Lilyana Mikova

Johan Olofsson

Advisor: Seethu Seetharaman, W. Patrick McGinnis Professor of Marketing

Investors' Attitudes Toward Complexity in Financial Markets

Senior Honors Thesis

Introduction

In April of 2009, Andrew Haldane, Bank of England's Executive Director of Financial Stability, gave a speech for the Financial Student Association in Amsterdam titled *Rethinking the Financial Network* . In it, he describes the prerequisites for properly understanding some of the products that contributed to massive trading losses during the financial crisis of 2008-2009.

"An investor in a CDO^2 would need to read in excess of 1 billion pages to understand fully the ingredients. With a PhD in mathematics under one arm and a Diploma in speed-reading under the other, this task would have tried the patience of even the most diligent investor. With no time to read the small print, the instruments were instead devoured whole. Food poisoning and a lengthy loss of appetite have been the consequences" (Haldane, 17).

This pre-crisis ignorance of the complexity present in some securities such as CDO^2s resulted in a consistent mispricing and subsequent overvaluation of these assets. In today's post-crisis world, the sentiment seems to be that while very few of us are able to price this complexity, we are aware that the risk is present, and as such, all else equal, discount the price of these assets relative to simpler assets. It is with this backdrop that we wrote this paper and designed a discrete choice experiment to see how individuals pick investment instruments and what attributes are valued.

Research Question

Our hypothesis is that besides attributes of expected return and the accompanying riskiness of the instrument, investors would also prefer simplicity. We believe that the recent crisis has increased the demand for straightforward, relatively simple instruments at the expense of the highly complex derivatives that became infamous in 2008.

We believe that applying discrete choice analytical methodology to strip out investor preferences provides some unique insights into why certain securities and asset classes may be preferred over other investments. With that being said, conducting an experiment in a laboratory setting with no money at risk for making poor investment decisions is a key distinction to the real investment environment. On the other hand, the laboratory setting of our experiment allows for us to control all the inputs that the investor is using to make his or her decision, which allows us to analytically determine the investor's beliefs about each attribute of the investment.

Justification for Research

The results from this research could be useful for private wealth managers in determining the preferences of their clients. Many private wealth managers give their clients a survey that attempts to isolate the individual's preference for risk taking, asset classes, and more. Our survey uses a sophisticated analytical methodology to strip out individual preferences, which could be very valuable in the world of private wealth management.

In addition to providing useful data on the retail investor for decision making within private wealth management, our sample group provides a collection of both future retail and institutional investors. Our discrete choice survey was administered to students via social media, professors, and our personal network. We recorded the student's gender, major, and academic program, which allowed us to control for these variables and determine if there is any difference in preferences across demographic groups. While we anticipate participants' preferences will change over time, this study gives a snapshot of the preferences of students towards financial complexity, all else equal. Unless they have significant future exposure to certain areas of finance, we do not expect their preferences to change significantly, making our results quite applicable to a future generation of retail and institutional investors.

Literature Review

Brunnermeier and Oehmke look at why complexity matters and what defines a complex security. They point out that this is a relatively unexplored topic, since modern asset pricing theory, as well as economic theory in general, assume a rational paradigm. In a rational world the complexity of a security does not matter, since, given sufficient information, the investor has unlimited ability to comprehend that information, conduct complicated calculations, and find the fair value of any security almost instantly.

The authors then introduce the concept of Bounded Rationality. Consider an investor, who has limited ability to process information because of computational limits or inability to keep track of an infinite number of variables. Complexity matters to this investor. Moreover, providing additional information would not solve this problem, but rather lead to information overload, as Mr. Haldane points out in the quote in the beginning of our paper. Besides the amount of information released, the way it is presented and simplified becomes crucial. High complexity could also lead to information asymmetry that occurs when some investors are better at processing complex information.

The article discusses the difficulty of formulating a definition of complexity as pertaining to financial instruments. A simple comparison between two securities, the notoriously complex CDO tranches, and a simple equity share, illustrates this difficulty. CDOs have a complicated cash flow structure, mainly because of the difficulty of estimating the correlation of defaults among the loans that comprise it. An equity share, on the other hand, is perceived as relatively simple security with a straight-forward payoff. However, if an investor is to value the equity share using a bottom-up approach, he or she needs to consider all lines of business of the company and to estimate the cash flows of all projects. If the firm is a financial institution, it might very well have CDOs on its balance sheet. Therefore, valuing equity shares is not necessary simpler than valuing CDOs.

We consulted the article Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples by Goodman, Cryder and Cheema to make a more informed decision on how to select participants for our surveys. While MTurk makes it easy to administer a survey with a high number of participants in a short amount of time, there are concerns that the registered participants might be unusual and so their responses cannot be used to gauge the attitudes of the larger population. The article strives to compare MTurk participants to student samples on a set of personality dimensions and classic decision-making biases. The authors review existing research and administer two studies. They find many similarities between MTurk participants and traditional samples, but also some differences. For example, MTurk participants tend to pay less attention to study materials, reducing statistical power. Furthermore, they are more risk averse than traditional samples and tend to value money more and their time less than the average community participant.

We found the article *Quick and Easy Choice Sets: Constructing Optimal and nearly Optimal Choice Experiments* by Street, Burgess and Louviere very helpful in guiding us when we constructed our discrete choice experiment. The authors compare a few common strategies of constructing an experiment and discuss their effectiveness in capturing consumer preferences. DCEs are used in marketing to estimate the effect of attributes on the overall “attractiveness” of a particular product. How accurate of a depiction of consumers’ attitudes researchers construct depends on what options have been used in the choice experiment, as well as how those options were grouped into choice sets. The authors present the optimal designs for a number of levels and choice sets. Most experiments created using the approach described in the paper have a level of efficiency of 90% or more.

Survey Formulation and Strategy

Our research question has two components: we want to investigate investors’ attitudes towards the securities that became infamous during the financial crisis, as well as perception and tolerance for complexity. While CDOs could be perceived as both undesirable because of the publicity they received during the crises, as well as complex, we wanted to separate the two factors in the eyes of the investor and to present them as two separate characteristics for consideration. To do this, we administered two initial control surveys, one testing for a perception for complexity and the other for negative associations stemming from the name of the security as a reminder of the bad publicity it received during the financial crisis.

We followed the general logic of the control surveys in the paper *Can Nervous Nelly Negotiate* by Brooks and Schweitzer. The authors set out to test how emotional state influences a participant’s ability to negotiate. Before they perform their main set of experiments, however, they administered surveys with a separate group of subjects to see if the stimuli they had chosen were effective in inducing the emotions they were seeking.

Similarly, in one control survey we set out to test if the description of the security we are providing is associated with complexity. In another, we set out to test for negative bias. Both surveys contained four securities. The respondents were asked to rate each based on their perceived complexity on a scale of 1 to 100. In light of Brunnermeier and Oehmke’s discussion of the many different ways of defining complexity, we came up with the following simple working definition: the complexity of an instrument depends on the difficulty of predicting cash flows in different states of the world. In this way, we limited the evaluation of complexity to only one characteristic. In the first control survey we provided descriptions for the securities taken from Investopedia and replaced the name of the security CDO with ABC. We wanted to test if participants would rate security ABC as more complex based only on its description and not on any negative bias towards the name. In the second control survey, we only provided the names of the securities, without descriptions, using the CDO name without replacing it. We were testing for bias stemming only from the name of the security, without influencing the survey takers by giving them security descriptions that vary in complexity.

We chose to administer our survey using Qualtrics and targeted the college population. We think that this is the best segment, since this generation's perceptions about investing have been shaped by the financial crisis. They are also future investors, so knowing their preferences would have direct benefits. Furthermore, while Mechanical Turk would have given us more results, we believe that the opinions expressed by those participants would have skewed our results, as there are cases when they have reported fundamentally different perceptions towards money and investing. Our surveys also contain detailed instructions, as well as complex definitions of securities, which could be problematic with Mechanical Turk participants, who have been reported to pay less attention to study materials than college survey takers. That is why we decided to use Qualtrics and distribute the survey by asking professors to send it out to their undergraduate or graduate classes.

We had a total of 100 participants take our surveys. To make sure we do not let the same participant take both surveys, we set a cap on Qualtrics for the first survey. Once the cap was filled, participants were automatically redirected to the second survey.

The following table summarizes our results:

#	Answer	Min Value	Max Value	Average Value	Standard Deviation
1	Bond	0.00	73.00	27.55	18.41
2	Stock	0.00	100.00	46.23	27.66
3	CDO	20.00	100.00	63.95	23.01
4	ETF	8.00	100.00	63.09	24.85

#	Answer	Min Value	Max Value	Average Value	Standard Deviation
1	Stock	0	85	39.48	26.66
2	ETF	5	93	48.59	24.4
3	Bond	1	76	22.79	17.59
4	ABC	1	100	53.09	26.81

While CDO was rated closely to an ETF in the second control survey, ABC was rated as significantly more complex than the other securities in the first control survey. This result shows that the complexity associated with this security is objective and stems from its description rather than from negative connotations of the name. Since we want to test for investors' attitudes towards complexity and not their preconceived bias against CDOs, we decided to use the security descriptions and to designate the CDO as ABC. We used the same security descriptions in our discrete choice survey, since our control survey showed that those descriptions portray the ABC security as more complex than the others. With the discrete choice model, we set out to test whether that complexity influences investor choices.

Discrete Choice Experiment

Based on the designs described in *Quick and Easy Choice Sets: Constructing Optimal and nearly Optimal Choice Experiments* by Street, Burgess and Louviere, we created a discrete choice experiment with four attributes: security type, return, liquidity and accessibility of information. We used four levels for security: bond, stock, ETF and ABC, and four levels for return: 2%, 4%, 6%, and 8%. Liquidity and accessibility of information have two levels each, high and low. We created a 16 question survey using the optimal design for this number of attributes and levels as described in the paper. The survey takers were asked to read the descriptions for the securities first and then to choose between three choice sets in each question, consisting of a combination of security, return and liquidity or accessibility of information level.

We set up a timer for the survey, since we wanted to make sure that participants take the time to consider the different options. Participants could not proceed, unless they have spent at least 45 seconds on the directions and security descriptions page. They needed to spend 15 seconds on each subsequent question. We believe those timers are the reason a large portion of our participants did not finish answering all questions. While a few of those participants who gave up might have been fast readers who got bored waiting for the timers, we think a significant number were survey takers who wanted to finish as quickly as possible without considering their answers carefully. The survey timers seem like an effective way of filtering through those participants, even if we ended up with less survey takers.

Data Analysis

Given that our intention was to create a discrete choice experiment, we created a series of maximum likelihood estimators (MLE) to find the model that best fits our collected data. With the guidance of Professor Seetharaman we set out to compute manually the maximized log likelihood of a fully heterogeneous, fully homogeneous, and segmented models. Using the method prescribed in the Street, Burgess and Louviere paper described above, we created the following sixteen choice sets:

	Media	Liquidity	Security	Return	Media	Liquidity	Security	Return	Media	Liquidity	Security	Return
1	High	High	CDOs	2%	Low	Low	ETF	8%	High	Low	Stock	4%
2	High	Low	CDOs	4%	Low	High	ETF	6%	High	High	Stock	2%
3	Low	High	Stock	2%	High	Low	Bond	8%	Low	Low	CDO	4%
4	Low	Low	Stock	4%	High	High	Bond	6%	Low	High	CDO	2%
5	Low	Low	CDO	6%	High	High	ETF	2%	Low	High	Stock	8%
6	Low	High	CDO	8%	High	Low	ETF	4%	Low	Low	Stock	6%
7	High	Low	Stock	6%	Low	High	Bond	2%	High	High	CDO	8%
8	High	High	Stock	8%	Low	Low	Bond	4%	High	Low	CDO	6%
9	Low	Low	Bond	2%	High	High	CDO	8%	Low	High	ETF	4%
10	Low	High	Bond	4%	High	Low	CDO	6%	Low	Low	ETF	2%
11	High	Low	ETF	2%	Low	High	Stock	8%	High	High	Bond	4%
12	High	High	ETF	4%	Low	Low	Stock	6%	High	Low	Bond	2%
13	High	High	Bond	6%	Low	Low	CDO	2%	High	Low	ETF	8%
14	High	Low	Bond	8%	Low	High	CDO	4%	High	High	ETF	6%
15	Low	High	ETF	6%	High	Low	Stock	2%	Low	Low	Bond	8%
16	Low	Low	ETF	8%	High	High	Stock	4%	Low	High	Bond	6%

Coded into a useful format the table looks like this:

O1		CDO is excluded			2% is excluded			O2		CDO is excluded			2% is excluded			O3		CDO is excluded			2% is excluded	
Low Media	Low Liquidity	ETF	Stock	Bond	8%	4%	6%	Low Media	Low Liquidity	ETF	Stock	Bond	8%	4%	6%	Low Media	Low Liquidity	ETF	Stock	Bond	8%	4
A11	A21	A31	A32	A33	A41	A42	A43	A1	A2	A31	A32	A33	A41	A42	A43	A1	A2	A31	A32	A33	A41	A42
0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	0	0	1	0	1	0	0	
0	1	0	0	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	
1	0	0	1	0	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0	0	0	
1	1	0	1	0	0	1	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	
1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	1	
1	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	1	1	0	1	0	0	
0	1	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	1	
0	0	0	1	0	1	0	0	1	1	0	0	1	0	1	0	0	1	0	0	0	0	
1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	
1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	
0	1	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	1	0	
0	0	1	0	0	0	1	0	1	1	0	1	0	0	0	1	0	1	0	0	1	0	
0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0	1	
0	1	0	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	
1	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0	1	1	0	0	1	1	
1	1	1	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	0	

For each participant we coded their sixteen decisions such that there were three columns, and the participant received a one if they choose object one, and zeros for columns two and three. Translating the raw data into this cleaner format helped us greatly when we later needed to utilize the participants' decision making to provide us their maximized log likelihood.

In the heterogeneous analysis we assumed that participants were all individuals with unique betas. As we will discuss later, the betas act as a sort of preference parameter relative to the excluded level within each attribute. The betas are multiplied by the attributes in the above table to give us the individual's utility for picking a certain choice within that choice set. We use the standard maximum likelihood, discrete choice experiment methodology for finding the individual betas, but since we are computing them in Excel, we need to make a few calculations to get there.

First we use the following formula to get the probability that an individual makes a certain choice in a given choice set:

$$\text{Equation 1: } \Pr_1 = \frac{e^{v_1}}{e^{v_1} + e^{v_2} + e^{v_3}}$$

V1 in this equation is the sum product of the betas and the attributes, formally defined as:
Equation 2:

$$V_1 = \beta_0 + \beta_{LM} A_{11} + \beta_{LL} A_{21} + \beta_{ETF} A_{31} + \beta_{Stock} A_{32} + \beta_{Bond} A_{33} + \beta_{8\%} A_{41} + \beta_{4\%} A_{42} + \beta_{6\%} A_{43}$$

High access to information, high liquidity, security ABC, and 2% are the excluded, baseline parameters.

This process is then repeated for V2 and V3, and we get the probabilities that the individual chooses those options in a given choice set.

As mentioned previously, we coded participants' choices such that the variable *I* was an indicator of choice. An individual's log likelihood was then computed using the formula:

$$\text{Equation 3: } \sum_{c=1}^{16} \sum_{j=1}^{3} I_{cj} \ln(\Pr_{cj})$$

Where c is the choice sets 1-16 and j is the options 1-3 within each choice set.

Once we had manually designed the spreadsheet to feed us individualized log likelihoods we used solver to maximize this log likelihood by changing the eight preference parameters, or betas. We repeated this process 64 times for all of our participants.
When running this optimization 64 times we ran into a calculation issue for seven participants. In the optimization some of the probabilities were so close to zero that Excel could not read them and gave us a #NUM! error. The issue remained even after adding machine zero, the number closest to zero that Excel recognizes as a positive number. After consulting with Professor Seetharaman we decided to simply eliminate these seven participants from our calculation of the heterogeneous maximized log likelihood and BIC calculation. We therefore have 57 participants.

Heterogenous	Whole Data	In Sample	Out Sample
Sum LL	-673.21	-549.58	-363.37
BIC	4454.35	4146.21	N/A

We calculated the Bayesian Information Criterion or BIC, which can be thought of as the adjusted R^2, penalizing the LL by accounting for the number of parameters estimated. Formally, the adjustment is defined as:

$$\text{Equation 4: } BIC = (-2 * LL) + k * \ln(T)$$

Where k is the number of parameters, in this case 57 times 8 betas per person and T is the number of observations, or 57 times 16 decisions per person.

In addition to running the heterogeneous analysis on what we refer to as our "Whole Dataset," we also ran In Sample and Out of Sample analysis to see how well our model predicted actual choices made by participants.

The In Sample methodology was identical to our Whole Dataset methodology, but the summed probabilities and choices were made from 1-14, instead of 1-16:

$$\text{Equation 5: } \sum_{c=1}^{14}\sum_{j=1}^{3} I_{cj}\ln(\Pr_{cj})$$

After running solver another 57 times we arrived at predicted betas. We then used these predicted betas in our Out of Sample model to analyze choice sets 15 and 16 for each of the 57 participants.

The results of these three analyses can be found in the table above labeled heterogeneous analysis.

While finding individual, unique preference parameters is interesting, it does not allow us to generalize our conclusions beyond what that individual survey taker believes about investing. In order to find a model that we can attempt to generalize further we must start on the opposite side of the modeling spectrum and treat our participants uniformly, using the homogenous model.

The homogeneous model is identical in methodology to the heterogeneous model, the only exception being that we run solver to optimize only once, on the summed LL across all 64 participants. Since we are not optimizing each individual, the optimization error that we encountered previously does not take place, and we can use all 64 observations in our whole dataset. However, when comparing the strength of the homogeneous model to the heterogeneous model we needed to compare apples to apples, and therefore only looked at 57 observations. The following two tables summarize how the homogeneous model performed:

Homogeneous	Whole Data	In Sample	Out Sample
Sum LL	-823.64	-681.86	-156.32
BIC	1701.80	1417.17	N/A
	Excludes 7	Excludes 7	Excludes 7

Homogeneous	Whole Data	In Sample	Out Sample
Sum LL	-903.99	-746.34	-167.01
BIC	1863.44	1547.07	N/A
	Full	Full	Full

The same methodology was used to calculate the In Sample and Out of Sample results.

Once we computed the maximum likelihood for the homogeneous model we asked ourselves how differentiated the participants were. If they were very differentiated, the best model might be a four or five segment latent class logit model. That is, there are four or five distinct groups that value certain attributes in their investment strategy. A less differentiated crowd might only be distinguished on two attributes, in which case a two segment model would be sufficient. In order to compute how differentiated our participants are we need to create a latent class logit model and run constrained optimization across the segments to find the appropriate weights for the segments as well as the preference parameters for each segment.

In order to create a 2 segment latent class logit model we needed to replicate the previous methodology. In the homogenous model we were only changing 8 parameters, the 8 betas, to find the maximized log likelihood for the participants as a whole. We are now maximizing the summed log likelihood by changing 17 parameters, 8 for each set of betas (segment 1 and segment 2) and then an additional parameter that we shall refer to as f1, or the weight of segment 1.

A few slight modifications were made to the homogeneous model to ensure that the 2 segment analysis was properly executed.

The purpose of a 2 segment model is to allow for a certain degree of differentiation and to observe the probability that a certain participant would fall into a specified segment. In order to calculate the probability that a participant falls into a certain segment, Prh1 or Prh2, we first needed to calculate Pr1 and Pr2, defined as:

$$\text{Equation 6: } \Pr_1 = P_{11}{}^{I_1} P_{21}{}^{I_2} P_{31}{}^{I_3}$$

$$\text{Equation 7: } \Pr_2 = P_{12}{}^{I_1} P_{22}{}^{I_2} P_{32}{}^{I_3}$$

Note that the only difference in generating these two probabilities comes from Pi, not I. I is the individual's choices, which remain identical across the two segments, whereas Pi is the probability of choosing a given option within a choice set. Pr1 and Pr2 and thereby calculated for the 16 choice sets and then Prh1 and Prh2 are defined as simply the product across 16 choice sets:

$$\text{Equation 8: } \Pr h_1 = \prod_{i=1}^{16} \Pr_{1i}$$

$$\text{Equation 9: } \Pr h_2 = \prod_{i=1}^{16} \Pr_{2i}$$

We now have the probability that an individual falls within a certain segment. Next we calculate the weighted average of the two segments, f1 and f2, to compute Prh:

$$\text{Equation 10: } \Pr h = f_1 \Pr h_1 + f_2 \Pr h_2$$

Now that we have Prh we sum take the natural log of it and sum it across all 64 participants. This summed LL is the new target cell for Excel.

Another key difference with the two segment model is that there is no global maximum, and therefore we need to pay attention to the starting values that we insert for betas during the optimization. With Professor Seetharman's guidance we slightly altered the homogeneous starting values by constants that can be seen in the table below:

Previous	Changes
0.00	0.1
0.00	0.05
0.00	0.075
0.13	0.1
0.00	0.15
0.27	0.2
2.45	0.5
1.57	0.5
1.96	0.5

Beta 1 starting values were therefore: Previous Betas - Changes
Beta 2 starting values were therefore: Previous Betas – Changes

During the optimization we maximized the target cell while changing 17 parameters and constraining f1 and f2 to vary between 0 and 1 and also constrained B0 to equal zero.

The table below summarizes the results for the 2 segment model. Since we are not running individual optimizations we were able to use all 64 data points.

2 Segment	Full, Whole Data
Sum LL	-865.12
BIC	1848.07
F1	89.07%
F2	10.93%
F3	N/A

2 Segment	Full, In Sample
Sum LL	-722.46
BIC	1560.49
F1	89.18%
F2	10.82%
F3	N/A

2 Segment	Full, Out Sample
Sum LL	-115.01
BIC	312.50
F1	84.67%
F2	15.33%
F3	N/A

The same methodology as before was used to calculate the In Sample and Out of Sample results.

To determine how differentiated preferences are we needed to keep segmenting the data until we did not achieve marginal benefit for adding another segment. We therefore decided to create a 3 segment model that employed the same methodology as our 2 segment model, but this time we changed 26 parameters. Three sets of 8 betas plus 2 weight variables, f1 and f2 are changed while f3 = 1 - f1 - f2.

The table below summarizes the results for the 3 segment model. Since we are not running individual optimizations we were able to use all 64 data points.

3 segment	Full, Whole Data
Sum LL	-851.35
BIC	1882.92
F1	34.37%
F2	10.91%
F3	54.73%

3 Segment	Full, In Sample
Sum LL	-715.28
BIC	1607.31
F1	49.35%
F2	10.67%
F3	39.98%

3 Segment	Full, Out Sample
Sum LL	-110.59
BIC	347.33
F1	18.66%
F2	14.97%
F3	66.37%

As was mentioned previously, while a log likelihood closer to zero implies a better model, it does not accurately take into account the increased use of parameters. The BIC calculation defined in equation 4 takes this into account. For the two segment model k = 17 while for the three segment model k = 26. Comparing the two models' BIC we see that the two segment model is superior for all three datasets, meaning that it has a lower BIC. We can therefore confidently say that we have found our best model, the two segment model.

BIC	2 Segment	3 Segment
Full, Whole Data	1848.07	1882.92
Full, In Sample	1560.49	1607.31
Full, Out Sample	312.50	347.33

To quantify the effectiveness of our two segment model, we computed a hit rate for the last two questions of the survey. We used the betas calculated in the In Sample two segment model to generate probabilities for the Out of Sample model. We then compared those probabilities to the ones from the In Sample model. If both models showed that a participant had the highest probability of selecting the same option, we counted that as a hit. On question 15 the hit rate was 89% and on question 16 the hit rate was 67%, for an average hit rate of 78%.

Discussion of Results

As seen above, the two segment model outperforms the three segment model in every dataset. This implies that for the investing public that we surveyed, 64 students, are differentiated in two major ways and a two segment model is optimal to describe their preferences.

Taking a closer look at the two segment betas we can draw some conclusions as to why that might be and as to what investment attributes these students prefer. The table below shows the betas for the two segment model. As mentioned previously, low information accessibility, low liquidity, security ABC, and 2% have been excluded as the baselines for the estimation.

Starting with segment one, we can see that this segment prefers high information accessibility and high liquidity over the baselines, given that they are both zero. This makes sense, as both of these are strictly better than the low baselines. Skipping down to the bottom of the table we can see that people made rational decisions when taking our survey, given that they seemed to prefer high returning securities over low returning securities, all else equal. Higher return also means higher risk, yet in this experiment environment investors seem to be comfortable with the risks associated moving up the yield spectrum. In a real-world scenario, it is possible that the marginal costs of additional risks associated with the marginal benefit of additional yield will be too great for the investor to stomach. As far as security type is concerned, we can see that segment one prefers all three security types over the excluded, more complex security, ABC. Given that these values are essentially an order of magnitude smaller than the preference parameters associated with return, we can draw the conclusion that investors care less about the type of security than they do the return associated with that security. However, as our hypothesis earlier in the paper suggests, investors in segment one seem to prefer simpler securities over the more complex ABC security, all else equal. It should be noted that we did not run the standard errors on the table below, so we cannot draw any conclusions about the statistical significance of our results.

The second segment is much smaller and includes only seven participants. After we ran the allocation part of our analysis we realized that those participants that had been allocated to segment two were those that received #NUM! errors in Excel when we attempted to optimize their individual betas in the heterogeneous analysis. It appears that these seven individuals cared only about return, hence the very large beta values for 8% and 6% and hence why the solver failed to optimize their individual betas, they were simply too large. Ex-ante, it would seem that low information accessibility and low liquidity should always be zero since the baseline should always be preferred for a rational individual; however, in segment two we see that both of these are greater than zero. It could be the case that segment two individuals are willing to sacrifice liquidity and information accessibility in order to capture higher yield. Additionally, the only security that has a preference parameter greater than zero is the ETF. It seems clear that investors in segment two do not care whether a security is simple or complex, they are simply looking for yield.

	Segment 1	Segment 2
beta0	0.00	0.00
betaLM	0.00	0.03
betaLL	0.00	0.02
betaETF	0.13	1.17
betaStock	0.01	0.00
betabond	0.30	0.00
beta8%	2.22	23.91
beta4%	1.60	0.00
beta6%	1.90	3.56

Now that we found the optimal model, we can see how our participants are allocated across the two segment model. As was previously reported, approximately 89% of our participants would fall in segment 1 and 11% would fall in segment 2. We allocated our participants by checking if f1*Prh1 > f2*Prh2, if that was the case they were allocated to segment 1, if not segment 2.
So who exactly are these people? The table below summarizes this information.

Allocation	Gender		
	Male	Female	Unreported
Segment 1	36	18	3
Segment 1	63.16%	31.58%	5.26%
Segment 2	6	0	1
Segment 2	85.71%	0.00%	14.29%

	Major									
	Engineering and Science	Finance	OSCM	Economics	Marketing	Accounting	ArtSci	Undecided	Other	Unreported
Segment 1	4	16	3	13	6	2	4	5	9	3
Segment 1	6.15%	24.62%	4.62%	20.00%	9.23%	3.08%	6.15%	7.69%	13.85%	4.62%
Segment 2	0	4	0	1	1	0	0	0	1	1
Segment 2	0.00%	50.00%	0.00%	12.50%	12.50%	0.00%	0.00%	0.00%	12.50%	12.50%

	Program					
	BS	BSBA	MBA	BA	Masters	Unreported
Segment 1	3	41	3	4	3	3
Segment 1	5.26%	71.93%	5.26%	7.02%	5.26%	5.26%
Segment 2	0	6	0	0	0	1
Segment 2	0.00%	85.71%	0.00%	0.00%	0.00%	14.29%

Looking at gender first we can see that no women are found in the second segment. While we only had seven people allocated to this second segment, six of which being men and one unreported, can we attempt to draw some conclusions about why this might be? For the survey as a whole we had 64 participants, 18 of which were women, meaning that 11% of 18, or approximately two women should be in segment two if these were evenly distributed. Given our small sample size we hesitate to draw generalizing conclusions, yet one might speculate what those would be given similar results in a larger sample.

As mentioned previously, the segment two betas signify that people in this segment exclusively valued return. Since the 4% preference parameter has a zero, we can infer that those in segment two only picked 8% or 6% returns. Note, that since our survey methodology had only three options per choice set, either 8% or 6% were excluded for some choice sets, but both of them were never excluded at the same time, meaning that participants that only valued return could thereby choose either 8% or 6% for all 16 choices. Given that zero women were allocated to segment two, it seems that men are much more likely to be "return only" investors. This fits into a certain biological investing literature that deems men more risk tolerant, since higher return implies higher risk. If our allocation results are an indicator of a larger trend and could be repeated with a larger sample size, we would be comfortable stating that this may warrant further study and may have implications for wealth managers and investing professionals.

Looking next at the type of major the different segments studied, we can see that only one person in the second segment studied something outside the business school, in the category "other." Again, small sample issues remain, but we see a much wider distribution of majors in segment one. A full 6 out of 8 reported majors for segment two are business school majors, perhaps indicating that these students have a higher risk tolerance, or at the very least, a higher preference for risk. Note that the sum of major responses does not equal 64 since some participants reported having more than one major.

Lastly, the program type confirms many of the conclusions from the major analysis. Business school students are the only students that are allocated into the second segment. On a very surface level analysis we do not see differences in age, in a sense that younger people are not necessarily more risk tolerant, as we do not see any undergraduate engineering or arts and sciences student allocated into segment two. On the contrary, the field of study is what seems to play a role in segment allocation. In order to be more confident in our findings we recommend running a second study on a larger and more diverse group of individuals.

We also performed an Importance calculation for the two segment In Sample and Whole Dataset models. The importance calculation tells us how much additional utility would this segment of investors get from their most preferred level of an attribute than from their least preferred. For the In Sample Model, segment one would have gained additional 36% of total utility from getting their most preferred security, while segment two would have gained only 4.5%. For the Whole Dataset Model, segment one would have gained an additional 20% from their most preferred security, while segment two would have gained only 5.4%. For both models, the rest of the utility is mostly dependent on the return the consumers get. Factors like liquidity and accessibility of information seem to have very little impact on utility.

Limitations

The main limitation to our research is the low number of participants. While we offered an incentive for survey takers, utilized professors' help in sending it to their classes, and allowed sufficient time for gathering results, we ended up short of the initial goal we set, mainly because of the amount of time it took to take the survey. We believe our analysis still offers a general idea of investors' preferences. More responses, however, would have allowed us to gain a better understanding of the investors comprising the different segments and to generalize about their demographic characteristics.

We also have not verified the statistical significance of our results. We did not calculate standard errors for the control surveys, or the DCE survey, as those calculations would have been too complex and time consuming and would have gone beyond the scope of this project.

Another potential limitation is participants taking our Discrete Choice Experiment survey with prior knowledge of the purpose of our research, skewing our results. We tried to prevent that by waiting a month before releasing the DCE survey and by using a different channel of distribution- a different group of professors sending it to their classes. Still, we have no way of ensuring that a student was not sent both surveys and did not remember taking the first survey when filling out the second one.

Additional Questions of Interest

While we initially wished to study whether investors had a preference for simpler securities over more complex ones, our results indicate that there are two types of investors, those that take factors such as complexity, liquidity, and accessibility of information into account, and those that ignore all of these factors in a search for yield. With this conclusion in mind, we think the following areas of study would be interesting extensions to our current research.

We are concerned about the small sample size and limited demographic distribution of our study and would recommend further research to see whether our results apply across a larger cross section of the population. Of particular interest are demographic variables such as income, occupation, education, and previous knowledge of financial markets. Would the results hold if we studied a group of coal miners? What about a group of English teachers? Answering these questions would be useful for the literature and in particular private wealth managers and investment professionals that are attempting to optimize their customers' investment portfolios.

In addition we think that some of the demographic curiosities found in the allocation part of our analysis warrant further analysis. For instance, if we expanded this to a larger dataset would we find that women are still not allocated into segment two? Would we still find that business minded students are the only ones to be allocated into segment two? These questions are important to answer before we make broader generalizations.

We would also recommend that future research study the statistical validity of our results. While are conclusions are interesting, we are unsure about their statistical significance. A broader statistical analysis is necessary in order to justify policy recommendations.

Lastly, we believe that our methodology could be used to study other obscure, not well understood parts of the financial industry. For instance, one could test for an investor's preferences of what exchanges they prefer to trade equities on. All else equal, do they prefer the NYSE, the NASDAQ, the AMEX, or perhaps they prefer trading in broker-dealer hosted exchanges or dark pools? There are pros and cons to each type of transaction mechanism and understanding investors' preferences is important for business decisions at large financial institutions.

Conclusion

In conclusion, we believe our analysis shows that investors in segment one clearly prefer simpler securities over more complex ones, all else equal, while investors in segment two only take into consideration return. We believe this could be the basis for a more extensive survey that sheds additional light on investor preferences and perceptions about complexity. Furthermore, our paper demonstrates that the Discrete Choice Experiment methodology can be successfully applied to a number of fields outside the realm of marketing.

Works Cited

Brooks, Alison. *Can Nervous Nelly Negotiate? How Anxiety Causes Negotiators to Make Low First Offers, Exit Early, and Earn Less Profit*. Diss. Philadelphia: The Wharton School, University of Pennsilvania, 2011. Print.

Brunnermeier, Markus, and Oehmke Martin. "Complexity in Financial Markets." (2009): n. pag. Print.

Goodman, Joseph, Cynthia Cryder, and Amar Cheema. "Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples." Journal of Behavioral Decision Making (2012): n. pag. Print.

Haldane, Andrew. "RETHINKING THE FINANCIAL NETWORK." Financial Student Association. Amsterdam. 01 Apr. 2013. Speech.

Street, D., L. Burgess, and J. Louviere. "Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments." International Journal of Research in Marketing 22.4 (2005): 459-70. Print.