

THEORIES OF LEARNING IN GAMES AND HETEROGENEITY BIAS

BY NATHANIEL T. WILCOX¹

Comparisons of learning models in repeated games have been a central preoccupation of experimental and behavioral economics over the last decade. Much of this work begins with pooled estimation of the model(s) under scrutiny. I show that in the presence of parameter heterogeneity, pooled estimation can produce a severe bias that tends to unduly favor reinforcement learning relative to belief learning. This occurs when comparisons are based on goodness of fit and when comparisons are based on the relative importance of the two kinds of learning in hybrid structural models. Even misspecified random parameter estimators can greatly reduce the bias relative to pooled estimation.

KEYWORDS: Learning in games, nonlinear dynamic models, panel data, heterogeneity, biased estimation, experimental design.

1. INTRODUCTION

LEARNING IN GAMES is of interest to many theorists, and two varieties of learning have received the most scrutiny in experiments. In belief learning models, players form beliefs on the basis of opponents' past decisions and tend to play strategies today that have relatively high expected payoffs given those beliefs. In reinforcement learning models, players tend to play strategies that have paid off relatively well in the past. Both models can be combined in flexible hybrid models. With few exceptions, empirical comparisons between learning principles begin with pooled estimation: That is, a single shared vector of learning model parameters is assumed for all subjects in a sample. I show that if subjects are in fact heterogeneous, this assumption tends to bias empirical comparisons of these theories in favor of reinforcement learning models.

The intuition behind this is simple. Suppose, for example, that all players are belief learners, but that they still learn and choose somewhat differently from one another, due to differences in their sensitivity to expected payoff differences, rates of discounting past experience, and so forth. Pooled estimation of the (otherwise) correctly specified belief learning model will yield prediction errors that are correlated with past strategy choices of players, because those past choices carry idiosyncratic parameter information for which the estimation does not account. Therefore, adding players' own past strategy choices—

¹Warm thanks to Lisa Rutström for stimulating collaborations that led to this paper; to Christopher Murray for help with econometric literature; to Dmytro Hryshko for computational insights; and to Antonio Cabrales and Ed Hopkins for comments on formal arguments. Comments and/or help from Emek Basker, Colin Camerer, Gary Charness, David Cooper, Mahmoud El-Gamal, Nick Feltovich, Dan Friedman, Shelby Gerking, Glenn Harrison, Teck-Hua Ho, Kyung-So Im, Charles Manski, Hervé Moulin, Xiaoguang Ni, Thomas Palfrey, Timothy Salmon, David Shanks, Roger Sherman, Charles Shimp, Robin Sickles, Robert Slonim, Kenneth Train, Roberto Weber, three referees, and many conference and seminar participants are also happily acknowledged. Any remaining errors are solely my responsibility.

or an informative function of those choices—to the pooled estimation of the belief model should improve its fit if players are heterogeneous. However, this is, in large part, just what reinforcement models do: Because players' past payoffs are a function of players' own past strategy choices, models based on the reinforcement principle are by definition directly conditioned on a function of players' own lagged choices.

As a result, heterogeneity alone gives pooled estimations of reinforcement learning models (and hybrid models) an automatic fit advantage over pooled estimations of belief learning models (if the model of beliefs is free of players' own lagged strategy choices, which is true of most such models). Heterogeneity also frequently causes pooled estimation of hybrid models to overstate the relative structural importance of reinforcement principles. Monte Carlo studies show that this bias can be dramatic and overwhelming. Heterogeneity of one particular parameter of learning models (known as “sensitivity” or “precision”) is the strongest source of the bias. Random parameter estimators are a promising (if imperfect) approach to the problem.

2. THE MODELS AND THE ECONOMETRIC PROBLEM

Camerer and Ho's (1999) experience-weighted attraction (EWA) model is a hybrid model that encompasses several belief and reinforcement models as special cases,² so I use it as the basis for this study.³ I present it in an unusual form (and just for row players in a 2×2 game) that helps illuminate the cause of the bias. Row players choose a strategy $r \in \{1, 0\}$ (up or down); similarly, column players choose $c \in \{1, 0\}$ (left or right). Let $\pi[r, c]$ be row players' payoffs, and let r_t and c_t be strategy choices of players in period t . In this two-strategy case, EWA may be expressed in terms of a single variable z_t that summarizes how the history of play through period t determines the relative attractiveness of a row player's two strategies, determining the probability P_{t+1} that $r_{t+1} = 1$:

$$(1) \quad P_{t+1} = \Lambda[\lambda z_t(\delta)] = [1 + \exp[-\lambda z_t(\delta)]]^{-1},$$

$$\text{where } z_t(\delta) = IC_t + F_t + \delta C_t \quad \forall t > 0;$$

$$IC_t = \phi^t(N_0/N_t)z_0 \quad \forall t > 0;$$

$$F_t = N_t^{-1} \sum_{j=1}^t \phi^{t-j} [r_j \pi(1, c_j) - (1 - r_j) \pi(0, c_j)] \quad \forall t > 0;$$

²See Camerer and Ho (1999) for a discussion of the behaviorist and game-theoretic roots of EWA. Cheung and Friedman (1997) and Fudenberg and Levine (1998) give good discussions of belief-based models, while Erev and Roth (1998) is the best-known contemporary description of the roots of reinforcement learning models.

³In particular, I use EWA *not* because it is an especially wanting model nor because its econometric treatment has been particularly lax. In fact, EWA is a quite creative synthesis with potentially great empirical usefulness, and Camerer, Ho, and their collaborators have been tirelessly cautious as far as its empirical treatment is concerned.

$$C_t = N_t^{-1} \sum_{j=1}^t \phi^{t-j} [(1 - r_j)\pi(1, c_j) - r_j\pi(0, c_j)] \quad \forall t > 0;$$

$$N_t = (1 - \kappa)\phi N_{t-1} + 1 \quad \forall t > 0 \quad \text{and} \quad N_0 = \eta[1 - (1 - \kappa)\phi]^{-1}.$$

The parameters of EWA are $z_0 \in \mathfrak{R}$, $\lambda \in \mathfrak{R}^+$, and δ , ϕ , κ , and η all in $[0, 1]$.⁴ Assume $\eta < 1$, $\phi < 1$, $0 < \lambda < \infty$, and a nontrivial game for all players⁵—all necessary for asymptotic consistency to hold without heterogeneity (Cabrales and Garcia-Fontes (2000), Ichimura and Bracht (2001)).

Camerer and Ho (1999, p. 836) note that δ is “...most important in EWA because it shows most clearly the different ways in which EWA, reinforcement and belief models capture two basic principles of learning—the law of actual effect and the law of simulated effect.” In Equations (1), F_t is “factual reinforcement” (of $r = 1$ relative to $r = 0$)—a discounted average difference between past payoffs to $r = 1$ and $r = 0$; C_t is “counterfactual reinforcement” (of $r = 1$ relative to $r = 0$)—what F_t would have been with exactly the opposite history of r_t ; and IC_t is the period t impact of initial conditions. When $\delta = 0$, only F_t determines z_t : Only the “law of actual effect” operates, which is the hallmark of pure reinforcement learning. If instead $\delta = 1$, F_t and C_t equally determine z_t and the “law of simulated effect” operates as strongly as the law of actual effect. Camerer, Ho, and Chong (2002) show that when $\delta = 1$ and $\kappa = 0$, EWA is the weighted fictitious play (WFP) model of Fudenberg and Levine (1998), a well-known belief learning model. They also show that when $\delta = 0$ and $\eta = 1$, EWA is a generalized reinforcement learning (GRL) model: When $\kappa = 1$, it resembles the cumulative reinforcement model of Erev and Roth (1998), and when $\kappa = 0$, it resembles the averaged reinforcement model of Mookerjee and Sopher (1997). Thus, EWA nests a variety of learning principles (and hybrids of them) within one parametric model, with the parameter δ largely responsible for the balance between them.

Learning model parameters may vary across players, but comparisons of EWA, WFP, and GRL models usually ignore this, and specify a single, shared parameter vector for estimation and figuring goodness of fit. Some papers allow individual or class variation of parameters, but other finite-sample biases are likely to influence these estimates.⁶ However, pooled estimation of models

⁴Camerer, Ho, and Chong (2002) specify $N_0 \in [0, [1 - (1 - \kappa)\phi]^{-1}]$ as a structural parameter. This is equivalent to specifying $\eta \in [0, 1]$ as a parameter and writing $N_0 = \eta[1 - (1 - \kappa)\phi]^{-1}$ as I do here. In my experience, optimization subject to the simple linear constraint $\eta \in [0, 1]$ was more robust (in terms of frequency of convergence problems).

⁵This last assumption is that every player’s payoff function depends on choices made by at least one other player. As Cabrales and Garcia-Fontes (2000) note, $\delta = 1$ would then cause identification problems. However, this assumption is true of all experimental games of which I am aware.

⁶Cheung and Friedman (1997), Camerer, Ho, and Wang (1999), and Broseta (2000) estimate various models for individual subjects. However, some Monte Carlo evidence (Cabrales and Garcia-Fontes (2000), Wilcox (2005)) suggests that individual estimation can suffer from strong

with lagged dependent variables can be biased and inconsistent in panel data with heterogeneity (Hsiao (1986), Heckman (1991)). Although several solutions to this problem exist for linear models (Anderson and Hsiao (1982), Ahn and Schmidt (1995)), no wholly general, distribution-free solution exists for nonlinear models like EWA (Wooldridge (2005)) and biases can be particularly severe in them (Heckman (1981)).⁷

Monte Carlo studies below will suggest that heterogeneity of λ is a most serious source of downward bias in estimation of δ .⁸ Focusing on this cause of the bias, consider a simplified estimation situation in which ϕ , κ , η , and z_0 are all known and constant across players i . Suppose δ too is known to be constant across players, but has unknown true value δ_0 that we wish to estimate, and that although it is not known whether λ^i varies across players i , it has a known true mean value λ_0 across players. If it were true that $\lambda^i = \lambda_0 \forall i$, it would be correct to write terms of a pooled log likelihood function and its derivative with respect to δ as

$$(2) \quad \ell_{t+1}^i(\delta) = r_{t+1}^i \ln[\Lambda[\lambda_0 z_t^i(\delta)]] + (1 - r_{t+1}^i) \ln[1 - \Lambda[\lambda_0 z_t^i(\delta)]]$$

and

$$(3) \quad \partial \ell_{t+1}^i / \partial \delta = \lambda_0 [r_{t+1}^i - \Lambda[\lambda_0 z_t^i(\delta)]] C_t^i.$$

Given M row players playing for T periods, the maximum likelihood estimator (MLE) $\hat{\delta}$ of δ_0 would then solve

$$(4) \quad m(\hat{\delta}) = [M(T-1)]^{-1} \sum_{i, T > t > 0} [r_{t+1}^i - \Lambda[\lambda_0 z_t^i(\hat{\delta})]] C_t^i = 0$$

(≤ 0 or ≥ 0 if $\hat{\delta} = 0$ or 1 , respectively).

finite-sample biases too. Studies of individual learning (rather than game learning) exist where finite-sample biases are less likely: Kitzis, Kelley, Berg, Massaro, and Friedman (1998) estimate five different learning models on an individual basis with 480 observations per subject, finding that subject-by-subject fits are much better than pooled fits and that variance due to subject heterogeneity dwarfs that due to treatment variations. Rutström and Wilcox (2005) have estimated both the EWA model and an extended version of Cheung and Friedman's gamma-weighted belief model using the "mixed random estimator" recommended later in this paper and find that heterogeneity is highly significant in their game.

⁷Even in dynamic linear models, dealing with heterogeneity beyond additive random effects is subtle (Pesaran, Smith, and Im (1996)). Distribution-free transformation solutions exist for certain classes of nonlinear models where heterogeneity affects dependent variables in special ways (Wooldridge (1997)), but EWA does not fall into this class.

⁸Heterogeneity of λ is essentially subject-specific heteroscedasticity conditional on given z_t . This is known to be an inferential nuisance in other discrete choice situations where lagged dependent variables are not an issue (Ballinger and Wilcox (1997)), but the resulting problems are qualitatively different and more severe here.

Consider the case of a repeated game experiment where play converges to a mixed steady state different from any mixed-strategy Nash equilibrium for both players. This is roughly observed in many experiments on games with unique mixed-strategy Nash equilibria (Erev and Roth (1998)).⁹ Also assume that the experiment uses two large populations of row and column players, with players randomly rematched to new partners in every period t . This idealized design is purposely mimicked by many experimental designs for theoretical reasons¹⁰ and is also econometrically convenient.¹¹ Let \tilde{r}^i , \tilde{F}^i , and \tilde{C}^i denote unconditional (in the time series sense) steady-state random variables that correspond to the conditional random variables r_t^i , F_t^i , and C_t^i ; note that their distributions depend on each player's true λ^i . Also let $\tilde{z}^i(\delta) = \tilde{F}^i + \delta \tilde{C}^i$ (with $\phi < 1$, IC_t vanishes asymptotically). Cabrales and Garcia-Fontes (2000) show that in the single-player case (and by extension the case where $\lambda^i = \lambda_0 \forall i$) the MLE $\hat{\delta}$ is consistent as $T \rightarrow \infty$.¹² Part of their proof involves showing that r_t and c_t

⁹This is part of the empirical inspiration for stochastic choice generalizations of Nash equilibrium that accommodate both descriptive successes and failures of Nash equilibrium (Goeree and Holt (2001)).

¹⁰The specific "adaptive EWA" specification above (and its special cases) assumes that subjects learn about stage game strategies and do not consider dynamic strategies. For this reason, many experimental designs that test such models either randomly rematch a large number of row and column players at the outset of each period or arrange a matching protocol whereby each pair of row and column players meets only once during play ($T < M$ in such designs) to undermine the appeal of dynamic strategies. Even in those designs, cooperative norms may emerge for certain kinds of games (Kandori (1992)), but Duffy and Ochs (2003) find no evidence of this in large enough random rematching designs using the prisoner's dilemma ("large enough" turns out to be surprisingly small too).

¹¹In particular, it allows one to treat r_t as statistically independent across row players and to treat successive values of c_t as statistically independent of a row player's own history of choices (vastly simplifying econometric analysis). By contrast, fixed pairing designs imply that a row player's own past strategy choices "return" to her in the lagged choices of her partner. Even in a belief-based model, this implies that row's belief model is indirectly conditioned on her own lagged choices and, hence, on traces of her own idiosyncratic parameter information. That information is, however, noisier than in a reinforcement model. This is because idiosyncratic parameter information passes through only one randomization (the row player's own choice) before returning as a direct conditioning variable in the row player's own reinforcement history; by contrast, such information content is randomized twice (first by the row player's own choice, and then subsequently by the partner's choice conditioned on it) before it enters into the row player's belief process in a belief model. As a result, even in a fixed pairing design, pooled estimations of belief models "benefit" (in terms of fit) much less than reinforcement models do from idiosyncratic parameter information contained in conditioning variables. In large random rematching designs, however, belief learning models get no such benefit at all, whereas reinforcement learning models continue to benefit.

¹²Cabrales and Garcia-Fontes (2000) also note the potential for bias from heterogeneity and conduct limited Monte Carlo studies of the consequences of heterogeneity in ϕ . However, they do not explain the likely direction of the bias, its central importance in comparisons of learning principles, or how the bias also plagues goodness-of-fit comparisons between learning principles. These points are the unique contribution of this paper.

are mixing sequences, and that functions of them in (4) are sufficiently well behaved for consistency to hold. This implies that the conditional moments that make up (4) converge to their unconditional steady-state expectations if $\lambda^i = \lambda_0 \forall i$. Temporarily suppressing player superscripts to consider this case, suppose that $\delta_0 \in (0, 1)$ so that (4) converges to an equality. Then $\hat{\delta}$ converges to δ_0 , and $m(\hat{\delta})$ converges to

$$(5) \quad m(\delta_0) = E[(\tilde{r} - \Lambda[\lambda_0 \tilde{z}(\delta_0)])\tilde{C}] \\ = [E(\tilde{r}) - E(\Lambda[\lambda_0 \tilde{z}(\delta_0)])]E(\tilde{C}) + W_C(\delta_0) = 0, \\ \text{where } W_C(\delta) = \text{Cov}(\tilde{r} - \Lambda[\lambda_0 \tilde{z}(\delta)], \tilde{C}).$$

The source of the bias is an extra term that appears in an analogous expression when λ^i varies over players. Restoring superscripts and attending to all sources of variance, the analogue of (5) with heterogeneity of λ^i becomes a question, rather than an equality:

$$(6) \quad m(\delta_0) = \{E^\lambda[E^i(\tilde{r}^i)] - E^\lambda[E^i(\Lambda[\lambda_0 \tilde{z}^i(\delta_0)])]\}E^\lambda[E^i(\tilde{C}^i)] \\ + E^\lambda[W_C^i(\delta_0)] + B_C(\delta_0) \stackrel{?}{=} 0, \\ \text{where } W_C^i(\delta) = \text{Cov}^i[\tilde{r}^i - \Lambda[\lambda_0 \tilde{z}^i(\delta)], \tilde{C}^i], \\ B_C(\delta) = \text{Cov}^\lambda\{[E^i(\tilde{r}^i) - E^i(\Lambda[\lambda_0 \tilde{z}^i(\delta)])], E^i(\tilde{C}^i)\}.$$

In (6), E^i and Cov^i are expectation and covariance with respect to the unconditional steady-state distributions of \tilde{F}^i , \tilde{C}^i , and $\tilde{z}^i(\delta)$ given λ^i , whereas E^λ and Cov^λ are expectation and covariance with respect to the distribution of λ^i . The systematic, signable troublemaker in (6) is the new final term $B_C(\delta_0)$. This is the between-players (that is, across players with idiosyncratic λ^i) steady-state covariance of $E^i(\tilde{r}^i) - E^i(\Lambda[\lambda_0 \tilde{z}^i(\delta)])$ (expected pooled model prediction errors given λ^i) and $E^i[\tilde{C}^i]$ (expected counterfactual reinforcement given λ^i).¹³ The first argument in the Appendix shows that $B_C(\delta)$ is negative for all δ in steady states of the kind discussed here. Hence, satisfaction of the asymptotic pooled model first-order condition (6) has to do battle with a heterogeneity-induced term whose asymptotic expectation is negative regardless of the choice of $\hat{\delta}$ or its true value. This creates the *potential* for an asymptotic downward bias of $\hat{\delta}$.

¹³This is the estimation counterpart of my introductory argument about bias in goodness of fit, that is, that pooled model prediction errors would be correlated with an informative function of lagged choices in the presence of heterogeneity, and that this would give reinforcement and hybrid models an automatic goodness-of-fit advantage.

However, the sum of the first and second terms in (6) is not necessarily zero, even though their analogues in (5) are zero. In (5), the first term is zero, because $E(\tilde{r}) \equiv E(\Lambda[\lambda_0 \tilde{z}(\delta_0)])$, so that the last term in (5) is zero as well (this sensible condition is that prediction errors are orthogonal to the regressor \tilde{C}). However, $E^i(\tilde{r}^i) \equiv E^i(\Lambda[\lambda^i \tilde{z}^i(\delta_0)])$ is nonlinear in λ^i and not equal to $E^i(\Lambda[\lambda_0 \tilde{z}^i(\delta_0)])$ unless $\lambda^i = \lambda_0$, so these conclusions fail in (6). Moreover, exact signs for the first two terms in (6) are elusive, but when δ and λ are estimated together (as they must be in practice) using pooled estimation, the second argument in the Appendix shows that when $\delta_0 = 1$, these terms approach zero as the variances of \tilde{F}^i and \tilde{C}^i (and hence $\tilde{z}^i(\delta)$ for all δ) approach zero. This is not true of $B_C(\delta)$, which is bounded away from zero whenever λ^i has positive variance across subjects and $E^\lambda(E^i[\tilde{z}^i(\delta)]) > 0$. Therefore, when $\delta_0 = 1$, $B_C(\delta)$ will dominate and make the sign of (6) negative if the steady-state variances of \tilde{F}^i and \tilde{C}^i are small enough relative to the between-players variance of λ^i , implying that the probability limit of $\hat{\delta}$ will be less than 1 when $\delta_0 = 1$. Put differently, if cross-sectional heterogeneity is large enough relative to expected steady-state within-player variances of reinforcement and its components, pooled estimation of the EWA model must asymptotically reject the WFP model when it is the true model.

Although this argument does not prove that the bias is always negative, it may play a role in explaining two things. First, Camerer, Ho, and Chong (2002, p. 9) note that “values of $[\hat{\delta}]$ tend to be between 0.5 and 1 in most studies except those in which games have only mixed-strategy equilibria, where $[\hat{\delta}]$ is close to zero.” The foregoing argument suggests one reason for this previously puzzling feature of pooled EWA estimations: Monte Carlo evidence presented subsequently shows that downward heterogeneity bias in $\hat{\delta}$, as well as fit comparisons biased in favor of GRL models relative to the WFP model, can be overwhelming in such games. Second, it also suggests that the disappointing variability of $\hat{\delta}$ across games and studies (evident in Camerer, Ho, and Chong (2002)) may in part be a statistical artefact caused by pooling, heterogeneity, and subtle structural differences between seemingly similar games. Monte Carlo evidence presented subsequently will illustrate this, too, in a collection of stag hunt games.

3. MONTE CARLO STUDIES OF POOLED ESTIMATION

Consider the asymmetric matching pennies game shown in Figure 1. Assume an experimental design in which $M = 40$ subject pairs (40 row players and 40 column players) play this game for $T = 36$ consecutive periods. Also assume the common experimental design in which each row player meets each column player in at most one period of play. Finally, assume that the data-generating process for all players is the EWA model. In what follows, a “simulation” refers to 100 computer-generated samples with these properties.

| | | |
|------------------|---------------------|----------------------|
| | $c=1$ (column left) | $c=0$ (column right) |
| $r=1$ (row up) | (19,0) | (0,1) |
| $r=0$ (row down) | (0,1) | (1,0) |

FIGURE 1.—Asymmetric matching pennies game.

Within each simulation, δ is constant across players and player types (as in the previous section, there is no heterogeneity of δ itself). However, δ varies *between* simulations, taking one value $\delta \in \{1, 0.75, 0.50, 0.25, 0\}$ in each simulation. Moreover, row and column players in all simulations will have the same values (or equivalent distributions of values) of the EWA model parameters ϕ , κ , and η , but different values (or distributions of values) of λ and z_0 (denoted λ_r , λ_c , z_{r0} , and z_{c0}) because these parameters arguably vary with the payoff functions of each player type.

The EWA parameters ϕ , κ , η , λ_r , λ_c , z_{r0} , and z_{c0} are created in three ways, which are summarized in Table I. First, there are high heterogeneity (HIGH) simulations where every player has an idiosyncratic parameter vector drawn from a common distribution just prior to the first period of simulated play. The top half of Table I shows the distributions used and the resulting mean, stan-

TABLE I
DISTRIBUTIONAL ASSUMPTIONS FOR MONTE CARLO SIMULATIONS

| Parameter | Distribution | Parameters of Distribution | Moments of Assumed Distributions | | |
|-------------------------|--------------------|-----------------------------|--------------------------------------|-----------|--------------------------|
| | | | Mean (and Value in NONE Simulations) | Std. Dev. | Coefficient of Variation |
| HIGH simulations | | | | | |
| $1-\phi$ | Beta(a, b) | $a = 1.178, b = 6.676$ | 0.150 | 0.12 | 0.80 |
| κ | Beta(a, b) | $a = 1.30625, b = 11.75625$ | 0.100 | 0.08 | 0.80 |
| λ_r | Gamma(a, b, c) | $a = 1, b = 0.08, c = 0.02$ | 0.100 | 0.08 | 0.80 |
| λ_c | Gamma(a, b, c) | $a = 1, b = 1.52, c = 0.38$ | 1.90 | 1.52 | 0.80 |
| η | Beta(a, b) | $a = 1, b = 1$ (uniform) | 0.50 | 0.289 | 0.577 |
| z_{r0} | $N(\mu, \sigma)$ | $\mu = 15, \sigma = 12$ | 15 | 12 | 0.80 |
| z_{c0} | $N(\mu, \sigma)$ | $\mu = 0, \sigma = 0.60$ | 0 | 0.60 | — |
| LOW simulations | | | | | |
| $1-\phi$ | Beta(a, b) | $a = 5.1625, b = 29.254$ | 0.150 | 0.06 | 0.40 |
| κ | Beta(a, b) | $a = 5.525, b = 49.725$ | 0.100 | 0.04 | 0.40 |
| λ_r | Gamma(a, b, c) | $a = 4, b = 0.02, c = 0.02$ | 0.100 | 0.04 | 0.40 |
| λ_c | Gamma(a, b, c) | $a = 4, b = 0.38, c = 0.38$ | 1.90 | 0.76 | 0.40 |
| η | Beta(a, b) | $a = 5.5, b = 5.5$ | 0.50 | 0.144 | 0.289 |
| z_{r0} | $N(\mu, \sigma)$ | $\mu = 15, \sigma = 6$ | 15 | 6 | 0.40 |
| z_{c0} | $N(\mu, \sigma)$ | $\mu = 0, \sigma = 0.30$ | 0 | 0.30 | — |

Notes: Beta(a, b), Gamma(a, b, c), and $N(\mu, \sigma)$ denote Beta, generalized Gamma, and Normal distributions on $[0, 1]$, $[c, \infty]$, and $[-\infty, \infty]$, respectively, with distribution parameters as given.

standard deviation and coefficient of variation of each parameter in these HIGH populations. The bottom half of Table I shows similar information for low heterogeneity (LOW) simulations. These distributions have the same means as in HIGH simulations, but half as much heterogeneity (as measured by the coefficient of variation, where it exists). Finally, there are NONE simulations, where parameters are constant across all players of each type and equal to the means shown for the HIGH and LOW simulations. These control simulations have no heterogeneity, so pooled estimators should perform relatively well in these samples.

Is the heterogeneity in Table I characteristic of real subject populations? As mentioned earlier, heterogeneity of λ is the main source of bias. Subsequently, I recommend a “mixed random estimator” that estimates the variance of λ in the population using parametric distributional assumptions. Rutström and Wilcox (2005) applied this estimator to data on the asymmetric matching pennies game used here and found a coefficient of variation of λ_r that ranged from 0.33 to 1.00 (depending on the model estimated and the experimental treatment). I have also applied that estimator to Battalio, Samuelson, and Van Huyck’s (2001) stag hunt game data and McKelvey, Palfrey, and Weber’s (2000) asymmetric matching pennies game data: The estimated coefficient of variation of λ in these studies ranges from 0.44 to 0.72 (depending on the game), so the coefficient of variation of λ_r assumed in the LOW and HIGH populations (0.40 and 0.80) lies within a range of prior estimates from several data sets. The coefficient of variation and/or variance of most other parameters is deliberately matched to that of λ_r to compare the degree of heterogeneity bias caused by each parameter (to the extent this is possible).

Table II shows the results of pooled estimations of the EWA model in the simulations. Results are shown for row players only (see footnote 16). In the NONE simulations (the top five rows of Table II), the design and pooled estimator can distinguish between $\delta = 1$ and $\delta = 0$ models fairly well. The pooled MLE $\hat{\delta}$ correctly estimates δ across the 100 samples of each NONE simulation, although there is a noticeable and understandable bias away from bounds when the true δ is at a bound. When $\delta = 1$, WFP fits better than GRL in all 100 samples; EWA fits significantly better than WFP in just 2 of 100 samples; and the mean $\hat{\delta}$ is about 0.91.¹⁴ *Mutatis mutandis*, the same is true when $\delta = 0$: GRL fits better than WFP in 98 of 100 samples; EWA fits significantly better

¹⁴The last three columns of Table II show the number of samples (out of the 100 samples in each simulation) in which one model fits better than another. The WFP and GRL models have the same number of free parameters (four each), so these comparisons simply report the number of samples in which the log likelihood of WFP exceeds that of GRL. However, WFP and GRL both result from imposing two parameter restrictions on the EWA model. Therefore, EWA is said to fit better than WFP or GRL in a sample only if the standard likelihood ratio test statistic (distributed χ^2 with $df = 2$) indicates that these restrictions on EWA are significant at the 5% level.

TABLE II
 POOLED ESTIMATES OF ROW PLAYERS' EWA PARAMETERS AND MODEL FIT COMPARISONS IN VARIOUS SIMULATIONS

| | | Pooled Estimates of EWA Parameters—Mean, Median, and Standard Deviation of Estimates Across 100 Simulated Samples | | | | | | | | | | | | Number of 100 Samples where First Model Outperforms Second (Log Likelihood Comparison) | | |
|------------------------------------|---------------------------------------|---|------|------|------------------------------|------|------|--------------------------------|------|------|-----------------------------------|------|------|--|---------------------|---------------------|
| Heterogeneity in Simulated Samples | True δ in Simulated Population | δ | | | ϕ (True Mean = 0.85) | | | κ (True Mean = 0.10) | | | λ_r (True Mean = 0.10) | | | WFP Better than GRL | EWA Better than WFP | EWA Better than GRL |
| | | Mean | Med. | S.D. | Mean | Med. | S.D. | Mean | Med. | S.D. | Mean | Med. | S.D. | | | |
| NONE | 1.00 | 0.91 | 1.00 | 0.13 | 0.85 | 0.84 | 0.04 | 0.20 | 0.09 | 0.28 | 0.10 | 0.11 | 0.04 | 100 | 2 | 99 |
| | 0.75 | 0.79 | 0.82 | 0.17 | 0.86 | 0.86 | 0.04 | 0.17 | 0.06 | 0.26 | 0.11 | 0.11 | 0.04 | 98 | 8 | 96 |
| | 0.50 | 0.56 | 0.57 | 0.22 | 0.85 | 0.85 | 0.04 | 0.23 | 0.10 | 0.32 | 0.10 | 0.09 | 0.04 | 74 | 27 | 76 |
| | 0.25 | 0.28 | 0.30 | 0.20 | 0.85 | 0.85 | 0.04 | 0.16 | 0.06 | 0.25 | 0.11 | 0.11 | 0.04 | 25 | 69 | 31 |
| | 0.00 | 0.08 | 0.02 | 0.12 | 0.84 | 0.84 | 0.04 | 0.23 | 0.10 | 0.31 | 0.10 | 0.10 | 0.04 | 2 | 95 | 12 |
| LOW | 1.00 | 0.58 | 0.58 | 0.25 | 0.86 | 0.87 | 0.04 | 0.29 | 0.18 | 0.31 | 0.09 | 0.08 | 0.05 | 66 | 34 | 68 |
| | 0.75 | 0.44 | 0.42 | 0.25 | 0.87 | 0.87 | 0.04 | 0.20 | 0.10 | 0.24 | 0.10 | 0.11 | 0.04 | 54 | 52 | 55 |
| | 0.50 | 0.24 | 0.22 | 0.20 | 0.89 | 0.89 | 0.04 | 0.19 | 0.07 | 0.25 | 0.11 | 0.11 | 0.05 | 23 | 76 | 40 |
| | 0.25 | 0.13 | 0.06 | 0.17 | 0.87 | 0.87 | 0.04 | 0.18 | 0.08 | 0.27 | 0.11 | 0.11 | 0.05 | 7 | 92 | 21 |
| | 0.00 | 0.02 | 0.00 | 0.06 | 0.87 | 0.87 | 0.05 | 0.18 | 0.10 | 0.23 | 0.10 | 0.10 | 0.04 | 0 | 99 | 13 |
| HIGH | 1.00 | 0.07 | 0.00 | 0.13 | 0.92 | 0.92 | 0.04 | 0.17 | 0.08 | 0.23 | 0.12 | 0.11 | 0.06 | 2 | 96 | 17 |
| | 0.75 | 0.03 | 0.00 | 0.08 | 0.91 | 0.91 | 0.04 | 0.20 | 0.11 | 0.26 | 0.11 | 0.10 | 0.06 | 0 | 100 | 22 |
| | 0.50 | 0.02 | 0.00 | 0.08 | 0.91 | 0.91 | 0.04 | 0.16 | 0.07 | 0.24 | 0.12 | 0.12 | 0.05 | 2 | 99 | 28 |
| | 0.25 | 0.00 | 0.00 | 0.01 | 0.91 | 0.91 | 0.04 | 0.12 | 0.07 | 0.16 | 0.11 | 0.11 | 0.05 | 0 | 100 | 29 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.90 | 0.90 | 0.04 | 0.17 | 0.10 | 0.23 | 0.11 | 0.11 | 0.05 | 0 | 100 | 35 |
| HIGH[ϕ] | 1.00 | 0.72 | 0.71 | 0.24 | 0.89 | 0.89 | 0.04 | 0.25 | 0.14 | 0.27 | 0.10 | 0.09 | 0.05 | 84 | 33 | 86 |
| HIGH[κ] | 1.00 | 0.84 | 0.92 | 0.19 | 0.85 | 0.85 | 0.04 | 0.21 | 0.06 | 0.30 | 0.10 | 0.11 | 0.05 | 99 | 7 | 98 |
| HIGH[η] | 1.00 | 0.91 | 1.00 | 0.13 | 0.86 | 0.86 | 0.04 | 0.22 | 0.07 | 0.31 | 0.10 | 0.11 | 0.05 | 100 | 4 | 100 |
| HIGH[λ_r] | 1.00 | 0.20 | 0.13 | 0.23 | 0.89 | 0.89 | 0.04 | 0.21 | 0.11 | 0.27 | 0.10 | 0.10 | 0.05 | 20 | 86 | 38 |
| HIGH[z_0] | 1.00 | 0.83 | 0.86 | 0.19 | 0.84 | 0.86 | 0.03 | 0.25 | 0.11 | 0.32 | 0.10 | 0.10 | 0.05 | 96 | 8 | 94 |

than GRL in just 12 of the 100 samples; and the mean $\hat{\delta}$ is about 0.08. Hence the design used in the simulations does not suffer hugely from power problems (Salmon (2001)), at least if one means the design to discriminate “pure” belief learning and reinforcement learning models.¹⁵

Table II also shows that the mean $\hat{\kappa}$ is upward-biased across all simulations (regardless of their type or their true value of δ). Mean and median $\hat{\kappa}$ are generally quite distinct, and the standard deviation of $\hat{\kappa}$ across samples exceeds that of other parameter estimates. That is, $\hat{\kappa}$ is quite variable and, since the true $\kappa = 0.1$ is near its lower bound, this is manifest in highly skewed distributions of estimates and hence a substantial upward bias of mean estimates. Although not shown in Table II, $\hat{\eta}$ is also quite variable. The variability of $\hat{\kappa}$ and $\hat{\eta}$ probably occurs because most useful sample information about them is provided early in each subject’s time series: N_t reaches a constant value equal to $[1 - (1 - \kappa)\phi]^{-1}$ and in practice this occurs rather quickly. It is only before this that information directly useful to separately identifying κ and η is to be had. For similar reasons, Ichimura and Bracht (2001) find that increased numbers of players (larger M) is generally much more helpful than longer time series (larger T) for estimating EWA with precision in nonheterogeneous simulations (see also Cabrales and Garcia-Fontes (2000)). Many researchers could live with bias in $\hat{\kappa}$, though: Interest centers on δ and, in the NONE simulations, the mean $\hat{\delta}$ is not misleading in any scientifically significant way.

Unfortunately, this is not true in simulations with heterogeneity. Table II shows that in LOW simulations, $\hat{\delta}$ is dramatically downward-biased: For true $\delta = 1, 0.75, 0.50,$ and 0.25 , the corresponding mean $\hat{\delta}$ are 0.46, 0.33, 0.29, and 0.05, respectively. In the HIGH simulations, the bias in $\hat{\delta}$ is simply shocking: No mean estimate exceeds 0.07, regardless of the true value of δ , and even in the $\delta = 1$ simulation, $\hat{\delta}$ is zero in 77 of 100 samples.¹⁶ The bias also plagues goodness-of-fit comparisons. To see this, consider the $\delta = 1$ simulations where WFP should fit much better than GRL. As noted earlier, this is the case in all 100 of the NONE simulation samples with $\delta = 1$. However, in the LOW simulation with $\delta = 1$, WFP fits better than GRL in 66 of 100 samples. In the HIGH simulations with $\delta = 1$, the goodness-of-fit comparison has been fully reversed: Here, WFP fits better than GRL in only 2 of 100 samples. Additionally, the relative fit of EWA does not consistently hint at the truth: In the HIGH simulation with $\delta = 1$, EWA fits significantly better than GRL in just 17 of 100 samples. In a highly heterogeneous population, for this game,

¹⁵In fact, Rutström and Wilcox (2001) arrived at this design under just this intention, using Monte Carlo methods that resemble those used here to choose a game, and using M and T capable of producing high power of simple two-sample nonparametric tests to detect a wholesale shift between reinforcement and belief learning models.

¹⁶The bias in $\hat{\delta}$ is similar, though not quite as strong, for column players. For true values of $\delta = 1, 0.75, 0.5, 0.25,$ and 0 , mean values of $\hat{\delta}$ are 0.29, 0.14, 0.04, 0.00, and 0.00, respectively, in HIGH simulations.

goodness-of-fit comparisons based on pooled estimators would almost always lead us to conclude that GRL was a better model than WFP, even when the true data-generating process (EWA with $\delta = 1$ and mean $\kappa = 0.1$) is very similar to the WFP model for every player in the population!

Heterogeneity of λ_r is the most serious source of bias, but variations in ϕ significantly contribute to bias as well. The last five rows of Table II show pooled estimation results in five new simulations, all with true $\delta = 1$. Each of these simulations draws a single parameter $\beta \in \{\phi, \kappa, \eta, \lambda, z_0\}$ from its HIGH distribution for each player while all other parameters are fixed at their NONE values shown in Table I. These simulations are denoted HIGH[β] in Table II. Notice that little bias results from heterogeneity in either η or z_0 : These determine initial conditions only and so are transitory sources of heterogeneity. Only permanent sources of heterogeneity (variability in ϕ , κ , and λ_r) should matter much in games that are repeated for many periods.¹⁷ Variability of κ causes little bias in $\hat{\delta}$, but variability of ϕ causes a noticeable bias and variability in λ_r is clearly the major source of bias.

Now consider nine stag hunt games of the form shown in Figure 2. This will show that heterogeneity bias in $\hat{\delta}$ is found even in games with no stable interior mixed-strategy Nash equilibrium, and reveals that the strength of the bias varies quite substantially across seemingly similar games. If a player believes her opponent will play X with probability q , the difference between the player's payoff from playing X versus Y is $r(q) = (45 + \pi_2 - \pi_1)(q - q^*)$, where $q^* = \pi_2 / (45 + \pi_2 - \pi_1)$ is the interior (unstable) mixed-strategy Nash equilibrium of the game. Battalio, Samuelson, and Van Huyck (2001) call $r(q)$ the *optimization premium* and call $(45 + \pi_2 - \pi_1)$ the *optimization premium parameter* in these games. Methodologically, $|r(q)|$ could be called the opportunity cost of misbehavior or foregone expected income (Harrison (1989)) given belief q .

Figure 3 shows nine versions of the game. Games in the same column have the same optimization premium parameter (falling from left to right), while games in the same row have the same q^* (falling from top to bottom). The center row contains Battalio, Samuelson, and Van Huyck's (2001) three games. Given any fixed $q < 0.6$, $|r(q)|$ falls from top to bottom as well as from left to

| | | |
|---|---------------|-----------------------|
| | X | Y |
| X | (45,45) | (0, π_1) |
| Y | (π_1 ,0) | (π_2 , π_2) |

FIGURE 2.—Stag hunt game form.

¹⁷However, massive heterogeneity of initial conditions could obviously have lasting effects, as in Kitzis, Kelley, Berg, Massaro, and Friedman (1998).

Optimization Premium Parameter

50

25

15

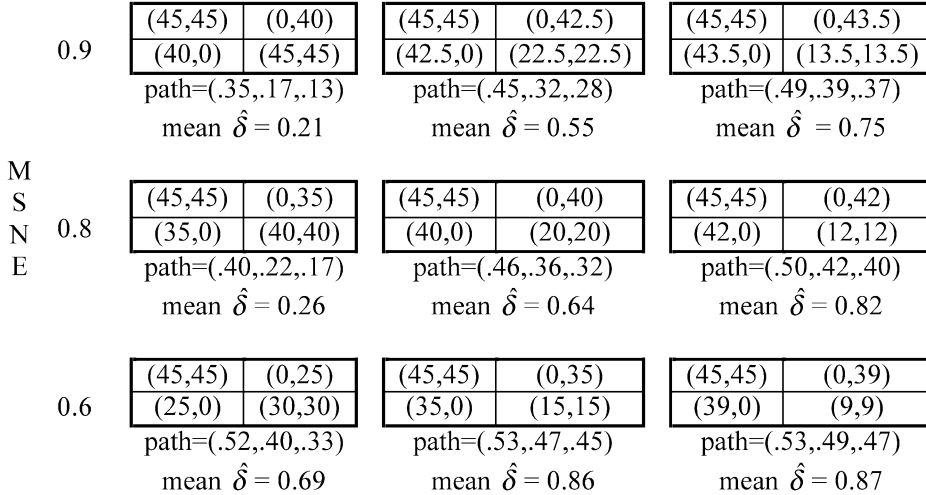


FIGURE 3.—Mean adjustment path and $\hat{\delta}$ in various stag hunt game simulations with $\delta = 1$.

right: Put differently, optimization incentives for belief-based play are greatest in the upper left game and smallest in the lower right game, given the same belief q across games. Below each game, Figure 3 shows two things: An adjustment path ($P_{1-12}, P_{13-24}, P_{25-36}$), where $P_{t,t+1}$ are 12-period average proportions of X choices across the one-hundred 36-period samples in each simulation, and the mean $\hat{\delta}$ in each simulation (true $\delta = 1$ in these simulations).¹⁸ First, note the large differences in the strength of heterogeneity bias across these games. This suggests that a part of the disappointing variability of $\hat{\delta}$ across games and studies may also result from subtle structural differences between seemingly similar games that mediate the magnitude of heterogeneity bias. In the upper left game, where optimization incentives are greatest, the bias is maximal: The mean $\hat{\delta}$ is 0.21. In the lower right game, where optimiza-

¹⁸Simulations are mostly identical to the HIGH simulations in Table I. However, the payoff range here is roughly double that in the asymmetric matching pennies game, so the λ distribution is given a mean of 0.05 rather than 0.10. Using McKelvey, Palfrey, and Weber's (2000) data on asymmetric matching pennies games where payoff ranges are varied systematically across treatments, the "mixed random" estimator I will recommend later supports this rough inverse proportionality of mean λ and payoff ranges within a given subject population. Additionally, the mean value of z_0 is set so that the expected probability of X in period 1 is 0.55—below the separatrix for all nine games, so that play is expected to evolve toward the pure Nash equilibrium (Y, Y) regardless of the size of λ .

tion incentives are weakest, there is hardly any bias: The mean $\hat{\delta}$ is 0.87 (nearly as close to the true $\delta = 1$ as in NONE simulations of the asymmetric matching pennies game). Notice that in this latter game the path is very flat, never distant from 0.6, and always near 0.5: Weak optimization incentives yield a very slow adjustment path toward the risk-dominant pure Nash equilibrium (Y, Y) , and $|z_t|$ stays relatively small most of the time. As a result, lagged choices carry relatively little useful information here about idiosyncratic λ^i , so heterogeneity bias is weak.¹⁹ Experimentalists will quickly see the apparent design quandary: Weak optimization incentives that keep choice probabilities relatively close to indifference can minimize heterogeneity bias.²⁰

The next section shows that even misspecified random parameters estimators can greatly reduce heterogeneity bias, so this unpleasant choice between weak incentives and biased estimation may not be necessary. First, though, note that various robustness checks and approaches to certain forms of heterogeneity already appear in the literature on learning in games, and these may mitigate the heterogeneity bias discussed here (although none was meant to do this). For instance, Camerer and Ho (1999) examine out-of-sample fits to guard against overfitting and also examine “latent class” mixture models that capture discrete forms of heterogeneity. Cooper and Stockman (2002) use first-order autoregressive specifications to allow for high-frequency persistence not explicitly treated by learning models. Other econometric approaches might help, too. A two-step approach to estimating δ should be asymptotically consistent (Wooldridge (2002)), and estimation on a subject-by-subject basis avoids pooling and is also asymptotically consistent, but the length of the time series required for such consistency to be of practical importance is very long indeed, much longer than existing or even imaginable experiments (Cabrales and Garcia-Fontes (2000)). Wilcox (2005) discusses all of these options at length and shows that none of them mitigates heterogeneity bias as well as the random parameters approach described in the next section.

¹⁹Although the [Appendix](#) analyzes steady-state asymptotics rather than finite-sample estimation along adjustment paths, (11) and (13) provide some intuition. When $\delta_0 = 1$, the size of these derivatives with respect to λ^i is mostly governed by the size of the term $A[\lambda^i z^i(\delta_0)](1 - A[\lambda^i z^i(\delta_0)])z^i(\delta_0)$. This term is maximized at $\lambda^i z^i(\delta_0) \approx 1.54$ (probability of playing Y of roughly 0.82), is zero when $z^i(\delta_0)$ equals 0 (probability of playing Y of 0.5), and is close to zero when $z^i(\delta_0)$ is very large (probability of playing Y near 1). Thus, the between-player covariance $B_C(\delta_0)$ is near zero when choice probabilities are mostly near indifference or, alternatively, mostly close to certainty.

²⁰This fact does *not* imply that symmetric matching pennies would be a good game choice. As Ichimura and Bracht (2001) point out, choice probabilities that are not significantly different from 0.5 create an identification problem, because $\lambda = 0$ will fit such data well and when $\lambda = 0$, no other parameters of EWA are identified. As is well known (see, e.g., Goeree and Holt (2001)), symmetric matching pennies happens to be a game where mixed-strategy Nash equilibrium predicts behavior well; therefore, this identification issue is crucial in that game. In the stag hunt game, the distinct trend in behavior is poorly explained by $\lambda = 0$, so the identification issue does not arise.

4. A SOLUTION

As mentioned earlier, no general solution to this particular brand of heterogeneity bias is known that is wholly free of distributional assumptions about unobserved heterogeneity (Wooldridge (2005)). Given prior knowledge of the underlying distributional form that governs parameter heterogeneity, a random parameters estimator based on that distributional form is asymptotically consistent and permits valid asymptotic inferences given an appropriate experimental design (Cabrales and Garcia-Fontes (2000)).²¹ However, finite sample properties could be poor, and misspecified distributions in random parameter estimators may also lead to serious biases (Heckman and Singer (1984)). Almost surely, distributional assumptions will be wrong and samples are certainly finite. Therefore, we need to know whether misspecified random parameters estimators perform well in finite samples that resemble practical experimental designs.

To examine this, I compare the performance of three random parameters estimators, all deliberately misspecified, on the five HIGH simulations of the asymmetric matching pennies game. The first estimator, the “lognormal λ ” estimator, is based on the assumption that $\ln(\lambda_r) \sim N(\mu_\lambda, \sigma_\lambda)$ in the underlying population while other parameters are constant.²² The second estimator assumes that all EWA parameters are constant across players, but also assumes that z_t differs across players i due to a time-invariant random effect $\alpha \sim N(0, \sigma_\alpha)$ so that $z_t^i = z_t + \alpha^i$ for each player i —an *ad hoc* “random effects” estimator. The third estimator, which I call a “mixed random” estimator blends these two estimators.²³ These random parameter estimators are implemented in the usual way: The EWA likelihood for each player is numerically integrated with respect to the assumed distribution of $\ln(\lambda_r)$ or α (both for the mixed random estimator).²⁴ (The log of this is summed across players in each sample, and this sum is maximized in $\delta, \phi, \kappa, \eta, z_0$, and the parameters of the distribution(s) assumed under each estimator.)

²¹When there is heterogeneity of parameters that determine initial conditions, such as η and z_0 , there is of course still the “initial conditions” problem in small T samples, even when permanent sources of heterogeneity are integrated out of the likelihood. Scattered Monte Carlo evidence suggests, however, that T may not need to be too large for this to be negligible (e.g., Heckman (1981)), and the HIGH[η] and HIGH[z_0] reported in Table II suggest this for EWA.

²²In applications, a generalized (three parameter) gamma distribution for λ is a good choice because this can take a wider variety of shapes (including near symmetric ones) than the lognormal distribution (but Rutström and Wilcox (2005) found that this alternative specification made little difference to estimates of δ). Because this is the distribution that generates λ in my simulations and because I want to examine misspecified estimators, I do not use it here.

²³The two distributions are assumed to be independent here. However, a bivariate normal distribution could easily be used to relax this assumption.

²⁴Numerical integration is by 14-node Gauss–Hermite quadratures for normal and lognormal distributions; spot checks of these estimations using 20-node quadratures reveal no noticeable differences in parameter estimates.

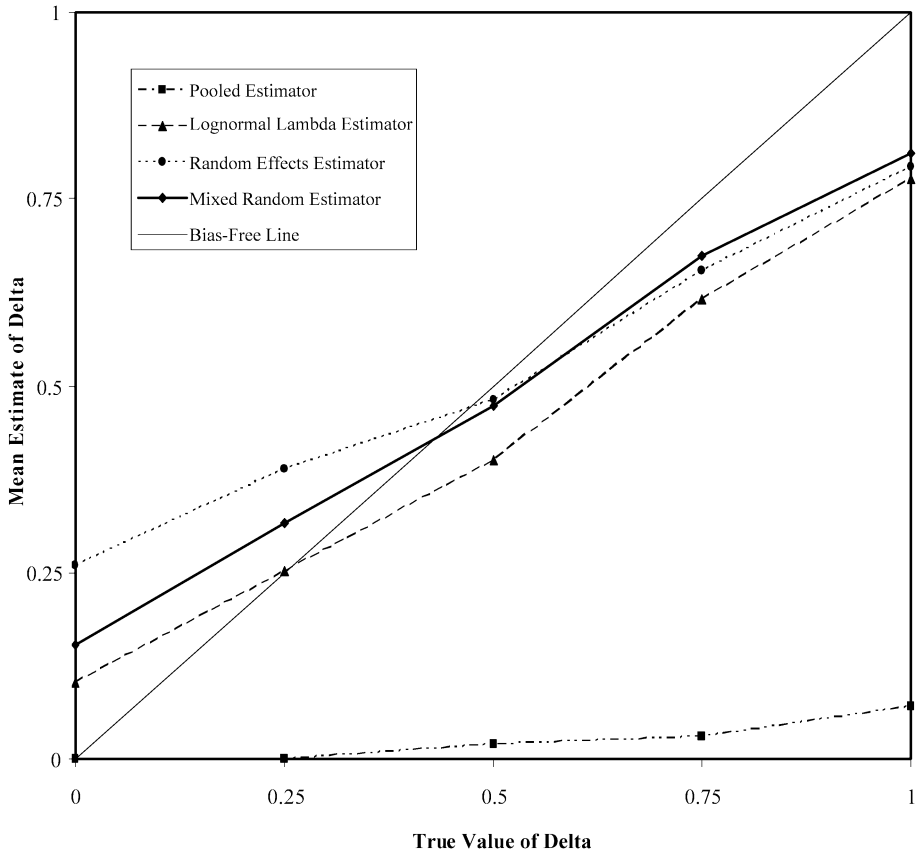


FIGURE 4.—Bias characteristics of three random parameter estimators in HIGH simulations.

Figure 4 illustrates estimation results in a unit square. Horizontal coordinates are true values of δ and vertical coordinates are mean estimates of δ across 100 simulated samples using each estimator. Two lines in the square are for reference: The diagonal line represents a bias-free relationship and the line near the bottom of the square, marked with square icons, represents the pooled estimator results. The other three lines represent the behavior of the three estimators described previously, and all improve greatly on the pooled estimator. Notice that the random effects and mixed random estimators are both nearly skew-symmetric about the square's center. If we are forced to choose from various biased estimators and the issue at hand is whether δ is closer to zero or one, this skew-symmetry is a good property (other things, such as magnitude of bias, equal). The lognormal λ estimator does not have this property: It is downward-biased even at $\delta = 0.5$. The mixed random estimator is the best of the bunch, having both the desirable skew-symmetry of the random effects esti-

mator and less bias than it. The mixed random estimator performs even better in HIGH simulations of the stag hunt game with $q^* = 0.9$ and optimization premium parameter 50, the game where heterogeneity bias was strongest: Mean estimates of δ , given the true values 1.0, 0.5, and 0.0, are 0.89, 0.50, and 0.038, respectively.

5. CAVEATS

The mixed random estimator probably performs well because the random effect α picks up relatively persistent heterogeneity in $z_i(\delta)$ caused by parameters like ϕ (a relatively minor problem) whereas the lognormal distribution of λ_r picks up the (relatively serious) heterogeneity of λ . However, heterogeneity can be more complex than that found in HIGH populations. Parameters may be correlated in the subject population. Estimators based on bivariate distributions may account for this well, but I offer no evidence on this. Additionally, some parameters that determine z_i may have sharply multimodal distributions. For instance, a population might be a mixture of Cournot best responders ($\phi = 0$), classical fictitious players ($\phi = 1$), and weighted fictitious players with ϕ in some relatively small interval strictly between these extremes. A hybrid of Camerer and Ho's (1999) latent class approach and the lognormal λ estimator might work best in such a case. It is certainly not clear at all that a smooth, unimodal distribution of fixed effects α will adequately capture sharply multimodal heterogeneity like this. I examined simulations of this kind and found that the mixed random estimator consistently *overestimates* δ by about 0.10 in such simulations (except when $\delta = 1$). Therefore, I stress that the mixed random estimator is not a wholly reliable fix for heterogeneity bias. Yet it outperforms all other solutions I have examined and is much less misleading than pooled estimation or individual estimation (Wilcox (2005)). Obviously, random parameter approaches will convince more readers if users show that conclusions are robust to several alternative distributional assumptions.

The precision of the random parameter estimators of δ leaves much to be desired: Their standard deviations are distressingly large in the simulated samples (0.20 to 0.35, depending on the estimator and true δ ; see also Cabrales and Garcia-Fontes (2000)). Yet similar problems characterize pooled estimation in homogenous populations (Salmon (2001)), so this does not distinguish random parameters estimators from pooled ones (particularly given heterogeneity bias). It will probably take very large numbers of subjects or meta analysis of data from many experiments to get high precision estimates of δ for the EWA model regardless of the estimator.

6. CONCLUSIONS

Heterogeneity bias is a serious problem in empirical comparisons of learning models. This is true whether comparisons are based on goodness-of-fit

comparisons or parametric evaluation within hybrid models like EWA. All results here suggest that this bias has a known direction. When heterogeneity is present (especially heterogeneity of λ), estimations and fit comparisons that ignore it unduly favor reinforcement models over belief models. The bias can be dramatic and overwhelming, although it can also be quite weak when optimization incentives are relatively weak over most periods. Fortunately, even misspecified random parameter estimators appear to be much less biased than pooled estimators. Although I do not examine this here, demographic variables or more theoretically grounded conditioning variables might instrument some heterogeneity; this could reduce the inferential burden placed on distributional assumptions in random parameters estimators and might improve their precision as well.

We experimentalists must remember that in *any* pooled estimation where models contain lagged dependent variables, a form of the bias described here will be present (although its seriousness and character will vary across contexts). Experiments on intertemporal optimization and the conditional cooperator hypothesis both come to mind as other settings where (depending on model specification and design) crucial questions of interest can boil down to the relationship between current decisions and lagged decisions or a function of them. Classical experimental design techniques can finesse heterogeneity in some cases (this is what randomized assignment is for), and in experimental economics we induce certain parts of preferences and beliefs, hence eliminating heterogeneity in them. There are probably limits to what experimental design and method alone can achieve when models contain lagged dependent variables and subjects are heterogeneous. In addition, heterogeneity is ubiquitous. Risk attitudes, other-regarding preferences, memory, sensitivity to payoff differences, problem-solving ability, depth of reasoning, intrinsic motivation, and other uncontrolled subjective motivations are among those things that may vary across subjects, and almost everyone believes that one or more of those things do vary.

Dept. of Economics, University of Houston, 4800 Calhoun Road, Houston, TX 77204-5019, U.S.A.; nwilcox@mail.uh.edu, www.uh.edu/~nwilcox.

Manuscript received December, 2003; final revision received March, 2006.

APPENDIX

First Argument

Let σ_{FF}^i , σ_{CC}^i , and $\sigma_{zz}^i(\delta)$ be variances of \tilde{F}^i , \tilde{C}^i , and $\tilde{z}^i(\delta)$, respectively, given λ^i . Assume that these variances are small enough that approximations described in the arguments below are adequate. Let $F^i \equiv E^i(\tilde{F}^i)$, $C^i \equiv E^i(\tilde{C}^i)$, $z^i(\delta) \equiv E^i[\tilde{z}^i(\delta)]$, and $P^i(\lambda, \delta) \equiv E^i(\Lambda[\lambda\tilde{z}^i(\delta)])$. Let steady-state unconditional expectations of $\pi(1, c_t)$ and $\pi(0, c_t)$ be π_1 and π_0 , respectively. Assume

that row's strategy choice labels are chosen so that $\pi_1 - \pi_0 > 0$ (so that $r = 1$ is more frequent in any steady state), implying that $z^i(\delta) > 0 \forall i$, and also assume that payoffs have been scaled so that π_1 and π_0 are nonnegative, implying that $\pi_1 + \pi_0 > 0$. Assume that $\tilde{z}^i(\delta)$ is finite, implying that $E^i(\Lambda[\lambda\tilde{z}^i(\delta)]) < 1 \forall i$. Also assume that the distribution of $\tilde{z}^i(\delta)$ is not extremely right-skewed for any i , so that $E^i(\Lambda[\lambda\tilde{z}^i(\delta)]) > 0.5 \forall i$. Under these assumptions, we have $1 > E^i(\Lambda[\lambda\tilde{z}^i(\delta)]) > 0.5 \forall i$, the case discussed in the text.

Also assume that

$$\beta^i(\lambda, \delta) = \gamma(1 - \delta)\lambda\Lambda[\lambda z^i(\delta)](1 - \Lambda[\lambda z^i(\delta)])(\pi_1 + \pi_0) < 1$$

$\forall \lambda$ in the support of λ^i and $\delta \in [0, 1]$, where $\gamma = [1 - (1 - \kappa)\phi]/(1 - \phi)$. This is a necessary condition for local steady-state stability of $z^i(\delta)$ at all true and estimated values of δ , because in that case it is approximately true (given small $\sigma_{zz}^i(\delta)$) that

$$E[z_{t+1}^i(\delta)|z_t^i(\delta)] - z_t^i(\delta) \approx (1 - \phi)[1 - \beta^i(\lambda, \delta)][z^i(\delta) - z_t^i(\delta)].$$

Thus if $\beta^i(\lambda, \delta) \geq 1$, $E[z_{t+1}^i(\delta)|z_t^i(\delta)]$ would not move toward $z^i(\delta)$ for small $z^i(\delta) - z_t^i(\delta)$.

Assuming that σ_{FF}^i , σ_{CC}^i , and $\sigma_{zz}^i(\delta)$ are small enough, approximately correct reasoning about steady-state expected values of nonlinear functions of $\tilde{z}^i(\delta)$ can be done in terms of $z^i(\delta)$ rather than the random variable $\tilde{z}^i(\delta)$. If so, F^i , C^i , and $z^i(\delta)$ are implicitly and approximately

$$(7) \quad F^i \approx \gamma[\Lambda[\lambda^i z^i(\delta)]\pi_1 - (1 - \Lambda[\lambda^i z^i(\delta_0)])\pi_0] > 0,$$

$$(8) \quad C^i \approx \gamma[(1 - \Lambda[\lambda^i z^i(\delta_0)])\pi_1 - \Lambda[\lambda^i z^i(\delta_0)]\pi_0],$$

$$(9) \quad z^i(\delta) \approx \gamma[(\delta + (1 - \delta)\Lambda[\lambda^i z^i(\delta_0)])\pi_1 - (1 - (1 - \delta)\Lambda[\lambda^i z^i(\delta_0)])\pi_0] > 0.$$

Evaluate (9) at δ_0 and differentiate with respect to λ^i to approximate

$$(10) \quad \frac{\partial z^i(\delta_0)}{\partial \lambda^i} \approx \frac{\beta^i(\lambda^i, \delta_0)}{\lambda^i[1 - \beta^i(\lambda^i, \delta_0)]} z^i(\delta_0) > 0.$$

One may then substitute (10) into derivatives of (7), (8), and (9) to get

$$(11) \quad \frac{\partial C^i}{\partial \lambda^i} = -\frac{\partial F^i}{\partial \lambda^i} \approx -\frac{\gamma\Lambda[\lambda^i z^i(\delta_0)](1 - \Lambda[\lambda^i z^i(\delta_0)])(\pi_1 + \pi_0)}{1 - \beta^i(\lambda^i, \delta_0)} z^i(\delta_0) < 0,$$

$$(12) \quad \frac{\partial z^i(\delta)}{\partial \lambda^i} \approx \frac{\gamma(1 - \delta)\Lambda[\lambda^i z^i(\delta_0)](1 - \Lambda[\lambda^i z^i(\delta_0)])(\pi_1 + \pi_0)}{1 - \beta^i(\lambda^i, \delta_0)} z^i(\delta_0) \geq 0$$

(with equality if and only if $\delta = 1$).

Continue to assume small $\sigma_{zz}^i(\delta)$, so that

$$(13) \quad \frac{\partial[E^i(\tilde{r}^i) - E^i(\Lambda[\lambda_0 \tilde{z}^i(\delta)])]}{\partial \lambda^i} \approx \frac{\Lambda[\lambda^i z^i(\delta_0)](1 - \Lambda[\lambda^i z^i(\delta_0)])[1 - \beta^i(\lambda_0, \delta)]}{1 - \beta^i(\lambda^i, \delta_0)} z^i(\delta_0) > 0.$$

The covariance of $E^i(\tilde{r}^i) - E^i(\Lambda[\lambda_0 \tilde{z}^i(\delta)])$ and C^i with respect to the distribution of λ^i is $B_C(\delta)$. According to (11) and (13), those two terms move in opposite directions with changes in λ^i , so their covariance with respect to the distribution of λ^i is negative, completing the first argument.

Second Argument

Suppose δ_0 and λ_0 are estimated simultaneously, using pooled MLE estimators $\hat{\delta}$ and $\hat{\lambda}$. Assume that $\text{plim}(\hat{\lambda}) = \lambda^*$ exists and that $\lambda^* > 0$ (necessary for asymptotic identification of δ). Then also assume that $\text{plim}(\hat{\delta}) = \delta^*$ exists. Finally, assume that $\delta_0 = 1$. By generalizing (4) and (6), and constructing similar equations that pertain to the first-order condition with respect to λ as well, we have the following asymptotic system under these assumptions:

$$(14) \quad m_{\delta}(\delta^*, \lambda^*) = \Delta P(\delta^*, \lambda^*) E^{\lambda}(C^i) + E^{\lambda}[W_C^i(\delta^*, \lambda^*)] + B_C(\delta^*, \lambda^*) = 0$$

(≤ 0 or ≥ 0 if $\delta^* = 0$ or 1 , respectively),

$$(15) \quad m_{\lambda}(\delta^*, \lambda^*) = \Delta P(\delta^*, \lambda^*) E^{\lambda}[z^i(\delta^*)] + E^{\lambda}[W_z^i(\delta^*, \lambda^*)] + B_z(\delta^*, \lambda^*) = 0,$$

where $\Delta P(\delta, \lambda) = E^{\lambda}[E^i(\tilde{r}^i)] - E^{\lambda}[E^i(\Lambda[\lambda \tilde{z}^i(\delta)])]$,

$W_C^i(\delta, \lambda) = \text{Cov}^i[\tilde{r}^i - \Lambda[\lambda \tilde{z}^i(\delta)], \tilde{C}^i]$,

$W_z^i(\delta, \lambda) = \text{Cov}^i[\tilde{r}^i - \Lambda[\lambda \tilde{z}^i(\delta)], \tilde{z}^i(\delta)]$,

$B_C(\delta, \lambda) = \text{Cov}^{\lambda}\{[E^i(\tilde{r}^i) - E^i(\Lambda[\lambda \tilde{z}^i(\delta)])], C^i\}$,

$B_z(\delta, \lambda) = \text{Cov}^{\lambda}\{[E^i(\tilde{r}^i) - E^i(\Lambda[\lambda \tilde{z}^i(\delta)])], z^i(\delta)\}$.

Now suppose that $\delta^* = 1$, contrary to what is claimed in the text. By definition, $\tilde{z} \equiv \tilde{z}^i(1)$ is independent of λ^i , implying $B_z(1, \lambda) = 0 \forall \lambda$. (However, C^i and $\lambda^i \tilde{z}$ still vary with λ^i when $\delta^* = 1$, so $B_C(1, \lambda^*) < 0$ is still true.) Making these substitutions into (14) and (15), (15) then implies that $\Delta P(1, \lambda^*) = -E^{\lambda}[W_z^i(1, \lambda^*)]/E(\tilde{z})$. Therefore, (14) may be rewritten as

$$(16) \quad m_{\delta}(1, \lambda^*) = -E^{\lambda}[W_z^i(1, \lambda^*)] E^{\lambda}(C^i) / E(\tilde{z}) + E^{\lambda}[W_C^i(1, \lambda^*)] + B_C(1, \lambda^*) = 0.$$

Note that $E^\lambda(C^i)/E(\tilde{z})$ and $B_C(1, \lambda^*) < 0$ are largely unaffected by within-player variances σ_{CC}^i and $\sigma_{zz}(1)$: In particular, because $B_C(1, \lambda^*)$ is a between-players covariance of player-specific expected values, it does not approach zero as the within-player variances σ_{CC}^i and $\sigma_{zz}(1)$ approach zero. By contrast, $E^\lambda[W_z^i(1, \lambda^*)]$ and $E^\lambda[W_C^i(1, \lambda^*)]$ converge to zero as $\sigma_{zz}(1)$ and σ_{CC}^i approach zero. So if heterogeneity is large enough relative to the expected steady-state within-player variances of reinforcement and its components, $B_C(1, \lambda^*) < 0$ will imply that (16) is contradicted, so that $\delta^* < 1$ as claimed in the text.

REMARK: Note that although $E^\lambda(C^i)/E(\tilde{z})$ is bounded above by $1/2$, it can be made arbitrarily large and negative by choosing $\gamma(\pi_1 - \pi_0) \equiv E(\tilde{z}) > 0$ arbitrarily close to zero, and enough of the λ^i large enough to make $E^\lambda(C^i) < 0$; and it is apparently impossible to sign $E^\lambda[W_z^i(1, \lambda^*)]$. This is another reason (additional to footnote 19) why downward bias of $\hat{\delta}$ might be mitigated and even reversed when $E(\tilde{z}) \approx 0$ (when and if $E^\lambda[W_z^i(1, \lambda^*)] > 0$ as well). Yet as discussed in footnote 20, parameters of EWA are then poorly identified, so, arguably, good experimental designs will try to avoid games where $E(\tilde{z}) \approx 0$.

REFERENCES

- AHN, S. C., AND P. SCHMIDT (1995): "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, 68, 5–27. [1274]
- ANDERSON, T. W., AND C. HSIAO (1982): "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, 67–82. [1274]
- BALLINGER, T. P., AND N. T. WILCOX (1997): "Decisions, Error and Heterogeneity," *Economic Journal*, 107, 1090–1105. [1274]
- BATTALIO, R., L. SAMUELSON, AND J. VAN HUYCK (2001): "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games," *Econometrica*, 69, 749–764. [1279,1282]
- BROSETA, B. (2000): "Adaptive Learning and Equilibrium Selection in Experimental Coordination Games: An ARCH(1) Approach," *Games and Economic Behavior*, 32, 25–50. [1273]
- CABRALES, A., AND W. GARCIA-FONTES (2000): "Estimating Learning Models from Experimental Data," Unpublished Manuscript, Universitat Pompeu Fabra. [1273,1275,1281,1284,1285,1287]
- CAMERER, C., AND T.-H. HO (1999): "Experience Weighted Attraction Learning in Normal-Form Games," *Econometrica*, 67, 827–874. [1272,1273,1284,1287]
- CAMERER, C., T.-H. HO, AND J.-K. CHONG (2002): "Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games," *Journal of Economic Theory*, 104, 137–188. [1273,1277]
- CAMERER, C., T.-H. HO, AND X. WANG (1999): "Individual Differences and Payoff Learning in Games," Unpublished Manuscript, University of Pennsylvania. [1273]
- CHEUNG, Y., AND D. FRIEDMAN (1997): "Individual Learning in Games: Some Laboratory Results," *Games and Economic Behavior*, 19, 46–76. [1272,1273]
- COOPER, D., AND C. STOCKMAN (2002): "Fairness and Learning: An Experimental Examination," *Games and Economic Behavior*, 41, 26–45. [1284]
- DUFFY, J., AND J. OCHS (2003): "Cooperative Behavior and the Frequency of Social Interaction," Unpublished Manuscript, University of Pittsburgh. [1275]
- EREV, I., AND A. E. ROTH (1998): "Predicting how People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88, 848–881. [1272,1273,1275]

- FUDENBERG, D., AND D. K. LEVINE (1998): *The Theory of Learning in Games (Economics Learning and Social Evolution)*. Cambridge, MA: MIT Press. [1272,1273]
- GOEREE, J., AND C. A. HOLT (2001): "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review*, 91, 1402–1422. [1275,1284]
- HARRISON, G. (1989): "Theory and Misbehavior of First-Price Auctions," *American Economic Review*, 79, 749–762. [1282]
- HECKMAN, J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. Manski and D. McFadden. Cambridge, MA: MIT Press, 179–195. [1274,1285]
- (1991): "Identifying the Hand of Past: Distinguishing State Dependence from Heterogeneity," *American Economic Review Papers and Proceedings*, 81, 75–79. [1274]
- HECKMAN, J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320. [1285]
- HSIAO, C. (1986): *Analysis of Panel Data*. Cambridge, U.K.: Cambridge University Press. [1274]
- ICHIMURA, H., AND J. BRACHT (2001): "Estimation of Learning Models on Experimental Game Data," Unpublished Manuscript, Hebrew University of Jerusalem. [1273,1281,1284]
- KANDORI, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, 59, 63–80. [1275]
- KITZIS, S., H. KELLEY, E. BERG, D. MASSARO, AND D. FRIEDMAN (1998): "Broadening the Tests of Learning Models," *Journal of Mathematical Psychology*, 42, 327–355. [1274,1282]
- MCKELVEY, R., T. PALFREY, AND R. A. WEBER (2000): "The Effects of Payoff Magnitude and Heterogeneity on Behavior in 2×2 Games with Unique Mixed Strategy Equilibria," *Journal of Economic Behavior and Organization*, 42, 523–548. [1279,1283]
- MOOKERJEE, D., AND B. SOPHER (1997): "Learning and Decision Costs in Experimental Constant-Sum Games," *Games and Economic Behavior*, 19, 97–132. [1273]
- PESARAN, H., R. SMITH, AND K.-S. IM (1996): "Dynamic Linear Models for Heterogeneous Panels," in *The Econometrics of Panel Data*, ed. by L. Mátyás and P. Sevestre. Dordrecht, Netherlands: Kluwer Academic, 145–195. [1274]
- RUTSTRÖM, E. E., AND N. T. WILCOX (2001): "Power Planning Appendix to 'Learning and Belief Elicitation: Observer Effects'," Unpublished Manuscript, University of Houston. [1281]
- (2005): "Learning and Belief Elicitation: Observer Effects," Unpublished Manuscript, University of Houston. [1274,1279,1285]
- SALMON, T. (2001): "An Evaluation of Econometric Models of Adaptive Learning," *Econometrica*, 69, 1597–1628. [1281,1287]
- WILCOX, N. T. (2005): "Learning in Games and Heterogeneity Bias: Some Nonsolutions," Unpublished Manuscript, University of Houston. [1273,1274,1284,1285,1287]
- WOOLDRIDGE, J. M. (1997): "Multiplicative Panel Data Models Without the Strict Exogeneity Assumption," *Econometric Theory*, 13, 667–678. [1274]
- (2002): *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press. [1284]
- (2005): "Simple Solutions to the Initial Conditions Problem for Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, 20, 39–54. [1273,1274,1284,1285,1287]