3-1-2013

Testing for Fictive Learning in Decision-Making under Uncertainty

Oliver D. Bunn

Caterina Calsamiglia

Donald J. Brown

TESTING FOR FICTIVE LEARNING IN DECISION-MAKING UNDER UNCERTAINTY

By

Oliver Bunn, Caterina Calsamiglia and Donald Brown

March 2013

COWLES FOUNDATION DISCUSSION PAPER NO. 1890

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Testing for Fictive Learning in Decision-Making under Uncertainty

Oliver Bunn, Caterina Calsamiglia and Donald Brown

March 20, 2013

Abstract

We conduct two experiments where subjects make a sequence of binary choices between risky and ambiguous binary lotteries. Risky lotteries are defined as lotteries where the relative frequencies of outcomes are known. Ambiguous lotteries are lotteries where the relative frequencies of outcomes are not known or may not exist. The trials in each experiment are divided into three phases: pre-treatment, treatment and post-treatment.

The trials in the pre-treatment and post-treatment phases are the same. As such, the trials before and after the treatment phase are dependent, clustered matched-pairs, that we analyze with the alternating logistic regression (ALR) package in SAS. In both experiments, we reveal to each subject the outcomes of her actual and counterfactual choices in the treatment phase. The treatments differ in the complexity of the random process used to generate the relative frequencies of the payoffs of the ambiguous lotteries. In the first experiment, the probabilities can be inferred from the converging sample averages of the observed actual and counterfactual outcomes of the ambiguous lotteries. In the second experiment the sample averages do not converge.

If we define fictive learning in an experiment as statistically significant changes in the responses of subjects before and after the treatment phase of an experiment, then we expect fictive learning in the first experiment, but no fictive learning in the second experiment. The surprising finding in this paper is the presence of fictive learning in the second experiment. We attribute this counterintuitive result to apophenia: "seeing meaningful patterns in meaningless or random data." A refinement of this result is the inference from a subsequent Chi-squared test, that the effects of fictive learning in the first experiment are significantly different from the effects of fictive learning in the second experiment.

JEL Classification: C23, C35, C91, D03
Keywords: Uncertainty, Counterfactual Outcomes, Apophenia

1 Introduction

Conditioning current choices under uncertainty on counterfactual outcomes of previous choices, i.e., fictive learning, is well documented in $fMRI$ studies of gambling behavior in humans — see Lohrenz et al. (2007) and decision-making under uncertainty in monkeys — see Hayden et al. (2009). Recently Boorman et al. (2011) identified neural circuits for counterfactual outcomes and fictive learning. A common practice in experimental studies of decision-making under uncertainty, such as the $fMRI$ studies in Huettel et al. (2006) and Levy et al. (2009) is to posit a cross-sectional model for the experimental data. Unfortunately, a cross-sectional analysis ignores that each subject's binary responses are correlated. In fact, this is the generic property of most revealed preference experiments in neuroeconomics.

Recently, Li et al. (2008) proposed longitudinal analysis of general linear models with general estimating equations (GEE), due to Liang and Zeger (1986) and Zeger and Liang (1986), for neuroimaging data. Li et al. argue that the existing statistical methods for analyzing neuroimaging data are primarily developed for cross-sectional neuroimaging studies and not for panel neuroimaging data. We find this critique of the current practice in neuroimaging studies equally compelling as a critique of the current statistical practice in neuroeconomic studies of revealed preferences for risk and ambiguity.

Consequently, we conduct two experiments where subjects make a sequence of binary choices between risky and ambiguous binary lotteries. The trials in each experiment are divided into three phases: pre-treatment, treatment and post-treatment. The trials in the pre-treatment and posttreatment phases are the same. As such, the trials before and after the treatment phase are dependent, clustered matched-pairs. For correlated binary data, Lipsitz, et al. (1991) introduced odds ratios as a measure of the within-subject association of binary responses. We use alternating logistic regression (ALR) with constant log odds ratios (LOR) as the within-subject association of binary responses, proposed by Carey et al. (1993), to estimate discrete binary choice models of revealed preferences for risk and ambiguity. There is an important difference between the application of longitudinal analysis to neuroimaging data, where the within-subject association of responses is considered a nuisance and our application of longitudinal analysis. In our application of ALR, the within-subject association of responses is not a nuisance but an essential part of our analysis. It is the within-subject association of responses as constant log odds ratios that we use to test for fictive learning in the revealed preferences derived from the dependent, clustered responses of subjects.

The treatment phases in each experiment differ in the complexity of the random process used to generate the relative frequencies of the payoffs of the ambiguous lotteries. In the first experiment, the probabilities can be inferred from the converging sample averages of the observed actual and counterfactual outcomes of the ambiguous lotteries. In the second experiment the sample averages do not converge. If we define fictive learning in an experiment as statistically significant changes in the responses of subjects before and after the treatment phase of an experiment, then we expect fictive learning in the first experiment, but no fictive learning in the second experiment. The surprising finding in this paper is the presence of fictive learning in the second experiment. We attribute this counterintuitive result to apophenia: "seeing meaningful patterns in meaningless or random data." A refinement of this result is the inference from a subsequent Chi-squared test, that the effects of fictive learning in the first experiment are significantly different from the effects of fictive learning in the second experiment.

The second section of the paper is an exposition of the experimental protocols. The final section is a presentation of the model and the analysis of the data from the experiments.

2 Experimental Protocols

To test for the presence of fictive learning in revealed preferences for risk and ambiguity, we conduct two experiments on revealed preferences for choices under uncertainty, consisting of 36 students as subjects randomly chosen from the 2011 Yale Fall term. Each subject makes a sequence of 100 binary choices between risky and ambiguous lotteries. Risky lotteries are defined as lotteries where the relative frequencies of outcomes are known. Ambiguous lotteries are lotteries where the relative frequencies of outcomes are not known or may not exist. As such, our model of decision-making under risk and ambiguity has its origins in the following quote of Keynes (1937):

> By uncertain knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty; nor is the prospect of a Victory bond being drawn. Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know.[1]

The experiments are divided into three phases: The pre-treatmemt phase (phase 1), the treatment phase (phase 2) and the post treatment phase (phase 3). Subjects face the same sequence of 30 binary choices between risky and ambiguous lotteries in the first and third phase of each experiment. That is, the trials in phases 1 and 3 are clustered matched-pairs, but the lotteries in phases 1 and 3 of the two experiments are independent. To test for fictive learning, we reveal to each subject the outcomes of her 40 actual and counterfactual choices in phase 2. In the first experiment, the relative frequencies of counterfactual ambiguous outcomes in phase 2 are relatively easy to learn, using sample averages of the outcomes of the ambiguous lotteries. In the second experiment,the relative frequencies of counterfactual ambiguous outcomes

[1] Uncertainty in this quote means ambiguity.

in phase 2 are quite difficult, if not impossible, to learn, since the sample averages do not converge.[2] The binary choices in phase 2 are the same in both experiments and independent of the binary choices in phases 1 and 3. Subjects are unaware that they will be exposed to counterfactual outcomes in phase 2 before they are presented binary choices in phase 3. In particular, subjects do not know if the relative frequencies of counterfactual outcomes of ambiguous lotteries in phase 2 is a sample average of the probabilities of ambiguous outcomes in phases 1 and 3. In fact, they are in the first experiment, but not in the second experiment.

No outcomes are revealed to subjects in the first and third phase of the experiments. The lotteries are displayed as pie graphs on each subject's computer screen, where we used the Psychtoolbox in Matlab for all displays. Probabilities for the risky lotteries are displayed. The probabilities determining the payoffs of ambiguous lotteries are constant in phases 1 and phase 3 of both experiments, but never revealed to the subjects. We randomly vary the placement and colors of the lotteries on the computer screen to control for positional bias. We randomly choose one group of 17 students from the 36 students as subjects for the first experiment. At the end of each experiment, a trial is randomly chosen for each subject participating in that experiment and the subject is given the payoff of her choice.

We define fictive learning in each experiment as statistically significant changes in the responses of subjects before and after exposure to the counterfactual outcomes in phase 2 of the experiment. This exposure to counterfactual outcomes constitutes the treatment phase of the experiments. In each experiment, we estimate a constant log odds ratio (LOR) of the odds of choosing the risky lottery in a trial in phase r, conditional on the choice in a trial in phase s. We use $GENMOD$ in SAS with the $LOGOR$ option to estimate the regression equations for both the first and second moments of the marginal model. We assume the LOR is constant in phase 1; phase 2; phase 3; between phases 1 and 2; between phases 1 and 3 and between phases 2 and 5. In ALR, the odds ratio for each pair of trials is

$$\frac{\Pr(Y_{ij}=1;Y_{ik}=1)\Pr(Y_{ij}=0;Y_{ik}=0)}{\Pr(Y_{ij}=1;Y_{ik}=0)\Pr(Y_{ij}=0;Y_{ik}=1)} = \frac{\frac{\Pr(Y_{ij}=1|Y_{ik}=1)}{\Pr(Y_{ij}=0|Y_{ik}=1)}}{\frac{\Pr(Y_{ij}=1|Y_{ik}=0)}{\Pr(Y_{ij}=0|Y_{ik}=0)}}$$

where i is the subject index and j and k are the indices of the trials.$Y_{ij}=1$ means subject i choose the risky lottery in trial j.

3 A Marginal Analysis of Fictive Learning

The most frequently used models for discrete repeated measurements of experimental outcomes are: random-effects models, used extensively in econometrics, and marginal models, where the regression parameters are computed using general estimating equations (GEE), the methodology of choice in biostatistics. The GEE approach has a number of appealing properties for estimation of the regression parameters in marginal models: First, we need only make assumptions on the first two moments of

[2] See Appendix B for a formal proof.

the distribution of the vector of responses. The GEE estimates of the regression coefficients are consistent and asymptotically normal, where the covariance matrix is consistently estimated using a sandwich estimator, even if the within subject associations among the repeated measurements have been misspecified. In many cases, GEE is almost as efficient as maximum likelihood estimation. We interpret the parameters in the marginal model as population averages in a given experiment.

We follow section 13.6 in Fitzmaurice et al., for GEE using $PROC\ GENMOD$ in SAS and the $LOGOR$ option, where the log odds represent the within-subject association among pairs of binary responses. The log odds are used to estimate the regression equations for both the first and second moments of the marginal model.[3]

Recently, Bracha and Brown (2013) proposed a new class of non-expected utility functions for subjective evaluations of state-contingent claims, i.e., ambiguous lotteries. In their model there is a proxy for risk-aversion, β, and a proxy for ambiguity-aversion, α. The paradigmatic case of the preferences they propose is the composition of quadratically concave (convex) utility functions representing preferences for risk and preferences for ambiguity. As such, Bracha and Brown define preferences represented by the composition of preferences for risk and preferences for ambiguity as Keynesian utility functions. The concavity of the utility functions in this class of non-expected utility functions depends on the ratio of α and β. In our model we restrict attention to the parametric class of linear-quadratic convex (concave) utility functions introduced by Rockafellar (1988). This specification allows a reduced form expression of Keynesian utility functions. We assume that subjects evaluate risky lotteries, X, using expected utility, $E(U(X))$, and evaluate ambiguous lotteries,Y, using Keynesian utility, $J(Y)$. In each binary choice between a risky and an ambiguous lottery, we assume that subjects choose the lottery that maximizes subjective value. The important technical aspect of the linear-quadratic formulation is that for any pair of risky and ambiguous lotteries, the difference in the utility of the risky lottery and the utility of the ambiguous lottery is linear in the parameters α and β. Hence the log likelihood for the associated discrete binary choice model is strictly concave in the parameters. In the parametric specification of expected utility

$$E_\beta(U(X)) \equiv \pi_1 u_\beta(x) + \pi_2 u_\beta(x)$$

$$u_{\beta,K}(w) \equiv Kw + \tfrac{\beta}{2}w^2 \text{ and } \beta \text{ is the proxy for risk.}$$

$$\text{If } X = (x_1, x_2; \pi_1, \pi_2) \text{ and } Y = (y_1, y_2)$$

$$\text{then } u_{\beta,K}(x_1) = Kx_1 + \tfrac{\beta}{2}x_1^2 \text{ and } u_{\beta,K}(x_2) = Kx_2 + \tfrac{\beta}{2}x_2^2$$

[3] The regression equations for the first moment of the marginal model are in Appendix A.
For an excellent survey of applications of the GEE approach in biostatistics and econometrics, we suggest "GEE: An Annotated Bibliography" by Zeigler et al. (1998). They observe that in panel or longitudinal studies the classical assumptions of statistics such as independence or normality of observations may be invalid. As examples they cite count data or binary data as lacking normal distributions and repeated observations on the same subject as violating the independence assumption. Ignoring the dependent nature of the data can lead to incorrect inferences. They recommend the article by Sherman and Le Cressie (1997) for a discussion of the implications of correlated observations on assessing the precision of estimated parameters.

$$E_{\beta,K_R}(U(X)) \equiv \pi_1 u_{\beta,K_R}(x_1) + \pi_2 u_{\beta,K_R}(x_2) = K_R(\pi_1 x_1 + \pi_2 x_2) + \tfrac{\beta}{2}[\pi_1 x_1^2 + \pi_2 x_2^2]$$

In the parametric specification of $J_{\alpha,\beta,K}(Y)$, the subjective valuation of the ambiguous lottery Y, where β is the proxy for risk and α is the proxy for ambiguity.

$$J_{\alpha,\beta,K_A}(Y) \equiv K_A(y_1 + y_2) + \tfrac{[\alpha-\beta]}{2}[y_1^2 + y_2^2]$$

If A is a symmetric $N \times N$ matrix, then $A \lesssim 0$ means A is negative semidefinite. Hence

$$\nabla_Y^2 J_{\alpha,\beta,K_A}(U(Y)) \lesssim 0 \Leftrightarrow \alpha \leq \beta$$

The binary choice model is a generalized linear model where the link function is a cdf. In this paper, the link function is the logistic cdf. The argument of the link function is the difference of the parametric nonrandom components of the random utility of a risky and an ambiguous lottery. If the nonrandom component of the random utility function is linear in the parameters, then the log-likelihood is strictly-concave in the parameters defining the nonrandom components of the random utility function. $\Phi_X(\alpha, \beta, K_R, K_A)$, the argument of the logistic cdf, is the difference of the expected utility of the risky lottery $X = (x_1, x_2; \pi_1, \pi_2)$ and the subjective valuation of the ambiguous lottery $Y = (y_1, y_2)$. Hence the choice probability for X, $p_X(\alpha, \beta, K_R, K_A)$, is implicitly defined by the logistic cdf

$$\Lambda[\Phi] \equiv \frac{\exp \Phi}{1 + \exp \Phi}$$

where

$$\Phi_X(\alpha, \beta, K_R, K_A) \equiv \log \frac{p_X(\alpha, \beta, K_R, K_A)}{1 - p_X(\alpha, \beta, K_R, K_A)} = [E_{\beta,K_R}(U(X)) - J_{\alpha.\beta,K_A}(U(Y))]$$

is the log-odds of choosing X.

$$\begin{aligned}[E_{\beta,K_R}(U(X)) - J_{\alpha,\beta,K_A}(U(Y))] \quad &= \quad K_R(\pi_1 x_1 + \pi_2 x_2) + \tfrac{\beta}{2}[\pi_1 x_1^2 + \pi_2 x_2^2] \\ &\quad - \left\{ K_A(y_1 + y_2) + \tfrac{[\alpha-\beta]}{2}[y_1^2 + y_2^2] \right\}\end{aligned}$$

$$p_X(\alpha, \beta, K_R, K_A) = \frac{\exp[E_{\beta,K_R}U(X) - J_{\alpha,\beta,K_A}(U(Y))]}{(1 + \exp[E_{\beta,K_R}U(X) - J_{\alpha,\beta,}(U(Y))])},$$

where $p_X(\alpha, \beta, K_R, K_A)$ is the explicit probability of choosing the risky lottery X in the pair-wise comparison between X and the ambiguous lottery Y. In each experiment, let $\theta_{i,j} = 1$ if the risky lottery is chosen by subject i on trial j and 0 otherwise, then the probability density of $\theta_{i,j}$ is

$$[p_X(\alpha, \beta, K_R, K_A)]^{\theta_{i,j}} [1 - p_X(\alpha, \beta, K_R, K_A)]^{1-\theta_{i,j}}.$$

We estimate the regression parameters in the marginal model, using generalized estimating equations (GEE). Z_i^k is a binary vector of length 100 denoting the sequence of binary choices of subject i in experiment k. Each entry Z_{ij}^k of Z_i^k is a Bernoulli random variable with mean μ_{ijk},where j is the trial index. We divide the

100 trials into three phases: Phase 1: $1 \leq j \leq 30$, Phase 2: $31 \leq j \leq 70$ and Phase 3: $71 \leq j \leq 100$. There are 12 equations for $k = 1$ and 12 equations for $k = 2$. Fixing k, for each phase there are four simultaneous nonlinear equations in the four parameters $\alpha_{l(j)k}, \beta_{l(j)k}, K_{Rl(j)k}$,and $K_{Al(j)k}$. where $l(j) = 1$, for j in phase 1; $l(j) = 2$ for j in phase 2; $l(j) = 3$ for j in phase 3.

$$\begin{aligned}\text{If logit } (\mu_{ijk}) \quad = \quad & \log \tfrac{\mu_{ijk}}{1-\mu_{ijk}} = K_{l(j)R}(p_{ijk}x_{1(ijk)} + (1 - p_{ijk})x_{2(ijk)}) \\ & + \tfrac{\beta_{l(j)k}}{2}(p_{ijk}x^2_{1(ijk)} + (1 - p_{ijk})x^2_{2(ijk)}) \\ & - \left\{ K_{l(j)A}(y_{1(ijk)} + y_{2(ijk)}) + \left(\frac{\alpha_{l(j)k} - \beta_{l(j)k}}{2} \right) (y^2_{1(ijk)} + y^2_{2(ijk)}) \right\}\end{aligned}$$

$$\begin{aligned}\text{then logit } (\mu_{ijk}) \quad = \quad & K_{l(j)R}(p_{ijk}x_{1(ijk)} + (1 - p_{ijk})x_{2(ijk)}) \\ & + \tfrac{\beta_{l(j)k}}{2}(p_{ijk}x^2_{1(ijk)} + (1 - p_{ijk})x^2_{2(ijk)}) \\ & - K_{l(j)A} \left\{ (y_{1(ijk)} + y_{2(ijk)}) + \left(\frac{\alpha_{l(j)k} - \beta_{l(j)k}}{2} \right) (y^2_{1(ijk)} + y^2_{2(ijk)}) \right\}.\end{aligned}$$

$$\begin{aligned}\text{If } N_{ijk} \quad & \equiv \quad p_{ijk}x_{1(ijk)} + (1 - p_{ijk})x_{2(ijk)}); \; R_{ijk} \equiv p_{ijk}x^2_{1(ijk)} + (1 - p_{ijk})x^2_{2(ijk)} \text{ and} \\ M_{ijk} \quad & \equiv \quad (y_{1(ijk)} + y_{2(ijk)}); \; A_{ijk} \equiv (y^2_{1(ijk)} + y^2_{2(ijk)})\end{aligned}$$

$$\text{then logit}(\mu_{ijk}) = K_{l(j)R}N_{ijk} + \frac{\beta_{l(j)k}}{2}R_{ijk} - \left\{ K_{l(j)A}M_{ijk} + \left(\frac{\alpha_{l(j)k} - \beta_{l(j)k}}{2} \right) A_{ijk} \right\}$$

$$\text{Hence } \mu_{ijk} = \text{logit}^{-1} \left[K_{l(j)R}N_{ijk} + \frac{\beta_{l(j)k}}{2}R_{ijk} - \left\{ K_{l(j)A}M_{ijk} + \left(\frac{\alpha_{l(j)k} - \beta_{l(j)k}}{2} \right) A_{ijk} \right\} \right]$$

$$\pi^k \equiv (\alpha_{1k}, \alpha_{2k}, \alpha_{3k}; \beta_{1k}, \beta_{2k}, \beta_{3k}; K_{A1k}, K_{A2k}, K_{A3k}; K_{R1k}, K_{R2k}, K_{R3k}) \in R^{12}$$

is the column vector of parameters in experiment k. For each k, we obtain an initial estimate of π^k by fitting a nonlinear logistic regression in which the binary choices at trial j of subject i in experiment k, Y_{ijk}, are assumed independent. The details are in the Appendix A. Box plots of the parameter estimates for each experiment are in Figure 1. Tables 1 and 2 are the estimates of parameters in the regression model for the first moment of the distribution of the vectors of responses in experiments 1 and 2.

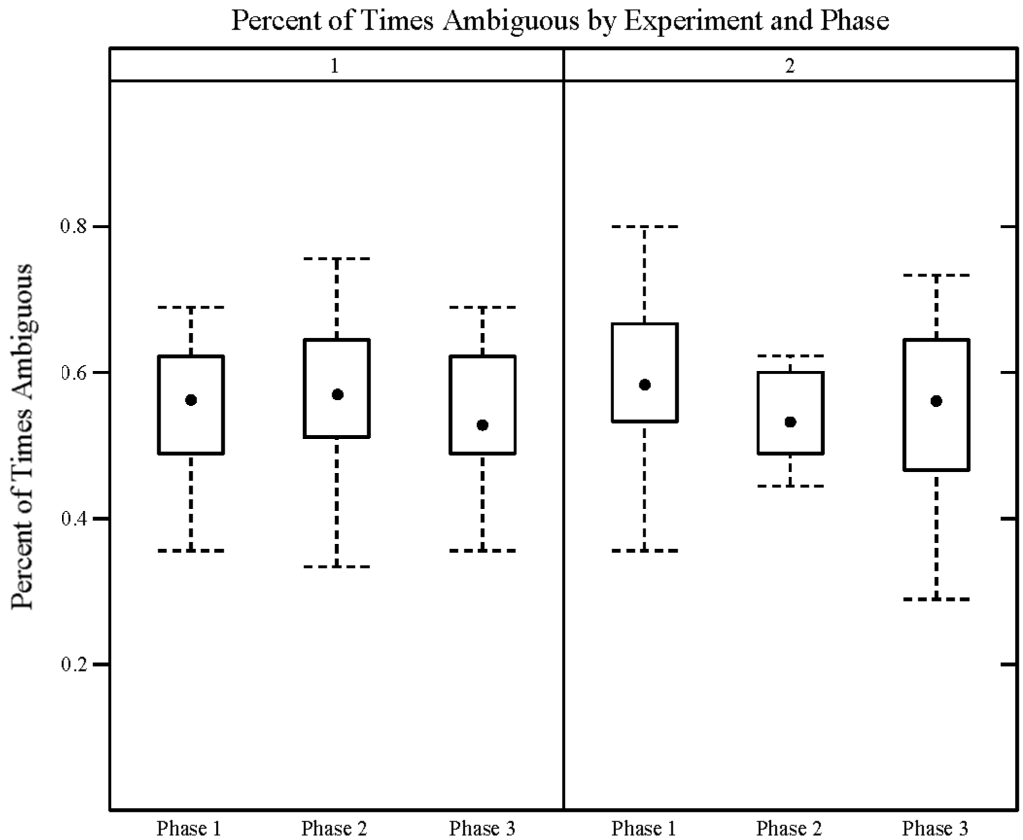

Table 1. Analysis of GEE Parameter Estimates: Experiment I

	Estimates	Standard error	95% confidence limits		Z	$\Pr > \lvert Z \rvert$
$kr1$	2.1074	0.3118	1.4863	2.7186	6.76	< 0.0001
$kr2$	1.4113	0.2397	0.9416	1.8810	5.89	< 0.0001
$kr3$	1.7255	0.2616	1.2127	2.2383	6.59	< 0.0001
$\beta1$	−0.0955	0.0172	−0.1292	−0.0617	−5.54	< 0.0001
$\beta2$	−0.0507	0.0125	−0.0752	−0.0261	−4.04	< 0.0001
$\beta3$	−0.0697	0.0150	−0.0992	−0.0402	−4.64	< 0.0001
$ka1$	0.8134	0.1418	0.5355	1.0914	5.74	< 0.0001
$ka2$	0.6000	0.1217	0.3615	0.8385	4.93	< 0.0001
$ka3$	0.6250	0.1303	0.3695	0.8804	4.80	< 0.0001
$\alpha1$	−0.0206	0.0072	−0.0347	−0.0065	−2.87	< 0.0041
$\alpha2$	−0.0103	0.0068	−0.0236	0.0030	−1.51	< 0.1300
$\alpha3$	−0.0090	0.0071	−0.0229	0.0048	−1.28	< 0.2016

Table 2. Analysis of GEE Parameter Estimates: Experiment II						
	Estimates	Standard error	95% confidence limits		Z	Pr $> \|Z\|$
$kr1$	1.6809	0.3058	1.0815	2.2803	5.50	< 0.0001
$kr2$	1.4511	0.2580	0.9454	1.9568	5.62	< 0.0001
$kr3$	1.4122	0.3367	0.7513	2.0711	4.19	< 0.0001
$\beta 1$	−0.0661	0.0156	−0.0966	−0.0355	−4.24	< 0.0001
$\beta 2$	−0.0525	0.0147	−0.0813	−0.0238	−3.58	< 0.0003
$\beta 3$	−0.0526	0.0172	−0.0863	−0.0189	−3.06	< 0.0022
$ka1$	0.7642	0.1503	0.4697	1.0587	5.09	< 0.0001
$ka2$	0.6261	0.1408	0.3501	0.9022	4.45	< 0.0001
$ka3$	0.0587	0.1821	0.2117	0.9256	3.12	< 0.0018
$\alpha 1$	−0.0195	0.0080	−0.0351	−0.0039	−2.44	< 0.0145
$\alpha 2$	−0.0134	0.0083	−0.0296	0.0028	−1.62	< 0.1055
$\alpha 3$	−0.0094	0.0099	−0.0287	0.0100	−0.95	< 0.3442

If $\theta = 1, 2, 3$ then the estimated expected value of a risky lottery $X = (x_1, x_2; \pi_1, \pi_2)$ in phase θ is

$$kr_\theta(\pi_1 x_1 + \pi_2 x_2) + \frac{\beta_\theta}{2}[\pi_1 x_1^2 + \pi_2 x_2^2]$$

where for all θ: $kr_\theta > 0$ and $\beta_\theta < 0$. Hence the estimated Bernoulli utility of wealth is concave and monotone for the range of values of x_1 and x_2 in experiment 2. That is, the maximum payoff is \$15.00 and the partial derivatives of the estimated expected value are bounded below by

$$\begin{aligned}(\pi_1 \wedge \pi_2)[(\wedge_\theta kr_\theta) + (\vee_\theta \beta_\theta)(x_1 \vee x_2)] &\geq (\pi_1 \wedge \pi_2)[1.4112 - (15.0)(0.0661)] \\ &= (\pi_1 \wedge \pi_2)[1.4112 - 0.9915] = 0.4197 > 0.\end{aligned}$$

Hence the estimated Bernoulli utility of wealth is concave and monotone for the values of x_1 and x_2 in experiment 1. Here the worse case is phase 2, the maximum payoff is \$15.00 and the partial derivatives of the estimated expected value are bounded below by

$$\begin{aligned}(\pi_1 \wedge \pi_2)[(kr_2) + (\beta_2)(x_1 \vee x_2)] &\geq (\pi_1 \wedge \pi_2)[1.4113 - (15.0)(0.0525)] \\ &= (\pi_1 \wedge \pi_2)[1.4113 - 0.7875] = 0.6238 > 0\end{aligned}$$

If $\theta = 1, 2, 3$ then the estimated subjective value of an ambiguous lottery $Y = (y_1, y_2)$ in phase θ is

$$ka_\theta(y_1 + y_2) + \frac{[\alpha_\theta - \beta_\theta]}{2}[y_1^2 + y_2^2]$$

where for all θ: $ka_\theta > 0$, $\alpha_\theta > 0$ and $[\alpha_\theta - \beta_\theta] > 0$. Hence the estimated Keynesian utility in both experiments is convex and monotone for all values of $y_1 \geq 0$ and $y_2 \geq 0$

The odds ratio in ALR is

$$\frac{\Pr(Y_{ij} = 1; Y_{ik} = 1)\Pr(Y_{ij} = 0; Y_{ik} = 0)}{\Pr(Y_{ij} = 1; Y_{ik} = 0)\Pr(Y_{ij} = 0; Y_{ik} = 1)} = \frac{\frac{\Pr(Y_{ij}=1|Y_{ik}=1)}{\Pr(Y_{ij}=0|Y_{ik}=1)}}{\frac{\Pr(Y_{ij}=1|Y_{ik}=0)}{\Pr(Y_{ij}=0|Y_{ik}=0)}}$$

$$\left[\frac{\Pr(Y_{ij}=1;Y_{ik}=1)\Pr(Y_{ij}=0;Y_{ik}=0)}{\Pr(Y_{ij}=1;Y_{ik}=0)\Pr(Y_{ij}=0;Y_{ik}=1)}\right]^{-1} = \frac{\frac{\Pr(Y_{ij}=1|Y_{ik}=0)}{\Pr(Y_{ij}=0|Y_{ik}=0)}}{\frac{\Pr(Y_{ij}=1|Y_{ik}=1)}{\Pr(Y_{ij}=0|Y_{ik}=1)}}$$

where i is the subject index and j and k are the indices of the trials. $Y_{ij} = 1$ means the subject choose the risky lottery in trial j. We assume the LOR is constant in phase 1; phase 2; phase 3; between phases 1 and 2; between phases 1 and 3 and between phases 2 and 5. In the following tables for Experiment I and II, the estimated constant LOR are denoted αQ for $Q = 1, 2, ..., 5$. We test the null hypothesis H_0: the log odds ratio $= 0$, against the alternative hypothesis H_A: the log odds ratio is $\neq 0$. Tables 3 and 4 are the estimates of parameters in the regression model for the second moment of the distribution of the vectors of responses in experiments 1 and 2.

Table 3. Analysis of GEE Parameter Estimates: LOR for Experiment I

	Estimates	Standard error	95% confidence limits		Z	$\Pr > \lvert Z\rvert$
$\alpha 1$ (Phase 1)	−0.0096	0.0468	−0.1013	0.0822	−0.20	0.8381
$\alpha 2$ (Phase 2)	0.1124	0.0641	−0.0132	0.2380	1.75	0.0796
$\alpha 3$ (Phase 3)	0.0334	0.0379	−0.1077	0.0408	−0.88	0.3778
$\alpha 4$ (Phase 1 & 2)	0.0208	0.0376	−0.0529	0.0944	0.55	0.5803
$\alpha 5$ (Phase 1 & 3)	0.1057	0.0389	0.0293	0.1820	2.71	0.0067
$\alpha 6$ (Phase 2 & 3)	0.0112	0.0454	−0.0778	0.1003	0.25	0.8046

Table 4. Analysis of GEE Parameter Estimates: LOR for Experiment II

	Estimates	Standard error	95% confidence limits		Z	$\Pr > \lvert Z\rvert$
$\alpha 1$ (Phase 1)	0.1217	0.0744	−0.0242	0.2675	1.64	0.1020
$\alpha 2$ (Phase 2)	−0.0463	0.0222	−0.0898	−0.0027	−2.08	0.0374
$\alpha 3$ (Phase 3)	0.1349	0.0832	−0.0282	0.2980	1.62	0.1050
$\alpha 4$ (Phase 1 & 2)	−0.0061	0.0304	−0.0656	0.0535	−0.20	0.8420
$\alpha 5$ (Phase 1 & 3)	0.2051	0.0779	0.0525	0.3577	2.63	0.0084
$\alpha 6$ (Phase 2 & 3)	−0.0239	0.0303	−0.0833	0.0356	−0.79	0.4315

$\alpha 2$ and $\alpha 5$ are the only significant statistics in each experiment. The LOR in phase 2 of experiment 2 is

$$\frac{\Pr(Y_{ij}=1 \mid Y_{ik}=0)}{\Pr(Y_{ij}=0 \mid Y_{ik}=0)} > \frac{\Pr(Y_{ij}=1 \mid Y_{ik}=1)}{\Pr(Y_{ij}=0 \mid Y_{ik}=1)}$$

and the LOR in phase 2 of experiment 1 is

$$\frac{\Pr(Y_{ij}=1 \mid Y_{ik}=1)}{\Pr(Y_{ij}=0 \mid Y_{ik}=1)} > \frac{\Pr(Y_{ij}=1 \mid Y_{ik}=0)}{\Pr(Y_{ij}=0 \mid Y_{ik}=0)}$$

For $\alpha 5$, the LOR between phases 1 and 3 in each experiment is

$$\frac{\Pr(Y_{ij}=1 \mid Y_{ik}=1)}{\Pr(Y_{ij}=0 \mid Y_{ik}=1)} > \frac{\Pr(Y_{ij}=1 \mid Y_{ik}=0)}{\Pr(Y_{ij}=0 \mid Y_{ik}=0)}.$$

Our null hypothesis is the absence of fictive learning in phase 2 of both experiments. In each experiment we reject the null hypothesis that the LOR is zero, i.e., no fictive learning in phase 2. In each experiment these findings are significant at the 1% level. The significant fictive learning in the first experiment is not surprising, given the sample data in phase 2 on the relative frequency of the payoffs of actual and counterfactual choices of ambiguous lotteries. The surprising, counterintuitive finding is that despite Keynes dictum "About these matters there is no scientific basis on which to form any calculable probability whatever" there is significant evidence of fictive learning in the second experiment. How can we reconcile the contradiction between Keynes theory of "uncertainty" and fictive learning in the second experiment? As is well known, patterns are often perceived in random data. That is, a type 1 error in the sense of statistics or apophenia: "seeing meaningful patterns in meaningless or random data."[4] A subject's perceptions of patterns in the factual and counterfactual payoffs of ambiguous lotteries in the treatment phase of the second experiment may well be the cause of fictive learning in the second experiment.

Whatever subjects "learn" in the treatment phases of the two experiments, we can ask if the effects of the treatments are significantly different. To compare the effects of the treatment phase in each experiment, we use the log-odds-ratio test proposed in chapter 10 of Fleiss et al. (2004). The analysis begins with the calculation of

$$\chi^2_{Total} = \chi^2_I + \chi^2_{II} = w_I LOR^2_I + w_{II} LOR^2_{II},$$

where χ^2_{Total} has 2 df, given the independence of the subjects in experiment I and the subjects in experiment II, where each group of subjects has a χ^2 distribution with $1 df$.[5]

$$w_I = \frac{1}{[SE(LOR_I)]^2} \text{ and } w_{II} = \frac{1}{[SE(LOR_{II})]^2}$$

$$\left[\frac{LOR_I}{SE(LOR_I)}\right]^2 = \left[\frac{0.1057}{0.0389}\right]^2 = 7.3833$$

$$\left[\frac{LOR_{II}}{SE(LOR_{II})}\right]^2 = \left[\frac{0.2051}{0.0779}\right]^2 = 6.9319$$

$$\chi^2_{Total} = \chi^2_I + \chi^2_{II} = \left[\frac{0.1057}{0.0389}\right]^2 + \left[\frac{0.2051}{0.0779}\right]^2 = 14.3152$$

significant at the 0.001 level. We decompose χ^2_{Total} into two orthogonal components

$$\chi^2_{Total} = \chi^2_{Homog} + \chi^2_{Assoc}.$$

where

$$\chi^2_{Assoc}. = \frac{[w_I LOR_I + w_{II} LOR_{II}]^2}{w_I + w_{II}} = \frac{[0.7804 + 1.4217]^2}{660.84 + 164.78} = \frac{4.849}{825.62} = 0.00587$$

[4] For an evolutionary rational of this behavior, see Shermer's article "Patternicity: Finding Meaningful Patterns in Meaningless Noise" in Scientific Anmerican (2008).

[5] df is degrees of freedom.

$$\chi^2_{Homog} = \chi^2_{Total} - \chi^2_{Assoc} = 14.3152 - 0.00587 = 14.30933$$

The LOR between phases 1 and 3 in the two experiments are significantly different. That is, the null hypothesis

$$LOR_I = LOR_{II}$$

is rejected at the 1% level.

4 Acknowledgements

Our largest debt is to Laurie Santos, who has given us sage advice on every aspect of this project. Le-Minh Ho provided excellent research assistance. The insightful comments of Anat Bracha helped to clarify our exposition. The Whitebox Foundation and SOM provided generous financial support for several summers. We thank them all. The data from the experiments, are available on request from the lead author, Oliver Bunn. Finally, we are once more indebted to Glena Ames for her flawless type setting.

5 Appendix A

In logistic regression,where we assume the independence of observations, we maximize the log likelihood L_N. In alternating logistic regression, the initial step is to ignore the dependence of the clustered observations and maximize the log likelihood L_N. The log-likelihood for experiment k:

$$L_{Nk}(\pi^k) \equiv \sum_i \sum_j Y_{ijk} \log \mu_{ijk}(\pi^k) + (1 - Y_{ijk}) \log(1 - \mu_{ijk}(\pi^k))$$

$$\begin{aligned} L_{Nk}(\pi^k) \quad &= \quad \sum_i \sum_j Y_{ijk} \left[K_{Rk} N_{ijk} + \tfrac{\beta_{l(j)k}}{2} R_{ijk} - \left\{ K_{Ak} M_{ijk} + \left(\tfrac{\alpha_{l(j)k} - \beta_{l(j)k}}{2} \right) A_{jk} \right\} \right] \\ &\quad + \sum_i \sum_j \log \left[1 - \exp \left[K_{Rk} N_{ijk} + \tfrac{\beta_{l(j)k}}{2} R_{ijk} - \left\{ K_{Ak} M_{ijk} + \left(\tfrac{\alpha_{l(j)k} - \beta_{l(j)k}}{2} \right) A_{jk} \right\} \right] \right] \end{aligned}$$

To maximize $L_{Nk}(\pi^k)$, we solve the first order conditions for the optimal π^k, where $T_1 = 17$ and $T_2 = 19$:

$$\begin{aligned} (I) \;\; 0 \quad &= \quad \partial_{\alpha_{1k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i=1}^{T_k} \sum_{j=1}^{j=30} Y_{ijk} A_{ijk} \\ &= \quad \sum_{i=1}^{T_k} \sum_{j=1}^{j=30} \left[\frac{A_{ijk} \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{ K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk} \}]}{1 - \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{ K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk} \}]} \right] \end{aligned}$$

$$(II) \quad 0 = \partial_{\beta_{1k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i=1}^{T_k} \sum_{j=1}^{j=30} Y_{ijk}[R_{ijk} + A_{ijk}]$$

$$= \sum_{i=1}^{T_k} \sum_{j=1}^{j=30} \frac{[R_{ijk} + A_{ijk}] \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk}\}]}{1 - \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk}\}]}$$

$$(III) \quad 0 = \partial_{K_{R1k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i} \sum_{j} Y_{ijk} N_{ijk}$$

$$= \sum_{i=1}^{T_k} \sum_{j=1}^{j=30} \left[\frac{N_{ijk} \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk}\}]}{1 - \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk}\}]} \right]$$

$$(IV) \quad 0 = \partial_{K_{A1k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i} \sum_{j} Y_{ijk} M_{ijk}$$

$$= \sum_{i=1}^{T_k} \sum_{j=1}^{j=30} \left[\frac{M_{ijk} \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk}\}]}{1 - \exp[K_{R1k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A1k} M_{ijk} + (\frac{\alpha_{1k} - \beta_{1k}}{2}) A_{jk}\}]} \right]$$

$$(V) \quad 0 = \partial_{\alpha_{2k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i=1}^{T_k} \sum_{j=31}^{j=70} Y_{ijk} A_{ijk}$$

$$= \sum_{i=1}^{T_k} \sum_{j=31}^{j=70} \left[\frac{A_{ijk} \exp[K_{R2k} N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]}{[1 - \exp[[K_R N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_A M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]} \right]$$

$$(VI) \quad 0 = \partial_{\beta_{2k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i=1}^{T_k} \sum_{j=31}^{j=70} Y_{ijk}[R_{ijk} + A_{ijk}]$$

$$= \sum_{i=1}^{T_k} \sum_{j=31}^{j=70} \left[\frac{[R_{ijk} + A_{ijk}] \exp[K_{R2k} N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]}{1 - \exp[K_{R2k} N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}} \right]$$

$$(VII) \quad 0 = \partial_{K_{R2k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i} \sum_{j} Y_{ijk} N_{ijk}$$

$$= \sum_{i=1}^{T_k} \sum_{j=31}^{j=70} \left[\frac{N_{ijk} \exp[K_{R2k} N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]}{[1 - \exp[[K_{R2k} N_{ijk} + \frac{\beta_{1k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]} \right]$$

$$(VIII) \quad 0 = \partial_{K_{A2k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i} \sum_{j} Y_{ijk} M_{ijk}$$

$$= \sum_{i=1}^{T_k} \sum_{j=31}^{j=70} \left[\frac{M_{ijk} \exp[K_{R2k} N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]}{[1 - \exp[[K_{R2k} N_{ijk} + \frac{\beta_{2k}}{2} R_{ijk} - \{K_{A2k} M_{ijk} + (\frac{\alpha_{2k} - \beta_{2k}}{2}) A_{jk}\}]} \right]$$

$$
\begin{aligned}
(IX)\ \ 0 &= \partial_{\alpha_{3k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i=1}^{T_k} \sum_{j=71}^{j=100} Y_{ijk} A_{ijk} \\
&= \sum_{i=1}^{T_k} \sum_{j=71}^{j=100} \left[\frac{A_{ijk} \exp[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]}{[1 - \exp[[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]} \right]
\end{aligned}
$$

$$
\begin{aligned}
(X)\ \ 0 &= \partial_{\beta_{3k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i=1}^{T_k} \sum_{j=71}^{j=100} Y_{ijk} [R_{ijk} + A_{ijk}] \\
&= \sum_{i=1}^{T_k} \sum_{j=71}^{j=100} \frac{[R_{ijk} + A_{ijk}] \exp[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]}{[1 - \exp[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{Ak3} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]}
\end{aligned}
$$

$$
\begin{aligned}
(XI)\ \ 0 &= \partial_{K_{R3k}} L_{Nk}(\pi^k) \Rightarrow \sum_{i} \sum_{j} Y_{ijk} N_{ijk} \\
&= \sum_{i=1}^{T_k} \sum_{j=71}^{j=100} \left[\frac{N_{ijk} \exp[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]}{[1 - \exp[[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]} \right]
\end{aligned}
$$

$$
\begin{aligned}
(XII)\ \ 0 &= \partial_{K_{Ak}} L_{Nk}(\pi^k) \Rightarrow \sum_{i} \sum_{j} Y_{ijk} M_{ijk} \\
&= \sum_{i=1}^{T_k} \sum_{j=71}^{j=100} \left[\frac{M_{ijk} \exp[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]}{[1 - \exp[[K_{R3k} N_{ijk} + \frac{\beta_{3k}}{2} R_{ijk} - \{K_{A3k} M_{ijk} + (\frac{\alpha_{3k}-\beta_{3k}}{2}) A_{jk}\}]} \right].
\end{aligned}
$$

We now follow section 13.6 in Fitzmaurice et al., for GEE using $PROC\ GENMOD$ in SAS

6 Appendix B

In the second experiment, the relative frequency of outcomes of ambiguous lotteries in phase 2 simply does not converge to any probability. The proof is an immediate consequence of Kolmogorov's Strong Law of Large Numbers ($SLLN$): Let $\{X_{j,}\}$, $j = 1, 2, ...$ be a sequence of independent and identically distributed random variables, then a necessary and sufficient condition that $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j \rightarrow^{a.s.} \mu$ is that $E[X_j] = \mu$, for $j = 1, 2, ...$

Define Z_j, where $Z_j = 1$ with probability p_j if the realized payoff of the ambiguous lottery $y_j = (y_{n,1}, y_{n,2})$ is $y_{n,1} \vee y_{n,2}$, the higher payoff and $Z_n = 0$ otherwise. Let $\bar{Z}_n = \frac{1}{n}\sum_{j=1}^{j=n} Z_j$, then it follows from Kolmogorov's $SLLN$ for the independent and identically distributed random variables Z_n that $\bar{Z}_n \rightarrow^{a.s} p$ iff $p = p_j$ for all j — see Rao (2002, p. 115) for the proof of the $SLLN$. We call the $\bar{Z}_n$ counterfactual probabilities, where $\bar{Z}_n$ is the counterfactual probability that the payoff of $y_n = (y_{n,1}, y_{n,2})$ is $y_{n,1} \vee y_{n,2}$ and $(1 - \bar{Z}_n)$ is the counterfactual probability that the payoff is $y_{n,1} \wedge y_{n,2}$. That is, $\bar{Z}_n$ converges in the first experiment but they do not converge in the second experiment, where $\bar{Z}_{2n}$ and $\bar{Z}_{2n+1}$ converge to different limits.

References

[1] Boorman, Erie D., Timothy E. Behrens, and Matthew F. Rushworth (2011). "Counterfactual Choice and Learning in a Neural Network Centered on Human Lateral Frontopolar Cortex."

[2] Bracha, A., and D. J. Brown (2013). "Keynesian Utilities: Bulls and Bears," forthcoming CFDP.

[3] Carey, V., S.L. Zeger and P. Diggle (1993). "Modelling Multivariate Binary Data with Alternating Logistic Regressions," *Biometrika*, 517–526.

[4] Fitzmaurice, Garrett M., Nan M. Laird, James H. Ware (2011). *Applied Longitudinal Analysis* (Wiley Series in Probability and Statistics).

[5] Fleiss, Joseph L., Bruce Levin, and Myunghee Cho Paik (2003). *Statistical Methods for Rates and Proportions*, Third Edition, John Wiley & Sons, Inc.

[6] Hayden, Benjamin Y., John M. Pearson, and Michael L. Platt (2009). "Fictive Reward Signals in Anterior Cingulate Cortex," *Science,* 948–950.

[7] Huettel, Scott A., C. Jill Stowe, Evan M. Gordon, Brent T. Warner, and Michael L. Plat (2006). "Neural Signatures of Economic Preferences for Risk and Ambiguity," *Neuron* 49, 765–775.

[8] Keynes, J.M. (1937). "The General Theory of Employment," *Quarterly Journal of Economics*, 209–223.

[9] Levy I,, J. Snell, A.J. Nelson, A. Rustichini, and P.W. Glimcher (2010). "Neural Representation of Subjective Value under Risk and Ambiguity," *Journal of Neurophysiology*, 1036–1047.

[10] Li, Yimei, Hongtu Zhua Lia, Hongtu, Yasheng Chenb, Hongyu, John Gilmorec, Weili Lin and Dinggang Shenb (2009) "STGEE: Longitudinal Analysis of Neuroimaging Data"Proceeding," Proceeding of the SPIE.

[11] Liang, K.-Y. and S.L. Zeger (1986). "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, 13–22.

[12] Lipsitz, Stuart R., Nan M. Lair and David P. Harrington (1991). "Generalized Estimating Equations for Correlated Binary Data: Using the Odds Ratio as a Measure of Association," *Biometrika*, 153–160.

[13] Lohrenz, Terry, Kevin McCabe, Colin F. Camerer, and P. Read Montague (2006). "Neural Signature of Fictive Learning Signals in a Sequential Investment Task," PNAS, 9493-9498.

[14] Platt, M., and Hayden,B. (2011). "Learning: Not Just the Facts, Ma'am, but the Counterfactuals as Well," PLoS Biol 9(6): e1001092.

[15] Rao, C.R. (2002). *Linear Statistical Inference and its Applications*, John Wiley and Sons.

[16] Rockafellar, R.T. (1970). *Convex Analysis*, Princeton University Press.

[17] Sherman, M., and S. le Cressie (1997), "A Comparison Between Bootstrap Methods and Generalized Estimating Equations for Correlated Outcomes in Generalized Linear Models,"Communications in Statistics.

[18] Zeger, S.L., and Liang, K.-Y. (1986). "Longitudinal Data Analysis for Discrete and Continuous Outcomes," *Biometrics*, 121–130.