Recoverability of Parameters from Learning Models

Duncan James[a] and Derrick Reagle[a]

[a]*Department of Economics, Fordham University, New York, USA*
Email: dujames@fordham.edu

Abstract: Econometric estimation of underlying parameters of learning models has been shown to be problematic (Salmon (2001)). We delve into the causes of that phenomenon. We do so by means of two complementary investigations, using data generated in bidding problems (first price sealed bid auction and Becker-DeGroot-Marschak) by Experience Weighted Attraction (EWA) simulated agents (with true parameters thus known to the researchers). First, we set out to determine how much each of the parameters in EWA affects the agent's attraction to possible strategies, holding constant the random draws involved in the environment in which the agent operates. For the first price sealed bid auction there are large regions of the strategy space where the underlying parameters do not significantly affect the graph of attractions. Our findings regarding the recoverability of EWA parameters in Becker-DeGroot-Marschak are even less promising: large parameter changes are needed for any change in the representation of attractions to register as statistically significant. Second, a further way of examining the sources of potential difficulties in accurately estimating EWA parameters is to see whether one can reject the similarity of attractions that are all generated by the same underlying EWA parameters (but in an environment where the random draws earlier held constant are now redrawn in every trial). In Becker-DeGroot-Marschak, we find that for all but a handful of parameter combinations, one would (incorrectly) reject a null that the same parameters generated the data. A similar statement holds true for the first price sealed bid auction for parameter combinations involving δ near 1, or δ near 1, or both.

Keywords: Estimation, learning, auctions

1. INTRODUCTION

Econometric estimation of underlying parameters of learning models has been shown to be problematic (Salmon (2001)). We seek in this paper to explore why that is the case. In pursuit of this objective, we have at our disposal a data set for which we have both observations on choice, and knowledge of the underlying (i.e. true or population) parameters of the learning model by which those choices were made. This data set – from James and Reagle (2009) - consists of bid data from the first price sealed bid auction (FPSB) and the Becker-DeGroot-Marschak procedure (BDM), where simulated Experience Weighted Attraction (EWA) agents produced the bids.

2. EFFECTS ON ATTRACTIONS (TO STRATEGIES) OF VARYING EWA PARAMETERS WHILE HOLDING VALUE DRAWS CONSTANT

We first seek to assess how much each of the parameters in EWA affects the agent's attraction to each strategy. After all, the attractions are the determinant of agent behavior for any given choice, since in EWA the agent forms a probability distribution function (pdf) over the domain of possible strategies, with strategies having higher attractions receiving more probability mass. Choice in a given auction is then made stochastically by means of a draw from this pdf. The pdf changes from one auction to the next as attractions to different possible bids are updated in light of prior auction results, as modulated by the parameters of the model. (Camerer and Ho (1999) give an exhaustive description of EWA. James and Reagle (2009) give details on how EWA was applied to the FPSB and BDM[1].)

We use the Kolmogorov-Smirnov (K-S)[2] goodness of fit test as a way to compare different sets of probability cdf's obtained from the pdf on strategies from equation (1) below.

$$P_i^j(a, t+1) = e^{f(A(a,t))}/\Sigma e^{f(A(a,t))} \qquad (1)$$

We do so for sets of attractions built up after the agent has participated in a sequence of 50 auctions. In this way, we compare sets of attractions for agents having different underlying values of the EWA parameters, holding constant the random draws necessitated by each institution. This allows us to judge if the graph of attractions to strategies is significantly affected by changing the EWA parameters. We do this for both the FPSB, and BDM.

2.1 The case of the First Price Sealed Bid Auction

Table 1: FPSB, λ=1, Value=1, K-S goodness of fit test of listed parameters versus δ=1, ϕ=1

		$\phi\rightarrow$0	0.2	0.4	0.6	0.8	1	kappa=0.0
$\delta\downarrow$	0.0	0.1492*	0.1147*	0.0810	0.0484	0.0307	0.0319	
	0.2	0.1486*	0.1060*	0.0648	0.0284	0.0295	0.0485	
	0.4	0.1476*	0.0921	0.0398	0.0238	0.0540	0.0943	
	0.6	0.1465*	0.0645	0.0126	0.0718	0.1247*	0.1693*	
	0.8	0.1416*	0.0239	0.1354*	0.2125*	0.2664*	0.3048*	
	1.0	0.1505*	0.6607*	0.6394*	0.5464*	0.4921*	0.4970*	
	0.0	0.1503*	0.1225*	0.0951	0.0684	0.0426	0.0316	kappa=0.2

---

[1] The objective of James and Reagle (2009) was to see whether both violations of the Vickrey (1961) model's predictions for bidding behavior in the FPSB (as observed with human subjects by Cox, Roberson, and Smith (1982), and typical (human) behavior in BDM could be rationalized by a single set of EWA parameters. They found that $\delta\approx$0 and $\varphi\approx$1 (where δ governs apprehension of opportunity cost (of strategies not employed in prior play), and φ governs memory) were able to reproduce the stylized facts on human bidding in these two institutions (overbidding in each). $\delta\approx$0 corresponds to ignoring or failing to imagine opportunity cost in past play, while $\varphi\approx$1 corresponds to a "long" memory of past play. They were thus able to use an approach analogous to Gode and Sunder (1993) to explain the cross-institutional patterns of bidding observed in Isaac and James (2000), wherein subjects who behaved as if risk averse in FPSB then behaved as if risk-seeking in BDM. The explanation is that overbidding in each of FPSB and BDM, mapping to risk aversion and risk-seeking interpretations respectively, could be an explanation common across the two institutions, and that such overbidding might in turn be due to the common parameterization of EWA uncovered by James and Reagle (2009).

[2] We opt for the K-S test instead of the Likelihood Ratio (LR) test due to the additional data available with the Monte Carlo simulations. We are working with the entire cdf on all strategies for a single point in time rather than a time series of observed bids used in the LR test.

0.2	0.1498*	0.1154*	0.0819	0.0495	0.0298	0.0306	
0.4	0.1490*	0.1041*	0.0612	0.0254	0.0254	0.0533	
0.6	0.1481*	0.0817	0.0211	0.0340	0.0819	0.1238*	
0.8	0.1447*	0.0145	0.0908	0.1667*	0.2230*	0.2649*	
1.0	0.0759	0.6536*	0.5529*	0.4928*	0.4796*	0.4863*	
0.0	0.1512*	0.1301*	0.1094*	0.0890	0.0689	0.0493	kappa=0.4
0.2	0.1508*	0.1248*	0.0993*	0.0743	0.0501	0.0308	
0.4	0.1503*	0.1163*	0.0833	0.0517	0.0268	0.0265	
0.6	0.1495*	0.0991*	0.0519	0.0172	0.0328	0.0695	
0.8	0.1475*	0.0476	0.0400	0.1092*	0.1646*	0.2087*	
1.0	0.1055*	0.3495*	0.4261*	0.4455*	0.4598*	0.4698*	
0.0	0.1519*	0.1378*	0.1238*	0.1099*	0.0962*	0.0827	kappa=0.6
0.2	0.1517*	0.1342*	0.1169*	0.0999*	0.0832	0.0667	
0.4	0.1513*	0.1285*	0.1061*	0.0842	0.0628	0.0421	
0.6	0.1509*	0.1169*	0.0842	0.0529	0.0232	0.0173	
0.8	0.1493*	0.0812	0.0191	0.0373	0.0854	0.1269*	
1.0	0.1257*	0.2639*	0.3638*	0.4043*	0.4263*	0.4416*	
0.0	0.1525*	0.1454*	0.1384*	0.1313*	0.1243*	0.1173*	kappa=0.8
0.2	0.1524*	0.1436*	0.1349*	0.1262*	0.1176*	0.1090*	
0.4	0.1522*	0.1407*	0.1293*	0.1180*	0.1068*	0.0958	
0.6	0.1519*	0.1348*	0.1180*	0.1015*	0.0853	0.0694	
0.8	0.1512*	0.1165*	0.0830	0.0511	0.0209	0.0109	
1.0	0.1413*	0.1226*	0.2516*	0.3161*	0.3531*	0.3777*	
0.0	0.1530*	0.1523*	0.1516*	0.1508*	0.1501*	0.1494*	kappa=1.0
0.2	0.1530*	0.1521*	0.1512*	0.1503*	0.1494*	0.1486*	
0.4	0.1530*	0.1518*	0.1506*	0.1495*	0.1483*	0.1472*	
0.6	0.1529*	0.1512*	0.1495*	0.1478*	0.1460*	0.1443*	
0.8	0.1528*	0.1493*	0.1458*	0.1423*	0.1388*	0.1353*	
1.0	0.1519*	0.1187*	0.0867	0.0562	0.0273	0.0000	

*significant at the 0.05 level

For the FPSB our findings are as follows. Table 1 reports the K-S test statistic for combinations of δ, ϕ, and κ versus a null of $\delta=\phi=\kappa=1$.[3,4] The areas of insignificance suggest difficulty in estimating parameters of EWA using maximum likelihood since there are large regions of the strategy space where the parameters do not significantly affect the graph of attractions. For example, $\delta=0$, $\phi=.4$, and $\kappa=0$ is statistically indistinguishable from $\delta=\phi=\kappa=1$.

In order aid the reader in developing an intuitive idea of the attractions, Figure 1 shows the graph of the attractions after the 50th period of bidding under three different parameter specifications, with a starting bid of $0.80.
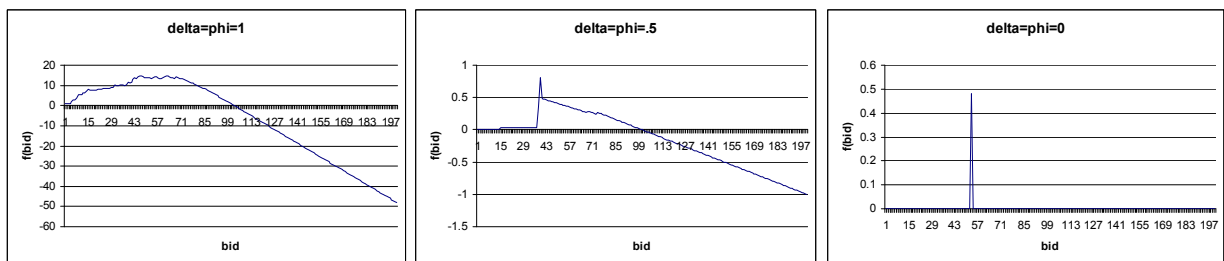

Figure 1: FPSB, Graph of Attractions

[3] The K-S results are reported for an initial bid of $0.80. We calculated the results for other initial bids. Bids farther away from optimal produced comparable results, while bids closer to optimal showed even less effect of parameter changes on the graph of attractions.
[4] The same set of random draws (as appropriate to the institution) is used throughout each table of Kolmogorov-Smirnoff results. As such the K-S results can be said to reflect differences in combinations of model parameters, and not just sampling variation in the draws.

For ϕ=δ=1 (the leftmost panel of the figure) the EWA attractions reflect the expected payoffs relatively well. As the parameters ϕ and δ each get farther from a value of 1 (in successive panels, from left to right, in Figure 1) information about opportunity cost is progressively ignored to a greater degree by EWA agents, and we see a spike in attraction for strategies that are used more often (conversely, a withering of other attractions to other strategies that could have been chosen in given rounds, but were not). The middle panel in Figure 1, showing the case where ϕ=δ=0.5, starts to show a noticeable asymmetry in attractions (i.e. the basis for probability mass which will then govern subsequent choice realizations), such that low bids do not build up attractions in the same manner as higher "win the auction, but at a low profit" bids (further to the right, in the same panel). Finally, as ϕ=δ=0, all strategies other than the one actually employed in a given round are ignored, and only the initial bid shows positive attraction.

2.1 The case of the Becker-DeGroot-Marschak procedure

Table 2: BDM, λ=1, value=1, K-S goodness of fit test of listed parameters versus δ=1, ϕ=1

		ϕ→0.0	0.2	0.4	0.6	0.8	1.0	kappa=0.0
δ↓	0.0	0.0763	0.0562	0.0388	0.0479	0.0844	0.1208*	
	0.2	0.0759	0.0547	0.0353	0.0468	0.0818	0.1167*	
	0.4	0.0752	0.0461	0.0197	0.0576	0.0955	0.1332*	
	0.6	0.0757	0.0255	0.0423	0.0889	0.1343*	0.1773*	
	0.8	0.0772	0.0287	0.1034*	0.1610*	0.2159*	0.2602*	
	1.0	0.3633*	0.3954*	0.4573*	0.4778*	0.4945*	0.5067*	
	0.0	0.0769	0.0604	0.0456	0.0314	0.0552	0.0844	kappa=0.2
	0.2	0.0765	0.0594	0.0435	0.0280	0.0531	0.0811	
	0.4	0.0759	0.0526	0.0299	0.0340	0.0644	0.0949	
	0.6	0.0762	0.0359	0.0222	0.0602	0.0976*	0.1338*	
	0.8	0.0772	0.0114	0.0750	0.1259*	0.1724*	0.2146*	
	1.0	0.1494*	0.3365*	0.4192*	0.4549*	0.4777*	0.4928*	
	0.0	0.0773	0.0649	0.0534	0.0426	0.0314	0.0479	kappa=0.4
	0.2	0.0771	0.0641	0.0517	0.0400	0.0283	0.0455	
	0.4	0.0766	0.0590	0.0417	0.0250	0.0333	0.0562	
	0.6	0.0766	0.0464	0.0181	0.0310	0.0596	0.0878	
	0.8	0.0772	0.0178	0.0429	0.0881	0.1251*	0.1607*	
	1.0	0.0666	0.2775*	0.3765*	0.4189*	0.4488*	0.4688*	
	0.0	0.0778	0.0694	0.0612	0.0537	0.0464	0.0388	kappa=0.6
	0.2	0.0776	0.0689	0.0603	0.0521	0.0443	0.0365	
	0.4	0.0773	0.0655	0.0538	0.0423	0.0309	0.0201	
	0.6	0.0769	0.0570	0.0377	0.0189	0.0207	0.0399	
	0.8	0.0773	0.0369	0.0091	0.0418	0.0730	0.1004*	
	1.0	0.0703	0.1872*	0.3034*	0.3640*	0.3992*	0.4237*	
	0.0	0.0781	0.0739	0.0697	0.0656	0.0615	0.0576	kappa=0.8
	0.2	0.0780	0.0737	0.0693	0.0650	0.0607	0.0564	
	0.4	0.0779	0.0720	0.0661	0.0602	0.0543	0.0485	
	0.6	0.0777	0.0677	0.0578	0.0481	0.0385	0.0290	
	0.8	0.0774	0.0568	0.0371	0.0183	0.0084	0.0248	
	1.0	0.0749	0.0728	0.1767*	0.2465*	0.2932*	0.3285*	
	0.0	0.0784	0.0780	0.0776	0.0771	0.0767	0.0763	kappa=1.0
	0.2	0.0784	0.0780	0.0775	0.0771	0.0767	0.0762	
	0.4	0.0784	0.0778	0.0772	0.0766	0.0760	0.0754	
	0.6	0.0784	0.0774	0.0764	0.0754	0.0744	0.0734	
	0.8	0.0783	0.0762	0.0742	0.0721	0.0700	0.0680	
	1.0	0.0778	0.0616	0.0456	0.0300	0.0148	0.0000	

* significant at the 0.05 level

Our findings regarding the recoverability of EWA parameters in BDM are even less promising than the findings from the case of the FPSB. Large parameter changes are needed for any change in the cdf to register as statistically significant. In fact, ϕ=1, δ=0 (reinforcement learning) is not statistically distinguishable from ϕ=δ=1 (fictitious play) for κ from 1 to 0.2. For κ=1, there is no combination of ϕ and δ that are statistically significant from ϕ=δ=1.

We now show examples of EWA attractions in BDM for different parameter sets.
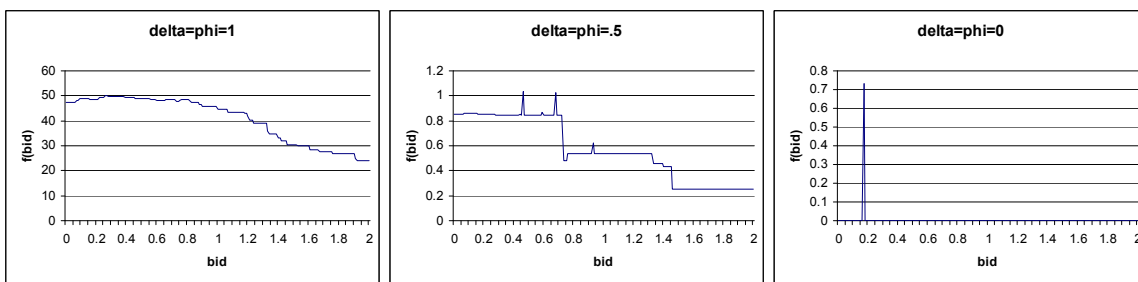

Figure 2: BDM, Graph of Attractions

These graphs of attractions illustrate that changes in EWA parameterization appear to make a difference to the attractions, which in turn leads to different behavior, but that fairly extreme changes in parameters are needed to guarantee that the resulting cdf can be distinguished statistically once passed through the logit transformation.

3. EFFECTS ON ATTRACTIONS (TO STRATEGIES) OF VARYING VALUE DRAWS WHILE HOLDING EWA PARAMETERS CONSTANT

One way of further examining the sources of potential difficulties in estimating EWA parameters is to see whether a K-S test rejects the similarity of cdf's that are all generated by the same underlying EWA parameters. To do this, we ran a Monte Carlo simulation of 999 trials, wherein the relevant random draws for each institution are redrawn at each trial, but the EWA parameters stay constant throughout all 999 trials. Probabilities over 5% indicate a greater Type I error than specified by K-S test, and that the cdf on strategies is not unique for given parameter values. The results for each institution are as follows in Tables 3 and 4, respectively.

For BDM, we see that except for a handful of parameter sets with ϕ or δ equal to1, K-S rejects similarity of graphs of attractions even though those attractions were generated subject to the same EWA parameters. Put another way, the random draws dominate the EWA parameters in shaping the cdf. This suggests that estimating EWA parameters from BDM data generated in an environment involving use of random draws might be problematic. FPSB shows similar concerns, but in another region of the parameter set where ϕ or δ equals1.

Table 3: FPSB, K-S goodness of fit test of listed parameters versus 999 Monte Carlo simulations

		φ→0.0	0.2	0.4	0.6	0.8	1.0	kappa=0.0
δ↓	0.0	0%	0%	0%	0%	5%	11%	
	0.2	0%	0%	0%	0%	4%	10%	
	0.4	0%	0%	0%	0%	0%	4%	
	0.6	0%	0%	0%	0%	0%	1%	
	0.8	0%	0%	0%	0%	0%	0%	
	1.0	0%	0%	0%	0%	0%	0%	
	0.0	0%	0%	0%	0%	2%	18%	kappa=0.2
	0.2	0%	0%	0%	0%	0%	15%	
	0.4	0%	0%	0%	0%	0%	8%	
	0.6	0%	0%	0%	0%	0%	6%	
	0.8	0%	0%	0%	0%	0%	0%	
	1.0	0%	0%	0%	17%	39%	54%	
	0.0	0%	0%	0%	0%	11%	15%	kappa=0.4
	0.2	0%	0%	0%	0%	9%	12%	
	0.4	0%	0%	0%	0%	6%	7%	
	0.6	0%	0%	0%	0%	1%	9%	
	0.8	0%	0%	0%	0%	1%	10%	
	1.0	0%	5%	42%	67%	88%	87%	
	0.0	0%	0%	0%	0%	9%	18%	kappa=0.6
	0.2	0%	0%	0%	0%	8%	15%	
	0.4	0%	0%	0%	0%	7%	8%	
	0.6	0%	0%	0%	1%	6%	19%	
	0.8	0%	0%	0%	0%	9%	16%	
	1.0	0%	32%	72%	84%	90%	89%	
	0.0	0%	0%	0%	0%	8%	12%	kappa=0.8
	0.2	0%	0%	0%	0%	9%	19%	
	0.4	0%	0%	0%	2%	9%	19%	
	0.6	0%	0%	0%	2%	11%	19%	
	0.8	0%	0%	0%	3%	22%	42%	
	1.0	0%	47%	80%	97%	93%	97%	
	0.0	0%	0%	0%	0%	9%	14%	kappa=1.0
	0.2	0%	0%	0%	1%	6%	17%	
	0.4	0%	0%	0%	2%	18%	16%	
	0.6	0%	0%	0%	4%	19%	31%	
	0.8	0%	0%	0%	8%	31%	54%	
	1.0	0%	75%	96%	95%	96%	99%	

Table 4: BDM, K-S goodness of fit test of listed parameters versus 999 Monte Carlo simulations

		φ→0.0	0.2	0.4	0.6	0.8	1.0	kappa=0.0
δ↓	0.0	100%	100%	100%	100%	81%	59%	
	0.2	100%	100%	100%	100%	100%	55%	
	0.4	100%	100%	100%	100%	100%	49%	
	0.6	100%	100%	100%	100%	100%	27%	
	0.8	100%	100%	100%	100%	100%	11%	
	1.0	100%	100%	100%	100%	100%	0%	
	0.0	100%	100%	100%	100%	77%	62%	kappa=0.2
	0.2	100%	100%	100%	100%	100%	55%	
	0.4	100%	100%	100%	100%	100%	55%	
	0.6	100%	100%	100%	100%	100%	43%	
	0.8	100%	100%	100%	100%	100%	35%	
	1.0	100%	100%	100%	100%	100%	99%	
	0.0	100%	100%	100%	100%	77%	72%	kappa=0.4
	0.2	100%	100%	100%	100%	100%	74%	
	0.4	100%	100%	100%	100%	100%	54%	
	0.6	100%	100%	100%	100%	100%	62%	
	0.8	100%	100%	100%	100%	100%	58%	
	1.0	100%	97%	100%	100%	100%	100%	
	0.0	100%	100%	100%	100%	100%	63%	kappa=0.6

0.2	100%	100%	100%	100%	100%	68%	
0.4	100%	100%	100%	100%	100%	64%	
0.6	100%	100%	100%	100%	100%	69%	
0.8	100%	100%	100%	100%	100%	78%	
1.0	77%	19%	100%	39%	100%	100%	
0.0	100%	100%	100%	100%	100%	56%	kappa=0.8
0.2	100%	100%	100%	100%	100%	62%	
0.4	100%	100%	100%	100%	100%	67%	
0.6	100%	100%	100%	100%	100%	63%	
0.8	100%	100%	100%	100%	100%	89%	
1.0	91%	63%	100%	99%	100%	100%	
0.0	100%	100%	100%	100%	70%	66%	kappa=1.0
0.2	100%	100%	100%	100%	100%	64%	
0.4	100%	100%	100%	100%	100%	70%	
0.6	100%	100%	100%	100%	100%	80%	
0.8	100%	100%	100%	100%	100%	88%	
1.0	4%	100%	96%	86%	100%	100%	

4. CONCLUSIONS

The overall results of this paper support the conclusion reached by Salmon (2001) that identifying which learning model parameters underlie the generation of a particular data set may not be generally possible. In particular, we show why this difficulty in recovery of learning model parameters occurs. Our results from section 2 demonstrate that attractions are often unresponsive to changes in EWA parameters, conditional on the agent facing the same set of random draws (appropriate to the institution). Put another way, different parameter sets could map to much the same behavior. Conversely, our results from section 3 show that quite different probability distributions on strategies can evolve from the same EWA parameter sets, provided that sufficiently different sets of random draws (appropriate to the institution) confront the agent. That is, the same parameter sets could map to very different behavior. Clearly parameter estimation is a daunting task under such conditions.

REFERENCES

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science*, 1964, 9, 226-232.

Camerer, Colin and T. H. Ho. "Experience-Weighted Attraction Learning in Normal Form Games." *Econometrica*, 1999, 67, 827-874.

Cox, James C., Bruce Roberson, and Vernon L. Smith. "Theory and Behavior of Single Object Auctions." In: Vernon L. Smith, ed., *Research in Experimental Economics*, Vol. 2, Greenwich, CT: JAI Press, 1982.

Gode, D. K. and Shyam Sunder. "Allocative Efficiency of Markets with Zero Intelligence Traders: Market as a Partial Substitute for Rationality." *Journal of Political Economy*, 1993, 101, 119-137.

Isaac, R. Mark and Duncan James. "Just Who Are You Calling Risk Averse?" *Journal of Risk and Uncertainty*, 2000, 20, 177-187.

James, Duncan and Derrick Reagle. "Experience Weighted Attraction in the First Price Auction and Becker DeGroot Marschak ." *Modsim Conference Proceedings*, 2009, 18(1), pp. 1437-1441. (Available at http://www.mssanz.org.au/modsim09/D8/james.pdf.)

Salmon, Timothy C. "An Evaluation of Econometric Models of Adaptive Learning." *Econometrica*, 2001, 69, 1597-1638.

Vickrey, William. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, 1961, 16, 8-37.