How to Use Decision Theory to Choose Among Mechanisms

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

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

August 20, 2009

Abstract

We extend a recently introduced approach to the positive problem of game theory, Predictive Game Theory (PGT Wolpert (2008)). In PGT, modeling a game results in a *probability distribution* over possible behavior profiles. This contrasts with the conventional approach where modeling a game results in an equilibrium *set* of possible behavior profiles. We analyze three PGT models. Two of these are based on the well-known quantal response and epsilon equilibrium concepts, while the third is entirely new to the economics literature. We use a Cournot game to demonstrate how to use our extension of PGT, concentrating on model combination, modeler uncertainty, and mechanism design. In particular, we emphasize how PGT allows a modeler to perform prediction and mechanism design in a manner that is fully consistent with decision theory. We do this even in situations where conventional approaches yield multiple equilibria, an ability that is necessary for a fully decision theoretic mechanism design. Where possible, PGT results are compared against equilibrium set analogs.

We would like to thank George Judge, Julian Jamison, Alan Isaac and audience members at the American University Economics Seminar.

1 Introduction

Predictive Game Theory (PGT), as first described in Wolpert (2008), is the practice of using what is known about a strategic situation, including utility information, player rationality, focal points, symmetry, equality, player honesty, etc., to formulate a probability distribution over all possible behaviors[1]. This distribution reflects the uncertainty that the external *modeler* of the strategic situation has about that situation. It is distinct from the uncertainty of the *participants* in that situation. For example, players in a non-cooperative game might know one another's utility functions, while the modeler does not. Conversely, the modeler might know those utility functions, while the players do not.

Whereas conventional game theory modeling produces a set of possible behaviors, PGT produces a full distribution over possible behaviors. Wolpert suggests a general Bayesian form for such a modeler's probability distribution over all possible behaviors. In this paper, we consider three statistical models that fall under that general Bayesian form. Two of these models are based on the well-known conventional game theory concepts of quantal response equilibrium (QRE) [see McKelvey and Palfrey (1995)] and epsilon equilibrium [see Radner (1980)], while the third comes from the multi-agent systems literature [see Wolpert (2003)].

In this paper we develop these models, demonstrate their properties and discuss computational methods for their implementation. Using the example of a Cournot duopoly, we illustrate how PGT seamlessly incorporates modeler uncertainty about utilities, rationalities or other aspects of the game. This allows PGT to model aspects of strategic situations that conventional approaches cannot consider, e.g., the relative probabilities of multiple Nash Equilibria (NE) in a given game, and therefore the relative probabilities of the social welfare values associated with those NE. We demonstrate how this allows a modeler to use PGT to perform prediction and mechanism design in a manner that is fully consistent with decision theory. We do this even in situations where conventional approaches yield multiple equilibria, an ability that is necessary for a fully decision theoretic mechanism design.

In the remainder of this section we motivate and describe the PGT approach in broad terms. We then present a roadmap of the rest of the paper.

[1] By "all possible behavior" we mean all possible choice profiles. If players can only choose pure strategies, it means all possible pure strategy profiles, while if they can randomize, it means all mixed strategy profiles

1.1 The predictive game theory approach

The traditional game theoretic approach to prediction is to choose some best-response equilibrium concept that is appropriate for the scenario in question (e.g., Nash Equilibrium in Nash (1959), Bayesian Nash Equilibrium in Harsanyi (1967), Markov Perfect Equilibrium in Maskin and Tirole (1988) or Fairness Equilibrium in Rabin (1993)). However, research in behavioral game theory, particularly in empirical studies, suggests that players are boundedly rational [see Arthur (1994); Conlisk (1996)].

With bounded rationality in mind, McKelvey and Palfrey developed the Quantal Response Equilibrium (QRE) to provide predictions based on the concept of better-responses rather than best-responses.

However all such predictions of behavior are point predictions, in that they predict the exact randomization each agent will use, and they assign probability zero to all other randomizations. (In addition to equilibrium concepts like the NE and QRE concepts, this is also true for "strategic thinking" concepts like Level-k-level thinking [see Costa-Gomes and Crawford (2006)] or cognitive hierarchy models [see Camerer et al. (2006)].) The situation is not as extreme with set-valued equilibrium concepts, like the epsilon equilibrium of Radner. However even these concepts do not provide a probability distribution over all possible behavior profiles. In particular, they do not give relative probabilities of the behavior profiles in the equilibrium sets they predict. They also implicitly maintain that all are other behaviors have probability zero, in manifest contradiction to the real world, where *any* behavior can occur with *some* non-zero probability.

If one wishes to predict some characteristic of interest y concerning some physical system, based on some information $\mathscr{I}$ concerning the system, then statistics provides many ways to convert such a $\mathscr{I}$ into a probability distribution over y. *A priori*, there is no reason that this standard approach to predicting the behavior of physical systems is not appropriate when the physical system in question is some human beings playing a game. PGT is just that standard statistical approach to prediction applied to situations where human beings are playing a game. The characteristic of interest, y, can be anything from a mixed strategy profile q to social welfare $w(q)$. In this paper, we generally regard $\mathscr{I}$ as information about player utility functions and about player rationalities.

The Bayesian PGT approach starts with a prior distribution over behavior profiles. This prior quantifies the modeler's beliefs about the relative probabilities of different mixed strategy profiles without regard to the strategic setting in question. As in Wolpert

(2008), here we employ an entropic prior. Based on Shannon's entropy [see Shannon (1948)], this prior favors mixed strategy profiles that contain less information over those that contain more information. It also embodies the principal of insufficient reason [see Mackay (2003), Cover and Thomas (1991)].

The other modeling choice in the Bayesian PGT approach of this paper is to specify the likelihood function. This is where information about the precise strategic environment comes in to the modeling. In this paper we consider three likelihood functions. The first, introduced in Wolpert (2008), is based on a logit quantal response. The second is based on an epsilon equilibrium-like concept. A third likelihood function uses a concept new to the game theory community, that we call "intelligence" Wolpert (2003). All three likelihoods assign relative probabilities to all mixed strategy profiles on the basis of information about player utility functions and about player rationalities. So while PGT is a new positive approach to game theory, PGT modeling relies heavily on concepts that already exist in the game theory literature. (In future work, we intend to consider more experimentally-grounded likelihoods, incorporating concepts like focal points, symmetry, equality and player honesty.)

Given a prior and likelihood, in Bayesian PGT they are combined in the usual way. This specifies the posterior distribution over mixed strategy profiles conditioned on the information about the strategic environment. This posterior is the focus of the PGT analysis, as it is the source of all predictive information.

With each of the likelihood models we discuss, there is no closed form for the posterior distribution. However modelers can apply well-known Monte Carlo techniques such as accept-reject and importance sampling to this posterior, to estimate its important characteristics [see Robert and Casella (2004)]. For example, we demonstrate how to estimate the marginalizations of the posterior distribution over mixed strategies — which is just the posterior distribution over *pure* strategies. We also show how to estimate the associated covariance over pure strategies. Similarly, we show how to estimate the distributions over expected utility profiles and over social welfare.

The information provided by these characteristics of the PGT posterior have no real analog in conventional equilibrium concepts like the NE or QRE. The closest analog with those concepts arises when there is a unique equilibrium. In such cases one can, for example, take the "distribution over expected utility profiles" to be a Dirac delta function about the expected utilities at the equilibrium. However when there are multiple equilibria, one cannot even do this; there is not a meaningful way to assign relative probabilities to expected utilities at the multiple equilibria. Yet without these probabilities it is not possible to perform prediction or mechanism design in a manner that is compatible with decision theory.

For real-world applications of mechanism design, the probabilities returned by a PGT model allow the researcher to answer the basic types of questions that stakeholders need answered. For example:

- “Which mechanism is most efficient?”
- “Which mechanism produces the least variance in efficiency?”
- “What is the probability that mechanism A produces greater welfare than mechanism B?”
- “Which mechanism has a greater probability of producing welfare below x?”

Without a proper statistical model, these questions simply cannot be answered.

Though the PGT posterior contains far more information than does a single point prediction, it can easily provide such a point prediction if desired. One *ad hoc* way to make such a prediction is to return the global maximizer of the posterior distribution, called the Maximum A Posterior (MAP) prediction (assuming that maximizer is unique). However, how to distill a distribution to a point prediction is a choice ultimately made by the modeler external to the game; it is not part of the game specification itself. PGT enables the modeler to conform with decision theory when making this choice. To do so, the modeler first needs to clarify her objective in making the point prediction. In general, this objective can be interpreted as minimizing some real-valued loss function whose arguments are the prediction and the actual outcome. In particular, Savage's decision theoretic axioms say that to make an optimal rational prediction, the modeler must choose the prediction that minimizes expected posterior loss, where the expectation is taken over the PGT posterior [see Savage (1954)]. For example, if the researcher uses an all-or-nothing loss function, then the MAP is the prediction she should make. Alternatively, for a quadratic loss function, the posterior mean is the appropriate prediction. Note that regardless of the loss function, with PGT there is no equilibrium selection problem. In general, the prediction that minimizes expected loss is unique.

There are many advantages of PGT that arise from its statistical nature. In particular, with PGT we can incorporate other statistical information besides a conventional game specification in predicting player strategy profiles, by averaging over that information to form the posterior over strategy profiles. For example, suppose the players know another's utility functions — but the modeler does not know those utility functions. However say that the modeler has data from an experiment that says half of the players have preferences $\mathscr{I}'$ and the other half have preferences $\mathscr{I}''$. Then the modeler can — indeed, should — account for her uncertainty by averaging over these preferences in forming the posterior over mixed strategy profiles. In fact, any relevant information of a statistical nature can be incorporated by averaging over it in the posterior. This includes real-world and experimental data on preferences, focal points, rationality, etc. It even includes information on the relative probabilities of various PGT likelihood models. See Wolpert (2008) for more on the advantages of PGT over equilibrium concepts as a way of modeling player behavior.

1.2 Roadmap of paper

We proceed as follows. First, we formally introduce the precise PGT approach investigated in this paper. This includes a detailed description of three likelihood models and an example that demonstrates how they compare. We briefly discuss the basic properties of these models. Then we introduce our prior distribution based on Shannon entropy.

Next we introduce a simple Cournot duopoly setting. This setting serves as a backdrop for the exposition of key PGT concepts. In particular we generate various PGT distributions for the Cournot setting, including posteriors for pure strategies, profits and social welfare. We show how modeler uncertainty regarding utility information is seamlessly accounted for due to the statistical nature of PGT. We also show that point prediction in the PGT framework is a well-defined decision problem, because PGT formally quantifies uncertainty over states of the world as a probability distribution. This contrasts with the point prediction of conventional equilibrium-based approaches which do not provide a distribution over all possible behaviors.

We introduce the benefits of applying PGT to the problem of mechanism design. We show how PGT allows the social planner to formulate the mechanism design problem as an exercise in decision theory — which is not possible in conventional mechanism design based on equilibrium concepts. In this way, for the first time, PGT allows the social planner to be fully rational, in the decision theoretic sense. We demonstrate PGT mechanism design in the context of a production tax on the duopoly market and show that the PGT results contrast with QRE and NE.

The final section suggests a long list of future work, including PGT for mechanism design, dynamic games and repeated games as well as games of imperfect information, coalitional games and unstructured bargaining.

The appendix details our computational approach to sampling the posterior. This involves generating random mixtures of Gaussian distributions. We also outline our Monte Carlo procedures for estimating moments of the posterior. Here we include a brief discussion of computational issues, including a density of states phenomenon that arises as the complexity of the game grows.

2 The PGT Model

We are interested in formulating a distribution over the space of mixed strategy profiles. The set of pure actions for player i is X_i. The set of mixed strategies for player i is $\Delta(X_i)$. A generic element of $\Delta(X_i)$ is q_i, a mixed strategy. The set of mixed strategy profiles is $\Delta_{\mathcal{X}} = \times_i \Delta(X_i)$. A generic element of $\Delta_{\mathcal{X}}$ is $q = \prod_i q_i$, a mixed strategy profile.

The central focus of the PGT approach, from which all predictive information is derived, is the posterior distribution, $P(q|\mathscr{I})$, over mixed strategy profiles $q \in \Delta_{\mathcal{X}}$:

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

where $P(q)$ is the prior distribution over mixed strategy profiles, $\mathscr{I}$ is information about utility functions and $\mathscr{L}(\mathscr{I}|q)$ is the likelihood of $\mathscr{I}$ given q.

2.1 Likelihood

The likelihood function, written $\mathscr{L}(\mathscr{I}|q)$, gives greater weight to q's that better coincide with the utility information as determined by some external criteria. In general, that criteria is a modeling choice left to the modeler because it largely reflects concerns specific to the strategic environment of interest. In this paper, we focus on likelihoods that involve quantifications of bounded rationality. Specifically, we develop three models that give greater weight to q's which reflect greater rationality by the players. The first model, QR-rationality, is based on the idea of a logit quantal response. The second model, N-rationality, says that the likelihood of a player choosing a specific q_i when the other players choose q_{-i} depends on how close the corresponding payoff is to the best response payoff. This model is closely related to the epsilon equilibrium concept, as it uses the perfectly rational payoff as a target. The third model, intelligence, 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 payoff than q_i, given q_{-i}. These three models are detailed below.

2.1.1 QR-rationality

In most games it is reasonable to think that players seek to maximize utility. However, there are many reasons why an assumption of perfect rationality might not be suitable for a given situation. Rather, we would like to incorporate some notion of bounded rationality. Our first likelihood model, called QR-rationality (short for quantal response rationality), incorporates bounded rationality by borrowing from the concept of a logit quantal response. Under the logit quantal response, a player's rationality is given by the degree to which that player responds optimally to the other players' strategies. This degree of rationality is the criterion upon which our likelihood differentiates between q's.

Before we formally introduce our measure of QR-rationality, we first need more notation. Let $U^i_{q_{-i}}$ be the vector of expected utilities that player i gets from playing each of his 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

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

where $\mathbf{E}_q(u_i|x_{i,j})$ is player i's expected utility of playing his j'th pure strategy against the mixture q_{-i}. The constant β_i is a measure of i's rationality because as β_i increases, the mixed strategy $\mathbb{L}$ assigns greater probability to those pure strategies of i with greatest expected utility. As shown in McKelvey and Palfrey (1995), as $\beta_i \to \infty$, the logit mixed strategy is a best response to q_{-i}.

So given any q (with finite support), the question is how to calculate β_i for each i. One method of doing so is to find the β_i that minimizes the Kullback-Leibler (KL) distance from q_i to the logit distribution parameterized by β_i. The KL distance is a concept from information theory that is used to measure the difference between two distributions [see

Kullback and Leibler (1951); Kullback (1951, 1987)]. The KL distance is:

$$KL\left(q_i(x_i), \mathbb{L}_{U^i_{q_{-i}},\beta_i}(x_i)\right) = \sum_{x_{i,j}\in X_i} q(x_{i,j}) \ln\left(\frac{q(x_{i,j})}{\mathbb{L}_{U^i_{q_{-i}},\beta_i}(x_i)}\right)$$
$$= \sum_{x_{i,j}\in X_i} q(x_{i,j}) \ln\left(\frac{q(x_{i,j})\sum_{x_{i,l}\in X_i} e^{\beta_i \mathbf{E}_q(u_i|x_{i,l})}}{e^{\beta_i \mathbf{E}_q(u_i|x_{i,j})}}\right). \qquad (2)$$

By minimizing the KL distance from q_i to the logit distrbution parameterized by β_i, we are finding the logit distribution that most accurately models q_i in an information theoretic sense. Then we borrow the common interpretation of β_i as i's rationality when playing $\mathbb{L}_{U^i_{q_{-i}},\beta_i}(x_i)$ in response to q_{-i}. This gives us the following characterization of rationality.

Definition 2.1. The *QR-rationality* of q_i against q_{-i} is the value of β_i that minimizes the KL distance from q_i to $\mathbb{L}_{U^i_{q_{-i}},\beta_i}(x_i)$, equation 2.

One potentially worrisome property of the QR-rationality parameter, that is also shared by the logit-QRE, is that it is not invariant to positive rescalings of utility. In other words, player QR-rationality parameters depend on units.

In the special case where q_{-i} is such that all entries of $U^i_{q_{-i}}$ are identical, the QR-rationality parameter β_i can be any real number. This is the case in a mixed strategy NE with full support, as the following example illustrates.

Example: Consider a mixed strategy NE in which each of two players randomize over 2 pure actions $\{x_1, y_1\}$ for player one and $\{x_2, y_2\}$ for player two with respective probabilities $(q, 1-q)$ and $(p, 1-p)$. Theory states that move conditioned expected utilities must be equal across any pure strategies that receive positive weight in an NE. Hence q might be .0001, yet since the move conditioned expected utilities of x_1 and y_1 are the same, the logit has no choice but to assign them both probability .5, and any β will work for that. So even though (p, q) is a NE, β can be anything. This is because, according to QR-rationality, a mixed strategy NE with full support is at once perfectly rational, irrational, anti-rational and everything in-between $\diamondsuit$

For many settings, the set of q that feature the above problem is of zero measure in $\Delta_{\mathcal{X}}$. This means that when randomly sampling the posterior, as is outlined in the appendix, the above anomaly will not be an issue. However, for completeness we define $\beta_i = \infty$ when q_{-i} is such that all entries of $U^i_{q_{-i}}$ are identical. In other words, when i is indifferent among his pure strategies, he is perfectly rational by default.

The next question is, given the choice QR-rationality to measure how smart players are, what should the functional form of the likelihood $\mathscr{L}(\mathscr{I}|q)$ be? In other words, in the absence of data about the particular human beings playing a game, how strongly do we believe they are likely to be smart, as measured by QR-rationality? One simple parameterized form is the following:

$$\mathscr{L}(\mathscr{I}|q) \propto \prod_i [\tanh(\beta_i(q)) + 1]^{\alpha_i} \tag{3}$$

where each α_i measures how much more likely i is to be smart rather than dumb. Physically, $\mathscr{L}(\mathscr{I}|q)$ quantifies how likely it is that out of all games a set of real-world players could have just played, that they played the game with utility information $\mathscr{I}$, given that they chose joint mixed strategy q when they played that game.

When looking at equation 3, it may be tempting to think of it as a type of averaging of QRE's. This is not the case. Rather, not every $q \in \Delta_{\mathcal{X}}$ is a QRE for properly chosen β. That's because equation 3 is defined for every q, while only an infinitesimal subset of product distributions are logit distributions.

It should be emphasized that 3 is not the only reasonable choice for a QR-rationality likelihood.[2] This is just like in statistics in general (and econometrics in particular). When predicting player behavior in a game, ultimately the modeler 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.

The likelihood in equation 3 produces the following likelihood ratio

$$\frac{\mathscr{L}(\mathscr{I}|q)}{\mathscr{L}(\mathscr{I}|q')} = \frac{\prod_i [\tanh(\beta_i(q)) + 1]^{\alpha_i}}{\prod_i [\tanh(\beta_i(q')) + 1]^{\alpha_i}}.$$

This ratio is unchanged by removing mixed strategy profiles $q'' \neq q$ or q' from the underlying space $\Delta_{\mathcal{X}}$. That is because the quantity $\mathscr{L}(\mathscr{I}|q)$ from equation 3 does not depend

[2]For example, one could use $\mathscr{L}(\mathscr{I}|q)$ is:

$$\mathscr{L}(\mathscr{I}|q) \propto \prod_i g_i(\beta_i(q)) \tag{4}$$

where

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

on the set of possible q's. In addition, because β is not invariant to affine transformations of utility, neither is the likelihood ratio.

Since we are using KL rationality, the choice of the form of the likelihood function has implications for the convergence rate of Monte Carlo estimates of the posterior. This is because the QR-rationality parameter, β_i, can diverge to infinity for q_i that are best responses to q_{-i}. Infinite values of β are unlikely a problem in practice because best response correspondences are often of measure zero in the space of $\Delta(X_i)$. However, large β are quite possible. Therefore, if $\mathscr{L}(\mathscr{I}|q)$ is unbounded as $\beta(q)$ grows, Monte Carlo estimates of the posterior may never converge.

Ultimately, the QR-rationality likelihood describes the underlying distribution of QR-rationalities in the set of players. The true distribution cannot be known with certainty, so any functional form will be wrong. The important point is that a non-degenerate distribution over rationalities is, in many settings, an improvement over an assumption of perfect rationality (as in NE) or a point mass assigned to a specific imperfect rationality (as in QRE). These settings include those for which learning has not yet converged to equilibrium, multiple equilibria exist, or computational complexities are involved (which covers most real-world settings).

2.1.2 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 2.2. 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,

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

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 $\mathscr{L}(\mathscr{I}|q)$ based on N-rationality.[3]

$$\mathscr{L}(q) \propto \prod_i \eta_i(q)^{\alpha_i} \tag{5}$$

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.

2.1.3 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 2.3. The *intelligence* of q_i against q_{-i} is the proportion of $q_i' \in \Delta(X_i)$ such that $\mathbf{E}_q(u_i|q_i) \geq \mathbf{E}_q(u_i|q_i')$. Alternatively,

$$\xi_i(q) = \int_{q_i' \in \Delta(X_i)} dq_i' f(q_i) I(\mathbf{E}_q(u_i|q_i) \geq \mathbf{E}_q(u_i|q_i')) \tag{6}$$

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)].

[3] Another example is

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

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 6 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:

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

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 5 is

$$\mathscr{L}(\mathscr{I}|q) \propto \prod_i \xi_i(q)^{\alpha_i}. \tag{7}$$

The likelihood ratios $\frac{\mathscr{L}(q)}{\mathscr{L}(q')}$ for the likelihood in equation 7 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

$$P(\hat{\xi}|\mathscr{I}) = \int_{\Delta_{\mathcal{X}}} I(\xi(q) = \hat{\xi})\mathscr{L}(\mathscr{I}|q)dq \tag{8}$$

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.

2.1.4 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.

2.2 Prior

The role of the prior distribution, $P(q)$ is to quantify the modeler's subjective beliefs about the relative probabilities of mixed strategy profiles without regard to the utility information used by the likelihood function, $\mathscr{L}(\mathscr{I}|q)$. At first glance, the task of formulating any beliefs about a distribution of mixed strategy profiles without the benefit of utility information may seem difficult and/or unproductive.

However 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 (2003), Cover and Thomas (1991)]. 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 written as $P(q) \propto \exp(\delta S(q))$ for real-valued parameter δ.

For $\delta > 0$, the entropic prior assigns greater probability to mixed strategy profiles that are more diffuse. This is attractive from a modeling perspective because it represents an agnostic way of differentiating between q's that have the same likelihood. More precisely, say we have two mixed strategy profiles, q and q', that have the same QR-rationality, and therefore the same likelihood. With $\delta > 0$, the posterior then favors the mixed strategy profile that has a smaller influence on the distribution over the support of the q's,

$$P(x|\mathscr{I}) = \int_q q(x)P(q|\mathscr{I})dq. \tag{9}$$

Alternatively, setting $\delta < 0$ has an important behavioral interpretation. If, for example, the researcher believes that human beings are particularly poor at randomizing, then such a specification will reflect this by giving greater weight to q's that are "less random." We can call this the anti-entropic prior.

Naturally, the entropic prior and anti-entropic prior can be combined. To be clear, suppose $\delta_e > 0$ is the parameter for the entropic prior, and $\delta_a < 0$ is the parameter for the anti-entropic prior. Then the combined prior is $P(q) \propto k\exp(\delta_e S(q)) + (1-k)\exp(\delta_a S(q))$. In this way the researcher expresses that her prior beliefs involve ambiguity in addition to uncertainty. This ambiguity says that the researcher is uncertain about her prior beliefs. She believes with probability k that her beliefs should be modeled by the maximum entropy principal, and with probability $(1-k)$ that they should be modeled by human beings' general lack of skill in randomizing.

The entropic prior is not the only candidate for prior distribution. Indeed, it just one member of the Cressie-Read family of distributions [see Cressie and Read (1984); Read and Cressie (1988)]. However, we adhere to the entropic prior ($\delta > 0$) as it is consistent with the principle of maximum entropy [see Jaynes (1957)], which can itself be derived from the principal of insufficient reason [see Jaynes (2003)]. The principle of insufficient reason tends that when faced with a set of possibilities that are indistinguishable based on the data at hand, each possibility should be equally likely [see Poincare (1912)]. The entropic prior upholds that principle because it says that for a given $\hat{\beta}$ the mixed strategy profile q that comes closest to putting equal weight on each pure strategy (i.e. maximizes entropy), subject to $\beta(q) = \hat{\beta}$, is the most likely.

3 Cournot Duopoly

Here we introduce a familiar strategic setting, the Cournot duopoly, for the purpose of demonstrating the models and techniques of the previous section. We also use the Cournot setting to compare our models with NE and QRE predictions. Finally, we will discuss mechanism comparison under the PGT approach by introducing an externality to the duopoly market and comparing outcomes under various tax rates.

The Cournot duopoly we use is standard. There are two firms, A and B, that produce goods A and B respectively. They each decide simultaneously 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 and $X = X_A \times X_B$. The price that each firm receives for its own good is determined by market demand. Market demand for firm i's product is decreasing in both x_i and x_j, where $i \neq j$. That is, the price that firm i can charge for good i decreases as the quantity of good i and the quantity of good j increase. We write this price as $D_i(x_i, x_j)$ with the assumptions $D_i(x_i, x_j) \geq 0$, $\frac{\partial D_i}{\partial x_i} \leq 0$ and $\frac{\partial D_i}{\partial x_j} \leq 0$ for all $(x_i, x_j) \in X$. The total cost for either firm i of producing x_i units of good i are given by $C_i(x_i)$ with the assumptions $C_i(x_i) \geq 0$, $\frac{\partial C_i}{\partial x_i} \geq 0$ and $\frac{\partial^2 C_i}{\partial x_i^2} \geq 0$ for all $x_i \in X_i$. Therefore, firm i's profit function is written as $\Pi_i(x_i, x_j) = x_i D_i(x_i, x_j) - C_i(x_i)$.

For illustrative purposes, we study this model for specific parametric forms of $D_i(\cdot, \cdot)$ and $C_i(\cdot)$ for $i = A, B$. The form for $D_i(\cdot, \cdot)$ is

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

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

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

These parametric forms allow us to describe a very broad range of strategic settings while remaining clear and concise in our descriptions. For 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, $x_i^*(x_j)$, in figure 1 below. In this example

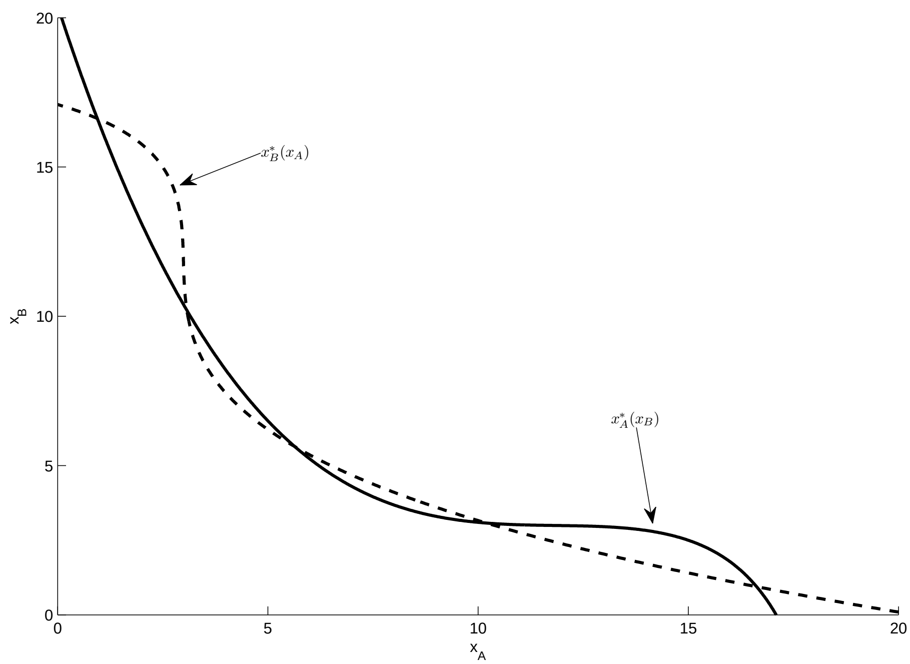

Figure 1: Best response functions for $\bar{x}_i = 20; d_{i1} = 20.4; d_{i2} = 2.165; d_{i3} = 0.12; d_{i4} = 0.0025; c_{i1} = 16,000,000$ for $i = A, B$.

there are five intersections of the best response functions, indicating five pure strategy NE.

Changing the parameter d_{A1} from 20.4 to 19.1 drastically changes the set of equilibria without drastically changing firm A's profit function. This situation is depicted in figure 2.

Finally, a completely new set of parameters $[\bar{x}_i = 9; d_{i1} = 7.1; d_{i2} = 0.8; d_{i3} = 0.15; d_{i4} = 0.0125; c_{i1} = 401.7]$ for $i = A, B$ drastically changes the set of NE and their relative locations. This is depicted in figure 3 below.

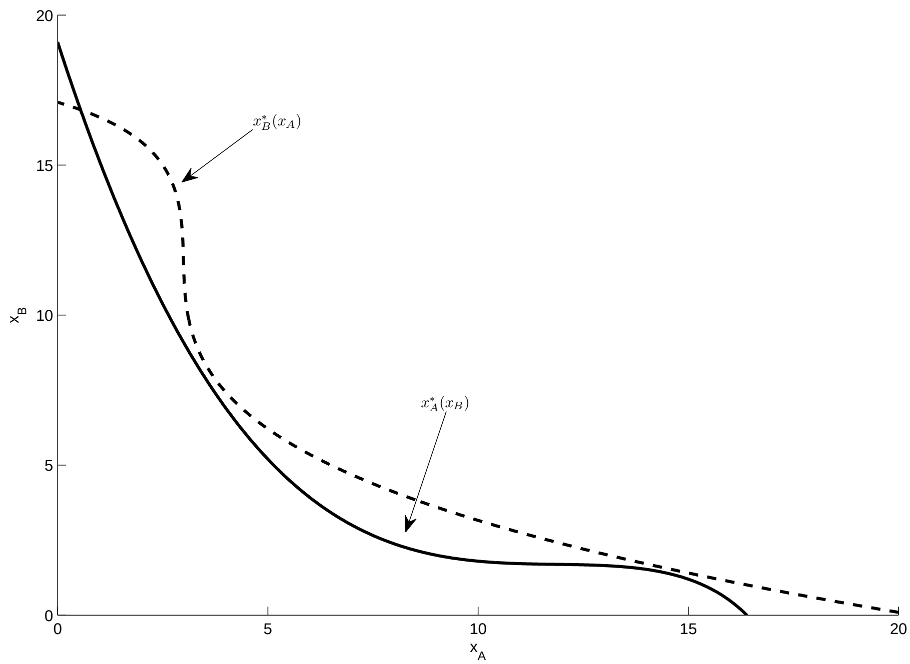

Figure 2: Best response functions for the same parameters as in figure 1, except that $d_{A1} = 19.1$ instead of 20.4.

We use these three strategic settings (i.e., three parameter vectors) as illustrative tools in discussing the results below. We refer to each by its corresponding figure number 1, 2 and 3 respectively.

4 Results

In order to make predictions about the outcome of our Cournot duopoly, we need to know $P(x|\mathscr{I})$ from equation 9, the posterior probabilities of each of the pure strategy profiles. More generally, researchers may want to know the expected value of any function $f(q)$ of the players' strategies

$$\mathbf{E}[f(q)] = \int_{\Delta_{\mathcal{X}}} f(q) P(q|\mathscr{I}) dq. \tag{10}$$

This includes expected profits, expected welfare, expected covariance, etc.

Unfortunately, we cannot evaluate the posterior in closed form for any of the likelihoods discussed in this paper. Therefore, we must numerically estimate it. We use the Monte Carlo method of importance sampling to do so. Importance sampling relies on

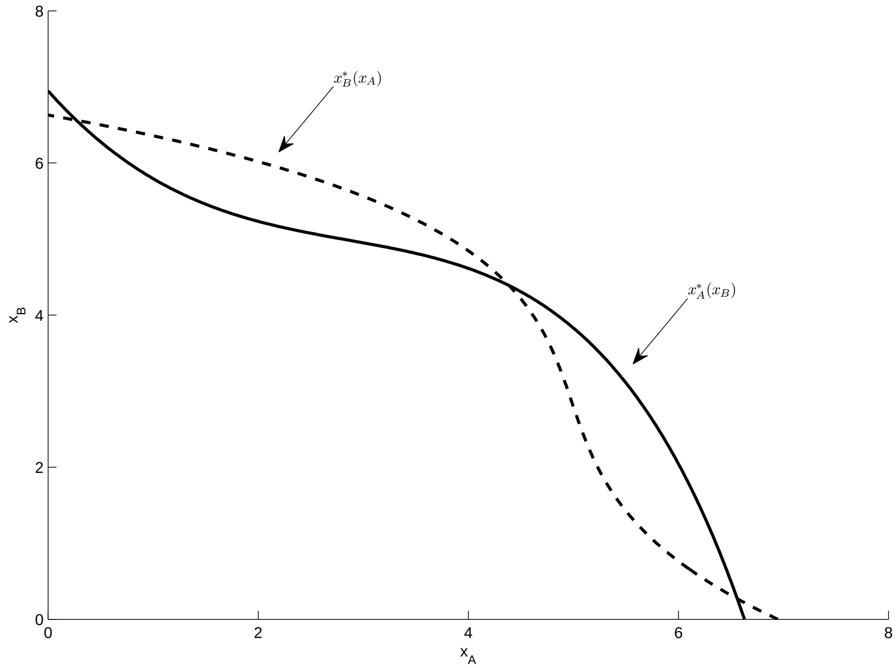

Figure 3: Best response functions for $\bar{x}_i = 9; d_{i1} = 7.1; d_{i2} = 0.8; d_{i3} = 0.15; d_{i4} = 0.0125; c_{i1} = 401.7$ for $i = A, B$.

taking draws from a known distribution $H(q)$ in order to estimate an unknown distribution $P(q|\mathscr{I})$. This means we need a population of mixed strategy profiles, q's, from the space of mixed strategy profiles, $\Delta_{\mathcal{X}}$.

In the Cournot duopoly application, quantities, x_i, can take on any value in the interval $[0, \bar{x}_i]$ for $i = A, B$. So a single q_i is a vector of infinite length. For obvious computational reasons, we cannot work directly with such vectors. Therefore, we need to discretize the space of mixed strategy profiles. That is, the actual value of $\mathbf{E}[f(q)]$ is an integral over an infinite-dimensional space, $\Delta_{\mathcal{X}}$, but we want to estimate this integral over a finite-dimensional space. However, we must be careful to do so in a way such that our estimate of the posterior approximates the actual posterior.

Our solution is to form a population of q's by randomly drawing mixtures of Gaussian distributions. The details of our sampling procedure are given in the appendix. The appendix also describes how we estimate the integral in equation 10.

As described above, $P(x|\mathscr{I})$ gives the posterior density over the space of pure strategy profiles, X. For the Cournot setting in figure 1, $P(x|\mathscr{I})$ is given below in figure 4 for

QR-rationality and figure 5 for N-rationality.[4]

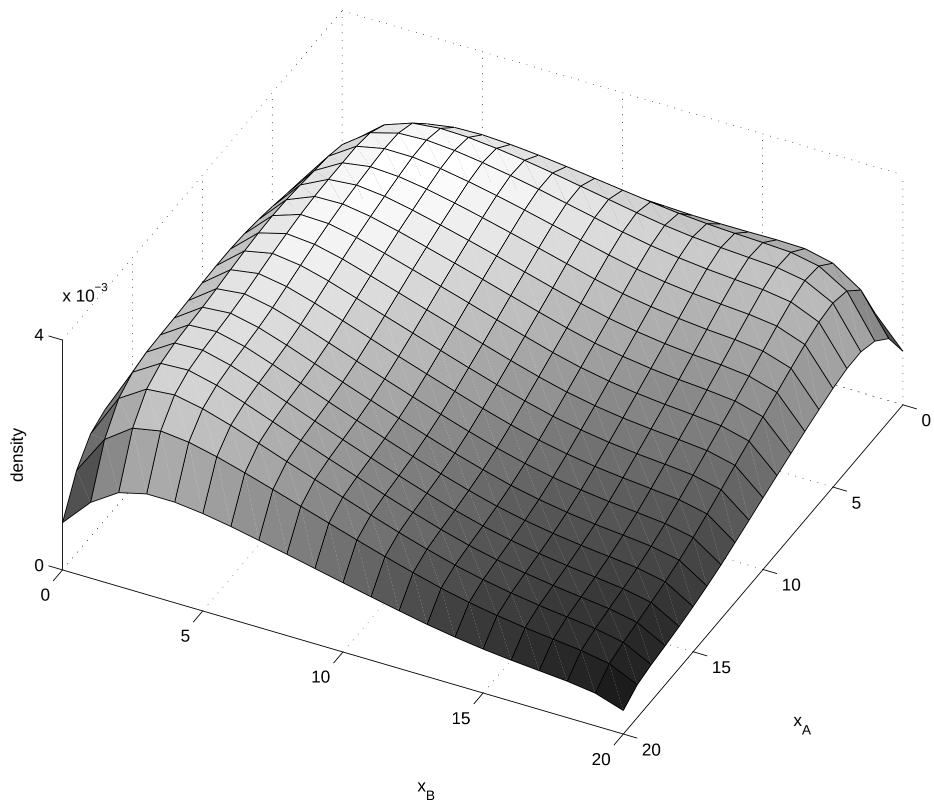

Figure 4: The PGT posterior over x's from the QR-rationality model with parameter $\alpha = 4$. Cournot parameters are $\bar{x}_i = 20; d_{i1} = 20.4; d_{i2} = 2.165; d_{i3} = 0.12; d_{i4} = 0.0025; c_{i1} = 16,000,000$ for $i = A, B$.

The issue of which likelihood to use, QR-rationality, N-rationality or intelligence, is similar to the issue of which equilibrium refinement to use. The answer depends on the setting in question. Ultimately, experiments and real-world data must decide which likelihood is best for the given setting. However, here we are merely illustrating results for some choices of likelihood, in a way that is similar to illustrating results for some choices of equilibrium refinement.

It is important to note that, with PGT, modelers are not constrained to make a hard choice among the likelihoods presented here. Because the PGT approach is fully

[4]Since the Cournot game move space is uncountable, the space of q's is infinite-dimensional. This means that our intelligence measure as discussed above is not well-defined. For that reason, we do not present a distribution over move profiles based on the intelligence likelihood.

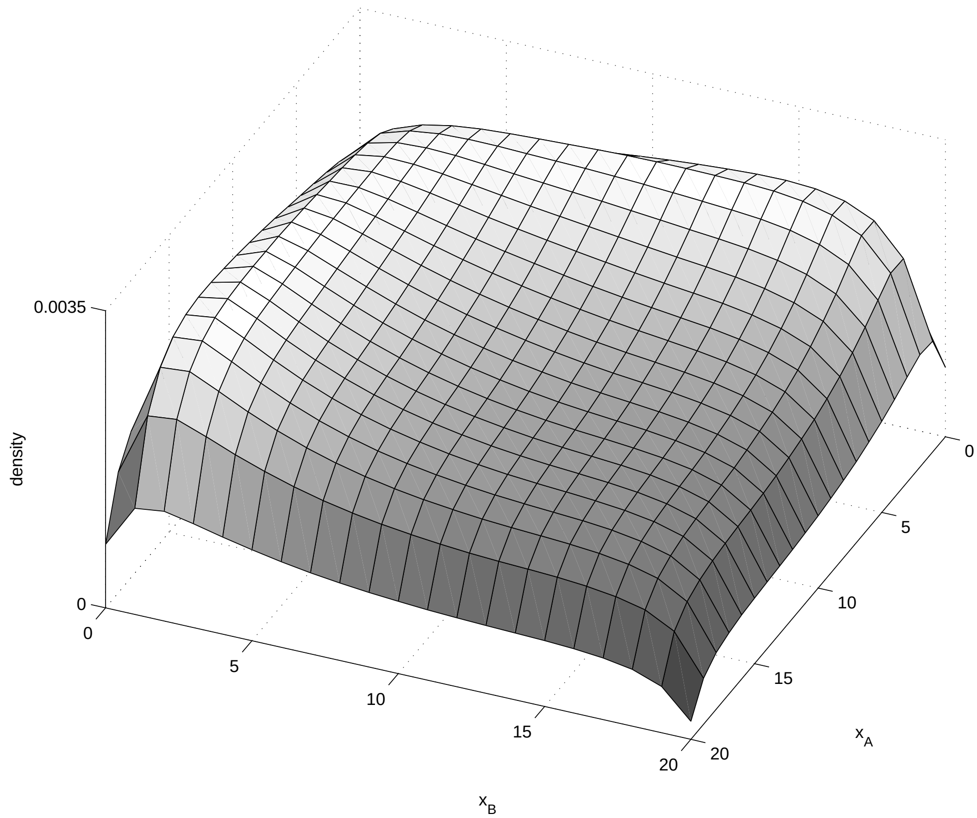

Figure 5: The PGT posterior over x's from the N-rationality model with parameter $\alpha = 15$. Cournot parameters are $\bar{x}_i = 20; d_{i1} = 20.4; d_{i2} = 2.165; d_{i3} = 0.12; d_{i4} = 0.0025; c_{i1} = 16,000,000$ for $i = A, B$.

statistical, it is trivial to combine likelihood models to reflect the modeler's uncertainty about which model is best in a given situation. When such uncertainty exists, the modeler can assign convex weights to each likelihood. The weights represent her beliefs about the relative explanatory power of each model. Then the full model is a weighted combination of the component models.

In the same way that we can express model uncertainty by combining different likelihoods, we can express uncertainty over firm profit functions by averaging over profit functions. To illustrate this, let k be the probability that the profit function parameters are those depicted in figure 1 ($\mathscr{I}'$), and let $1 - k$ be the probability that the profit function parameters are those depicted in figure 2 ($\mathscr{I}''$). Recall that $\mathscr{I}'$ has five NE, and $\mathscr{I}''$ has only one. Loosely speaking, by using PGT we can average those two sets of utility information to capture modeler uncertainty. Formally, we write $\mathscr{I} = \{\mathscr{I}', \mathscr{I}''\}$

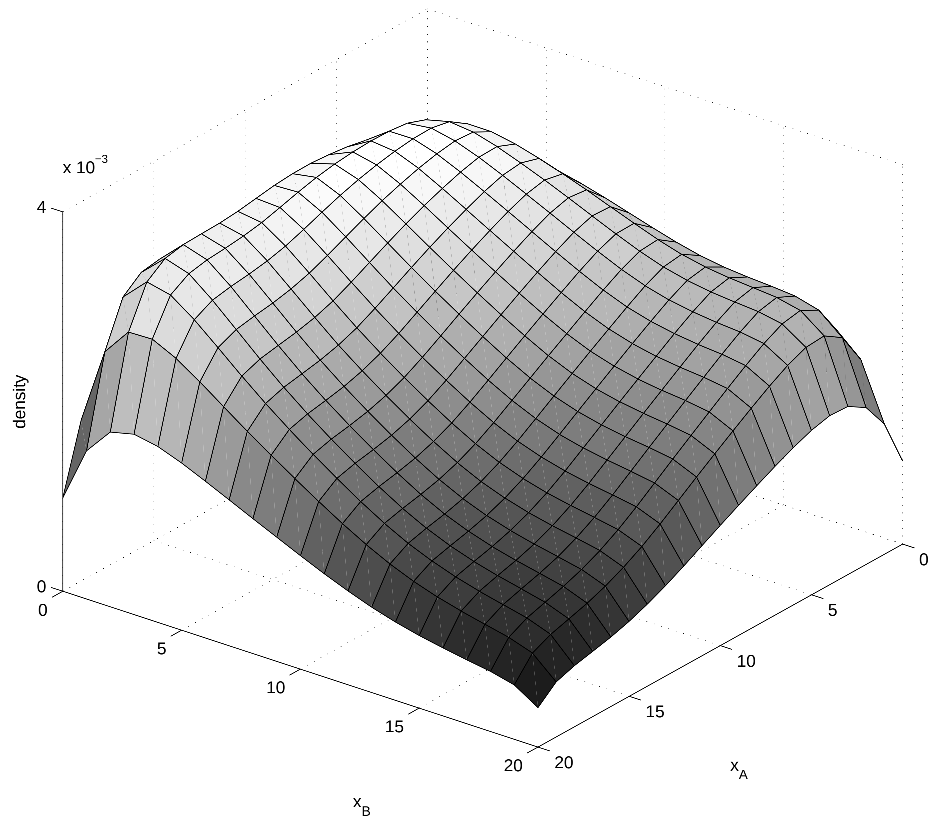

Figure 6: QR-rationality PGT distribution over moves when modeling modeler uncertainty about firm A's profit function ($\alpha = 2.75$).

and break the likelihood into two parts:

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

Figure 6 depicts a combination of $\mathscr{I}'$ and $\mathscr{I}''$ with $k = .5$ under the QR-rationality likelihood.

This averaging out modeler uncertainty over utility functions can be applied to any information concerning utility function values, $\mathscr{I}$. Note that we are concerned here with the *modeler*'s uncertainty of the utility functions. (Since the players have complete information in the Cournot game scenarios we are analyzing, they have no such uncertainty.) It is not clear how one might address such uncertainty using conventional equilibrium concepts. There does not seem to be a statistically meaningful way to combine the five

NE of $\mathscr{I}'$ with the one NE of $\mathscr{I}''$.

4.1 Correlation of pure strategies

Despite the fact that each $q \in \Delta_{\mathcal{X}}$ is a product distribution, $P(x|\mathscr{I})$ is generally not a product distribution. Therefore, there is coupling between x_A and x_B under $P(x|\mathscr{I})$. This coupling of pure strategies is different than the correlation that arises in a correlated equilibrium [see Aumann (1974)], as the result of learning, or as the result of pre-play communication. The coupling between x_A and x_B that arises in $P(x|\mathscr{I})$ is the result of averaging over $P(q|\mathscr{I})$. That is, there is coupling between the players' pure strategies *from the researcher's perspective* because she is 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, i.e. each q is in the space of product distributions.

As an example, consider an industry comprising many firms, where the firms repeatedly play two-player games with one another. Say that a regulator of that industry observes the joint moves of many different pairs of 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 the regulator* it would appear as though there is collusion in the industry.

Consider the duopoly setting from figure 1 with the QR-rationality likelihood where $\alpha = 4$. Our estimate of the correlation between x_A and x_B is small in magnitude, just -0.002, yet it is statistically significant at the 95% level. Changes to the likelihood, such as an increase in α, can increase the magnitude of this correlation. For instance, by setting $\alpha = 100$, we increase the correlation to 0.095. Naturally, the utility information also affects the degree of coupling in $P(x|\mathscr{I})$. For the duopoly setting from figure 2 with $\alpha = 100$, the correlation is -0.062.

Alternatively, a QRE is, by definition, a product distribution. Hence it does not exhibit coupling between x_A and x_B. For comparison with the PGT distributions, we present in figure 7 the QRE distribution for the duopoly setting from figure 1. Recall from equation 8 that the likelihood function implies a distribution over rationality or intelligence. For the PGT distribution from figure 4, we find that the mean of the implied distribution over QR-rationality is approximately 0.2.[5] Therefore, we use $\beta = 0.2$

[5]For more on the distribution over QR-rationality and how it relates to a density of states phenomenon, please see appendix C

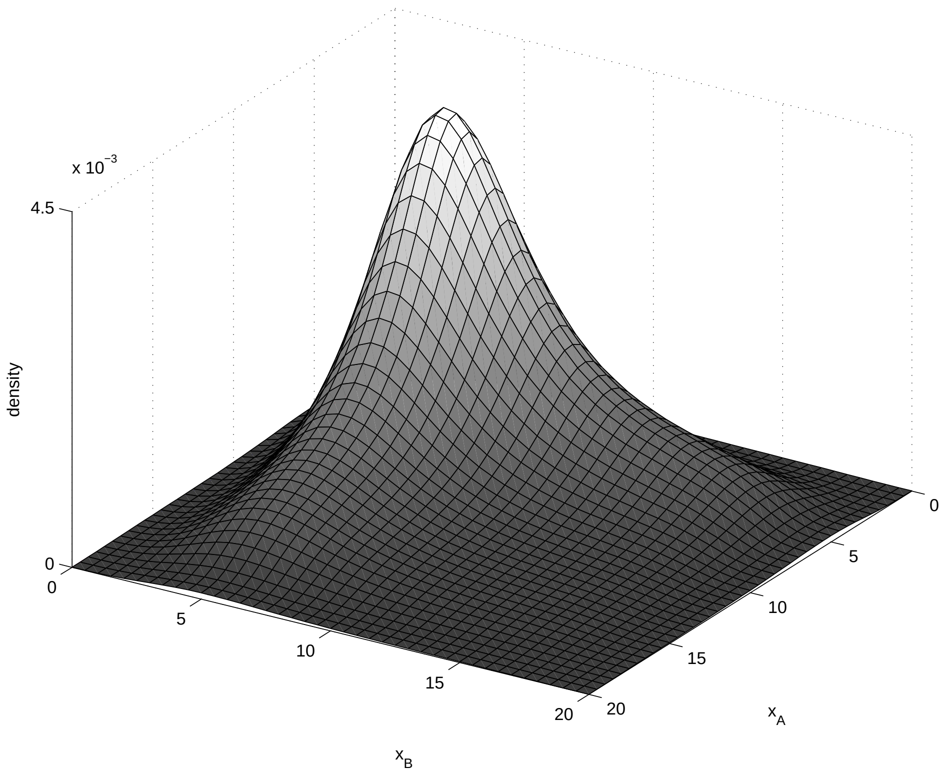

Figure 7: QRE distribution of moves with $\beta = 0.2$.

to generate the comparison QRE distribution in figure 7.

4.2 Predicting outcomes

A point prediction is merely a choice under uncertainty made by a modeler. There is a decision theory for making rational choices under uncertainty that is well-founded on basic axioms of rationality [see von Neumann and Morgenstern (1947); Savage (1954); Luce (1959)]. Game theoretic predictions should also adhere to decision theory if they are to be rational. By providing a probability distribution over behavior, PGT enables the modeler to use decision theory to make rational predictions.

To start, we need those elements of a decision under uncertainty that are prescribed by decision theory:

1. a set of alternatives - the set of mixed strategy profiles

2. a probability distribution over states of the world - the PGT posterior

3. an objective - a quantification of the modeler's preferences (i.e. a loss function)

While the objective and the set of alternatives are often determined by the statement of the research problem, the probability distribution over states of the world is precisely the reason we consult a model. If the model does not produce such a probability distribution, then we cannot properly use it for prediction.

Since the PGT approach yields a distribution over alternatives, q's, we can use it to make a rational prediction. Suppose a modeler wants to predict the quantities (x_A, x_B) that will be played in the Cournot duopoly. Suppose further that this modeler has a loss function. This loss function is not defined as a part of the game in question. Rather it is specified by the modeler to quantify the penalty she suffers for predicting x' when the realized profile is x. There are many ways to quantify this penalty. All-or-nothing loss functions report a zero when $x' = x$ and one otherwise. They are primarily appropriate when the modeler cares only about predicting the exact outcome, and proximity does not matter. In the Cournot duopoly example, quantities can take on any values between zero and $\bar{x}_i$. Therefore, the probability of predicting the exact outcome is very low, suggesting that an all-or-nothing loss function may not be appropriate for the Cournot setting. However, if the modeler does apply an all-or-nothing loss function to the Cournot setting, she should choose the most likely quantity profile. With the PGT posterior, the most likely profile is the maximum a posteriori (or MAP) prediction, and is given by:

$$x^* = \underset{x'}{\operatorname{argmax}} \int_q q(x) P(q|\mathscr{I}) dq = P(x|\mathscr{I}).$$

We can easily apply this loss function to the Cournot duopoly setting depicted in figure 1 or its slight variation in figure 2. Under the QR-rationality likelihood, the MAP from the first setting is approximately $(4.5, 4.5)$, while the MAP prediction from the second setting is approximately $(10, 2)$.

Other loss functions, such as the quadratic loss function, penalize based on the distance between x' and x. Therefore quadratic loss function may be appropriate when a modeler prefers a close prediction to a far-off prediction even when the close prediction is still not quite equal to the realized value. Suppose our loss function is quadratic, $L(x, x') = ||x' - x||^2$, where x' is the predicted profile and x is the realized profile. Then by decision theory, the modeler's prediction should be

$$x^* = \underset{x'}{\operatorname{argmin}} \int_q q(x)||x' - x||^2 P(q|\mathscr{I}) dq.$$

Using QR-rationality, and applying the quadratic loss function to the Cournot duopoly example depicted in figure 1, we predict $x^* \approx (8.97, 8.97)$. For the slight variation in figure 2 we predict $x^* \approx (9.4, 8.6)$.

We can contrast the above predictions with the NE and QRE counterparts. The Cournot duopoly setting from in figure 1 yields a QRE expected quantity profile of approximately $(7.4, 7.4)$ (for $\beta = 0.2$), and a set of pure NE profiles

$$\{(16.6, 0.95), (10.2, 3.1), (5.6, 5.6), (0.95, 16.6), (3.1, 10.2)\}.$$

The QRE prediction is less than the PGT prediction. It is not possible to compare the PGT and NE predictions because there are multiple NE. However, we do note that the total output predicted by the PGT model is greater than the total output under any pure NE.

The example depicted in figure 2 yields a QRE expected quantity profile of approximately $(6.3, 8.2)$ and a unique pure NE profile of $(0.55, 16.8)$. The fact that both equilibrium-based predictions have firm B producing more than firm A is in stark contrast to the PGT predictions, where firm A produces more than firm B. The divergence between PGT predictions and their equilibrium-based counterparts arises because the prior and likelihood assign nonzero weight to more than one q. This is precisely the reason that PGT prediction a well-defined decision problem.

Finally, we note that sometimes the researcher may be interested in predicting mixed strategy profiles rather than pure strategy profiles. This may be the case when the researcher is attempting to choose among several mechanisms to implement and cares about the distribution of pure strategies the players will employ rather than the outcome of any one instance of the game. The mechanics of prediction are the same in both situations. The only difference is that we use $P(q|\mathcal{I})$ instead of $P(x|\mathcal{I})$, and the researcher's loss function must be defined for q's rather than for x's. Otherwise, the researcher minimizes expected posterior loss in precisely the same way.

4.3 Mechanism Design

Given the above, there is no reason that we need to abandon decision theory when deciding among mechanisms to implement. This is one area in which the PGT approach opens a whole new level of analysis.

Consider, for example, the decision that a social planner faces when choosing a tax level for regulating a duopoly market with negative externalities. For simplicity, assume that there are two possible taxes that the social planner has to consider, τ_H and τ_L. Let $w_k(q)$ stand for social welfare (the social planner's objective function) at tax level k. If there are unique NE q^{H*} and q^{L*} under the respective taxes, and the players are known to be fully rational, then using conventional game theory, the social planner decides between τ_H and τ_L by comparing $w_H(q^{H*})$ and $w_L(q^{L*})$.

If there are multiple equilibria, or if the social planner has any uncertainty about the players' rationality (or payoffs), then the social planner simply cannot use such equilibrium approaches to make a rational decision. This is because equilibrium concepts do not provide a probability distribution over the multiple equilibria, so the social planner cannot compute expected welfare as decision theory prescribes.

To choose a mechanism using a PGT model, simply select the mechanism, m, that maximizes expected social welfare over the corresponding posterior. In the context of our duopoly market, consider the scenario represented by the best response functions in figure 3. We model a simple negative externality in this market by assuming external costs equal to $EC(x) = e_1 x$.[6] We also assume that the social welfare function equals firm profits plus tax revenue minus external costs. For a given behavior, q, and tax level k this is:

$$w_m(q) = \mathbb{E}_q[\pi_A + \pi_B] + \mathbb{E}_q[x_A + x_B]\,(\tau_m - e_1)\,.$$

Averaging the social welfare function for each tax level over the posterior, we determine the expected welfare of each tax level. This is written

$$\mathbb{E}_m[w_m(q)] = \int_q w_m(q) P_m(q|\mathscr{I}) dq.$$

Suppose $\tau_H = 4$ and $\tau_L = 2$. Then for our example, $\mathbb{E}[w_L(q)] \approx 6.1$ and $\mathbb{E}[w_H(q)] \approx 5.2$. Without taxes, expected social welfare is $\mathbb{E}[w_0(q)] \approx -0.3$. Hence, the social planner can choose the tax rate that yields greatest expected social welfare, τ_L. Note that,

[6] For more on equilibrium analysis of Cournot efficiency see Seade (1985)

because we are taking an expectation of social welfare over all behavior according to the distribution $P_m(q|\mathscr{I})$, the optimal mechanism may be very different from the mechanism that corresponds to the equilibrium with the highest social welfare. In other words, we may have

$$\mathbb{E}_L[w_L(q)] > \mathbb{E}_H[w_H(q)],$$

even though there are equilibria q^{H*} and q^{L*} such that $w_H(q^{H*}) > w_L(q^{L*})$. The implication is that a conventional approach to mechanism design ignores potentially consequential uncertainty about which strategies will be used.

There is also a subtle implication for risk aversion here. Suppose

$$g(w_m(q)) = (w_m(q))^r = \left(\mathbb{E}_q[\pi_A + \pi_B] + \mathbb{E}_q[x_A + x_B]\,(\tau_m - e_1)\right)^r$$

where $r \in [0, 1]$. This says that if the social planner knows the behavior q^m that each mechanism m will elicit, then she will simply choose the mechanism that maximizes $w_m(q^m)$. However, if she is uncertain about which behavior q^m will be elicited by mechanism m, then she can be risk averse. In other words, even if the social planner is not averse to the risk that a given behavior will produce bad outcomes, she may still be averse to the risk that the mechanism will systematically elicit bad behavior. Modeling the social planner's objective this way 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 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 a lower expectation of $w(q)$ with a tighter distribution rather than a mechanism that produces a higher expectation of $w(q)$ with a broader distribution.

Using PGT and standard decision theory, the social planner fully accounts for uncertainty when choosing mechanisms. She has a complete decision framework. If she finds that a particular PGT model does not incorporate relevant behavioral considerations, such as truth-telling, market exit/entry decisions, etc., then she simply quantifies these considerations and expands the PGT model to include them. In this way, the PGT approach, like all statistical modeling, is completely modular.

Because a PGT model produces a distribution over behavior profiles for each tax scheme, it also produces a distribution over the value of the social planner's objective for each tax scheme. These distributions allow us to fully compare mechanisms by answering basic questions that real-world stakeholders ask. For instance,

- "Which of the taxes has greatest expected social welfare?"
- "Which of the taxes produces greater variance in welfare?"
- "What is the probability that τ_H produces greater welfare than τ_L?"
- "Which of the taxes has a greater probability of producing welfare below some threshold value?"

Such questions simply cannot be answered using conventional approaches. Therefore, when advising a regulator on which tax to implement, we must use PGT.

Using the duopoly setting from figure 3 we get the posterior distributions over social welfare, $w(p)$, from τ_0 (i.e. no tax), τ_L and τ_H. These are displayed in figure 8. There are several features of these distributions to note. First, the high tax level $\tau_H = 4$ produces a strongly bi-modal distribution that leads to the highest variance in social welfare. The high tax level gives both the highest probability of low social welfare (below -5) and the highest probability of high social welfare (above 20). These features of the τ_H distribution of social welfare mean that a risk averse social planner is unlikely to select the high tax level.

The low tax level, $\tau_L = 2$, distribution over social welfare very nearly first-order stochastic dominates the zero tax distribution. Hence, it is unlikely that any social planner, no matter how risk averse, will ever choose zero tax. The low tax level also clearly gives the lowest probability of social welfare less than zero. This obviously means it gives the highest probability of social welfare greater than zero (i.e. loss avoidance). It also gives the highest probability of social welfare greater than 10.

Figure 8 also shows that NE and QRE expected social welfare is greatest for the high tax level, τ_H. However, as mentioned above, PGT expected social welfare is greatest under the low tax level, τ_L. Like the differences between PGT predictions and equilibrium-based predictions, this difference stems from the fact that PGT gives non-zero weight to multiple q's. Therefore, PGT expected social welfare is averaged over these multiple mixed strategy profiles, while with equilibrium-based approaches, expected social welfare comes from a single mixed strategy profile.

Finally note that because the mechanism design problem is couched fully within decision theory, it is straightforward to introduce constraints. For example, the social planner

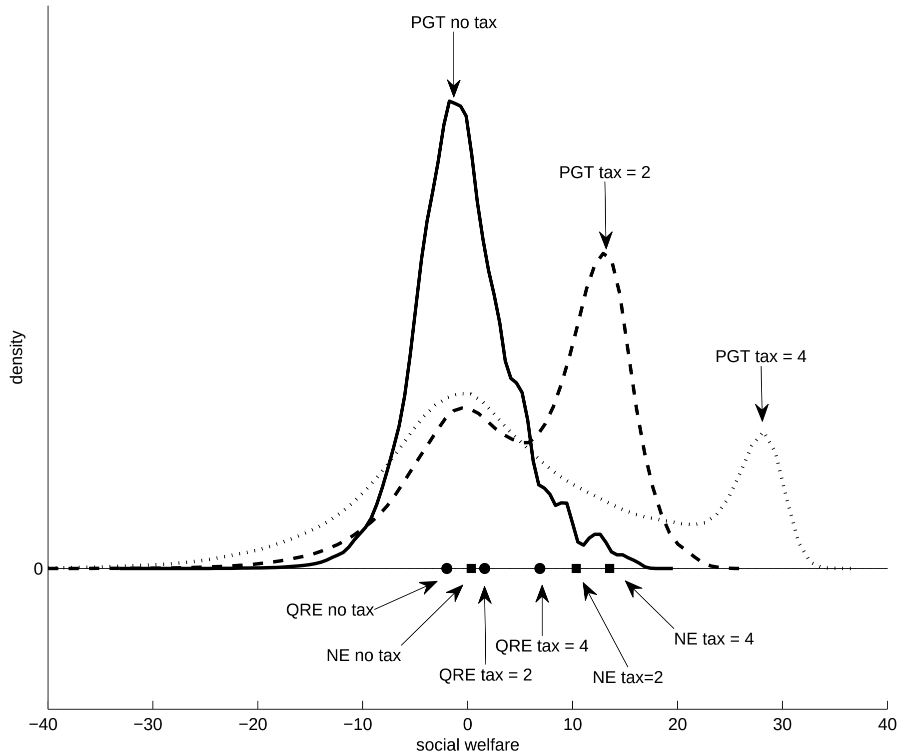

Figure 8: PGT (QR-rationality, $\alpha = .75$), QRE ($\beta = .5$) and NE distributions over expected welfare for tax rates 0, 2 and 4.

may not want to choose a mechanism for which the probability that firms make negative profits is greater than lower bound. To work with such a constraint, we must know the distribution of expected profits. With PGT, the information is readily available, and we display it in figure 9 for the duopoly setting from figure 1 with $\alpha = 4$.

Using the distribution of expected profits, it is straightforward to calculate the probability that the firms will both achieve some minimum profit, the probability that they will together achieve some minimum aggregate profit, or any other quantity of interest. Making these calculations for each mechanism further informs the social planner for the purpose of mechanism comparison.

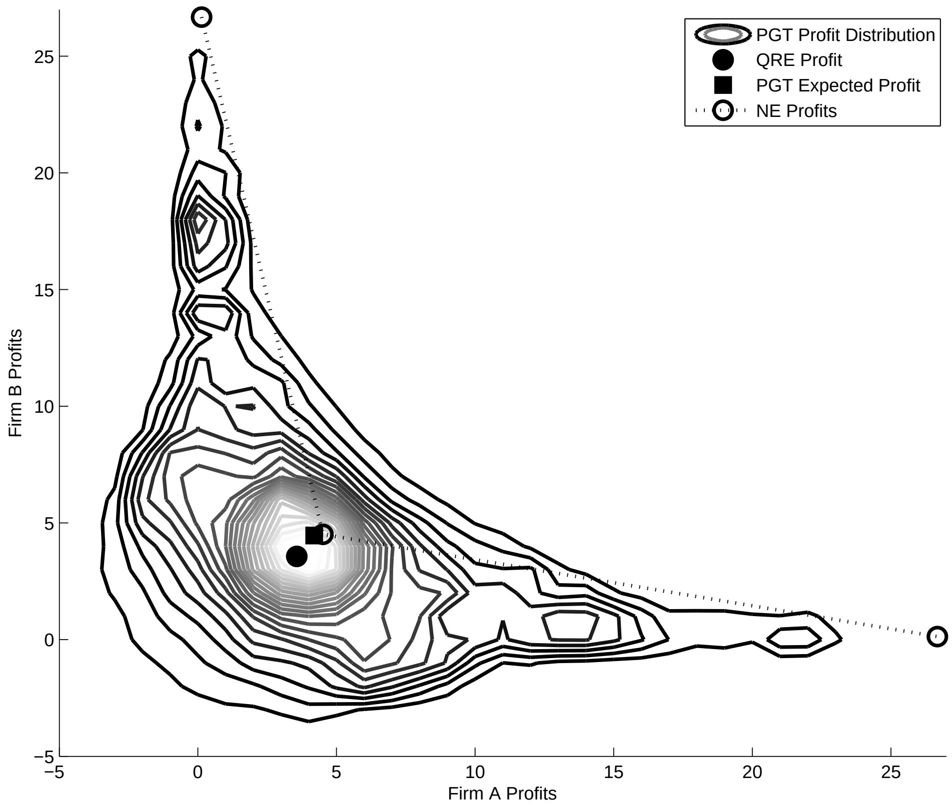

Figure 9: PGT distribution of profits versus NE profits, QRE profits, and PGT expected profits.

5 Future Work

PGT is an important step forward for game theory. By definition, PGT puts a probability distribution on all states of the world. Therefore, when the goal is prediction and/or mechanism design, PGT is the only approach that is compatible with decision theory. It allows for any degree of uncertainty with respect to payoff functions and model specification. It also seamlessly incorporates the important and growing body of data from the behavioral economics literature. In many cases, PGT models can be formulated to recover a statistical analog of conventional equilibrium concepts (i.e. likelihood assigns zero weight to non-NE q's).

There is a long list of additional issues yet to be fully addressed. In this paper we focused on the simplest formulation of PGT, a one-shot simultaneous-move game. However, the PGT approach translates directly to repeated and dynamic games. Future work should develop models for such situations, potentially allowing for learning. To be successful, such work must also focus on computational issues. Of primary importance is a "density of states phenomenon" that arises in the space of product distributions. In particular, as $\Delta_{\mathcal{X}}$ becomes larger, highly intelligent and/or rational q's are drawn less frequently from the proposal distribution, $H(\rho, \mu, \Sigma)$, that we describe in the appendix. Dynamic games naturally increase the dimensionality of the players' strategy spaces and therefore the dimensionality of $\Delta_{\mathcal{X}}$. As the dimensionality increases linearly, the convergence rates of our Monte Carlo estimations slow exponentially.

Although we described the way in which modelers can incorporate their own uncertainty regarding player payoffs into PGT, we avoided modeling each player's uncertainty about her opponents' payoff functions. This will be an important issue going forward, as most real-world settings involve such uncertainty. To predict behavior and choose mechanisms in such settings, we must have a PGT model that treats this uncertainty directly.

The PGT approach should also be adapted to coalitional and unstructured bargaining situations.[7] Like noncooperative games, conventional coalitional and unstructured bargaining models have focused on analytic or set-valued solutions rather than prediction. Therefore, they share the same shortcomings in terms of prediction and mechanism design. We anticipate that advances in this direction will borrow from existing ideas, such as the core and Shapley Value [see Aumann (1961); Shapley (1953)], in much the same way the analysis of this paper borrowed from QRE and epsilon-equilibrium.

Finally, the mechanism design problem discussed above was relatively simple because there were only three possible taxation levels.[8] In general, the problem will be to describe some set of variables that fall under the control of the social planner and affect the strategic environment. Then mechanism design is simply searching for the values of those variables that maximize the expectation of the social planner's objective over the PGT distribution induced by those variables. For example, if the social planner can choose m from some set M, then the PGT approach to mechanism design is simply

$$\underset{m \in M}{\operatorname{argmax}} \, \mathbb{E}_m[w_m(q)] = \int_q w_m(q) P_m(q|\mathscr{I}) dq.$$

[7] Predictive Coalitional Theory (PCT) and Predictive Unstructured Bargaining (PUB) respectively.
[8] Predictive Mechanism Design (PMD)

With most PGT models, we will not be able to solve for the gradient of expected social welfare. Therefore, successful work in this area will likely borrow Monte Carlo optimization techniques from statistics and control theory.

Although PGT is a way forward for game theory, it will always rely heavily on existing concepts (i.e. NE, QRE, level-k, behavioral economics, etc.) and problem domains (dynamic games, learning, incomplete information, etc.). These have proven to provide valuable insight into strategic human interaction, and are indispensable for the pursuit of better statistical models.

References

Arthur, W. Brian, "Inductive Reasoning and Bounded Rationality," *The American Economic Review*, 1994, *84* (2), 406–411.

Aumann, R.J., "The Core of a Cooperative Game without Side-Payments," *Trans American Math Society*, 1961, *98* (3), 539–552.

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

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

Conlisk, John, "Why Bounded Rationality?," *Journal of Economic Literature*, 1996, *34* (2), 669–700.

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.

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

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.

Harsanyi, John, "Games with Incomplete Information Played by Bayesian Players, I-III. Part I. The Basic Model," *Management Science*, 1967, *14* (3). Theory Series.

Jaynes, Edwin Thompson, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003.

Jaynes, E.T., "Information theory and statistical mechanics," *Phys. Rev.*, 1957, *106*, 620.

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.

Luce, D.R., *Individual Choice Behavior*, John Wiley & Sons, Inc., 1959.

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

Maskin, Eric and Jean Tirole, "A Theory of Dynamic Oligopoly: I & II," *Econometrica*, 1988, *56* (3), 549–600.

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

Nash, John F., "Equilibrium Points in N-person Games," *Proceedings of the National Academy of Sciences*, 1959, *36*, 48–49.

Poincare, H., *Calcul des probabilits* 1912. Reprinted Jacque Gabay, 1987.

Rabin, Matthew, "Incorporating Fairness Into Game Theory and Economics," *The American Economic Review*, 1993, *83*, 1281–1302.

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.

Savage, Leonard J., *Foundations of Statistics*, New York: John Wiley, 1954.

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.

Shapley, Lloyd S., "A Value for n-person Games," in H.W. Kuhn and A.W. Tucker, eds., *Contributions to the Theory of Games II*, Princeton University Press 1953.

von Neumann, John and Oskar Morgenstern, *Theory of games and economic behavior*, 2nd ed., Princeton, NJ: Princeton University Press, 1947. (Original work published 1944).

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, David H., "Statistical Prediction of the Outcome of a Noncooperative Game," October 2008. submitted to Journal of Artificial Intelligence Research.

A Sampling the Posterior

Our solution for sampling from $P(q|\mathscr{I})$ is to form a population of q's by randomly drawing mixtures of Gaussian distributions. 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

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

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.

B Estimating Statistics

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:

$$\begin{aligned} \mathbf{E}_{\rho,\mu,\sigma}[f(q)] &= \int_{\rho,\mu,\sigma} f(q^{\rho,\mu,\sigma}) P(q^{\rho,\mu,\sigma}|\mathscr{I}) d\rho d\mu d\sigma \\ &= \int_{\rho,\mu,\sigma} f(q^{\rho,\mu,\sigma}) \frac{V(q^{\rho,\mu,\sigma})}{Z} d\rho d\mu d\sigma \end{aligned} \tag{11}$$

where

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

and

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

is the normalizing constant.

As an example, choose $f(q) = q$. Then $\mathbf{E}_{\rho,\mu,\sigma}(f(q) \mid \mathscr{I}) = \mathbf{E}_{\rho,\mu,\sigma}(q \mid \mathscr{I})$ 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, $\mathbf{E}_{\rho,\mu,\sigma}(f(q) \mid \mathscr{I})$ is just the posterior expected pure strategy profile, $P(x \mid \mathscr{I})$.

We can estimate the numerator integral in equation 11 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

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

Similarly, we can estimate the denominator integral by

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

C Computational Issues

Here we briefly describe two computational issues that arise in PGT modeling. The first concerns a "density of states phenomenon" that arises as the complexity of the game grows. The second concerns the choice of QR-rationality likelihood so that Monte Carlo estimates converge.

C.1 Density of states

In many cases it will be very difficult to have any idea what the space of mixed strategy profiles, $\Delta_{\mathcal{X}}$, looks like. In particular, it will be difficult to know how to efficiently sample this space so that we draw with high probability the types of q's that get high probability under $P(q|\mathscr{I})$. Therefore, we resort to a proposal distribution $H(\rho, \mu, \Sigma)$ that is roughly uniform over the set of mixtures of Gaussians. This can be very inefficient. In addition, as the complexity of the game in question grows, the inefficiency of a uniform proposal distribution grows. In the Cournot setting from figure 1, we drew 120,000 q's from $H(\rho, \mu, \Sigma)$. The histogram of these rationalities for firm A is given below:

Note that the mass is tightly packed around zero, which represents complete non-rationality. The density drops off quickly when moving away from zero in either direction. Because our draws of q_A and q_B are independent under $H(\cdot)$ for each q, the joint distribution of (β_A, β_B) is the product distribution. This means that while we have a very low probability of drawing a high rationality q for *one* firm, the probability of drawing a high rationality q for *both* firms is far lower still.

As the number of players increases, the density of states problem gets worse. It

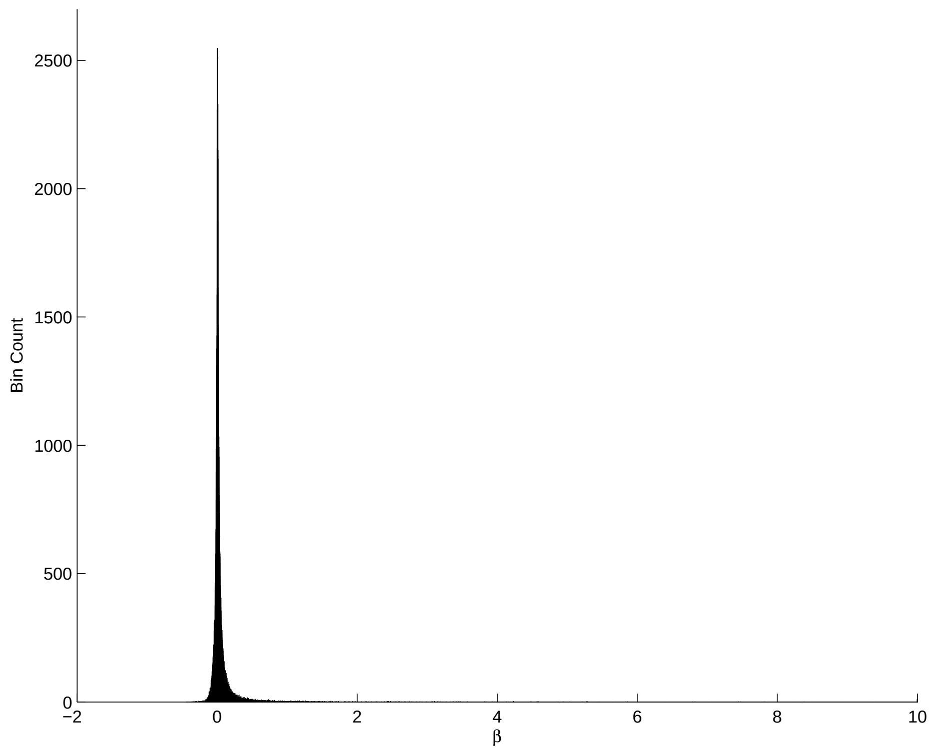

Figure 10: Histogram of firm A's QR-rationality parameters, β_A, from 120k random draws of q for Cournot setting from figure 1.

also gets worse as the dimensionality of each individual's strategy space increases. For example, if each firm were to choose a quantity *and* a price, then the associated histogram for QR-rationality would be much tighter around zero than even the histogram in figure 10.

For the intelligence criterion, the density of states means that most q's yield an intelligence of about 0.5. That is, given q_{-i}, most q_i's are better than roughly half of the $q'_i \in \Delta(X_i)$.

C.2 Bounded likelihood

The QR-rationality criterion is unbounded. The parameter β can vary from $-\infty$ to ∞. Therefore, if our Monte Carlo estimates of the posterior and its moments are to converge, then we must worry about the specific form of the likelihood function.

In particular, as we established in the discussion of density of states above, the probability of drawing a q with high QR-rationality under the proposal distribution $H(\cdot)$ can be vanishingly small. So if the likelihood is not bounded above for large β, then the ratio

$$\frac{\mathscr{L}(\mathscr{I}|q^{\rho(t),\mu(t),\sigma(t)})}{H(q^{\rho(t),\mu(t),\sigma(t)})}$$

will diverge for q that give rise to large β. This leads our Monte Carlo estimator to have infinite variance [see Robert and Casella (2004)].