ.)

Setting $\delta < 0$ rather than $\delta > 0$ has an important behavioral interpretation: it means that Predictor believes that human beings are particularly poor at randomizing. As usual, if Predictor is actually uncertain how to set δ, in the Bayesian framework she should average over it.

Just as there is more than one reasonable likelihood, there is more than one reasonable prior. Indeed, the entropic prior is just one member of the Cressie-Read family of distributions [see Cressie and Read (1984); Read and Cressie (1988)]. As usual, ultimately real-world behavior should be used to set the prior. For simplicity though, here we will focus on the entropic prior.

Given the choice of an entropic prior, the associated posterior for the QR rationality likelihood is

$$P(q \mid \mathscr{I}) \;\; \propto \;\; \exp(\delta S(q)) \prod_i [\tanh(\beta_i(q) - .5) + 1]^{\alpha_i} \tag{15}$$

It is important to emphasize that this posterior is not an average of QRE's having different exponents. The set of $q \in \Delta_{\mathcal{X}}$ that can be expressed as a QRE for some β has measure 0. In contrast, the posterior in equation 15 is non-zero for every q.

In addition, consider an alternative likelihood where Predictor knows the rationality of each player exactly:

$$\hat{P}_{\beta^*}(\mathscr{I} \mid q) = \prod_i \delta(\beta_i(q) - \beta_i^*)$$

for some set of player-indexed constants $\{\beta_i^*\}$. Combining this alternative likelihood with the entropic prior provides an alternative posterior, $\hat{P}_{\beta^*}(q \mid \mathscr{I})$. It turns out that for either games against Nature, or multi-player games where all the pure strategy spaces are binary, the mode of $\hat{P}_{\beta^*}(q \mid \mathscr{I})$ is the QRE for QRE exponents β_i^*. However for more general games this need not be the case. On the other hand, even in such games the mode of $\hat{P}_{\beta^*}(q \mid \mathscr{I})$ is often well-approximated by the associated QRE, and the correction terms can be calculated explicitly.

These results and others are established in Wolpert and Bono (2008). That paper also analyzes the relationship between the alternative posterior $\hat{P}_{\beta^*}(q \mid \mathscr{I})$ and regret (see Shoham et al. (2007)). This analysis establishes that the alternative likelihood, which stipulates that a player's rationality has the same constant value for all games, is not equivalent to a likelihood that instead stipulates that a player's expected regret has the same constant value in all games. Finally, that paper also explores the relation among the alternative posterior, statistical physics and information theory.

3 Cournot Duopoly

We now illustrate PGT for prediction and design using the Cournot duopoly. We start with a review, to introduce our notation for the Cournot duopoly.

There are two firms, A and B, that produce goods α and β respectively. They independently decide how much of their own good to produce. The produced quantities are $x_A \in X_A = [0, \bar{x}_A]$ and $x_B \in X_B = [0, \bar{x}_B]$, where $\bar{x}_i$ is the maximum quantity that firm i can produce. We define $X = X_A \times X_B$.

The market price for firm i's product is decreasing in both x_i and x_{-i}, i.e. goods α and β are substitutes. So if we write this price as $D_i(x_i, x_j)$, then $D_i(x_i, x_j) \geq 0$, $\frac{\partial D_i}{\partial x_i} \leq 0$ and $\frac{\partial D_i}{\partial x_j} \leq 0$ for all $(x_i, x_j) \in X$. The total cost for firm i of producing x_i units is written as $C_i(x_i)$. As is standard, we assume that $C_i(x_i) \geq 0$, $\frac{\partial C_i}{\partial x_i} \geq 0$ and $\frac{\partial^2 C_i}{\partial x_i^2} \geq 0$ for all i, $x_i \in X_i$. Combining, each firm i's profit function is

$$\Pi_i(x_i, x_j) = x_i D_i(x_i, x_j) - C_i(x_i).$$

We now analyze this duopoly model using PGT, using illustrative parametric forms of $D_i(\cdot, \cdot)$ and $C_i(\cdot)$, $i = A, B$. We will concentrate on D_i's having the form

$$D_i(x_i, x_j) = \begin{cases} d_{i1} - d_{i2}x_i + d_{i3}x_i^2 - d_{i4}x_i^3 - x_j, & \text{if greater than zero} \\ 0, & \text{otherwise} \end{cases}$$

We require $-d_{i2} + 2d_{i3}x_i - 3d_{i4}x_i^2 \leq 0$ for all $x_i \in X_i$ to ensure that $\frac{\partial D_i}{\partial x_i} \leq 0$. The parametric form for $C_i(\cdot)$ is

$$C_i(x_i) = \frac{e^{x_i}}{c_{i1}}.$$

These parametric forms describe a very broad range of strategic settings. As an example, the parameters $[\bar{x}_i = 20; d_{i1} = 20.4; d_{i2} = 2.165; d_{i3} = 0.12; d_{i4} = 0.0025; c_{i1} =$

$16,000,000]$ for $i = A, B$ produce the symmetric best response functions $x_i^*(x_j)$ illustrated in figure 1. In this example there are five intersections of the best response functions,

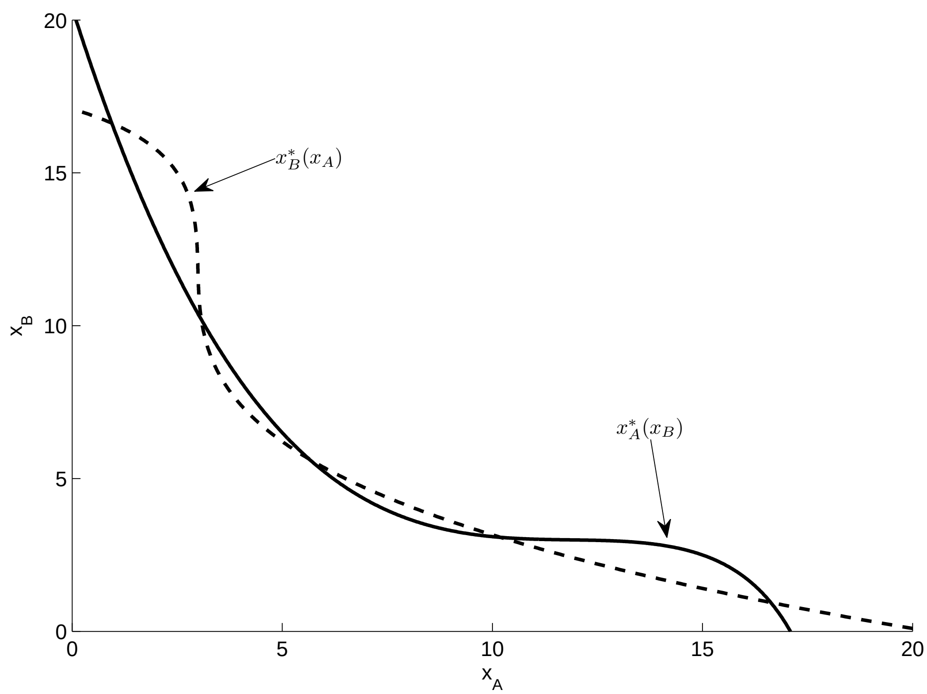

Figure 1: Best response functions for $\bar{x}_i = 20; d_{i1} = 20.4; d_{i2} = 2.165; d_{i3} = 0.12; d_{i4} = 0.0025; c_{i1} = 16,000,000$ for $i = A, B$.

indicating five pure strategy NE. That equilibrium set is

$$\Delta_1^{NE} = \{(16.6, 0.94), (10.2, 3.1), (5.6, 5.6), (0.94, 16.6), (3.1, 10.2)\}.$$

Changing the parameter d_{A1} from 20.4 to 19.1 represents a downward shift in firm A's inverse demand function of 1.7 dollars. This shift reduces the number of equilibria from five to one. The associated best response functions are depicted in figure 2. The unique equilibrium is $\Delta_2^{NE} = (0.55, 16.84)$.

4 Results

We formed our estimate of $\mathbb{E}(f \mid \mathscr{I})$ using conventional Monte Carlo importance sampling: We constructed a sampling distribution $H(\lambda)$, and then IID sampled H to generate many λ's, with our estimate of $\mathbb{E}(f \mid \mathscr{I})$ given by averaging the (importance sampling corrections to) the associated values $f(q^\lambda)P(q^\lambda \mid \mathscr{I})$. The details are given in the appendix.

As an illustration, $P(x \mid \mathscr{I})$ is depicted in figure 3 for the Cournot setting in figure 1. For comparison, we present in figure 4 the QRE distribution for the same Cournot

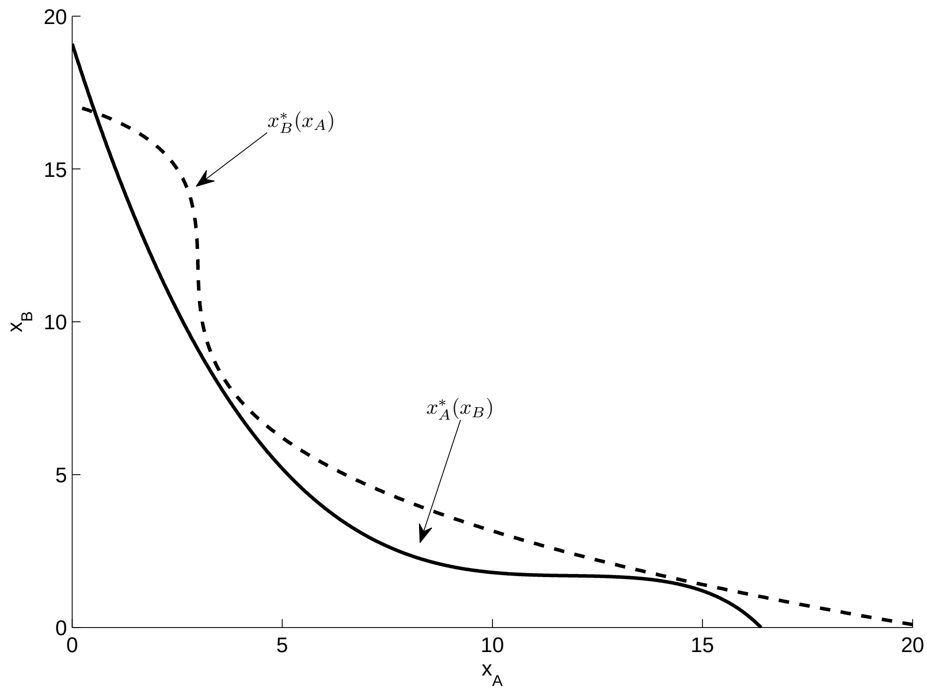

Figure 2: Best response functions for the same parameters as in figure 1, except that $d_{A1} = 19.1$ instead of 20.4.

setting.[12]

4.1 Model Combination

As was mentioned in regard to equation 2, if Predictor is uncertain about her modeling choices, she can simply marginalize out that uncertainty. As an example, to calculate the surface in figure 4, Predictor had to discretize the move spaces of the players, modeling each of them as though had a finite set of pure strategies, $\Lambda_i\{0, 0.1, 0.2, ..., 19.9, 20\}$. However Predictor is not forced to make a hard choice of one particular set of Λ_i's. She can instead average over such sets, according to how well she suspects each of them describe real-world behavior. As another example, Predictor can average over the values of the parameters in the likelihood function of equation 12. She can even average over rationality functions, or for that matter over entirely different types of likelihood functions.

Just as she can average over likelihoods, which involve the firms' profit functions, Predictor can average out her uncertainty about the firms' profit functions themselves. Note that in doing this Predictor is modeling *her* uncertainty concerning the player utility

[12] To specify this QRE we must choose the value of its exponent. We chose the mean of the distribution over QR-rationalities induced by $P(q \mid \mathscr{I})$, which is approximately 0.42. We must also discretize the move space. We do this by having the pure strategy sets of the players be the subspace $\Lambda \subset X$, where for all players i, $\Lambda_i = \{0, 0.1, 0.2, ..., 19.9, 20\}$.

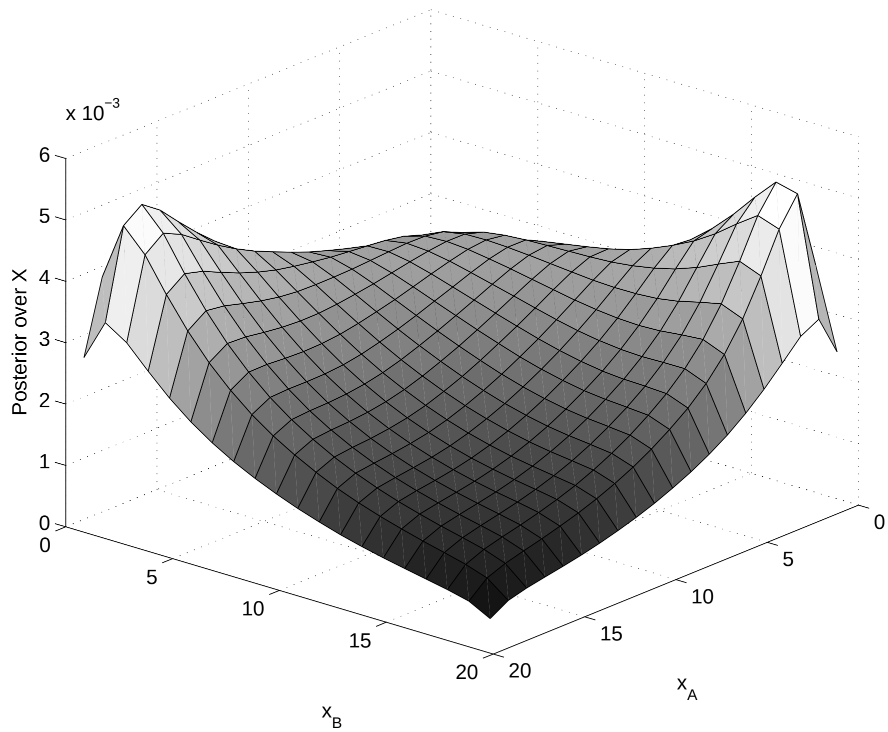

Figure 3: The Bayesian PGT posterior over x's with QR-rationality likelihood parameter $\alpha = 2.5$ and entropic prior parameter $\delta = 1$. Cournot profit parameters are as in figure 1.

functions. Since the players have complete information in the Cournot game scenarios Predictor is analyzing, they have no such uncertainty — which is reflected in the posterior Predictor creates.

To illustrate this, let m be the probability that the profit function parameters are those depicted in figure 1, comprising the set $\mathscr{I}'$, and let $1 - m$ be the probability that the profit function parameters are instead those depicted in figure 2, comprising the set $\mathscr{I}''$. Recall that $\mathscr{I}'$ has five NE, and $\mathscr{I}''$ has only one. If she believed the players were perfectly rational (as reflected in her choices of likelihood function), then by using PGT Predictor can average her model over those two sets of NE to properly capture her uncertainty about the profit functions. Formally, we write $\mathscr{I} = \{\mathscr{I}', \mathscr{I}'', m\}$ and break the likelihood into a sum of two terms:

$$\mathscr{L}(\mathscr{I} \mid q) = m\mathscr{L}(\mathscr{I}' \mid q) + (1 - m)\mathscr{L}(\mathscr{I}'' \mid q).$$

Alternatively, if Predictor does not assume perfect rationality, but instead assumes the QR-rationality likelihood, she can do the same sort of averaging over her uncertainty about profit functions. Figure 5 depicts such an averaging of $\mathscr{I}'$ and $\mathscr{I}''$ with $m = .5$.

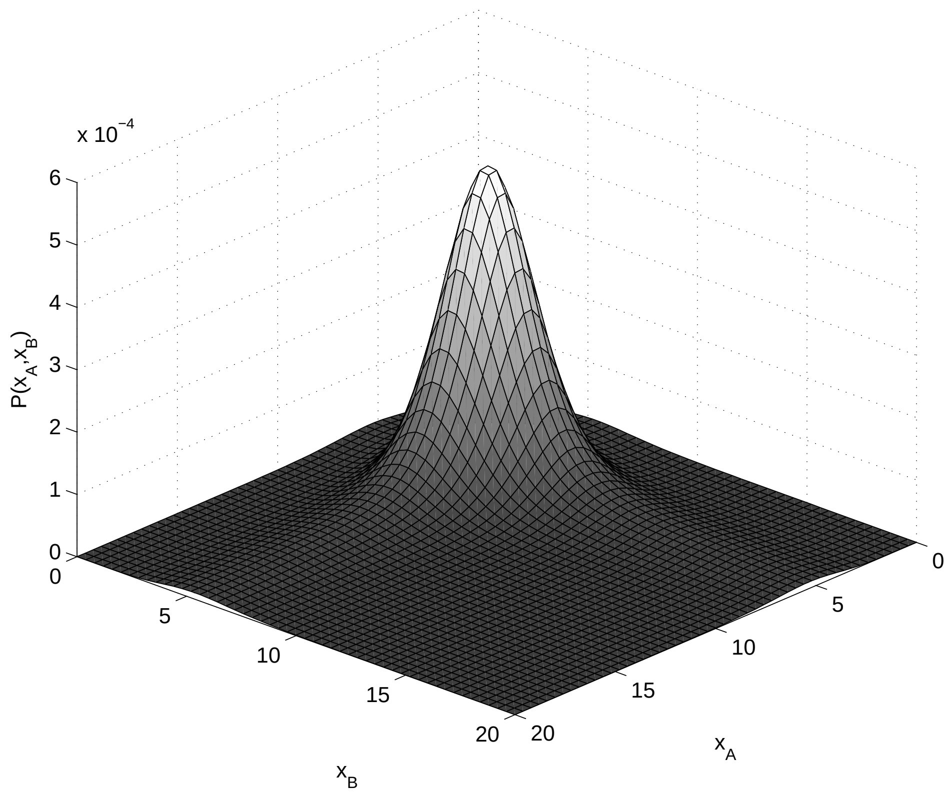

Figure 4: QRE distribution of moves with $\beta_i^q = 0.42$ $i = A, B$. The move space for each player i is $\Lambda_i = \{0, 0.1, 0.2, ..., 19.9, 20\}$.

4.2 Correlation of Pure Strategies

As was mentioned in Sec. 1.1, despite the fact that each $q \in \Delta_{\mathcal{X}}$ is a product distribution, $P(x \mid \mathscr{I})$ is generally not a product distribution. This is true in particular in our Cournot duopoly setting. As an example, consider an industry comprised of many firms, where subsets of those firms engage in noncooperative games with each another. To simplify the analysis, assume that all the games are the same, and all involve only two firms. (For example, the firms might be a set of many distinct duopolies, each duopoly controlling production of the same good, but in a different city.) Say that Predictor observes the joint moves of many different pairs of the firms engaged in such two-player games. Then even if there is no collusion — in each game, the moves of the two firms are independent — *to Predictor* it would appear as though there is collusion in the industry.

To illustrate this, consider the duopoly setting from figure 1 with the QR-rationality likelihood where $\alpha = 2.5$. The correlation between x_A and x_B is -0.29. Changes to the likelihood, such as an increase in α, can increase the magnitude of the correlation. For instance, by setting $\alpha = 4$, we increase the correlation from -0.29 to -0.53. Changes to the profit function also affect the degree of coupling in $P(x \mid \mathscr{I})$. For example, in the

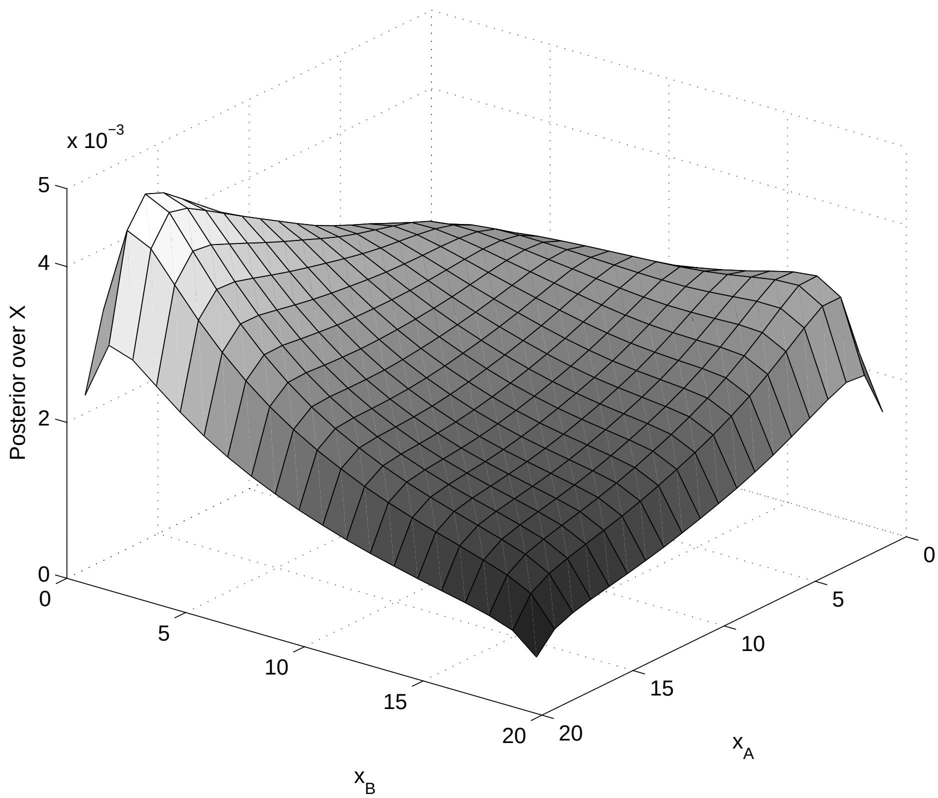

Figure 5: Posterior distribution over moves formed by integrating Predictor's uncertainty concerning firm A's profit function; $\alpha = 2.5$ and $\delta = 1$.

duopoly setting from figure 2 the correlation between x_A and x_B is -0.28. When $\alpha = 4$, the correlation is -0.59. In all these cases, the fact that the covariance is non-zero means that the quantity choice of firm A is, on average, informative about the quantity choice of firm B.

4.3 Decision-Theoretic Prediction

Suppose Predictor wants to make a point prediction of the quantities (x_A, x_B) that will be played in the Cournot duopoly. As discussed in Sec. 1.1, to do this she must specify a loss function, and then make the prediction that minimizes her posterior expected loss. That prediction is Bayes-optimal. (See equation 7.)

For our Cournot duopoly game, calculating that Bayes-optimal prediction is straightforward for both the zero-one loss function and the quadratic loss function. The results for the Cournot game parameters from figures 1 and 2 are reported in table 1. For each of those game parameters, there is a single QRE, i.e., a single profile over pure strategies. Accordingly we can also calculate the Bayes-optimal point-predictions for the QRE, for both game parameters. Those point predictions are also reported in the table.[13] Of course, if there were multiple QRE's for either set of game parameters, then we would not be able to make this calculation. Finally, for completeness, the table also reports the NE pure strategy profiles for the two game parameters.

One interesting point is that the PGT pure strategy prediction that results from a zero-one loss function applied to figure 1 is not unique. This is because $P(x|\mathscr{I})$ is bimodal *and* symmetric. In the PGT context, this is not problematic as it is with set-valued concepts because we know the associated probabilities of each mode, which are the same, i.e. both modes minimize the zero-one loss function. Therefore, a researcher with a zero-one loss function is indifferent between the two modes in exactly the same way that any decision-maker is indifferent between any two alternatives that yield the same expected utility.

Prediction	Figure 1	Figure 2
PGT zero-one loss	$(16.9, 1.1), (1.1, 16.9)$	$(0.8, 16.8)$
PGT quadratic loss	$(8.7, 8.7)$	$(7.9, 9.4)$
QRE zero-one loss	$(5.4, 5.4)$	$(4.0, 6.5)$
QRE quadratic loss	$(6.5, 6.5)$	$(4.8, 7.5)$
Pure NE	$\{(16.6, 0.94), (10.2, 3.1), \ldots$ $(5.6, 5.6), (3.1, 10.2), (0.94, 16.6)\}$	$(0.55, 16.8)$

Table 1: Predictions based on PGT & on set-values solution concepts. The QR-rationality likelihood parameter is $\alpha = 2.5$. The entropic prior parameter is $\delta = 1$. The QRE parameter is $\beta^q = 0.42$ for the game from figure 1 and $\beta^q = 0.50$ for the game from figure 2.

4.4 Decision-Theoretic Choice of a Mechanism

Suppose Predictor must choose a per-unit tax, τ, for our Cournot duopoly setting with negative consumption externalities. Let $W(q, \tau)$ be the social welfare under τ for player mixed strategy profile q. As an example, the negative externality could by a total external cost function given be $EC(x) = e(x_A + x_B)^2$ for some $e > 0$. We could then have the social welfare function be the social surplus, which equals consumer surplus (CS) plus firm profits plus tax revenue minus external costs.[14] For a given strategy profile q and tax τ, this social welfare is

$$W(\tau, q) = \mathbb{E}_q[CS] + \mathbb{E}_q[\pi_A + \pi_B] + \tau\mathbb{E}_q[x_A + x_B] - e\mathbb{E}_q[(x_A + x_B)^2].$$

Regardless of what the social welfare functions is, being a Bayesian decision theorist,

[13] As in figure 4, the QRE is computed with $\beta_i^q = 0.42$ $(i = A, B)$, for the Cournot game from figure 1. For the game from figure 2, $\beta^q = 0.3$, the mean of the distribution over β's from the corresponding Bayesian PGT posterior.

[14] For a discussion of analysis of Cournot efficiency using set-valued solution concepts, see Seade (1985)

Predictor wants to choose τ to maximize the associated expected social welfare,

$$\mathbb{E}(W \mid \tau) \;=\; \int W(q,\tau)P(q \mid \tau, \mathscr{I})dq \tag{16}$$

To illustrate this, consider the Cournot setting represented by the best response functions in figure 1, and set $e = .5$. Table 2 reports the posterior expected welfare values for five tax levels, $\tau = 0, 1, 2, 3, 4$. Also reported there are predictions of expected social welfare for the five tax values under the NE and QRE modeling choices. As before, the logit exponents β^q for the three QRE's correspond to the PGT values of α for the three PGT models, respectively.

Before using these results to compare the various solution concepts, it's worth making some general comments. First, note that for the PGT models with $\alpha > 1$ and the corresponding QRE models with $\beta^q > 0.13$, expected social welfare is highly nonlinear in the tax rate. In fact, even when the tax rate is restricted to the range $[0, 4]$, there is at least one local maximum that is not a global maximum. For instance, in the PGT model with $\alpha = 2.5$, the expected social welfare for $\tau = 1$, w_1, is greater than the expected social welfare for $\tau = 0$ and $\tau = 2$. However, w_1 is less than w_4. Similarly, the PGT optimal tax rate is highly nonlinear in α.

Note also that for several tax rates the QRE expected welfare values makes large jumps as the tax rate changes only slightly. This is because the QRE mapping encounters a bifurcation point as τ varies. For example, in the QRE model with $\beta^q = 1.21$, the expected social welfare is 97.17 for $\tau = 2.074$, then drops to 92.84 for $\tau = 2.0745$, then jumps to 137.61 for $\tau = 2.075$. This contrasts with the PGT model which, being an average over all q's, does not exhibit such sensitive dependence on exogenous parameters.

The advantages of using PGT extend substantially beyond such robustness however. For example, the optimal tax level under the NE model is not well-defined, since it depends on which of the three NE under each of the first two tax values, $\tau = 0, 1$, are selected. The fact that there *is* a well defined optimal tax level for the QRE models is serendipitous, given that there are multiple QRE's for $\tau \geq 1$. The reason the optimal tax level is well-defined despite this multiplicity is that the Cournot game considered here is symmetric, and therefore so are its QRE's. This means that all of the QRE's yield the same expected welfare.[15] In general, whenever there are multiple QRE's for some tax level but we are not so fortunate that they all give the same value of social associated social welfare values, there is no well-defined optimal tax level under the QRE solution concept. Such problems cannot arise with PGT.

Because PGT provides a full posterior distribution over welfare, $P(w \mid \tau, \mathscr{I})$, Predictor's model contains far more information for each τ than simply the associated posterior expected social welfare. Figure 6 illustrates this extra information by plotting the posterior distribution over social welfare for tax rates $\tau = 0, 2, 4$, where $\alpha = 2.5$. For comparison purposes, that figure also gives the QRE expected social welfare values and several of the NE expected social welfare values for the same three tax rates.

[15] We should note that although we think we have found all the QRE's numerically, we have not formally proven this.

Model	Welfare				
	w_0	w_1	w_2	w_3	w_4
PGT ($\alpha = 1$)	73.06	77.24	77.99	78.88	**80.36**
PGT ($\alpha = 2.5$)	82.10	88.93	88.76	88.31	**88.97**
PGT ($\alpha = 3.25$)	88.72	**95.88**	95.59	94.55	94.92
PGT ($\alpha = 4$)	95.86	102.09	**102.79**	101.63	102.13
QRE ($\beta^q = 0.13$)	68.37	71.75*	74.16*	75.81*	**78.54***
QRE ($\beta^q = 0.42$)	92.11	93.49*	91.29*	99.75*	**105.85***
QRE ($\beta^q = 0.74$)	94.72	95.77*	95.46*	123.92	**133.31***
QRE ($\beta^q = 1.21$)	95.80	96.84*	96.41*	146.35*	**152.63***
Pure NE	316.27* 210.10* 174.10	301.06* 234.19* 164.07	154.18	139.82	135.69

Table 2: Comparing welfare values across taxes and across models. For the PGT and QRE models, the optimal tax (among the choices $\tau = 0, 1, 2, 3, 4$) is indicated by boldface type for the associated expected welfare value. An asterisk indicates that multiple equilibria give rise to the same expected welfare. The QRE is computed for the subspace $\Lambda \subset X$ where $\Lambda_i = \{0, 0.1, ..., 19.9, 20\}$.

Predictor can use this extra information to use decision theory to analyze many issues that are often very important in practice. For example, Predictor's objective may not be interested in choosing the tax level τ that maximizes posterior expected social welfare, but rather interested in choosing the τ that maximizes the posterior expected value of some function $g(W(q, \tau)) = (W(q, \tau))^r$. For example this is the case if she is risk-averse, being more concerned that the social welfare not be too low than that it be large. Because PGT provides a full posterior distribution over social welfare values, Predictor can use PGT to choose the tax value that accounts for this risk aversion by maximizing the posterior expected value of g. In contrast, set-valued solution concepts, since they allow multiple equilibria, in general cannot be used in such a decision-theoretic way to accommodate risk-aversion.

Modeling Predictor's risk-averse objective in this decision-theoretic manner is particularly appealing in the case of a major market change like new taxes. Major market changes are the result of costly legislative processes, and are often very difficult to retract once in place. Therefore, a social planner (Predictor) may be averse to the risk that firms engage systematically in behavior that is detrimental to her objective. She may prefer a mechanism that produces lower expected welfare with less risk rather than a mechanism that produces greater expected welfare with more risk.

PGT's posterior over social welfare provides many other capabilities to Predictor that are similar to allowing her to incorporate risk aversion. Armed with such a posterior, Predictor can compare mechanisms by answering questions that real-world stakeholders often ask. The following calculations are for the PGT model of the Cournot duopoly from figure 1 with $\alpha = 2.5$ and $\delta = 1$.

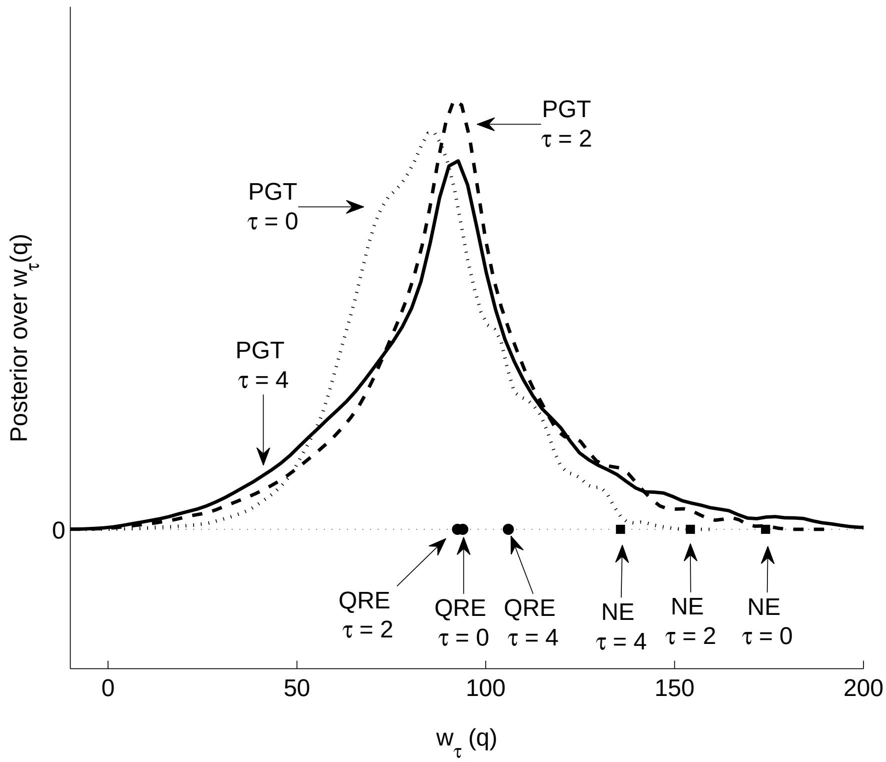

Figure 6: Bayesian PGT posterior distribution over social welfare $W(q, \tau)$ for tax rates 0, 2 and 4. QRE ($\beta = 0.42$) and NE social welfare points are provided for comparison.

- *"Which of the taxes minimizes the variance in welfare?"*
  Procedure:
$$\min_{\tau} \mathrm{var}[w_\tau] = \int \left(w - \mathbb{E}[w_\tau]\right)^2 P(w \mid \tau, \mathscr{I})dw.$$
  Solution: $\tau^* = 0$, $\mathrm{var}[w_0] = 376.00$.

- *"Which of the taxes minimizes the probability that welfare is below some threshold value $\underline{w} = 50$?"*
  Procedure:
$$\min_{\tau} Pr(w \leq \underline{w} \mid \tau, \mathscr{I}) = \int_{-\infty}^{\underline{w}} P(w \mid \tau, \mathscr{I})dw.$$
  Solution: $\tau^* = 0$, $Pr(w \leq 50 \mid \tau = 0, \mathscr{I}) = 0.045$.

- *"Which of the taxes maximizes social welfare subject to the condition that the expected profits of firms are positive with probability greater than some threshold value $p = 0.5$?"*
  Procedure:
$$\max_{\tau} \int wP(w \mid \tau, \mathscr{I})dw \quad \text{s.t.} \quad \int_0^{\infty}\int_0^{\infty} P(\pi_A, \pi_B \mid \tau, \mathscr{I})d\pi_A d\pi_B \geq p$$

where

$$P(\pi_A, \pi_B \mid \tau, \mathscr{I}) = \int I\left[\pi(q), (\pi_A, \pi_B)\right] P(q \mid \tau, \mathscr{I}) dq$$

and $I[a, b]$ is the indicator function that equals one when $a = b$ and zero otherwise.

Solution: $\tau^* = 1$, $Pr(\pi \geq 0 \mid \tau = 1, \mathscr{I}) = 0.779$, $\mathbb{E}(w|\tau = 1, \mathscr{I}) = 88.93$.

Note that depending on what question she wants to answer, Predictor gets a different answer for the optimal tax value. Note also that Predictor cannot answer any of these questions using a set-valued distribution concept.

5 Future Work

There are very many ways that the PGT analysis above can be extended. Some of them have to do with numerical issues (see the appendix). In this section we briefly describe some of the more theoretical future work.

Most set-valued solution concepts have been motivated by introspection. This is also true of the distribution-valued solution concept used in this paper. However since our purpose with PGT is to predict behavior in the real world, it would be preferable to use a distribution-valued solution concept motivated by experimental data. Constructing distribution-valued solutions concepts that are directly motivated by experimental data will be the focus of some future work.

The PGT analysis in this paper considered complete and perfect information single-shot strategic form games. Some future work involves extending PGT models to include noncooperative strategic scenarios that differ from these kinds of games. For example, it would be interesting to extend PGT to analyze extensive form games, Bayesian games, signalling games, repeated games, etc. (In regard to the latter, note that due to its stochastic formulation, PGT might be particularly well-suited to analyze Markov games.)

In addition to analyzing such noncooperative games though, PGT might also be extended to consider cooperative games. Such a "Predictive Cooperative game Theory" (PCT) would assign relative probabilities to all possible sets of coalitions and payoff vectors. Ideally, by providing this set of probabilities, PCT would resolve the difficulty that current cooperative game theory faces, of having several different (set-valued) solution concepts that all seem quite reasonable. According to PCT, the issue is not which such set-valued solution concept is "correct". Rather the issue is to determine the relative probabilities of the associated coalitions and payoff vectors.

PGT can also be extended to consider unstructured bargaining games, providing a Predictive Unstructured Bargaining (PUB). In PUB, one does not map a bargaining problem $S \subset \mathbb{R}^N$ to a single point in that problem's feasible set, $s \in S$. Rather one maps S into a density function over $\mathbb{R}^N$, $\mu(s)$. In particular, it is straight-forward to translate the Nash bargaining axioms into a form that concerns such a map taking any S to a $\mu(s)$. For example, the Nash scale invariance axiom gets translated into an axiom saying that if S is scaled by a certain amount, then $\mu(s)$ is scaled accordingly.

It turns out that the only map $f : S \to \mu(s)$ that obeys these translations of the Nash bargaining axioms is

$$
\begin{aligned}
[f(S)](s) &\propto \prod_{i=1}^{N} s_i^{\alpha_i} && \text{for } s \in S \\
&= 0 && \text{otherwise}
\end{aligned}
$$

for some set of player-indexed constants $\{\alpha_i\}$, where it is implicitly assumed that the default point is $s = 0$. (Interestingly, Nash's Pareto axiom is not used in deriving this result.) If all constants α_i are the same, the mode of $\mu(s)$ is just the Nash bargaining solution, while if they differ, the mode is the Harsanyi solution. Of course, in general the Bayes-optimal prediction of s by Predictor, specified by her loss function,[16] will differ from the mode of $\mu(s)$ in general.

More generally, the PUB distribution is perfectly-well defined even if the feasible set is non-convex, non-comprehensive and consists of only a finite number of values. (This is formally proven in work in preparation by the authors.) In addition, PUB can be used by Predictor to design the feasible set, to maximize the associated expected social welfare.

Finally, we note that PGT can be used to elaborate some subtleties in interpreting the physical scenario underlying single-shot normal form noncooperative games of perfect and complete information. The first distinction has to do with what it is that the players choose, and the second has to do with whether they have previously interacted. These subtleties can be important for deciding how best to model a strategic scenario using a distribution-valued solution concept. Since they are somewhat philosophical in nature, we defer discussion of them to the appendix.

References

Aguirregabiria, Victor and Pedro Mira, "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 2009, *In Press, Corrected Proof*, –.

Aumann, Robert, "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, 1974, *1*, 67–96.

Bajari, P., J. Hahn, H. Hong, and G. Ridder, "A Note on Semiparametric Estimation of Finite Mixtures of Discrete Choice Models with Application to Game Theoretic Models," *International Economic Review*, 2010. in press.

Camerer, Colin, Teck-Hua Ho, and Juin-Kuan Chong, "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics*, 2006, *119* (3), 861–898.

Costa-Gomes, Miguel and Vincent Crawford, "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, December 2006, *96* (5), 1737–1768.

[16] That loss function says how much Predictor loses if she predicts s' while the actual bargain ends up being s.

Cover, T. and J. Thomas, *Elements of Information Theory*, New York: Wiley-Interscience, 1991.

_ **and** _ , *Elements of Information Theory*, New York: Wiley-Interscience, 1991.

Crawford, V. and N. Iriberri, "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," *Econometrica*, 2007, *75*, 1721 – 1770.

Cressie, N. and T. R. C. Read, "Multinomial goodness of fit tests," *Journal of the Royal Statistical Society, Series B*, 1984, *46*, 440464.

Geweke, J., "Bayesian Inference in Econometric Models using Monte Carlo Integration," *Econometrica*, 1989, *57*, 1317–1340.

Gull, S. F., "Bayesian Inductive Inference and Maximum Entropy," in "Maximum Entropy and Bayesian Methods" Kluwer Academic Publishers 1988, pp. 53–74.

Kullback, S., *Information theory and statistics*, NY: John Wiley and Sons, 1951.

_ , "The Kullback-Leibler distance," *The American Statistician*, 1987, *41*, 340341.

_ **and R.A. Leibler**, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, 1951, *22* (1), 7986.

Mackay, D., *Information theory, inference, and learning algorithms*, Cambridge University Press, 2003.

Mackay, D.J.C., *Information theory, inference, and learning algorithms*, Cambridge University Press, 2003.

McKelvey, Richard D. and Thomas R. Palfrey, "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, July 1995, *10* (1), 6–38.

Radner, Roy, "Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives," *Journal of Economic Theory*, April 1980, *22* (2), 136–154.

Read, T. R. C. and N. A. C. Cressie, *Goodness-of-Fit Statistics for Discrete Multivariate Data*, New York: Springer-Verlag, 1988.

Robert, Christian P. and George Casella, *Monte Carlo Statistical Methods*, 2nd ed., Springer, 2004.

Rogers, B.W., T.R. Palfrey, and C.F. Camerer, "Heterogenous quantal response equilibrium and cognitive hierarchices," *Journal of Economic Theory*, 2009.

Seade, J., "Profitable cost increases and the shifting of taxation: equilibrium response of markets in Oligopoly," 1985. working paper, University of Warwick, Department of Economics, The Warwick Economics Research Paper Series (TWERPS) #260.

Shannon, Claude E., "A Mathematical Theory of Communication," *Bell System Technical Journal*, July, October 1948, *27*, 379–423, 623–656.

Shoham, Y., R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?," *Artificial Intelligence*, 2007, pp. 365–377.

Starmer, C., "Developments in Non-Expected Utility Theory: the Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, 2000, *38*, 332–382.

Strauss, C.E., D.H. Wolpert, and D. Wolf, "Alpha, Evidence, and the Entropic Prior," in A. Mohammed-Djafari, ed., *Maximum Entropy and Bayesian Methods 1992*, Kluwer 1994.

Train, K. E., *Discrete Choice Methods with Simulation*, Cambridge University Press, 2003.

Wolpert, D. H., "Theory of Collective Intelligence," in K. Tumer and D. H. Wolpert, eds., *Collectives and the Design of Complex Systems*, Springer New York 2003.

Wolpert, D.H. and J.W. Bono, "Statistical Prediction of the Outcome of a Noncooperative Game," 2008. http://ssrn.com/abstract=1184325.

_ **and** _ , "PGT: A Statistical Approach to Prediction and Mechanism Design," in S.K. Chai, J. Salerno, and J. Mabry, eds., *Proceedings of SBP 2010*, Springer 2010.

Appendix

A Advantages of Distribution-Valued Solution Concepts

There are many important advantages to using PGT models rather than set-valued equilibrium concepts. We mention two such arguments here:

1. One benefit of PGT is that, in general, it assigns non-zero probability to all mixed strategy profiles. This means that PGT respects the fact that in the real world, all mixed strategy profiles can occur with some non-zero probability. In contrast, set-valued solution concepts generally assigns probably zero to almost all profiles, in the sense that it treats all strategy profiles outside a measure-zero equilibrium set as physically impossible.[17]

   This feature of using set-valued solution concepts presents a well-known empirical problem. In particular, it means that all econometric studies of equilibrium concepts must first devise an error structure and append it to the equilibrium theory before estimation can be carried out. This error structure effectively converts the set-valued solution concept into a distribution-valued solution concept. In other words, this *ex post* error structure converts an equilibrium model into an instance of PGT.

   Such *ex post* theorizing carries its own assumptions about strategic behavior. These assumptions are in addition to the assumptions of the equilibrium theory, and the two sets of assumptions can be difficult to reconcile. Regardless of whether or not the two sets of assumptions can be reconciled, it is clear that empirical tests of equilibrium concepts are not direct tests of the equilibrium theory. Rather they are tests of the equilibrium theory *as modified by* the error structure.

   In contrast, when a researcher devises a PGT model, that model can be tested without modification. So empirical studies can be direct tests of the PGT theory, where the theory itself accounts for the inherently stochastic nature of the strategy observations. In this way, by explicitly modeling the researcher's uncertainty regarding which strategy profile will be played, PGT accounts for the systematic risk discussed above.

2. Another advantage of PGT is that, being a fully statistical model, it can combine multiple types of information / data into an associated posterior. This ability is necessary to properly express the uncertainty the game theoretician still has about the strategy profile after all that information. As an example, say the game theoretician is uncertain about the utility functions, so that $\mathscr{I}$ is a distribution over possible utility functions. (Note that the game theoretician may have such uncertainty about the players utility functions even for a complete information game, where the players have no such uncertainty about one another's utility functions.)

[17] One notable exception is the epsilon equilibrium concept of Radner (1980)

Then the proper way for the game theoretician to express her associated uncertainty over mixed strategy profiles is by averaging over that distribution.

As a simple illustration, suppose m is the probability that the utility functions are $\mathscr{U}'$, and $1 - m$ the probability that they are instead $\mathscr{U}''$. Then PGT says we must average over those two sets of utility information to properly express game theoretician uncertainty. Formally, we write $\mathscr{I} = \{\mathscr{U}', \mathscr{U}''\}$ and break the posterior into two terms:

$$P(q \mid \mathscr{I}) = mP(q \mid \mathscr{U}') + (1 - m)P(q \mid \mathscr{U}'').$$

In contrast, with set-valued solution concepts, trying to address uncertainty about the utility functions in a similar fashion would entail averaging over the associated equilibrium sets somehow. It is not at all clear that the axiomatic foundations of equilibrium concepts provide a principled way of doing such averaging.

Furthermore, often we will have types of information that are relevant to our prediction of the player profile but that do not directly concern the game specification. Examples of such information are demographic data, observational data concerning a particular player's idiosyncrasies (e.g., in the form of a Bayes net stochastic model of that player's behavior in the absence of utility functions), and empirical data about the relative probabilities of various focal points. Again, a statistical approach like PGT is necessary to use these types of information to refine our prediction in a principled manner. (For example, given a distribution over focal points,one should use it to average the posteriors given each possible focal point, in exact analogy to the average over utility functions described above.) In contrast, there is nothing in set-valued solution concepts that would allow us to incorporate this information in such a principled way. This is why the (possibly huge) benefits of integrating such information into predictions of strategic behavior has been largely unexplored in conventional game theory.

These alternative types of information have the potential of bringing true explanatory power to game theory modeling. However, because it is unclear how one might combine this data with the more traditional utility information in set-valued solution concepts, the usefulness of such data has largely gone unexplored. PGT represents one way to begin exploring the explanatory power of such data.

In fact, in PGT the principled integration of uncertainty extends even to uncertainty about what model of human strategic behavior to use. Just as a game theoretician does not need to make a choice between utility information $\mathscr{I}'$ and $\mathscr{I}''$, she also does not need to make a choice among models that describe player behavior. As an example, she does not need to choose between a posterior $P(q \mid \mathscr{I}')$ motivated by the QRE model (like the posterior analyzed below) and a posterior $P(q \mid \mathscr{I}'')$ motivated by a level-k model. In fact, she *should not* make such a choice. Rather she should average over both posteriors, according to the the relative probabilities that she assigns to the possibilities that each of those two models applies to her particular prediction problem. No such principled averaging is possible with set-valued solution concepts.

B Two Alternative Likelihoods and an Example

B1 N-rationality

Similar to QR-rationality, N-rationality says that the likelihood of q_i given q_{-i} increases as q_i gets closer to a best response. The difference is how we measure the distance to a best response. With N-rationality we borrow from the epsilon equilibrium concept to say that players differentiate between responses according to the payoffs they generate. Therefore, we measure the rationality of q_i given q_{-i} as the normalized distance between the payoff yielded by q_i and the payoff yielded by i's worst response.

Definition B.1. The *N-rationality* of q_i against q_{-i} is the normalized distance from the payoff to q_i to the payoff from i's worst possible response. Alternatively,

$$\eta_i(q) = \frac{\mathbb{E}_q(u^i) - \min_{x_i}[U^i_{q_{-i}}(x_i)]}{\max_{x_i}[U^i_{q_{-i}}(x_i)] - \min_{x_i}[U^i_{q_{-i}}(x_i)]}$$

where $\min_{x_i}[U^i_{q_{-i}}(x_i)]$ is the minimum expected utility achievable by player i when the other players are randomizing according to q_{-i}, and $\max_{x_i}[U^i_{q_{-i}}(x_i)]$ is similarly defined.

The following are is a general form of $\mathscr{L}(\mathscr{I} \mid q)$ based on N-rationality.[18]

$$\mathscr{L}(q) \propto \prod_i \eta_i(q)^{\alpha_i} \tag{B1}$$

Note that this formulation gives a likelihood ratio $\frac{\mathscr{L}(q)}{\mathscr{L}(q')}$ that is invariant to affine transformations of utility. It is also invariant to the deletion of strategies $q'_i \in \Delta_i(X_i)$ except the minimizers and maximizers. It should again be noted that choosing a specific functional form for the N-rationality likelihood is subject to the same considerations as were mentioned with respect to QR-rationality.

When using N-rationality the modeler must be careful that $U^i_{q_{-i}}$ is bounded for every q_{-i}. If one entry of $U^i_{q_{-i}}$ diverges, then N-rationality is not well-defined. Take for example a first price auction with 2 players where $x \in \mathbb{R}^2_+$ is the profile of bids and $v \leq \infty$ is the profile of valuations. If the modeler wants to use N-rationality here, then she cannot specify that $U_i(x_i, x_{-i}) = v_i - x_i$ whenever $x_i > x_j$ for all x. This is because if i is allowed to bid an infinite amount, then the minimum entry of $U^i_{q_j}$ is negative infinity for every "reasonable" q_j (i.e. q_j in which there exists some number $\mathcal{N}$ such that $q_j(n) = 0$ for all $n \geq \mathcal{N}$) and undefined for other q_j.

[18] Another example is

$$\mathscr{L}(q) \propto \prod_i \tanh\left(\alpha_i \eta_i(q)\right)$$

B2 Intelligence

As an alternative to the rationality criteria outlined above, an intelligence criterion is useful in capturing the relative likelihood of coming across good responses in a random search of one's strategy space.

Definition B.2. The *intelligence* of q_i against q_{-i} is the proportion of $q'_i \in \Delta(X_i)$ such that $\mathbb{E}_q(u_i \mid q_i) \geq \mathbb{E}_q(u_i \mid q'_i)$. Alternatively,

$$\xi_i(q) = \int_{q'_i \in \Delta(X_i)} dq'_i f(q_i) I(\mathbb{E}_q(u_i \mid q_i) \geq \mathbb{E}_q(u_i \mid q'_i)) \tag{B2}$$

where $I(a \geq a')$ is the indicator function that returns one if the argument is true and zero if it is false and $f(q_i) = 1$ is the area of the simplex $\Delta(X_i)$ [see Wolpert (2003)].

The intelligence of q is defined as the vectors of intelligences of each q_i against q_{-i} individually. We suggest one approach to estimating intelligence by importance sampling $\Delta(X_i)$ that is general enough to be applied to any game. However, more efficient methods for calculating intelligence in closed form may be available depending on the details of the game in question (see matching pennies example below).

Since it occurs in the associated likelihood function, we will want to estimate the integral B2 to investigate that likelihood. One way to do that is with Monte Carlo estimation. To do this we will choose a sampling density $h(\cdot)$ with full support on $\Delta(X_i)$. In our case, a sufficient condition for obtaining a finite variance estimator [see Geweke (1989)] is that $\frac{1}{h(q_i)}$ is bounded for all $q_i \in \Delta(X_i)$. (Formally, this is true because $\Delta(X_i)$ is compact, $\text{var}_f(I(\cdot))$ is bounded, and because our target density $f(q_i)$ is uniform, it is therefore bounded over $\Delta(X_i)$.)

Having selected a suitable distribution $h(\cdot)$, we can form T *i.i.d.* samples $\{q'_{i,t}\}_{t=1}^T$. The estimate of intelligence is then:

$$\xi_i(q) = \mathbb{E}_f(I(\cdot)) \approx \frac{1}{T} \sum_{t=1}^{T} \frac{I(\mathbb{E}_q(u_i \mid q_i) \geq \mathbb{E}_q(u_i \mid q'_{i,t}))}{h(q_i)}.$$

Repeating the above procedure for each player i yields a vector of player intelligences, $\xi(q)$, where $\xi_i(q)$ is the estimated intelligence of q_i. As usual, we want the likelihood function to assign more weight to q than q' if and only if q is more intelligent than q'.

For example, the intelligence analog to equation B1 is

$$\mathscr{L}(\mathscr{I} \mid q) \propto \prod_i \xi_i(q)^{\alpha_i}. \tag{B3}$$

The likelihood ratios $\frac{\mathscr{L}(q)}{\mathscr{L}(q')}$ for the likelihood in equation B3 are invariant under affine transformations of utility. However, it is clear from the definition of intelligence that the likelihood ratio between q and q' does not remain unchanged when deleting q'' from $\Delta_{\mathcal{X}}$.

Just as the choices of likelihood function for QR-rationality and N-rationality depend on the specifics of the strategic setting, so does the choice of likelihood function for intelligence. Ultimately, the likelihood implies a distribution over intelligence or rationality. For intelligence, this distribution is given by

$$P(\hat{\xi} \mid \mathscr{I}) = \int_{\Delta_{\mathcal{X}}} I(\xi(q) = \hat{\xi}) \mathscr{L}(\mathscr{I} \mid q) dq \tag{B4}$$

where $I(a = a')$ is the indicator function that returns one when the argument is true and zero otherwise. Therefore, changes in the likelihood imply changes in the distribution of intelligence or rationality.

B3 Example: comparing likelihood criteria

The following example illustrates the difference between QR-rationality, N-rationality, and intelligence.

Consider zero-sum matching pennies, where player 1 wants to match and player 2 wants to mismatch. Assume the environment where player 1 randomizes with $q_1 = .25$. Then for any given q_2, the proportion of alternatives $q_2' \in [0, 1]$ that give expected utility less than or equal to q_2 is simply q_2. In other words, when $q_1 = .25$, the intelligence of q_2 is $\xi_2(q) = q_2$. If q_1 increases to $q_1' = .4$, the intelligence of q_2 is still $\xi_i(q) = q_2$.

Now consider QR-rationality in both cases, $q_1 = .25$ and $q_1' = .4$. In the first case, where $q_1 = .25$, $\beta_2(q)$ solves

$$q_2 = \frac{\exp[\beta_2(-.25 + .75)]}{\exp[\beta_2(-.25 + .75)] + \exp[\beta_2(.25 - .75)]}$$

and in the second case, where $q_1' = .4$, $\beta_2(q)$ solves

$$q_2 = \frac{\exp[\beta_2(-.4 + .6)]}{\exp[\beta_2(-.4 + .6)] + \exp[\beta_2(.4 - .6)]}$$

Now consider N-rationality in both cases. In the first case, where $q_1 = .25$ we have

$$\eta_2(.25, q_2) = \frac{q_2}{2}.$$

In the second case, where $q_1' = .4$, we have

$$\eta_2(.4, q_2) = .2q_2.$$

In both cases, $q_1 = .25$ and $q_1' = .4$, intelligence equals $\xi_2(q) = q_2$. However, QR-rationality, $\beta_2(q)$, changes when q_1 changes from .25 to .4. Whether $\beta_2(q)$ increases or decreases depends on the value of q_2. N-rationality also changes when q_1 changes from .25 to .4, but the direction of the change is certain. It decreases.

C Sampling the Posterior

The formal PGT framework introduced in Sec. 2 assumes finite pure strategy spaces X_i. Accordingly, we must modify our Cournot duopoly setting to analyze it with PGT, by replacing each infinite-dimensional space of possible mixed strategy density functions q_i with a finite-dimensional subspace Λ_i. Once we choose the finite-dimensional subspaces Λ_i, we can then use Monte Carlo importance sampling to estimate the associated expectation values $\mathbb{E}(f \mid \mathscr{I})$ for any f of interest to Predictor, by estimating the numerator and denominator terms in equation 3 (see Robert and Casella (2004)). More precisely, say we parameterize elements of Λ_i with vectors λ, writing the associated density functions as $q_i^\lambda(x_i)$. Then following along with equation 3, motivated by the reasoning behind the posterior introduced in Sec. 2 we would write

$$\begin{aligned}
\mathbb{E}(f \mid \mathscr{I}) &= \int f(q^\lambda) P(q^\lambda \mid \mathscr{I}) d\lambda \\
&= \frac{\int f(q^\lambda) P(\mathscr{I} \mid q^\lambda) P(\lambda)\, d\lambda}{\int P(\mathscr{I} \mid q^\lambda) P(\lambda)\, d\lambda} \\
&= \frac{\int f(q^\lambda) \prod_i [\tanh(\beta_i(q^\lambda) - .5) + 1]^\alpha \exp(\delta S(q^\lambda))\, d\lambda}{\int \prod_i [\tanh(\beta_i(q^\lambda) - .5) + 1]^\alpha \exp(\delta S(q^\lambda))\, d\lambda} \qquad \text{(C1)}
\end{aligned}$$

(For pedagogical simplicity, we are assuming that the constant α is the same for all players.)

To carry out this procedure we must choose the players' mixed strategy spaces, $\{\Lambda_i\}$. How should we do that? To guide us, consider how the logit quantal response function became the focus of work on the QRE. The original choice of a logit quantal response function was not motivated by comparing it to other possible quantal response functions to see which best approximated experimental data. Nor was it derived from theoretical considerations. While it was pointed out that the logit response function arises for a Weibell distribution governing utility uncertainty, there was no effort to justify the Weibell distribution from first principles in the context of multi-player games. In short, the choice of a logit distribution was made because it was a reasonable model of real-world single-player choice behavior (and therefore hopefully also of multi-player choice behavior), and because it was computationally tractable.

These kinds of modeling choices, central to using set-valued solution concepts, are also central to modeling with distribution-valued solution concepts. In particular, they mean that we must choose Λ_i's for the players that meets two criteria. First, they must result in computationally tractable estimates of expectations $\mathbb{E}(f \mid \mathscr{I})$. Second, it must be that we would expect the values of the expression in equation C1 for the f's of interest to well-approximate the values given by real-world behavior.

As an example, say we chose Λ_i to be probability distributions over a (finite) discretization of $[0, \bar{x}_i]$, $A_i \equiv \{0, a, 2a, \ldots \bar{x}_i\}$.[19] Formally, choosing this Λ_i means we do not

[19] In fact, this choice for Λ_i is implicitly made in forming the two figures of the previous subsection; all computer code that uses "floating point arithmetic" uses such a discretization.

allow player i to choose any density function over $[0, \bar{x}_i]$, but rather only to choose those density functions given by a normalized weighted sum of square wave functions,

$$\begin{aligned} q_i(x_i) &= \sum_j \lambda_i(j) I(x_i \in [ja, j(a+1)) \\ &\equiv q_i^\lambda(x_i) \end{aligned} \tag{C2}$$

where $\sum_j \lambda_i(j) = \bar{x}_i/a$. If we made this choice for the Λ_i's, we would use it to complete the specification of the numerator and denominator integrals in equation C1, and then use Monte Carlo to estimate those two integrals.

Unfortunately, for a relatively fine discretization (i.e., for $a \ll \bar{x}_i$), the associated Λ_i is a very high-dimensional space. In such a situation, Monte Carlo estimation of either of the two integrals in equation C1 can be prohibitively slow to converge. On the other hand, for a coarse discretization, we might worry that our associated estimates of $\mathbb{E}(f \mid \mathscr{I})$ are poor approximations to real world behavior, since they amount to modeling mixed strategy profiles in terms of broad square waves.

As an alternative, here we parameterized mixed strategy profiles as mixtures of Gaussians, truncated to have no support outside the domain $[0, \bar{x}_A] \times [0, \bar{x}_B]$. So for us, the variable λ occurring in equation C1 is the parameter vector specifying a truncated mixture of Gaussians. Our choice of this Λ_i rather than one given by A_i's amounts to an assumption about real-world behavior: we are assuming that for a fixed number of degrees of freedom M, for most q arising in real world Cournot duopolies, we can form a better fit to that q by using the M degrees of freedom to specify a mixture distribution than we can by using those M degrees of freedom to specify a density in terms of sums of square waves (like in equation C2).

Given the choice of mixture distribution Λ_i's, we formed our estimate of $\mathbb{E}(f \mid \mathscr{I})$ using conventional Monte Carlo importance sampling: We constructed a sampling distribution $H(\lambda)$, and then IID sampled H to generate many λ's, with our estimate of $\mathbb{E}(f \mid \mathscr{I})$ given by averaging the (importance sampling corrections to) the associated values $f(q^\lambda)P(q^\lambda \mid \mathscr{I})$.

The q's are drawn from the sampling distribution $H(q) = H(\rho, \mu, \Sigma)$. Without much information about the space of joint distributions q, it is safest to explore the space of triples (ρ, μ, Σ) uniformly. Hence, each ρ_i is sampled uniformly from the $\mathcal{M}_i$-dimensional simplex, where $\mathcal{M}_i$ is the number of mixture components in q_i. The means, μ_i, are sampled uniformly from the hypercube given by lower and upper bounds μ_{il} and μ_{ih}. Finally, Σ_i^j is the covariance matrix of the j'th component of i's mixture distribution. It is determined by random Jacobi rotations of a diagonal matrix with eigenvalues λ. These eigenvalues are drawn from a uniform distribution with lower bound λ_l and upper bound λ_h. In order to guarantee positive definiteness of Σ_j^i, λ_l is non-negative.

Specifically, to obtain each q, we draw a mixture of truncated multivariate normal distributions for each player,

$$q_i(x_i) = \begin{cases} \sum_{j=1}^{\mathcal{M}_i} \frac{\rho_i^j \phi_i^j(x_i)}{Z_i} & \text{if } B_i \leq x_i \leq L_i \\ 0 & \text{otherwise} \end{cases}$$

where

$$\phi_i^j(x_i) = \frac{1}{2\pi^{\mathcal{D}_i/2} \mid \Sigma_i^j \mid^{.5}} \exp\left[-.5(x_i - \mu_i^j)'(\Sigma_i^j)^{-1}(x_i - \mu_i^j)\right].$$

and

$$Z_i = \int_{L_i}^{B_i} \sum_{j=1}^{\mathcal{M}_i} \rho_i^j \phi_i^j(x_i) dx_i.$$

The constant Z_i normalizes the mixture to the hypercube $[L_i, B_i]$, where L_i is the minimum of i's action and B_i is its maximum. $\mathcal{D}_i$ is the dimensionality of i's mixed strategy vector.

The question of whether to let $\mathcal{M}$ (the vector that gives the number of component distributions in each player's mixture) be fixed or allow it to be determined randomly remains. Aside from the obvious computational issues that arise by extending the dimension of our integral over all possible values of $\mathcal{M}$, there are strong behavioral reasons to fix the number of component distributions. Suppose $\mathcal{M}_i = \bar{\mathcal{M}}$. With $\bar{\mathcal{M}}$ components, a mixture of Gaussians can have any number of peaks less than or equal to $\bar{\mathcal{M}}$. In a behavioral model, it does not seem unreasonable to assign probability zero to situations in which a player has a mixed strategy with many multiple peaks. This restriction contradicts the QRE, which assumes that each q_i can have any number of peaks. However as shown in the results section, restricting the *sampling routine* to single-peaked q's does not rule out the possibility of a multi-modal posterior distribution over x_i's.

The Cournot duopoly in this paper involves only two players each with a one-dimensional move space. Therefore, importance sampling with a uniform proposal distribution is feasible. However, as more players are introduced, and the move spaces increase in dimension, the space of q's grows exponentially. With higher dimensional games, a uniform proposal distribution may not efficiently explore the space of q's. In such a case, it may be more appropriate to select a more targeted proposal distribution or to employ alternative sampling routines such as the Metropolis-Hastings algorithm.

D Estimating the posterior expected value of $f(q)$

Now that we have a method for sampling the posterior, it is possible to form Monte Carlo estimates of statistics that come from the posterior.

Let $q^{\rho,\mu,\sigma}$ be the parameterized mixed strategy profile distribution and $f(q^{\rho,\mu,\sigma})$ be any function of $q^{\rho,\mu,\sigma}$. The posterior expectation of $f(\cdot)$ is then:

$$\begin{aligned}\mathbb{E}_{\rho,\mu,\sigma}[f(q)] &= \int f(q^{\rho,\mu,\sigma}) P(q^{\rho,\mu,\sigma} \mid \mathscr{I}) d\rho d\mu d\sigma \\ &= \int f(q^{\rho,\mu,\sigma}) \frac{V(q^{\rho,\mu,\sigma})}{Z} d\rho d\mu d\sigma\end{aligned} \tag{D1}$$

where

$$V(q^{\rho,\mu,\sigma}) = e^{\alpha S(q^{\rho,\mu,\sigma})} \mathscr{L}(\mathscr{I} \mid q^{\rho,\mu,\sigma})$$

and

$$Z = \int V(q^{\rho,\mu,\sigma}) d\rho d\mu d\sigma$$

is the normalizing constant.

As an example, choose $f(q) = q$. Then $\mathbb{E}_{\rho,\mu,\sigma}(f(q) \mid \mathscr{I}) = \mathbb{E}_{\rho,\mu,\sigma}(q \mid \mathscr{I})$ is the expected mixed strategy profile. Now each mixed strategy profile q is a distribution $P(x \mid q)$. Accordingly, for this choice of f, $\mathbb{E}_{\rho,\mu,\sigma}(f(q) \mid \mathscr{I})$ is just the posterior expected pure strategy profile, $P(x \mid \mathscr{I})$.

We can estimate the numerator integral in equation D1 with T i.i.d. samples $\{\rho(t), \mu(t), \Sigma(t)\}_{t=0}^{T}$ from H. In the usual way with importance sampling Robert and Casella (2004), we write

$$\int f(q^{\rho,\mu,\sigma}) V(q^{\rho,\mu,\sigma}) d\rho d\mu d\sigma \simeq \frac{1}{T} \sum_{t=1}^{T} \frac{f(q^{\rho(t),\mu(t),\sigma(t)}) V(q^{\rho(t),\mu(t),\sigma(t)})}{H(q^{\rho(t),\mu(t),\sigma(t)})}$$

Similarly, we can estimate the denominator integral by

$$\int V(q^{\rho,\mu,\sigma}) d\rho d\mu d\sigma \simeq \frac{1}{T} \sum_{t=1}^{T} \frac{V(q^{\rho(t),\mu(t),\sigma(t)})}{H(q^{\rho(t),\mu(t),\sigma(t)})}.$$

E Large games

For larger games than the ones considered in this paper, it may not be feasible to use importance sampling Monte Carlo to perform the computations, even if one works in the space of mixtures of Gaussians. For such problems more sophisticated computational techniques would be needed. As an example, it might prove necessary to use Markov Chain Monte Carlo (MCMC) techniques, e.g., with the starting points for the Markov random walk being the QRE's of the system for several different values of the logit exponents. (Conceivably, by using MCMC we could even dispense with the mixture of Gaussians parameterization used for the experiments in this paper, and instead use a very fine discretization of the player's pure strategy spaces.) Future work involves investigating such more sophisticated MC techniques.

F Whether the players have previously interacted, and whether they choose pure or mixed strategies

Much of the earliest, pre-Nash work on game theory did not assume that players randomize. It was assumed that a player chooses a pure strategy, not a mixed strategy that is later randomly sampled. This seems a reasonable modeling choice for many real-world scenarios. Indeed, in the field outside of the laboratory, arguably humans choose pure strategies far more often than they purposely randomize. However there are some scenarios where it instead seems reasonable to model the humans as though they do indeed choose mixed strategies that they then sample. (For example, this seems reasonable in soccer where the shooter of a penalty kick must choose where to aim her kick.)

Most current game theory set-valued solution concepts consider the second type of scenario, by presuming the players choose mixed strategies. This is also the presumption underlying the analysis in this paper; each player i directly chooses q_i, and Predictor's uncertainty about which choice they make is given by $P(q_i \mid \mathscr{I})$.

However some more recently explored solution concepts always result in a unique predicted mixed strategy profile q, in contrast to set-valued solution concepts like the NE. For example, this is the case in Level-K reasoning (see Crawford and Iriberri (2007); Costa-Gomes and Crawford (2006)). Although not conventionally interpreted that way, such single-valued concepts can be interpreted as concerning scenarios where each player i chooses a single pure strategy, not a mixed strategy. Under this interpretation each player chooses a unique pure strategy, without any randomization. However Predictor is uncertain about that choice. So we interpret player i's distribution over their pure strategies under the solution concept, q_i, as Predictor's uncertainty about i's pure strategy choice, $P(x_i \mid \mathscr{I})$.[20] In this way a concept like Level-K reasoning can be reconciled with the desiderata forcing us to make predictions using probability theory (and thereby reconciled with PGT), in contrast to a set-valued solution concept like NE which cannot be reconciled with those desiderata.

We will refer to strategic scenarios where players choose mixed strategies as *mixed* scenarios, and to scenarios where players choose pure strategies as *pure* scenarios. Mixed vs. pure is the first major distinction among different strategic scenarios.

Next, note that in Level-K reasoning, cognitive hierarchy (see Camerer et al. (2006)) etc., it is implicitly presumed that the players have not had earlier, personalized (i.e., non-anonymous) interactions with one another. Due to this lack of earlier interactions, each player must choose their strategies based on population averages, or theoretical notions of how their opponents might reason, rather than based on knowledge of their human opponent's idiosyncratic tendencies. Such solution concepts are most appropriate for one-shot games, or for scenarios where the game has been repeated, but play is anonymous.

In contrast, in set-valued solution concepts like the NE, QRE, etc., the players implicitly *do* choose their strategies based on knowledge of one another's idiosyncracies. (This is necessary for them to coordinate in the choice of the same equilibrium out of a set of multiple equilibria.) To pertain to real-world behavior, such solutions concepts implicitly presume some form of personalized interactions among the players before start of play.[21]

This distinction among strategic scenarios based on whether the players have (not) had earlier personalized interactions can be formalized using distribution-valued solution concepts. We use the term *non-interacted* to refer to a distribution-valued solution concept where the posterior over the variable the players jointly choose (be it the pure or

[20] The more conventional interpretation of Level-K reasoning is that each Level-K player i chooses a unique mixed strategy q_i, so that $P(q \mid \mathscr{I})$ is a Dirac delta function, and $P(x_i \mid \mathscr{I})$ is given by marginalizing that delta function.

[21] Although these solution concepts must presume such earlier interactions, they do not model those interactions. The premise is that the details of those earlier interactions can be abstracted away.

mixed strategy profile) is a product distribution. So for example, for non-interacted-mixed (where the players have not previously interacted), we must have $P(q \mid \mathscr{I}) = \prod_i P(q_i \mid \mathscr{I})$. On the other hand, for interacted mixed (the case investigated in this paper), that equality is violated.

In the field, all four kinds of strategic scenario — interacted/mixed, non-interacted/mixed, interacted/pure, and non-interacted/pure — seem to be quite common. In this paper we only use PGT to consider interacted/mixed scenarios, to most directly match the bulk of the literature on set-valued solution concepts. However there are many ways to extend PGT to the other scenarios.

As an example, it is straight-forward to modify the PGT posterior introduced above for interacted/mixed scenarios to consider interacted/pure scenarios. The set of the states of the world in such scenarios is X. A natural prior over X is the uniform prior. We could then use essentially the same likelihood as the one introduced above:

$$\mathscr{L}(\mathscr{I} \mid x) \propto \prod_i [\tanh(\beta_i(x) - .5) + 1]^{\alpha_i} \tag{F1}$$

where $\beta_i(x)$ is just $\beta_i(q)$ for a profile q given by a product of Kronecker delta functions about the components of x.

The resultant posterior $P(x \mid \mathscr{I})$ could be viewed as a kind of a set-valued solution concept over $\Delta_{\mathcal{X}}$, just like the NE or QRE. There are some important advantages of using the PGT distribution $P(x \mid \mathscr{I})$ instead of the NE or QRE though. First, its motivation clarifies that it is appropriate only when the players are choosing *pure* strategies, not mixed strategies. So we know that in scenarios where we expect that players randomize, we should not use this $P(x \mid \mathscr{I})$, but should instead use the posterior $P(q \mid \mathscr{I})$. There is no corresponding sensitivity for what the choice space of the players is in conventional set-valued solution concepts.

More practically, whereas conventional set-valued solution concepts can have multiple equilibria, $P(x \mid \mathscr{I})$ is always unique. Another practical advantage is that in experiments of one-shot games, it is extremely difficult (if not impossible) to directly elicit the mixed strategies of the players.[22] This makes it a fraught exercise to statistically analyze a particular set-valued (mixed strategy) concept using experimental data. In contrast, since the PGT distribution-valued solution concept for interacted/pure scenarios directly predicts the probability of what *pure* strategies the player chooses, we have no such difficulty in analyzing it using experimental data.

[22] Typically one instead measures population averages, or has the players repeat the game, hoping that they do not learn as they do so. Both attempts to deal with the issue are quite problematic (see Starmer (2000)).