Outline

Timothy Schwieg

March 23, 2018

1 Reading list

References

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Prespectives 27*, 173–196.

Cripps, M. W. and J. M. Swinkels (2004). Efficiency of large double auctions.

Li, H. (2015). Nonparametric identification of k-double auctions using price data. Job Market Paper.

Li, H. and N. Liu (2015). Nonparametric identification and estimation of double auctions with bargaining. *Shanghai University of Finance and Economics*.

Paarsch, H. J. (2018). The foundations of empirical intuition. Version: Sunday 7th January, 2018.

Parsons, S., M. Macinkiewicz, J. Niu, and S. Phelps (2006). Everything you wanted to know about double auctions, but were afraid to (bid or) ask. Unpublished Paper.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative reperesentation of uncertainty. *Journal of Risk and Uncertainty 5*, 297–323.

2 Model

2.1 General Approach

Because of the nature of the data, as well as the complicated structure that Cumulative Prospect theory places upon the valuation function, it is difficult to use it within the structure used commonly in double auction literature. Almost all models of auctions specify a structure between the valuation and the price that takes the form of the valuation is the valuation for the item minus the price paid for the item at the auction. This difference cannot be reconciled with the framing effects required by Cumulative Prospect Theory.

To circumvent this issue, I use the result found in Cripps and Swinkels (2004) which states that a continuous double auction converges quickly, in order $O(N^2)$ to a competitive market. From this, I can consider that in large double auctions, such as is the case for the purchases of the lotteries in question. From this we can conduct our analysis as if a buyer of a lottery is at a competitive market, sees the price given by the market and makes a decision choice based upon his valuation, as a function of the price.

This leads us to one of two outcomes in the market, either there is a purchase made, indicating a valuation above the price seen by the consumer, or he leaves a buy order, indicating a price he would be willing to pay. Since in a double auction where the buyer pays the seller's price, it is dominant to tell the truth I will consider these buy orders as the actual valuations of the consumer Li (2015).

2.2 Discrete Choice

By Following the discrete choice model, we assert that decisions made by the individual are made such that if the valuations of lottery plus ξ_l is greater than the valuation of not purchasing which is given by ξ_n where $\xi_{\mathrm{j}} \sim$ Gumbel. This means that the purchase is dictated by if: $V(f)+\xi_b-\xi_n > 0$. Since the difference of two Gumbel distribution is distribution logistically, and V(f) is not random. Knowing that the Logistic Distribution is in the location scale family, we may consider the entire expression as: $V_i(f) \sim logit(V(f), s)$.

However, we observe some censored realizations of V. That is, we observe valuations based on the buy orders which are non-censored realizations, but we also observe purchases which are censored. Denote the censored purchases by $d_n = 1$ and uncensored by: $d_n = 0$. The likelihood function for censored data as shown by Paarsch (2018) is:

$$\sum_{j=1}^{J}\sum_{n=1}^{N_j} d_{n,j}\log(\frac{1}{4s}sech^2(\frac{x_{n,j}-V_j}{2s})) + (1-d_{n,j})\log(\frac{\exp(\frac{x_{n,j}-V_j}{s})}{1+\exp(\frac{x_{n,j}-V_j}{s})})$$

Where V_j is the valuation for the box of type j, $x_{n,j}$ is the price paid by observation n of type j, and $d_{n,j}$ is whether or not the data at observation n of type j was censored.

We may maximize this likelihood to find the parameters that best fit. What remains to be decided is the form of the function V. We will have to assume a structural form for the function, and while we will be able to test between structural forms using a validation set, there will still be assumptions held.

2.3 Cumulative Prospect Theory

Since one key aspect of cumulative prospect theory is that we are more strongly motivated by losses than by gains, it is immediately obvious that the value function of the contents of the lottery will not be symmetric, and the easiest way to handle this will be to estimate two separate functions, one of which is used to evaluate losses, and one of which estimates gains. By applying a piece wise function where the loss function is used on losses, and the gain function for gains, we arrive at a continuous function (since both are zero at zero) which we can use. The question of whether or not loss aversion is displayed is one for the empirics.

Since I am following the Discrete Choice model, the decision to purchase a create will be driven by the valuations of the crate against the alternative which is buying nothing. Since Cumulative prospect theory functions by looking at deviations from a reference point, which we will use as the price of the box combined with the costs of opening it (the key).

Following the Notation of Tversky and Kahneman (1992), we represent the valuation of the lottery as $V(f) = \sum_{i=-m}^{n} \pi_i v(x_i)$ where v is a function that is convex in losses and concave in gains. π is a weight function that will be defined in a later section.

The question of a reference point is commonly debated, however in this application it is clear that the reference point for the crate is the cost of opening such a crate, which is the sum of the "key" that must be purchased from Valve in order to open, and the price paid for the crate. The cost of a key is constant, and marks the direct cost of participation, they are very often purchased directly before opening the crate, as when you attempt to open it, you are prompted with the price and a link to purchase a key. I believe that it is therefore reasonable to ignore all effects time play play on the problem, and treat the issue as if the consumer simply buys the key and crate at the same time as he opens it.

2.3.1 Valuations

The question of how do we measure the gains of the lottery now looms. Since each box when opened contains an item that has a particular value to the consumer who opened it, and is unobserved, the only thing that can be observed is the market price of the item over time. Identification of buyer and seller valuations, even in a simplified situation where there is only one buyer and one seller, still requires more information than we are given. As shown in Li and Liu (2015), for their identification strategy, all the bids in the auction are required, and for the strategy shown in Li (2015), there must exist exclusive covariates that shift only one trader's value distribution, which are not given by the data.

Absent an ability to identify the valuations of the specific losses and gains in the lottery, an identifying assumption will have to be made. I will represent valuations of the contents of the lottery as the weighted average of the purchases made, weighted by the quantity purchased. This is effectively the sample mean of prices purchased, which I will assume to be average valuations of the good.

The valuations of the i^{th} possible element of the lottery will be given by: $v(\, p_i - p_l - p_{key}\,)$, where p_i is the average price of the i^{th} element at market, p_l is the price paid for at market for the lottery, and p_{key} is the price of the key required to open the lottery. Depending on whether or not this difference is positive or negative will result in different functions being used to evaluate the valuation. Using the specification suggested by Tversky and Kahneman (1992):

$$v(x) = \begin{cases} x^{\alpha} & x \geq 0 \\ -\lambda(-x)^{\alpha} & x < 0 \end{cases}$$

2.3.2 Probability weighting function

In Cumulative Prospect theory, the cumulative mass (distribution) function is weighted such that individuals overweight the tail probabilities. This is especially important in this model, as there are many high valued rare items, that if this part of the theory is correct, heavily influence the valuation of the box, despite their extremely low probability of occurrence.

Again, I will use the specification suggested by Tversky and Kahneman (1992), and use the cumulative transformation function of:

$$w(P) = \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{\frac{1}{\delta}}}$$

To define the decision weights π_i, we must first order the prospects of the lottery in ascending order of gains. the weight π_i then is defined by:

$$\pi_i = w(\sum_{j=-m}^{i} P(x_j)) - w(\sum_{j=-m}^{i-1} P(x_j))$$

2.3.3 The Valuation function

From these, we can create a valuation function for an individual facing a particular lottery:

$$V_i = \begin{cases} (w(\sum_{j=-m}^{i} P(x_j)) - w(\sum_{j=-m}^{i-1} P(x_j)))(p_i - p_l - p_{key})^\alpha & (p_i - p_l - p_{key}) \geq 0 \\ -\lambda(w(\sum_{j=-m}^{i} P(x_j)) - w(\sum_{j=-m}^{i-1} P(x_j)))(p_l + p_{key} - p_i)^\alpha & (p_i - p_l - p_{key}) < 0 \end{cases}$$

3 Data

3.1 The Data

The data is pricing data of items on the Steam Community Market for the game Counter Strike: Global Offensive. Players in game earn items random that they can sell on the market or open themselves. However most rare items are earned via opening of dropped "loot boxes" that are then opened by players via purchasing of a key. These boxes can be earned by playing or received randomly from players who are watching games of professionals play. The probabilities of the drops are not known or even estimated well, as they change depending on many factors including time.

However, once a box has been obtained, the probability of receiving an item is well documented as per Chinese Law. Each item has a certain grade of rarity, for example the Ak-47 Redline has a rarity level of Classified which means that there is a 3.2% chance of receiving a Classified item in the crate. All Classified items contained in the crate have the same probability of being dropped by the crate.

However there are many variants of each item. Each item has a quality ascribed to it, the float of the item. This describes the wear on the item, and is distributed uniformly on the interval 0-1. On the market the items are split into intervals: Battle-scared, well-worn, field-tested, minimal wear and factory new. Each quality is a separate listing on the market with a separate price. In addition to each item having a quality type there is also a 10% chance of each item being labeled as StatTrak, which also distinguishes the value of a weapon. This means that each item has 10 different variations all with different probabilities of being obtained. Some rare items, usually knives and gloves may have more or less variants, but the amount and probabilities are known.

Each box contains some subset of these items that is known, and the market value of each item at a particular time period is also known, so the expected value, or any other modified version of a valuation of the lottery can easily be calculated.

3.2 Sources of the Data

The data has been mined from the steam community market api, which provides a purchase history for every item on the market, down to the hour for the last thirty days and daily for the rest of the lifetime of the item. It does not provide a record of every purchase, just the quantity sold in that time period as well as the median price they were sold at. Obviously this is less than ideal, but I believe it will cause less problems than the inaccuracies introduced by the market only working in one cent intervals.

Also available is current buy and sell orders for each item. If a potential buyer wishes to buy on this market, he may either select a box directly and purchase from a particular seller, or he may put forward a buy order, which he stipulates a price, and as soon as a seller puts an item up for sale below that price, it is sold to the buyer, and he is charged the seller's price. This gives the valuations of people who have not yet obtained the item directly. However, it does not appear that there is a history available for these items. In some ways this is beneficial because it would impossible to determine the differences between buy orders that were fulfilled and buy orders that were removed because of changes in the prices of underlying assets. I have decided to treat all outstanding buy orders as valuations in the final time period that simply are below the market price. I will not consider the case that there are buy orders placed and forgotten about.