Year: 2011

Correlated individual differences and choice prediction

Lindsay, Luke

Abstract: This note briefly summarizes the consequences of adding correlated individual differences to the best baseline model in the Games competition, I-SAW. I find evidence that the traits of an individual are correlated, but refining I-SAW to capture these correlations does not significantly improve the model's accuracy when predicting average behavior.

games

Short Note

Correlated Individual Differences and Choice Prediction

Luke Lindsay

Department of Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland;
E-Mail: luke.lindsay@econ.uzh.ch

Abstract: This note briefly summarizes the consequences of adding correlated individual differences to the best baseline model in the Games competition, I-SAW. I find evidence that the traits of an individual are correlated, but refining I-SAW to capture these correlations does not significantly improve the model's accuracy when predicting average behavior.

Keywords: individual differences; choice prediction; I-SAW; modeling correlation

1. Introduction

Are individual differences correlated and can modeling them as such increase the accuracy of a model's predictions? Correlations between individual differences was one of the features of the model I entered in the *Games* choice prediction competition (CPC). This note briefly summarizes the consequences of adding correlated individual differences to the I-SAW model, the best baseline model in the CPC. It is assumed the reader is familiar with the experiments, I-SAW model, and competition described by Erev *et al.* [1].

One of the regularities observed in the results of the CPC estimation experiment and previous studies is that individuals differ in their behavior [2,3]. In the I-SAW model, individual differences in behavior are determined by five trait parameters, which are summarized in Table 1.

Table 1. The five trait parameters.

Parameter and Distribution	Description
$\varepsilon_i \sim U[0,0.24]$	Probability of exploration in trials after the first one.
$\pi_i \sim U[0,0.6]$	Tendency for inertia.
$\mu_i = \{1, 2, \text{ or } 3\}$	Number of samples taken in exploitation trials.
$\rho_i \sim U[0,0.2]$	Probability a sample draw is biased. If the draw is biased, the most recent trial is selected. If it is unbiased, a previous trial is selected at random.
$w_i \sim U[0,0.8]$	In exploitation trials, the sample mean is given weight $(1 - w)$ and the mean of all previous trials weight w.

In the standard version of I-SAW, it is assumed that the values an individual's trait parameters take are independent. This study is motivated by the conjecture that traits are correlated. For example, one might expect that exploration (ε_i), the tendency to choose at random, is negatively correlated with inertia (π_i), the tendency to repeat the previous choice. Conversely, one might expect that the tendency of participants to give more weight to average payoffs (w_i) is positively correlated with inertia (π_i).

Data from the CPC experiments were used to investigate correlations between trait parameters. First, the individual decisions made by the 120 participants in the estimation experiment were used to calculate maximum likelihood estimates of the five trait parameters for each participant. These estimates provide evidence that traits are correlated and not independent. Second, the estimation experiment was simulated with a refined I-SAW model. Parameters were constrained to have the same uniform distributions as in the baseline I-SAW model. Selected correlations between trait parameters were introduced and estimated using a grid search. Interestingly, adding these correlations between trait parameters while holding the model and distribution of parameters constant had only a very small effect on how accurately the model predicts average behavior in the estimation and competition experiments.

2. Estimating Trait Parameters and Identifying Correlations

Entry decisions are stochastic in the I-SAW model. In the first trial, each player enters with a fixed probability. In subsequent trials, the probability player i enters depends on the realized and forgone payoffs in the previous $t - 1$ trials and on player i's traits. Traits vary between individuals, but for a given individual are constant across trials and decision problems.

Each of the 120 participants in the CPC estimation experiment played 10 games, g, and each game had 50 trials, t. For each participant, the 500 observed entry decisions were used to calculate maximum likelihood estimates of a vector containing the five trait parameters $\boldsymbol{\theta_i} = (\varepsilon_i, \pi_i, \mu_i, \rho_i, w_i)$. The following log likelihood function was used (a similar approach has been used by Yechiam and Busemeyer [3,4]):

$$\ln \mathcal{L}(\boldsymbol{\theta_i} \mid data) = \sum_{g} \sum_{t} \ln\left(\Pr\left[G_g(t+1) \mid V_g(t), \boldsymbol{\theta_i} \right] \right) \qquad (1)$$

where $G_g(t + 1)$ denotes the participant's entry decision in trial $t + 1$ of game g and $V_g(t)$ is a matrix of the participant's payoffs in game g (including those forgone) from trials up to and including t. The I-SAW model has three response modes: exploration, inertia, and exploitation. The likelihood function

giving probability of entry conditional on individual traits and previous outcomes was constructed as follows

$$\Pr[\,enter\,] = \Pr[\,explore\,].\Pr[\,enter \mid explore] + \Pr[intertia].\Pr[\,enter \mid intertia]$$
$$+ \Pr[\,exploit].\Pr[\,enter \mid exploit] \tag{2}$$

For exploration and inertia, the exact probability of entry was calculated directly. For exploitation, to accommodate the internal stochastic component of I-SAW (drawing a small sample of previous outcomes of the current game), it was calculated as follows

$$\Pr[enter \mid exploit] = \sum_{sample \in S} \Pr[sample].\Pr[enter \mid exploit, sample] \tag{3}$$

where S is the set of all samples that it is possible for the player to draw.

The vector of parameters $\boldsymbol{\theta_i}$ was estimated using a nonlinear optimization method. The parameter values were constrained by the upper and lower bounds of the distributions shown in Table 1. The correlation coefficients and summary statistics of the estimates are shown in Table 2. The coefficients shown in the top half of the table suggest an individual's trait parameters are not independent. The following relationships between individual traits are apparent. Participants with a higher tendency to explore ε_i have a lower tendency to exhibit inertia π_i. They also give less weight to average payoffs over all previous trials w_i and more weight to a small sample of previous trials $(1 - w_i)$. This would cause them to have a greater tendency to underweight rare events. Participants with a higher tendency for inertia π_i, in contrast, give more weight to average payoffs w_i. Finally, in sampling during exploitation trials, participants with a greater tendency to bias their sample by selecting the most recent trial ρ_i, give less weight to average payoffs w_i.

Table 2. Correlation coefficients and summary statistics for estimated trait parameters.

Parameter	ε_i	π_i	μ_i	ρ_i	w_i
ε_i	1.00				
π_i	−0.38***	1.00			
μ_i	−0.03	0.14	1.00		
ρ_i	−0.22*	−0.14	−0.14	1.00	
w_i	−0.31***	0.24**	0.05	−0.27**	1.00
mean	0.12	0.31	1.87	0.12	0.29
variance	0.01	0.04	0.70	0.01	0.05
max	0.24	0.60	3.00	0.20	0.80
min	0.00	0.00	1.00	0.00	0.00

*:$p < .05$, **:$p < 0.01$, ***:$p < 0.001$

3. Adding Correlated Individual Differences to I-SAW

The only constraints on the parameter estimates reported in the previous section were the upper and lower bounds. To refine the I-SAW model to accommodate correlated traits, two additional constraints were imposed: traits were assumed to be uniformly distributed between the upper and lower bounds, and the correlations between traits were assumed to have a specific structure as described below.

The following procedure was used to generate $\boldsymbol{\theta}_i$, a 1×5 vector of correlated trait parameters, where each component θ_{ij} has a predefined distribution (specified in Table 1) with cumulative distribution function $F_{\theta j}$. First, a vector $\boldsymbol{x} = (x_1, \dots, x_5)$ of independent standard normal values was generated. This was used to generate a vector of correlated standard normal values

$$\mathbf{y} = \mathbf{xL} \tag{4}$$

where **L** is a 5×5 lower triangular matrix with strictly positive diagonal elements. The values in each column of **L** are subject to the constraint $\sum_j^5 l_j^2 = 1$. As a consequence, the values contained in **y** are from the standard normal distribution[1]. Finally, the vector $\boldsymbol{\theta}_i$ was produced by transforming each component of **y** using $F_{\theta j}$ and the standard normal cumulative distribution function Φ.

$$\theta_{ij} = F_{\theta j}^{-1}\left(\Phi(y_j)\right) \tag{5}$$

Notice that when **L** is the identity matrix, the generated values $\boldsymbol{\theta}_i$ are independent; when **L** has non-zero off-diagonal elements, the generated values $\boldsymbol{\theta}_i$ are not independent.

The matrix **L** was estimated using a simulation based grid search. To limit the time required to compute the estimates, attention was restricted (a) to elements of the matrix that give rise to the four correlations that Section 2 identified as statistically significant at the one percent level, and (b) to values of these elements around those that would produce the correlation coefficients estimated in Section 2. The remaining off-diagonal elements were fixed equal to zero. The matrix that minimized the normalized mean square deviation scores (the criteria used to judge models in the CPC) is shown below. The elements of the matrix that were varied during the grid search are shown in bold.

$$\hat{\mathbf{L}} = \begin{pmatrix} \mathbf{1.00} & 0 & 0 & 0 & 0 \\ \mathbf{0.00} & \mathbf{0.99} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{0.80} & 0 \\ \mathbf{0.00} & \mathbf{0.10} & 0 & \mathbf{-0.60} & 1 \end{pmatrix} \tag{6}$$

On the estimation set, using these estimates gives a score of 1.34 compared to 1.37 when there are no correlations. The respective figures for the competition set are 1.17 and 1.19. In both cases, introducing correlations leads to a slight increase in the accuracy of the model's predictions.

---

[1] The sum of independent normally distributed random variables has the following property: if $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ then $(X_1 + X_2) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Further, if $X \sim N(\mu, \sigma^2)$ and $Y = aX$, then $Y \sim N(a\mu, a^2\sigma^2)$. Let $\boldsymbol{x} = (x_1, \dots, x_n)$ be a vector of independent standard normal values. Let $Z = \sum_i a_i x_i$. If $\sum_i a_i^2 = 1$, it follows $Z \sim N(0,1)$.

4. Discussion

This study found that while correlations between traits matter for individual behavior, refining I-SAW to capture these correlations does not significantly improve the prediction of averages such as average entry rates, efficiency, and alternation rates. Hence, when the goal is prediction of average rather than individual behavior, assuming individual trait parameters are independently distributed appears to be a sound simplifying assumption. A natural question for future research is why the refinement of I-SAW did not produce better predictions. One possibility is that since participants interacted in groups, there may be group effects that carry over to the parameters obtained by fitting the model separately to each individual. Another direction for future research is testing models with fewer trait parameters. If, as this study suggests, correlations between trait parameters only have a small effect on the accuracy of the model's prediction of average behavior, it may be possible to achieve the same degree of predictive accuracy with a model that is simplified by combining some of the trait parameters.

References and Notes

1. Erev, I.; Ert, E.; Roth, A. A choice prediction competition for market entry games: An introduction. *Games* **2010**, *1*, 117-136
2. Ert, E.; Yechiam, E. Consistent constructs in individuals' risk taking in decisions from experience. *Acta Psychol.* **2010**, *134*, 225-232.
3. Yechiam, E.; Busemeyer, J.R. Evaluating generalizability and parameter consistency in learning models. *Games Econ. Behav.* **2008**, *63*, 370-394.
4. Yechiam, E.; Busemeyer, J.R. Comparison of basic assumptions embedded in learning models for experience based decision-making. *Psychonomic Bull.* **2005**. *12*, 387–402.