Level-n Bounded Rationality and Dominated Strategies in Normal-Form Games

by

Dale O. Stahl
Department of Economics
University of Texas at Austin

and

Ernan Haruvy
Department of Management
University of Texas at Dallas

April 11, 2003

Abstract

Dominated strategies play a crucial role in game theory and its solution concepts. While empirical studies confirm that humans generally avoid dominated strategies, they also suggest that humans seldom believe others will avoid such strategies. Hence, the iterated dominance solution is not likely to be a good predictor of one-shot behavior. We investigate how the salience of a dominated strategy affects the extent to which players believe that others will recognize and avoid it. Level-n theory serves as a useful tool in this empirical investigation, as it is able to classify behavior into levels of bounded rationality and provide clear statistical tests for model comparisons. We find that even the most obviously dominated strategies do not induce consistently significant behavioral differences in a variety of one-shot games. Nevertheless, the fit of the level-n model can be improved by hypothesizing that the level-0 choices and level-1 beliefs are tilted slightly away from the uniform distribution to the extent that the average payoff of a strategy falls below a threshold.

1. Introduction.

Dominated strategies have posed problems for game theory. No rational person should choose a strictly dominated strategy, so if rationality is common knowledge, then no rational person should choose an iterated strictly dominated strategy (e.g. Tan and Werlang, 1988). However, in the laboratory, the solution concept of iterated dominance has consistently been a poor predictor of behavior. Experimental investigations find that subjects generally avoid dominated strategies, but seldom iteratively eliminate dominated strategies. This has been shown in dominance solvable beauty contest games (Stahl, 1995; Nagel, 1995), in two-iteration dominance-solvable symmetric normal-form games (Stahl and Wilson, 1994,1995), in two- and three- iteration dominance solvable games (Costa-Gomes et al, 2001), and in nine-iteration dominance solvable games by Sefton and Yavas (1996) and Katok, Sefton, and Yavas (2002).

This paper explores how robust this empirical finding is in symmetric normal form games. We are motivated by the intuition that some dominated strategies are less obvious than others. For example, when a pure strategy is dominated only by a mixed strategy, it may be harder to detect. Similarly, when other strategies are perceived to be risky in the sense that their maximin payoff level is well below some of the payoffs of the dominated strategy, the dominated strategy may appear less obvious. In these cases, it is not surprising that humans may have doubts as to whether all other human players recognize and avoid the dominated strategy. On the other hand, some dominated strategies can be so obvious that no one should have doubts about others' ability to recognize and avoid them. By manipulating the salience of dominated strategies in an experiment, it may be possible to refine our understanding of when iterated dominance is a good predictor of behavior and when it is not.

To assess the effect of obviously dominated strategies on behavior in symmetric normal form games, we choose as a benchmark model the Stahl-Wilson (1995; hereafter SW95) level-n theory of bounded rationality. The theory and its extensions have been reasonably robust in a wide class of symmetric normal form games (Stahl-Wilson, 1995; Haruvy and Stahl, 1999; Haruvy, Stahl, and Wilson, 2001).[1] As we show here, the theory is not necessarily robust when games have "obviously" dominated strategies. To see this, recall that a *level-1* type is assumed

---

[1] Robustness of hierachical bounded rationality in various experimental settings has also been shown by Nagel, 1995; Duffy and Nagel, 1997; Sonsino, Erev, and Gilat, 1999; Costa-Gomes, Crawford and Broseta, 2001; Ho, Camerer, and Wiegelt, 1998; and Gneezy, 2002.

to believe that the other players are equally likely to choose among the available strategies, and to choose a logit best-reply to that belief. However, suppose one strategy profile is "obviously dominated" as in the following symmetric game (payoffs for row player):

	A	**B**	**C**
A	0	0	0
B	10	90	10
C	100	5	20

Surely strategy A is obviously so bad that no reasonable player would believe it is as likely to be chosen as B or C. Nonetheless, the level-1 type believes A is as likely and, therefore, is predicted to be most likely to choose C. But if A were first eliminated, leaving the 2×2 game with strategies B and C, then a level-1 type would be predicted to choose B. This reasoning strongly suggests that the level-n theory of bounded rationality should be modified for games with obviously dominated strategies. While the SW95 experiments included games with strictly dominated strategies, none were as obvious as the above example.

This modification of level-n theory is in keeping with game theory practice, where we routinely ignore hundreds of possible strategies in real-world situations in order to focus on the most strategically relevant aspects. Though the principle of ignoring obviously dominated strategies is hardly controversial, the extant literature has failed to identify clear empirical criteria for when a strategy is so obviously dominated as to warrant deleting it from consideration before further analysis of the game.

While strategy A in the above game might be obviously dominated even to unsophisticated players, consider modifying the payoffs of row A to (0, 0, x). Strategy A is still strictly dominated for all $x < 20$, but it is obvious for all $x < 20$? Noting that 10 is the maximin payoff of the game, perhaps strategy A is obviously dominated when $x < 10$, but not when $x > 10$. These are empirical issues that will be addressed in this paper.

Our benchmark data comes from three different experiments involving 47 symmetric 3×3 games, many with dominated strategies. We focus on models of the frequency of choices in the population of 155 experiment participants, rather than individual choices because the theory of one-shot complete information games rests fundamentally on beliefs about the population of other anonymous players. The experiments and data are described in Section 2. We present a nested sequence of bounded rational models in Section 3, beginning with a trimodal model of boundedly rational behavior, and culminating with a quadrimodal eight-parameter model that fits the data remarkably well.

A fourth experiment was designed to test how well the above fitted model predicts on games with strictly dominated strategies of varying degrees of salience. We find that even the most obviously dominated strategies do not induce major deviations from the level-n model of behavior in one-shot games. That is, though most players appear adept at avoiding dominated strategies, they still believe that others are likely to choose dominated strategies, even when such strategies are obviously dominated. Nevertheless, the fit of the level-n model can be improved by hypothesizing that level-0 choices and level-1 beliefs are tilted slightly away from the uniform distribution when the average payoff of a strategy falls below a threshold. Further, we reject the hypothesis that a dominated strategy is obviously dominated to the extent that its maximum payoff falls short of the maximin payoff of the game.

2. The Experiments and Data.

All of the experiments were conducted in a computer laboratory using a computer interface that presented the participants the row payoffs of a symmetric game (one game per screen), and provided a calculator that would compute the hypothetical payoffs to any hypothesis about the other's choices the participant cared to enter. The payoffs of each game were in terms of binary lotteries with a range of [0, 100] percent probability of winning a monetary prize. The first of these experiments is reported in Haruvy, Stahl and Wilson (2001) and entails 15 symmetric 3×3 games and 58 participants. While some of these games had dominated strategies, none had multiple pure-strategy Nash equilibria - see Appendix A. The second experiment is reported in Haruvy and Stahl (1999) entails 20 3×3 games and 50 participants. Of these 20 games, 14 were coordination games with multiple pure-strategy Nash equilibria for the purposes of testing equilibrium selection theories - see Appendix B. The third experiment was designed for this paper and entailed 15 games and 47 participants; 12 games were selected from the previous two experiments and 3 new games (numbered 48-50 in Appendic C) had obviously dominated strategies as the example in the Introduction.

For each game that appeared in more than one of these experiments, we performed Pearson Chi-square tests of the hypothesis that they shared a common data generation process. Since all of these tests passed, our subsequent analysis is based on the pooled data from these three experiments, with the exception of the three games with supposed obviously dominated strategies in the third experiment (48, 49, and 50), which are held back for the purpose of testing out-of-sample predictions. As a benchmark, the entropy of this 47-game data set is -1788.28; this is the upper bound on the maximum log-likelihood of the data.

A fourth experiment was designed to test a variety of conjectures about what makes a dominated strategy "obvious" in the sense that no rational person would assume it is as likely to be chosen as any other strategy. This experiment entailed 15 3×3 games (11 of which had strictly dominated strategies) and 75 participants - see Appendix D.

3. Population Models of Bounded Rationality.

Let U denote the J×J payoff matrix of the row player, so the transpose U' is the payoff matrix for the column player, and let U_j denote the j^{th} row of U. For notational convenience, let p^0 denote the uniform probability distribution over the J strategies, and let p^{NE} denote the unique symmetric Nash equilibrium when there is one, or a uniform probability distribution over the symmetric pure-strategy Nash equilibria otherwise.

We begin with a unimodal four-parameter probabilistic choice model in which a player has a three-parameter $(\mu,\varepsilon_1,\varepsilon_2)$ belief about the other players, given by

$$q^w(\mu,\varepsilon_1,\varepsilon_2) \equiv \varepsilon_2 p^0 + \varepsilon_1\, br(Up^0,\mu) + (1-\varepsilon_1-\varepsilon_2)p^{NE}, \tag{1}$$

where $br(y,\mu)$ denotes the logit best-reply to expected payoff y with precision μ. That is,

$$br_k(y,\mu) = \frac{\exp(\mu y_k)}{\sum_{j=1}^{J} \exp(\mu y_j)} \tag{2}$$

The modal player then chooses a logit best-reply to Uq^w with precision ν; i.e. the choice function is $br(Uq^w,\nu)$.

To allow for diversity in the population of participants, we assume that the actual beliefs of an individual player come from a normal distribution with mean $q^w(\mu,\varepsilon_1,\varepsilon_2)$ and standard deviation σ, truncated to the J-dimensional simplex, as in Haruvy, Stahl and Wilson (2001). Letting $f(q \mid \mu, \varepsilon_1, \varepsilon_2, \sigma)$ denote this distribution, the expected population choice probabilities are

$$P^e(\nu,\mu,\varepsilon_1,\varepsilon_2,\sigma) \equiv \int br(Uq,\nu) f(q \mid \mu, \varepsilon_1, \varepsilon_2, \sigma) dq \,. \tag{3}$$

In addition, we add the possibility that the player chooses randomly (a level-0 type), and that the player is a Maximax type with probabilistic choice function $br(\underline{m},\nu)$ where $m_j \equiv max_k U_{jk}$. Letting α_0 denote the proportion of the population that is level-0, and α_m denote the proportion that is Maximax, the combined probabilistic choice function is

$$P^*(\nu,\mu,\varepsilon_1,\varepsilon_2,\sigma,\alpha_0,\alpha_m) \equiv \alpha_0\, p^0 + \alpha_m\, br(\underline{m},\nu) + (1\text{-}\alpha_0\text{-}\alpha_m)\, P^e(\nu,\mu,\varepsilon_1,\varepsilon_2,\sigma) \,. \tag{4}$$

To gauge the specification error, we assume that the actual choice probabilities have a Dirichlet distribution with mean $P^*(\nu,\mu,\varepsilon_1,\varepsilon_2,\sigma,\alpha_0\ )$ and "strength" S.[2] Hence, the actual choice probabilities, p, have a density proportional to

$$\Pi_{j=1}^{J}(p_j)^{SP_j^*} \,. \tag{5}$$

The parameter S can be interpreted as a prior sample size; the larger S is, the more concentrated is the density at the mean P*. We call this the unimodal+ model ("+" for the level-0 and Maximax types and the Dirichlet parameter S).

We use a combination of simulated annealing and the Nelder-Meade algorithm to maximize the log-likelihood (LL) function of the 47-game data set described in Section 2. The maximized LL of the unimodal+ model is -1849.05. The nested hypotheses that $\alpha_0 = 0$ (no level-0 types), $\alpha_m = 0$ (no Maximax types), and $\sigma = 0$ (no population diversity) are individually and jointly rejected at all common acceptance levels. Without the level-0 or Maximax type, we reject the hypothesis that $S = \infty$ (no mis-specification as represented by a Dirichlet distribution); however, with the level-0 and Maximax types, we cannot reject $S = \infty$.

---

[2] Including S is analogous to computing the sum of squared errors of an ordinary least squares model, and as such does not reduce the effective degrees of freedom, but merely serves as an inverse measure of mis-specification.

The mean belief, $q^w(\mu,\varepsilon_1,\varepsilon_2)$, is estimated to have weight $\varepsilon_2 = 0.623$ on p^0, $\varepsilon_1 = 0.318$ on $br(Up^0,0.09)$, and 0.059 on p^{NE}. The preponderance of the weight on the uniform belief is most likely due to the absence of a pure level-1 type in the model. To test this conjecture, we add a level-1 type whose belief has a normal distribution with mean p^0 and standard deviation σ, and whose choice probabilities are logit best-replies with precision ν. One justification for using the same standard deviation parameter is that it is standard procedure in kernel density estimation. We use a single precision parameter to economize on parameters, later test this restriction, and fail to reject. The maximized LL increases to -1845.26, and this increase has a p-value of 0.0059. Moreover, the mean belief, $q^w(\mu,\varepsilon_1,\varepsilon_2)$, is estimated now to have weight $\varepsilon_2 = 0.030$ on p^0, $\varepsilon_1 = 0.817$ on $br(Up^0,0.062)$, and 0.153 on p^{NE}. As conjectured, the weight on the uniform belief declines dramatically. Indeed, we cannot reject the hypothesis that $\varepsilon_2 = 0$ (no weight on p^0), leaving us with the original SW95 specification of the Worldly type. Further, the MLE of the standard deviation declines from 0.146 to 0.088, indicating a distinctly bimodal distribution of beliefs for the population.

One potential shortcoming of the representation of diversity in the model so far is the thin tails of the normal distribution. A simple way to investigate this possible misspecification is to introduce a parameter for the tails: specifically, let ε_0 denote the proportion of rational players' beliefs that are uniformly distributed over the simplex. In other words, the beliefs are a convex combination of the uniform distribution over the simplex (with weight ε_0) and a truncated normal distribution with type-specific mean and standard deviation σ. With this modification (entailing one additional parameter), the maximized LL increases to -1838.78, which has a p-value of 0.0003. Moreover, the MLE of the standard deviation declines to 0.016, indicating that the beliefs can be approximately represented as two atoms (one for level-1 types at p^0 and one for Worldly types) plus a uniform distribution over the simplex with $\varepsilon_0 = 0.285$. The computational burden of our model stemming from eq(2) would be considerably reduced if this conjecture were true. The alternative model is equivalent to having $\sigma = 0$. Unfortunately, this restriction decreases the LL to -1841.78, which with one degree of freedom has a p-value of 0.035.

Previous studies have revealed some evidence for two additional archetypes: a level-2 type who believes other players are level-1 types, and naïve Nash types who believe p^{NE} to be the choice frequencies of others. For our pooled data set of 47 games and 155 participants, we found no statistically significant contribution from including such types. Apparently, the Worldly type whose belief puts weight of 0.843 on *br*(Up^0,0.065) and 0.157 on p^{NE} already captures the belief that other players might be level-1 types or Nash types (and less naïvely than the pure level-2 and Nash types).

To summarize, we have found that a mixture model with a level-0 type, a Maximax type, and level-1 and Worldly types with diverse thick-tailed priors emerges as the ML winner. Let us refer to this model as the *Boundedly Rational Population* (BRP) model. Table I presents the parameter estimates, variance-covariance matrix of these estimates, and the t-ratios.

How well does the BRP model fit the data? We consider three measures of goodness-of-fit: (1) the pseudo-R^2 (entropy/LL), (2) the Pearson Chi-square statistic (PCS), and (3) the root mean squared error (RMSE). The pseudo-R^2 is an impressive 0.973. The aggregated PCS is 91.90, which with 94 degrees of freedom has a p-value of 0.541. Further, only one game out of 47 fails the individual PCS test at the 5% level, and none fail at the 2.5% level. The RMSE is a modest 0.052. Therefore, we cannot reject the hypothesis that the fitted BRP model is the data generating process.

4. Confronting Obviously Dominated Strategies.

A fourth experiment was designed to test a variety of conjectures about what makes a dominated strategy "obvious" and what effect this salience has on the beliefs and behavior of boundedly rational players. Specifically, we are looking for evidence that some dominated strategies are so obvious that boundedly rational players' beliefs have lower probabilities that others will play these strategies, relative to undominated strategies. This experiment entailed 15 3×3 games (11 of which had strictly dominated strategies of various degrees of obviousness) and 75 participants - see Appendix D.

The design criterion was to have diversity across two dimensions: (1) the maximum payoff a dominated strategy offers, and (2) the maximin level of the game. The notion was that a strategy is obviously dominated to the extent that its maximum payoff is less than the maximin payoff of the game. We ran three sessions of 25 participants each at UT-Austin. The third session reversed the order of the 15 games. In addition to this experimental data, we also include the three games held back from the third experiment described in Section 2 (games 48, 49 and 50 of Appendix C). Chi-square tests fail to reject the hypothesis that the aggregate session behavior comes from the same data generating process. We therefore pool the data for further analysis. As a benchmark, the entropy of this data is -868.82.

A cursory look at the data provides no support for the above notion that a strategy is obviously dominated to the extent that its maximum payoff is less than the maximin payoff of the game. The eight games for which the maximum payoff of the dominated strategy is less than the maximin payoff (48, 49, 50, 52, 56, 60, 62, and 64) do not manifest dramatically different behavior: the original level-1 prediction still accounts for 292 out of 516 choices (56.6%), and it is the modal choice in six of the eight games. For the other six games with strictly dominated strategies (51,5 4, 57,58 , 59, and 65), the level-1 prediction accounts for 233 out of 450 choices (51.8%), and it is the modal choice in three of the six games.

To gauge the robustness of our BRP model to dominated strategies, we use the estimated BRP parameters to predict choice frequencies for the 18 games of this data set and compute three goodness-of-fit measures. The LL of the data is -912.55, giving a pseudo-R^2 of 0.952. The game-averaged RMSE is a disappointing 0.084. The aggregated PCS statistic for all 18 games is 82.74, with a p-value of 0.000015, clearly rejecting the hypothesis that this BRP model generated the data. On the other hand, only five of the 18 games fail the individual PCS test (48, 50, 51, 59, and 60). Game 51 in appendix D has only 35% choosing the level-1 action (A); deleting the dominated strategy (C), the level-1 choice would be action B. Thus, it appears that C was "obvious enough" for level-1 types to believe others will avoid it.

However, the same logic does not appear to extend to game 52. Though strategy A appears obviously dominated to us, the BRP model predicts the choices almost perfectly (PCS = 0.508, and RMSE = 0.010). As another troubling example, take game 56, in which A appears obviously dominated by B and yet the BRP model predicts the choices quite well (PCS = 0.885, and RMSE = 0.029). Fitting this data with a parsimonious model will not be an easy task.

4.1. Hypotheses of a Level-0 Tilt.

A natural and parsimonious modification of the BRP model is the hypothesis that level-0 types are less likely to choose a dominated strategy than an undominated strategy, but remain equally likely to choose any undominated strategy. Let q^0 denote such a "tilt" away from the uniform distribution (p^0). We define the logit best-reply function with precision μ and tilt q^0 as:

$$BR_k(y,\mu,q^0) = \frac{q_k^0 \exp(\mu y_k)}{\sum_{j=1}^{J} q_j^0 \exp(\mu y_j)} \tag{6}$$

Note that as the precision μ goes to 0, $BR_k(y,\mu,q^0)$ goes to q^0 instead of p^0. We then assume that level-0 and level-1 choice frequencies are q^0 and $br_k(Uq^0,\nu)$ respectively, and that the belief of a Worldly type is

$$q^w(\mu,\varepsilon_1,q^0) \equiv \varepsilon_1 BR(Uq^0,\mu,q^0) + (1-\varepsilon_1)p^{NE}. \tag{7}$$

A crude but simple hypothesis about q^0 is that it is a slight deviation from the uniform p^0 in the direction away from the strictly dominated strategy. In our case with three strategies for each game, suppose q^0 has probability $\eta \in [0, 1/3]$ on the strictly dominated strategy, and probability $(1-\eta)/2 \in [1/3, 1/2]$ on each of the other two strategies. We will refer to this as the *linear-tilt* model.

Fixing the eight parameters of the BRP model at the MLE values from Section 3,[3] and maximizing the log-likelihood of the new data set with respect to the one new parameter η, we find $\eta = 0.298$. This q^0 means that about 10.6% of the level-0 types are believed to avoid strictly dominated strategies. Despite this small tilt, the LL increases to -898.61; this 13.94 increase in LL has a p-value of 1.3×10^{-7}. Further, the RMSE declines to 0.061, the aggregated PCS declines to 53.06 (p-value of 0.033), and only 3 of the 18 games fail the individual PCS test (48, 60, and 62). While the increase in LL is very statistically significant, the PCS test indicates that the modified model is not the data generating process. These mixed results beg two questions. Why does such a small tilt in q^0 make such a large improvement, and why does the aggregate PCS test still fail (albeit not at the 2.5% level)?

One potential shortcoming of the linear-tilt model is its invariance to how dominated a strategy is. The motivating notion of the experiment design - that the comparison of the

---

[3] Since we are exploring "minimal" departures from the BRP model, we do not want to compound the effects of the uniform level-0 tilt with interactive effects with other parameters. Later in section 4.3, we will address the robustness of all the parameter estimates.

maximum payoff of a strategy with the maximin payoff of the game - suggests a measure of how bad a strategy is. Let $z_j \equiv min\{0, m_j - M\}$, where m_j is the maximum payoff of strategy j and M is the maximin payoff of the game, and suppose $q^0 = br(z,\gamma)$. All strategies whose maximum payoff is at least as great as the maximin payoff will be equally likely, while any strategy whose maximum payoff is less than M will have a probability less than 1/J. The parameter γ gauges how sensitive level-0 choice probabilities are to such shortfalls. Fixing the eight parameters of the BRP model at the MLE values from Section 3, and maximizing the log-likelihood of the new data set with respect to the one new parameter, we find $\gamma = 0.020$, LL = -908.67, PCS = 75.12, and RMSE = 0.076. Since all the goodness-of-fit measures are substantially worse than the simple linear-tilt model, we reject our conjecture that a dominated strategy is obvious to the extent that its maximum payoff is less than the maximin payoff of the game.

Rather than using the maximin payoff as the reference, we considered using a fixed reference level, Z, to be estimated. A single game-invariant reference level makes sense for our data since the range of payoffs for all games was [0,100]. Haruvy and Stahl (1999) demonstrated that one-parameter logit choice functions fit the data better after rescaling all payoffs to have the same range. Hence, the estimated value of Z should be interpreted in terms of the [0,100] range. Continuing, we let $z_j \equiv min\{0, m_j - Z\}$, and again suppose $q^0 = br(z,\gamma)$. Maximizing the log-likelihood of the new data set with respect to the two new parameters, we find $\gamma = 0.0054$, Z = 50, LL = -902.00, PCS = 61.01, and RMSE = 0.066. These goodness-of-fit measures are not as good as for the linear-tilt model, despite having an additional parameter. Moreover, now five games (48, 51, 52, 60, and 62) fail the individual PCS tests. Therefore, we reject this alternative in favor of the linear-tilt model.

Finally, instead of the maximum payoff (m_j) of a strategy, we considered the average payoff of strategy j ($a_j \equiv U_j p^0$), letting $z_j \equiv min\{0, a_j - Z\}$and $q^0 = br(z,\gamma)$. We call this the *logit-average-payoff* (LAP)-tilt model. Maximizing the log-likelihood of the new data set with respect to the two new parameters, we find $\gamma = 0.0080$, Z = 38.3, LL = -899.66, PCS = 56.85, and RMSE = 0.062. These goodness-of-fit measures are almost but not quite as good as for the linear-tilt model, despite having an additional parameter. Games 48, 59, 60, and 62 still fail the individual PCS tests.

From the point of view of a theory of bounded rationality, one disadvantage the linear-tilt model has over the LAP-tilt alternative is that the participant must identify all the dominated strategies before computing q^0, even in cases like game 59 in which it takes a mixture of A and B to strictly dominate C. While it may be reasonable to assume that level-1 and Worldly types can do this, it does not seem reasonable that level-0 types can. Rather, it is more plausible that level-0 types shy away from a dominated strategy simply because of its *unattractive* payoffs (a_j). Human cognitive processes may be better at eliminating bad choices than at finding optimal choices (e.g. Cosmides and Tooby, 1992; and Gigerenzer, et. al., 1999).

4.2 Predictions of Level-0 Tilts.

The BRP model predictions failed the individual PCS tests for games 48, 50, 51, 59, and 60, while the linear-tilt model predictions failed for games (48, 60, and 62). Thus, the level-0 tilt helped in games 50, 51, and 59, and hurt in game 62.

To investigate these ambiguous results further, we present here the predicted choice frequencies of both models for games 51 and 62. In the first column, p^1, p^m, and p^w denote the predicted choice frequencies for level-1, Maximax, and Worldly types respectively; q^w denotes the belief of the Worldly type, p* denotes the predicted population mixture, and freq gives the actual choice frequencies. The last row (labeled GOF) gives LL less entropy, PCS, and RMSE in that order.

Game 51							
	BRP Model:				Linear-tilt Model		
	A	B	C		A	B	C
q^0	0.333	0.333	0.333		0.351	0.351	0.298
p^1	0.787	0.213	0.000		0.644	0.356	0.000
p^m	1.000	0.000	0.000		1.000	0.000	0.000
q^w	0.480	0.469	0.050		0.433	0.519	0.048
p^w	0.160	0.840	0.000		0.160	0.840	0.000
p*	0.503	0.470	0.026		0.440	0.536	0.024
freq:	0.347	0.653	0.000		0.347	0.653	0.000
GOF:	-6.17	10.97	0.140		-3.38	5.20	0.088

Game 62							
	BRP Model:				Linear-tilt Model		
	A	B	C		A	B	C
q^0	0.333	0.333	0.333		0.351	0.298	0.351
p^1	0.897	0.000	0.103		0.871	0.000	0.129
p^m	0.950	0.000	0.050		0.950	0.000	0.050
q^w	0.661	0.040	0.300		0.616	0.037	0.347
p^w	0.409	0.001	0.590		0.269	0.000	0.730
p*	0.653	0.027	0.320		0.585	0.024	0.391
freq:	0.720	0.000	0.280		0.720	0.000	0.280
GOF:	-2.41	2.90	0.048		-4.04	6.47	0.102

For game 51, we can see that the tilt of q^0 away from the dominated strategy C causes a shift in the level-1 prediction away from A towards B, thereby improving all three goodness-of-fit measures. This tilt has no effect on the Worldly prediction because the effect of q^0 in $BR_k(y,\mu,q^0)$ offsets the effect of q^0 on the belief $q^w(\mu,\varepsilon_1,q^0)$. For game 62, the tilt away from the dominated strategy B causes a shift in both the level-1 and Worldly prediction away from A towards C, thereby deteriorating all three goodness-of-fit measures.

4.3. Confronting all Four Data Sets.

So far in this section we have fixed the parameter values of the BRP model at the MLE values for the first three data sets, and used the fourth data set to estimate the level-0 tilt. Since there are games in the first three data sets with strictly dominated strategies, the modified model will have an effect there as well. Therefore, we need to investigate this effect. Moreover, the effect may differ for the linear-tilt model and the LAP-tilt model and enable us to select one model over the other.

First, we fix the parameters at the values used in Section 4.2, and compute the three goodness-of-fit measures for the first three data sets and compare this to that obtained without the level-0 tilt. For the linear-tilt model, we find that all three measures deteriorate: LL decreases by 6.91, PCS increases by 11.68, and RMSE increases by 0.008. For the LAP-tilt, however, we find that all three measures remain essentially the same: LL increases by 0.32, PCS decreases by 0.23, and RMSE is unchanged. This finding is a good reason to prefer the LAP-tilt model over the linear-tilt model.

Second, we estimate all ten parameters of the LAP-tilt model on the four data sets pooled, and find that the LL increases by only 5.385, which (with 10 degrees of freedom) is insignificant at all common acceptance levels. In other words, the parameters of the LAP-tilt model are robust across data sets.

5. Conclusions.

We began this investigation with the intuition that we could find an empirical criteria to determine when a strategy is so obviously dominated that no reasonable human player would believe that others are as likely to choose it as an undominated strategy. One hypothesis was that a strategy is obviously dominated to the extent that its maximum possible payoff is less than the maximin level of the game. Our conjecture was not only that players would avoid an obviously dominated strategy but that minimally sophisticated players would be able to assign a negligible probability to that strategy being played by others. The data from our experiment reject this hypothesis.

On the other hand, we did find that the fit of the level-n boundedly rational population (BRP) model could be improved by hypothesizing a tilt in the level-0 behavior. Two candidates producing statistically significant improvements on the test set of data were the linear-tilt model and the LAP-tilt model. Comparing how these models predicted on the benchmark data sets, we concluded that the LAP-tilt model is better. In other words, a level-0 player is exponentially less likely to choose a strategy to the extent that its average payoff falls below a threshold that is about 40% of the range of payoffs in the game.

The above finding is critical in any setting where hierarchical bounded rationality is present. Though level-n thinking survives the most obviously dominated strategies we could devise, a tilt in the choices of level-0 players away from obviously dominated actions, as well as a corresponding tilt in the beliefs of the more sophisticated level-1 and worldly players shows that dominated strategies can be made obvious in some sense. However, the surprisingly small impact obviously dominated strategies had on subjects' behavior suggests that level-n thinking is very deeply rooted in subjects' belief formation and approach to strategic games to the point that it can represent a real bias in behavior and generate 'obviously unreasonable' beliefs.

Although we have found that our human subjects in one-shot games seldom believe others will recognize and avoid even obviously dominated strategies, it does not follow that dominated strategies will survive over time. Since humans do succeed in avoiding dominated strategies, when that empirical evidence is available to the players, any learning dynamic for which the beliefs are responsive to the history will eventually drive out iterated strictly dominated strategies. Nonetheless, the presence of dominated strategies at the start of the dynamic process can substantively alter the path of play, thereby affecting long-run behavior.

References

Cosmides, L. and J. Tooby (1992), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture,* Oxford Univ. Press.

Costa-Gomes, M., V. Crawford, and B. Broseta (2001), "Cognition and Behavior in Normal-Form Games: An Experimental Study,", *Econometrica*, **69**(5) 1193-1235.

Duffy, J. and R. Nagel (1997), "On the Robustness of Behavior in Experimental 'Beauty Contest' Games," *Economic Journal*, 107, 1684-1700.

Gigerenzer, G., P. Todd, and the ABC Research Group, (1999), *Simple Heuristics That Make Us Smart,* Oxford Univ. Press.

Gneezy, U. (2002), "On the Relation between Guessing Games and Bidding in Auctions," working paper, The University of Chicago Graduate School of Business.

Haruvy, E., and D. Stahl, (1999), "Empirical Tests of Equilibrium Selection Based on Player Heterogeneity," mimeo.

Haruvy, E., Stahl, D. and P. Wilson (2001), "Modeling and Testing for Heterogeneity in Observed Strategic Behavior," *Review of Economics & Statistics,* **83**, 146-57.

Ho, T-H., C. Camerer and K. Weigelt (1998), "Iterated Dominance and Iterated Best Response in Experimental p-Beauty Contests," *American Economic Review*, **88**, 947-969.

Katok, E. M. Sefton, and A. Yavas (2002), "Implementation by Iterative Dominance and Backward Induction: An Experimental Comparison," *Journal of Economic Theory*, **104**(1), 89-103.

Moulin, H. (1979), "Dominance Solvable Voting Schemes", *Econometrica*, **37**, 1337-1353.

Moulin, H. (1984). "Dominance Solvability and Cournot Stability," *Mathematical Social Sciences*, **7**, 83-102.

Nagel, R. (1995), "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, **85**, 1313-1326.

Sefton, M., and A. Yavas (1996), Abreu-Matsushima Mechanisms: Experimental Evidence," *Games and Economic Behavior*, **16**(2), 280-302.

Sonsino, D., I. Erev, and S. Gilat (1999), "On Rationality, Learning, and Zero-Sum Betting—An Experimental Study of the No Betting Conjecture," mimeo.

Stahl, D. (1996), "Boundedly Rational Rule Learning in a Guessing Game," *Games and Economic Behavior*, **16**, 303-330.

Stahl, D. and P. Wilson (1994), "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior and Organization*, **25**(3), 309-27.

Stahl, D. and P. Wilson (1995), "On Players' Models of Other Players: Theory and Experimental Evidence, Games and Economic Behavior," **10**, 218-254.

Tan, T. and S.Werlang (1988), "The Bayesian Foundation of Solution Concepts of Games, " *J. of Economic Theory*, **45**, 370-391.

Table I. Parameter Estimates and Statistics of BRP Model

			Var-Cov Matrix (x 10^{-4})								
	MLE	T-Ratio	σ	ε_0	ν	μ	ε_1	α_0	α_1	α_m	α_w
σ	0.0157	2.816	0.3100								
ε_0	0.285	5.183	0.1600	30.2000							
ν	0.295	8.744	-0.4490	8.1100	11.4000						
μ	0.645	18.733	-0.0287	-0.1690	-0.1100	0.1190					
ε_1	0.843	34.844	0.2130	-0.0666	-0.5890	-0.0971	5.8500				
α_0	0.079	4.822	-0.1660	-0.4900	2.6600	-0.0568	0.1430	2.6800			
α_1	0.451	14.135	-0.1860	-5.0700	-3.1800	0.5330	0.2340	-0.1480	10.2000		
α_m	0.0552	4.128	0.0868	-2.1100	-1.4400	-0.0296	-0.4140	-0.8130	-1.1700	1.7900	
α_w	0.415	12.857	0.2650	7.6700	1.9600	-0.4470	0.0371	-1.7200	-8.8800	0.1940	10.4000

Appendix A. The games (and choices) of Haruvy, Stahl, and Wilson (2001)

1)

	A	B	C
	25	33	0
A	25	30	100
B	40	45	65
C	31	0	40

2)

	A	B	C
	15	31	12
A	75	40	45
B	70	15	100
C	70	60	0

3)

	A	B	C
	15	28	15
A	10	100	0
B	5	60	70
C	80	30	10

4)

	A	B	C
	29	10	19
A	70	90	38
B	100	0	40
C	88	48	43

5)

	A	B	C
	45	10	3
A	30	50	100
B	40	45	10
C	35	60	0

6)

	A	B	C
	28	4	25
A	10	100	40
B	0	70	50
C	20	50	60

7)

	A	B	C
	31	8	19
A	25	30	100
B	60	31	51
C	95	30	0

8)

	A	B	C
	27	9	22
A	80	60	0
B	40	10	50
C	100	5	20

9)

	A	B	C
	3	14	41
A	75	0	45
B	80	35	45
C	100	35	41

10)

	A	B	C
	32	12	14
A	30	100	50
B	40	0	90
C	50	75	29

11)

	A	B	C
	15	35	8
A	0	100	50
B	90	63	50
C	46	82	52

12)

	A	B	C
	35	3	20
A	40	100	65
B	33	25	65
C	80	0	65

13)

	A	B	C
	13	10	35
A	45	50	21
B	41	0	40
C	40	100	0

14)

	A	B	C
	33	2	23
A	30	100	22
B	35	0	45
C	51	50	20

15)

	A	B	C
	17	2	39
A	40	15	70
B	22	80	0
C	30	100	55

Appendix B. Haruvy and Stahl (1999)

16)

	A	B	C
	38	4	8
A	60	50	90
B	50	75	40
C	25	0	100

17)

	A	B	C
	0	4	46
A	60	0	0
B	0	55	25
C	100	35	35

18)

	A	B	C
	21	27	2
A	25	30	100
B	40	45	65
C	31	0	40

19)

	A	B	C
	9	39	2
A	100	20	0
B	80	80	20
C	50	50	40

20)

	A	B	C
	44	2	4
A	50	100	50
B	0	85	0
C	35	0	80

21)

	A	B	C
	12	2	36
A	40	15	70
B	22	80	0
C	30	100	55

22)

	A	B	C
	42	0	8
A	80	20	80
B	60	10	70
C	20	0	100

23)

	A	B	C
	1	48	1
A	35	35	70
B	25	55	100
C	0	0	60

24)

	A	B	C
	39	6	5
A	30	50	100
B	40	45	10
C	35	60	0

25)

	A	B	C
	1	2	47
A	0	55	85
B	30	70	30
C	15	55	100

26)

	A	B	C
	33	17	0
A	80	35	90
B	50	50	100
C	0	0	85

27)

	A	B	C
	31	5	14
A	30	50	100
B	35	0	45
C	51	50	20

28)

	A	B	C
	35	4	11
A	80	80	20
B	20	100	0
C	100	10	30

29)

	A	B	C
	1	46	3
A	70	0	0
B	100	50	50
C	0	35	80

30)

	A	B	C
	17	22	11
A	75	40	45
B	70	15	100
C	70	60	0

31)

	A	B	C
	36	9	5
A	35	0	100
B	0	100	0
C	15	40	40

32)

	A	B	C
	2	5	43
A	50	0	0
B	0	55	25
C	100	35	35

33)

	A	B	C
	19	8	23
A	25	30	100
B	60	31	51
C	95	30	0

34)

	A	B	C
	10	34	6
A	80	60	50
B	60	70	90
C	0	0	100

35)

	A	B	C
	0	6	44
A	55	0	25
B	50	50	30
C	35	100	35

Appendix C. Third Experiment for this paper.

36)

	A	B	C
	16	30	1
A	25	30	100
B	40	45	65
C	31	0	40

37)

	A	B	C
	6	38	3
A	100	20	0
B	80	80	20
C	50	50	40

38)

	A	B	C
	36	7	4
A	30	50	100
B	40	45	10
C	35	60	0

39)

	A	B	C
	30	4	13
A	80	80	20
B	20	100	0
C	100	10	30

40)

	A	B	C
	1	42	4
A	71	0	0
B	100	50	50
C	0	35	80

41)

	A	B	C
	16	22	9
A	75	40	45
B	70	15	100
C	70	60	0

42)

	A	B	C
	39	5	3
A	33	0	100
B	1	100	1
C	15	40	40

43)

	A	B	C
	0	6	41
A	50	0	0
B	0	55	25
C	100	35	35

44)

	A	B	C
	11	33	3
A	80	60	50
B	60	70	90
C	0	0	100

45)

	A	B	C
	32	4	11
A	70	90	38
B	100	0	40
C	88	48	43

46)

	A	B	C
	7	31	9
A	0	100	50
B	90	63	50
C	46	82	52

47)

	A	B	C
	4	9	34
A	75	0	45
B	80	35	45
C	100	35	41

48)

	A	B	C
	36	11	0
A	75	10	100
B	5	90	5
C	0	1	1

49)

	A	B	C
	0	29	18
A	1	0	0
B	10	90	10
C	100	5	20

50)

	A	B	C
	13	0	34
A	10	100	10
B	0	0	0
C	5	5	90

Appendix D: Fourth Experiment for this paper.

51)

	A	B	C
	26	49	0
A	0	35	100
B	55	40	20
C	30	0	0

52)

	A	B	C
	1	36	38
A	1	0	0
B	100	10	5
C	5	5	90

53)

	A	B	C
	64	8	3
A	30	50	100
B	40	45	10
C	35	60	0

54)

	A	B	C
	62	12	1
A	55	10	100
B	5	90	5
C	15	0	0

55)

	A	B	C
	53	4	18
A	80	80	20
B	20	100	0
C	100	10	30

56)

	A	B	C
	1	41	33
A	10	5	5
B	100	30	35
C	0	80	30

57)

	A	B	C
	36	1	38
A	80	10	5
B	25	0	10
C	5	100	20

58)

	A	B	C
	0	42	33
A	15	0	0
B	0	90	10
C	100	0	20

59)

	A	B	C
	44	27	4
A	35	0	100
B	1	100	1
C	10	40	40

60)

	A	B	C
0	12	63	22
A	0	0	38
B	55	25	40
C	35	35	43

61)

	A	B	C
	17	52	6
A	80	60	50
B	60	70	90
C	0	0	100

62)

	A	B	C
	54	0	21
A	20	100	20
B	5	5	5
C	0	5	90

63)

	A	B	C
	39	14	22
A	70	90	38
B	100	0	40
C	88	48	43

64)

	A	B	C
	1	43	31
A	10	10	10
B	15	80	15
C	100	0	30

65)

	A	B	C
	30	44	1
A	20	0	100
B	10	90	0
C	0	0	5