Beliefs of traders in prediction markets

Bachelor's Project Thesis

Xabi Krant, s2955156, x.krant@ai.rug.nl,
Supervisor: prof. dr. D. (Davide) Grossi

Abstract: Prediction markets are a relatively new tool for aggregating dispersed personal information. The idea is that the price in these markets can directly be used as a prediction for an event. This research uses an agent-based model to test whether the price is a reflection of the beliefs of traders. The results shows that the price is a direct indicator of general patterns in the beliefs of traders, but only weakly correlates with exact differences. This research also sheds new light on why some markets do not produce trades. All in all this research supports the idea that prediction markets are able to aggregate dispersed information, with the price as a direct measure for this information.

1 Introduction

Many economists have come to believe that markets can be used as tools for aggregating information and predicting future events. A relatively new type of financial market, called prediction or information markets, are designed solely with this intent. A variety of processes for aggregating information can be found in organizations, such as committees, polling processes, networks of contacts, reporting (Plott & Chen, 2002). This process can however be troublesome. Individuals might have incentives to not reveal their personal information, or the mechanisms of aggregating the information are sub optimal. Prediction markets aim to improve the current practices.

In these prediction markets agents trade in contracts whose payoff depends on unknown future events (Wolfers & Zitzewitz, 2004). von Hayek (1945, 1948) was one of the pioneers stating that market prices reflect the relevant information of traders. "The proof that properly designed markets can aggregate information has existed in the experimental economics literature for over a decade" (Plott & Chen, 2002). Even when traders know little about other traders or the environment, market prices will accurately forecast an assets value (Smith, 1982). That a price reflects information is the basis for the prediction markets. "The bets establish a price for a contract tied to the outcome, a prediction, can be obtained directly or indirectly from this price" (Wolfers & Zitzewitz, 2010). One main advantage of prediction markets as opposed to previous practices is that this aggregation captures the confidence of the traders, through the amount they are willing to invest or bet.

A sizable body of research has been built over the last years. The earlier research has focused on how the market mechanisms work and on how a prediction market can be set up. Recent research mostly focuses on how accurately the market price predicts the future event. There has been a discussion in the literature on how the prices of markets are set. The Marginal Trader Hypothesis (MTH) states that not the average trader but "a small group of active and well-informed traders are responsible for steering market price to efficient levels" (McManus & Blackwell, 2011). Forsythe, Rietz, and Ross (1999) found support for this theory, stating: "There is substantial evidence, that agents, on average, act far less rationally within these markets than theory would demand." This research will try to attribute to this discussion.

Plott and Chen (2002) states that a properly designed market can aggregate information, but what is a properly designed prediction market? Laboratory research by Forsythe and Lundholm (1990), Forsythe et al. (1999) and Plott (2000) has shown that a prediction market can produce useful information. Although this does show the potential of markets, the results might not generalize to real markets as the circumstances were tightly controlled. Evidence from Berg, Forsythe, Nelson, and Rietz (2001), Pennock, Lawrence, Giles, and Nielsen (2001) and Plott and Chen (2002) has

shown that prediction markets perform very well in predicting future events. Prediction markets perform significantly better than other, more traditional, forms of aggregating information, like for instance polling and expert aggregation. This however only looks at the predictive power of the market price, and not at the aggregation of information. It could for instance be that the information held by the traders is wrong, but the market will accurately predict an event. Plott (2000) stated “Markets have the capacity to collect information and publish it, but that capacity is not perfect. The market can make mistakes.”
This research will test how the beliefs of traders influence the market price: “How are the beliefs of traders in a prediction market reflected in the price?”. The research will use simulations of the market with different market designs. The main advantage of this approach is that one can control the beliefs, which is impossible in a normal market setting. The research will abstract away from details, and look at how market design and traders influence the ability of the market to aggregate personal information.

2 Model

An agent-based model was developed to look at the influence of private information on the price of contracts in a prediction market. Here we discuss prediction markets and their working in more detail, and will discuss the implementation within the model.

2.1 Contracts

There are different types of contracts in prediction markets, Wolfers and Zitzewitz (2004) give an overview of some possible contracts. This overview is shown in 2.1. The contracts used in this model are winner-take-all contracts. These contracts pay an arbitrary amount of \$1 if and only if an event occurs. In these markets two contracts are traded, so that either one of the contracts will pay \$1. An example is a contract that pays 1\$ if Donald Trump will be elected as the president of the United States and another contract that pays \$1 if Donald Trump is not elected as the president of the United States. Owning both contracts is riskless, as one will pay \$1 and the other will pay \$0. The contracts will be called alpha and beta throughout the remainder of this paper.

2.2 Market

The market system is modeled similar to that of the Iowa Electronic Markets (IEM). The IEM is the most researched prediction market. The IEM website contains a traders manual discussing the structure of the markets in detail (*Traders manual IEM, Objects Traded in the IEM*, n.d). Forsythe, Nelson, and George R. Neumann (1992) describe the anatomy of the IEM and the traders in the market in detail. Here we will give a simple description and discuss the differences between the real IEM market and the model.
The prime mechanism is the double auction system, which is also used by many other prediction markets like the Hewlett Packard IAM. In a double auction buyers submit their bids to a bid queue and sellers submit their bids to an ask queue simultaneously. “An ask is an order to sell” and “a bid is an order to buy” (*Traders manual IEM, Objects Traded in the IEM*, n.d). Once a new bid is placed in the queue that is higher or equal to the lowest ask price, the market clears. In the IEM the price of the trade is the price of the oldest bid or ask involved in the trade. In the model we use a slightly different approach, the market does not explicitly clear at the price of the oldest bid. Agents will determine a price for their bids and asks. If there is a contract in the bid queue with a higher bid than their ask price, they will match the bid and the market will clear at this bid price. If there is no such bid, their ask will be placed in the ask queue. The same principle is used for placing bids. This results in non-overlapping queues; there is no combination in the queues that would cause the market to clear. Opposed to the IEM Market, all bids and asks are for one contract only and bids and asks do not expire. This does not influence the market, as agents can place multiple bids and asks at the same price, and all bids and asks are removed when the agent is chosen to perform actions.
Asks can only be placed when the trader has enough contracts to fulfill every ask. The number of placed asks should therefore be smaller or equal to the number of contracts owned. Bids should always be executable as well. This means that the total price of all bids placed by a trader can not be higher than the wealth owned by that agent. The same rules are used by the IEM.

Contract	Example	Details	Reveals market expectation of . . .
Winner-take-all	Event y: Al Gore wins the popular vote.	Contract costs $\$p$. Pays \$1 if and only if event y occurs. Bid according to value of $\$p$.	Probability that event y occurs, $p(y)$.
Index	Contract pays \$1 for every percentage point of the popular vote won by Al Gore.	Contract pays $\$y$.	Mean value of outcome y: $E[y]$.
Spread	Contract pays even money if Gore wins more than y^*% of the popular vote.	Contract costs \$1. Pays \$2 if $y > y^*$. Pays \$0 otherwise. Bid according to the value of y^*.	Median value of y.

Figure 2.1: Types of contracts, taken from Wolfers and Zitzewitz (2004)

Just as in the IEM there is an unlimited number of contracts. Contracts can be bought through unit portfolios. "A unit portfolio consists of one of each of the contracts in the market and has a price equal to the guaranteed aggregate payoff of this contract set" (*Traders manual IEM, Objects Traded in the IEM*, n.d). Unit portfolios can also be sold back to the market at any moment, for the same amount as for which they can be bought. The number of contracts in the market therefore fluctuates over time, with the number of each of the contracts being equal.
Contrary to the IEM we limited the queues to 200 orders. This was done to improve the processing speed. During testing no situations were observed in which this influenced the market. If the length of the queue exceeds 200, the highest asks and the lowest bids are excluded from the queue. The agents will always act on the most profitable orders, so this does not influence the results.

2.3 Agents

In the market there are 200 agents or traders. This is similar to the number of agents active in the market researched by Forsythe et al. (1992). Agents are initialized with a wealth, a strategy and a belief. Agents are selected uniformly to perform actions. When selected an agent determines whether it wants to sell or buy a unit portfolio or place a bid or ask. The market ends after 1000 steps, meaning on average each agent is selected five times to perform actions.

2.3.1 Wealth

With the wealth the agent can buy unit portfolios and trade in contracts. The wealth can only be increased by selling contracts or selling unit portfolios back to the market. The total wealth in the market is equal to 5000, which is similar to the total investment found by Forsythe et al. (1992).
The wealth of the agents is initialized randomly according to three different distributions: a discrete uniform distribution, equally among all agents and a Pareto distribution. The distributions are shown in A
For the Pareto distribution three different power values or Pareto indices are used. The Pareto indices used are 1.16, 1.42, 2.26. The power values match the "joint ratios" of respectively, 80/20 (high imbalance), 70/30 (moderate imbalance) and 60/40 (low imbalance). In the economics literature it is a well known principle that 80% of the wealth (or income) is owned by the richest 20% of the population, which is represented by the high imbalance and the power value of 1.16 (Dubay, 2009).

2.3.2 Beliefs

Each agent has some private information that the agent uses to buy and sell contracts. This information constitutes a belief. The belief of an agent determines the way an agent values each contract in the market. The belief is a number between 0 and 1. The number represent the likelihood that event alpha will happen according to the agent. 1 minus the same number represents the belief that event beta will happen, as either event alpha or beta will

happen. If the belief is 0.5, the agent beliefs the event occurring is as likely as the event not occurring. The beliefs are randomly initialized according to two distributions: a truncated normal distribution and a discrete uniform distribution.
For the truncated normal distribution three different values for the mean are used: 0.25 (low), 0.5 (medium) and 0.75 (high). Wolfers and Zitzewitz (2004) suggested that the traders should have different opinions for a prediction market to aggregate information. To test the influence of the variation in opinions, two different standard deviations are used. The low standard deviation is 0.125, the high standard deviation is 0.25. The distributions are shown in A

2.3.3 Agent strategies

The agents in the market have different strategies for buying and selling contracts. All agents act according to some basic rules:

- Agents never sell a contract for less than the value they give to a contract, and never buy for more than the value of a contract.
- Agents can either buy unit portfolios or buy contracts on the market. Agents will buy unit portfolios when they have enough wealth and when it is cheaper to buy a unit portfolio than buy an alpha contract and a beta contract on the market. Otherwise they place bids on the market.
- Agents can either sell unit portfolios or sell contracts on the market. When they have one or more of each contract and selling a unit portfolio earns more than selling an alpha and a beta contract on the market, the agent will sell a unit portfolio. Otherwise they place asks on the market.
- Agents use a utility rule to determine which contract they should try to buy, so they bid on the highest utility.

Given a wealth w and a belief value b, the utility for a contract is:

$$u^{\alpha} = \begin{cases} 0 & \text{if } p > b \\ 0 & \text{if } p > w \\ (1-p) * b/p & \text{otherwise} \end{cases} \tag{2.1}$$

where p is the calculated bid price for α

$$u^{\beta} = \begin{cases} 0 & \text{if } p > 1-b \\ 0 & \text{if } p > w \\ (1-p) * (1-b)/p & \text{otherwise} \end{cases} \tag{2.2}$$

where p is the calculated bid price for β

The agents differ in their bidding and asking strategies. There are four different types of traders: a fixed markup agent, a random markup agent, a clearing price agent and a strategic agent. These are an adaptation on the agents described by Park, Durfee, and Birmingham (2000). Their characteristics are described below. B contains the Python code for each agent strategy.
All agents within one market simulation have the same strategy. The Efficient-market hypothesis (Fama, 1969) states that agents can not make a substantial profit on a market. No agent should be able to make a substantial profit. Park et al. (2000) show that agents can increase their profits when their strategy is different from that of another agent. To control for this effect, we use homogeneous agent strategies.

Fixed markup agent
The fixed markup agent uses a simplistic strategy to determine their bidding and asking price. It uses the belief to value both the alpha and the beta contract, so that the value for alpha is the belief and the value for beta is one minus the belief. For the ask price it adds a fixed markup to the value of the contract. The bid price is determined by subtracting the markup off the value of the contract. The markup is set to 0.05 for all agents.

Random markup agent
The random markup agent is similar to the fixed markup agent. The only difference is that the random markup agent does not use a fixed markup of 0.05. The random markup agent determines a markup between 0 and 0.1 for every bid and ask.

Clearing price agent
The clearing price agent will bid a price that is 0.01 higher than the current highest bid in the queue, as long as it is lower or equal to the value of the contract. Its ask price will be 0.01 lower than the current ask price, as long as it is higher than or equal to the value of the contract. If the

agent already has the highest or lowest order in the queue, it will bid or ask the same price. This makes sure the agent does not compete with itself when placing multiple orders.

Strategic agent
The strategic agent is based on the p-strategy described by Park et al. (2000). The p-strategy is described in detail in an earlier paper by Park, Durfee, and Birmingham (1996). The p-strategy agent maximizes an expected utility function. They use a stochastic Markov Process (MP) model which captures factors that influence the utility function. "The MP-based mechanism enables a contractor to choose different optimal payments depending on the payment(s) of the other contractor(s) or the contractees costs of doing the task, and therefore to receive a better profit" (Park et al., 1996). The number of Markov Chain states is equal to $(m + 1) * (n + 1)n/2 + 2$, where m is the number of buyers and n the number of sellers. Park et al. (2000) limited the number of states by maximizing the number of contracts in the queues to 5. Whenever a new, higher, bid was placed, the lowest bid would be excluded from the queue. For the ask queue a lower ask would exclude the highest ask. As they were looking at how agent strategies influence profits, this was a suitable approach. The state space in this experiment is a lot larger, with a maximum of $4,040,102$ states. Using a MP model is not feasible for this state space. This is why the agents in this model will use a heuristic function to determine the price at which they bid and ask. Similar to the p-strategy agents these agents will take into account the demand and supply in the market. Next to that they try to find a price which has a good balance between the value of the transaction and the chance of a transaction completing. This captures the most important aspects of the p-strategy, but is computable.
The agent values all bid and asks that will not be placed in the queue as worthless. These orders will never result in a transaction and will therefore never produce any value. A bid above the value of a contract and an ask lower than the value of the contract have a negative valuation, and will never be placed.
The higher the number of bids in the queue, the more demand. This means that the price for bids placed by the agent is higher, so that the chance of a successful transaction becomes higher. If the demand is high the price for asks becomes higher, this maximizes the value of a transaction. The probability of a transaction is high because of the high demand.
When there are many ask orders in the queue, there is a high supply. The price for bids placed by an agent will decrease, to maximize the value of a transaction. The price for asks will be lower, to increase the chance of a transaction and compete with other asks.

3 Methods

3.1 Data collection

The model takes three parameters as the input: the agent type, the wealth distribution and the distribution of the beliefs. The combination of these parameters will be called the market configuration. There are four agent types, five wealth distributions and seven belief distributions, making a total of 140 different market configuration. For each market configuration $1,000$ data points were collected. Each data point consists of the market configuration, the beliefs of all agents and the transactions that are carried out. Although the market configuration also includes the distribution of the beliefs, the actual beliefs need to be collected as there is randomization within the distribution. This is not done for the distribution of the wealth, as the wealth is used a control variable and statistical notions of the beliefs are used as an independent variable.
The program was written in the programming language Python (*Python Programming Language*, n.d), using the framework mesa (*Mesa framework documentation*, n.d). The data was collected on a Microsoft Surface Pro Laptop with an Intel Core i5-6300U Processor (2×2.4 GHz) and 8 GB DDR3L RAM. Collecting a single data point takes approximately 1.5 seconds. As we collected $140,000$, this took a little over 58 hours.

3.2 Measures & data processing

The collected data consists of market configurations, beliefs and transactions. The beliefs will be used as the independent variable. The transactions will be transformed to a price for alpha and a price for beta and will then be used as the dependent variable. The market configuration is used as a control variable. C shows a summary of the variables and the sample characteristics.

3.2.1 Dependent variable

This research looks at how the beliefs of traders are reflected in the price of an prediction market. For a prediction market, the price is not clearly defined. The price of a prediction market is a time series, and can vary from moment to moment. Another problem is that within a double-auction market there are actually two prices: the lowest ask price and the highest bid price. D shows the time series of three simulations (with different configurations) to illustrate the problem.

Previous research has solved the second problem by using the transaction price instead of bid or ask prices. There is no general consensus on how to solve the first problem. One approach, for instance used by Forsythe et al. (1992), is to use the last-trade price on a given day. This approach is very sensitive to outliers as the result is based on a single trade. This is why other researchers, like Plott and Chen (2002), have chosen to use a volume averaged price. This takes the prices of multiple transaction, and uses the average of this as the price. Usually not all trades are included in this calculation, but a subset of trades towards the end of the market is used. "The rationale being the market achieved some sort of equilibrium towards the end as suggested by laboratory experiments". Plott and Chen (2002) tested five methods for getting a prediction from the trades. Their results were "robust with respective to different calculation methods". This research uses the average price of the last 50% of the trades as the dependent variable.

In some market runs, no trades were made for a particular contract. These were excluded from the data, as there is no average transaction price for these runs. In almost 5% of the runs it happened that there was no trade for a particular contract. This happens when the beliefs of the traders are very similar, and the traders are not willing to trade at prices close to their belief. This happened especially often in market configurations with a fixed markup strategy and a normal distribution of the beliefs with a medium mean and a low standard deviation. The discussion includes a more detailed explanation of why this happens.

3.2.2 Independent variable

The beliefs of the traders are used as the independent variable. The individual beliefs have to be aggregated to make any claims on the general belief of the traders. No objective criterion exists to choose a method for aggregating the individual beliefs. The most evident option is using a measure for the central tendency. This is why this research uses the mean and the median of all individual beliefs, as the independent variable.

3.2.3 Control variables

The market configuration is used as a control variable. The market configuration is made up of three variables: the agent strategy, the wealth distribution and the belief distribution. Each of these variables is a categorical variable. The description of the model included an explanation of the different categories per variable.

4 Results

4.1 correlations

The research question looks at the relationship between the price and the beliefs. To test for this relationship we present the correlations between the dependent and independent variables per market configuration. The overall correlations between the mean belief and prices and the median belief and prices are both close to perfect with values of around 0.9 ($\pm$0.005). 4.1 shows the correlations per group.

Agent strategy

For agent strategy we see that the correlations are very high. The correlation is between 0.88 and 0.98 for the different subgroups. The strongest correlation between the trade price and the mean belief is for the strategic agents and the clearing price strategy. Both have an extremely high correlation, there is almost a perfect linear relationship. The other two strategies have a similar correlation, slightly lower than the other two cases. The correlation between the median belief and the trade price are almost equal to the correlations between the mean belief and the trade price.

Wealth Distribution

The correlation for all wealth distribution cases is very similar. The correlation is around 0.90 for

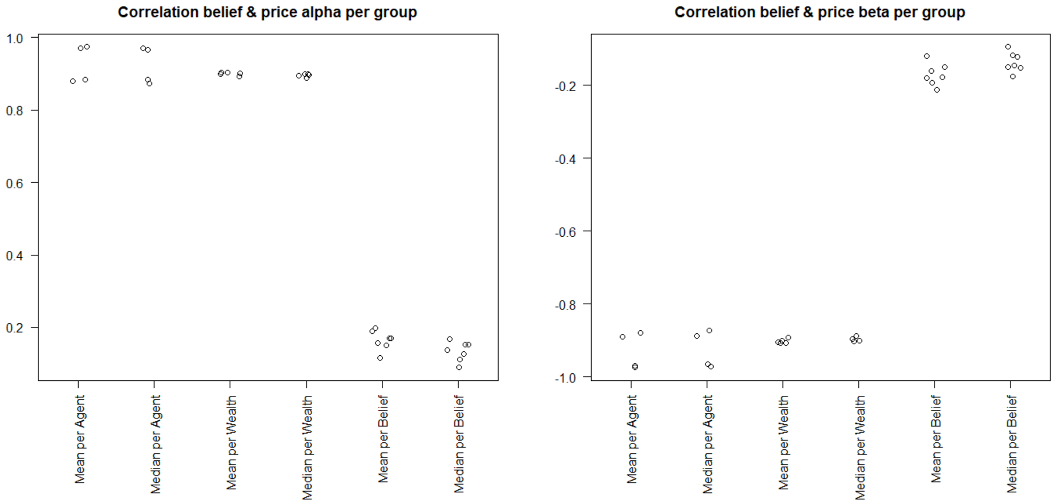


Figure 4.1: Correlations price and belief per group

the different subgroups. This holds for both the correlation between trade prices and the mean belief and the correlation between trade prices and the median belief.

Belief distribution
The correlation between the trade price and the mean belief is quite similar for all cases, but the correlation is a lot lower. The correlation is around 0.16 for the different subgroups.
That the correlation is lower within the different subgroups for the belief distribution also means that the price correlates better with the different means per group than within the groups. The price and the mean and median belief correlate mainly with the different means per group, and correlate with general patterns more than specific differences.
If we look at the correlation of all groups with a high standard deviation, we can see that the correlation is weaker. This is only the case when all three means are included, only including one mean will result in all correlations being weak.

Overall we can see a high correlation between the independent variables, average price alpha and average price beta, and the dependent variables, mean belief and median belief. The correlation is influenced by the different agent strategies and mainly by the belief distribution. The wealth distribution does not have an effect on the correlation between the dependent and independent variables.

4.2 Regression results

E shows the results of the "multiple linear regression" models. The models 1α to 5α show the results for the dependent variable average price alpha, models 1β to 5β show the results for the dependent variable average price beta. Model 1 shows the relationship between the control variables and the dependent variable. Model 2 adds the explanatory variable mean belief, model 3 uses the control variables and the variable median belief. Model 4 only uses the independent variables to to predict the prices. Model 5 combines all explanatory variables into one model.
Looking at the first model we can see that all control variables have a significant linear relationship with the dependent variables average price alpha and average price beta.

Agent strategy
The different categories for the agent strategies have different estimated effects on the average prices. For both alpha and beta the categories clearing price and random markup have a similar estimated value. For alpha the other two categories have a higher estimated value. For alpha the strategic category has a slightly higher estimated value. The estimated effect for the strategic cat-

egory on average price beta is lower than that of the fixed markup category.

Wealth distribution

The different categories for the wealth distribution all have similar estimated effects on both dependent variables. The pareto distribution with a high imbalance has the highest estimate, but it is only slightly higher than the other categories. We do see that the more imbalanced, the higher the estimated value.

Belief distribution

For the belief distribution the estimates vary strongly among the categories. The estimates mainly vary between categories with different means. As expected, the higher the mean the higher the estimate of alpha. Looking at the difference between the standard deviations, we also see some clear differences. These differences are caused by the actual mean of the categories with a high standard deviation being closer to 0.5, as the beliefs are distributed among a truncated normal distribution. The difference in estimates for the medium mean categories is therefore smaller. We do see that the estimate for the low standard deviation is further from 0 than for the high standard deviation, for both alpha and beta.

Mean belief

We see that for both alpha and beta the mean belief has a significant relationship with the dependent variable. In both cases the estimate is approximately the same, but inverse. Adding the mean belief to the explanatory variables, causes some belief distribution categories to have an insignificant effect. This means that the mean belief can explain some of the difference in prices caused by the belief distribution. The mean belief can not explain the effect of the medium mean categories. This is because the mean in these cases is the same, but there is variance in the prices, caused by the difference in the standard deviation. This also means that the standard deviation does have an effect on the price. The estimates of all the belief distribution categories become close to zero.

Median belief

The median belief does have a significant relationship with the average price. The estimates for both alpha and beta are similar, but inverse. The median belief does not explain the difference for the different belief categories. The belief categories do still have significant explanatory power.

Model 4 only uses the independent variables as explanatory variables. This is done to show how much of the variation can be explained by these variables alone. We see that these variables can predict the prices almost as well as the models that use the control variables.

Combining all explanatory variables into one model has the highest explanatory power. Over 82% of the variance in the price can be predicted using this linear model. We can also see that the mean belief has an estimate very close to one, especially for beta. This means that an increase of the mean belief will increase the price of the transactions by almost the same amount. We can also see that the estimate of the median belief is smaller and has been inverted.

4.3 Robustness

To check the robustness of the results additional tests were done. These experiments test the model with different configurations, to check for the influence.

4.3.1 Agent Strategies

In the original experiment we controlled for any effects that might be caused by heterogeneous agent strategies. To test whether this choice had any influence on the results, we did some experiments in which the market contained agents with different strategies. 0, 50, 100, 150 or 200 agents were initialized with the same strategy. The total amount of traders still remained 200, so the combination of the strategies had to add up to 200 agents. All correlations were similar to the correlations found with homogeneous agent strategies. Combining agents with a random markup strategy and agents with a fixed markup strategy had a slightly weaker correlation than any of the homogeneous strategies. Combining agents with a clearing price strategy and agents with a strategic strategy had a slightly stronger correlation than any of the homogeneous strategies. Initializing agents with homogeneous or heterogeneous strategies has little effect on the correlations.

4.3.2 Wealth

The distribution of wealth seemed to have little effect on the correlation between the price and the beliefs. Although the alpha value of 1.16 for the pareto distribution is according with a high imbalance, there is no reason to assume that the distribution of the wealth could not be more imbalanced in reality. We therefore tested a pareto distribution with alpha values of 1.01, 1.05 and 1.1. These values result in an extreme imbalanced distribution of the wealth. The correlations are as high for these distributions as they are for more balanced distributions. The way the wealth is distributed seems to have little effect on the way beliefs are represented in the price.

4.3.3 Belief Distributions

In the initial experiments we chose standard deviations of 0.125 and 0.25. The results show that the correlation was weaker for the higher standard deviation. To test whether this holds for other values we tested different standard deviations both lower and higher than the original values. The values used were 0.025, 0.075, 0.375 and 0.5.

The results show that there is indeed a negative relationship between the correlation and the standard deviation. The correlation for a standard deviation of 0.5 is quite a lot lower at 0.4.

We also see that the lower the standard deviation the less often the market produces trades. For the lowest standard deviation, none of the simulations with fixed markup agents produced transactions. The standard deviation of 0.075 did not produce trades for a mean belief of 0.5, but did produce trades for the other mean beliefs in some simulations.

The results in the initial experiment showed that the price reflects general patterns of the beliefs, but does not correlate with exact differences. To test this we used configuration with mean beliefs between 0 and 1 with steps of 0.1. The results show that the smaller the spread, the weaker the correlation. This shows that indeed the general patterns are better represented in the price than the small differences.

5 Discussion

5.1 Conclusion

This research aims to answer the question: "How are the beliefs of traders in a prediction market reflected in the price?" An agent-based model was developed to test for the relationship between the beliefs of traders and market prices. The agent-based model included different market configurations, with a difference in the distribution of beliefs, the distribution of wealth and the strategy of the agents. These different market configurations show the effect different market designs have on the relationship between the beliefs of traders and the market prices.

5.1.1 Lack of trades

The model did not produce trades in all runs. Previous research has done some minor analysis of why some markets do not produce trades. This was mainly attributed to a lack of participants, which was called the thin market problem (Wolfers & Zitzewitz, 2004). Another factor was the lack of disagreement about the probabilities and agents trading "too" rational. Wolfers and Zitzewitz (2004) state that in order for agents to trade, there needs to be a possibility to make a profit on taking risks. This research sheds new light on why markets may not produce trades. It shows three necessary conditions for agents to withhold from trading.

Agents only make trades when these are beneficial; the price for buying should be lower than the belief and the price for selling should be higher than the belief. When no agent believes it can make a profit from trading, no trade is made. This happens when the beliefs of agents are especially similar, and their trading strategy is to have a relatively high markup. This is shown in the configuration with a fixed markup agent and a low standard deviation, which produces trades in only a few simulations. For the other agent strategies, the agents place bids and asks of which the price is only slightly above/-below or even equal to their belief, causing trades even when beliefs are similar.

When the beliefs contain information, favoring either one of the contracts, this effect does not take place. In the cases where the beliefs were similar, but they were distributed around a higher or lower value, the markets did produce trades. This means that only when the agents have similar beliefs, and they believe both events are (almost) equally likely to happen, they will withhold from trading.

Wolfers and Zitzewitz (2004) and Forsythe et al. (1992) both discuss that trade requires disagreement about the outcomes. Our model shows that this is the case if and only if agents are unwilling to trade with low profits. The number of participants does not necessarily cause this problem, so the "thin market problem" seems to be the result of an underlying problem.

When there is little disagreement about outcomes, traders are unable to hedge against taking risk. Although it seems unlikely that human traders use a fixed markup strategy to hedge for this risk, it could very well be that a high profit is required to weigh up against the risk. We would expect that markets with lower risks, like markets with play money, less frequently produce no trades.

5.1.2 The relationship between beliefs and transaction prices

The results show a strong correlation between the beliefs and the prices. Although there is some difference per market configuration, the overall correlations are very close to perfect. The correlation per subgroup of belief distribution is a lot lower. This shows that the price is a good predictor for the general pattern in the beliefs, but not a good predictor for small, specific differences in beliefs. Prices should therefore be used as a general indicator, and not as an exact measure for the beliefs. The results of the regression models show similar results. The estimates of the independent variables show that the prices follow the mean belief almost perfectly, as the estimate is close to one. Although the median has a high correlation with the price, the price does not increase by the same amount when the median rises. The estimate has a value of around 0.5. Prices can therefore be used directly as an indicator of the mean belief, but the median belief is not equal to the price. This research shows that the beliefs are strongly reflected in the price of a prediction market.

5.2 Limitations

In this research agents only had a belief value, and were completely certain of this belief. Agents were willing to sell at any price above their belief, and buy at any price below their belief. This trade behavior is rational for agents that are completely certain of their belief. In reality however, traders will never be completely certain of the outcome of the market. There are too many factors that influence the outcome of the market, making it impossible to include all factors. Traders will trade with incomplete information and will more likely make bounded rational decisions. The model did not include uncertainty, due to the increased complexity that including another factor causes. There is no rational for how the uncertainty should be distributed, and there are many options available. The uncertainty within a market will highly depend on the event that is predicted in the market. In this research we decided to lay the focus on market mechanisms instead of the uncertainty within a market. Depending on how it is modeled, uncertainty will likely have a direct effect on the prices. One option would for instance be to adapt the markup based on the uncertainty and a profit margin. This option would have likely had an influence on the results, and could have caused the correlation between the beliefs and the prices to be weaker than in the current setup. Future research could include this factor to test the robustness of the results when uncertainty is included.

Another factor that could have been used was a bias. The Marginal Trader Hypothesis (McManus & Blackwell, 2011) states that not the average trader but the agents that are well informed and have little bias are the ones that determine the prices. Combining a bias and uncertainty could test this hypothesis.

Future research could also take a closer look at the situations in which a market does not produce any trades. In this study we found that there is evidence that markets do not produce trades because agents are unwilling to trade at prices close to their belief. Since this experiment was not setup with a focus on when markets produce trades, the information that can be deduced from this experiment is only limited. To test the conditions that were found in this study and in previous research, new experiments could be setup focusing on when markets produce trades.

Another limitation is the use of a heuristic rule for the strategic agent. The strategy is based on research by Park et al. (1996), who described the p-strategy. This p-strategy is a stochastic Markov Process (MP) model. In this experiment we were unable to use a MP-model. The reason for using a heuristic was the lack of processing power, in combination with the higher number of states possible in this experiment. Future experiments could test these results with the agent described by Park et al. (1996). Other agent strategies and their effect on the results could also be tested in future research.

References

Berg, J., Forsythe, R., Nelson, F., & Rietz, T. (2001). *Results from a dozen years of election futures markets research.* (Forthcoming in Charles Plott and Vernon Smith (Eds) Handbook of Experimental Economic Results)

Dubay, C. S. (2009, 5 4). *The rich pay more taxes: Top 20 percent pay record share of income taxes* (No. 2420). The Heritage Foundation. (Webmemo)

Fama, E. F. (1969). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Forsythe, R., & Lundholm, R. (1990). Information aggregation in an experimental market. *Econometrica*, *58*(2), 309–347.

Forsythe, R., Nelson, F., & George R. Neumann, J. W. (1992). Anatomy of an experimental political stock market. *The American Economic Review*, *82*, 1142–1161.

Forsythe, R., Rietz, T. A., & Ross, T. W. (1999). Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior & Organization*, *39*, 83–110.

McManus, T. C., & Blackwell, C. (2011). An exploration of market efficiency and the marginal trader hypothesis. *Undergraduate Economic Review*, *7*.

Mesa framework documentation. (n.d). Retrieved from `https://mesa.readthedocs.io/en/master/`

Park, S., Durfee, E. H., & Birmingham, W. P. (1996). Advantages of strategic thinking in multiagent contracts (a mechanism and analysis). *Proceedings of the Second International Conference on Multiagent Systems*, 259–266.

Park, S., Durfee, E. H., & Birmingham, W. P. (2000). Emergent properties of a market-based digital library with strategic agents. *Autonomous Agents and Multi-Agent Systems*, *3*, 33–51.

Pennock, D. M., Lawrence, S., Giles, C. L., & Nielsen, F. A. (2001). The real power of artificial markets. *Science*, *291*(5506), 987–988.

Plott, C. R. (2000). Markets as information gathering tools. *Southern Economic Journal*, *67*(1), 1–15.

Plott, C. R., & Chen, K.-Y. (2002). Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem. *California Institute of Technology Social Science Working Paper*(1131), 309–347.

Python programming language. (n.d). Retrieved from `https://www.python.org/`

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review*, *72*, 923–955.

Traders manual iem, objects traded in the iem. (n.d). Retrieved from `https://iemweb.biz.uiowa.edu/trmanual/IEMManual_2.html`

von Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, *35*, 519–530.

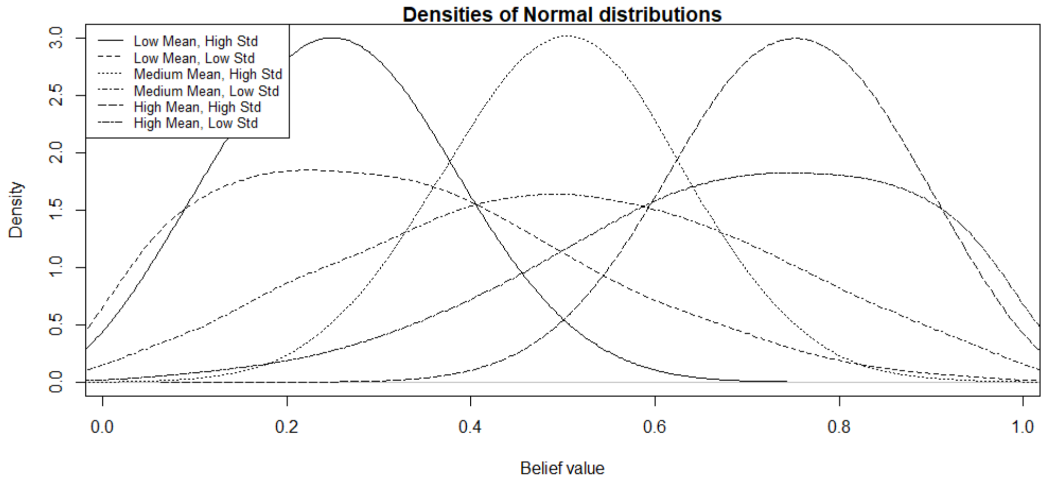
von Hayek, F. A. (1948). *Individualism and economic order.* University of Chicago Press.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, *18*(2), 107–126.

Wolfers, J., & Zitzewitz, E. (2010). Prediction markets. *California Management Review*, *52*(4), 125–141.

A Appendix

Density Pareto Distribution
High Imbalance
Moderate Imbalance
Low Imbalance
Density
0.0
0.1
0.2
0.3
0
50
100
150
200
250
300
350
Starting wealth



Densities of Normal distributions
Low Mean, High Std
Low Mean, Low Std
Medium Mean, High Std
Medium Mean, Low Std
High Mean, High Std
High Mean, Low Std
Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0.0
0.2
0.4
0.6
0.8
1.0
Belief value

B Appendix

Listing 1: Clearing Price Agent

```
def askPriceAlpha(self, model):
        price = 1
        if(len(model.askQueueAlpha)>0):
                low = model.lowestAskAlpha()
                if(self.belief<low.price):
                        if(low.agent.unique_id == self.unique_id):
                                price = low.price
                        else:
                                price = low.price-0.01
        else:
        price = 0.99
    return price

def askPriceBeta(self, model):
        price = 1
        if(len(model.askQueueBeta)>0):
                low = model.lowestAskBeta()
                if((1-self.belief)<low.price):
                        if(low.agent.unique_id == self.unique_id):
                                price = low.price
                        else:
                                price = low.price-0.01
    else:
        price = 0.99
    return price

def bidPriceAlpha(self, model):
        price = 0
        if(len(model.bidQueueAlpha)>0):
                high = model.highestBidAlpha()
                if((self.belief>high.price) & (self.wealth > high.price)):
                        if(high.agent.unique_id == self.unique_id):
                                price = high.price
                        else:
                price = high.price+0.01
    elif (self.wealth >= 0.01):
        price = 0.01
    return price
```

```
def bidPriceBeta(self, model):
        price = 0
        if(len(model.bidQueueBeta)>0):
                high = model.highestBidBeta()
                if(((1-self.belief)>high.price) & (self.wealth > high.price)):
                        if(high.agent.unique_id == self.unique_id):
                                price = high.price
                        else:
                                price = high.price+0.01
        elif (self.wealth >= 0.01):
                price = 0.01
        return price
```

Listing 2: Random Markup Agent

```
def askPriceAlpha(self, model):
        markup = random.randint(0,10)/100
        return self.belief + markup

def askPriceBeta(self, model):
        markup = random.randint(0,10)/100
        return (1 - self.belief) + markup

def bidPriceAlpha(self, model):
        markup = random.randint(0,10)/100
        return self.belief - markup

def bidPriceBeta(self, model):
        markup = random.randint(0,10)/100
        return (1 - self.belief) - markup
```

Listing 3: Fixed Markup Agent

```
def askPriceAlpha(self, model):
        return self.belief + 0.05

def askPriceBeta(self, model):
        return (1 - self.belief) + 0.05

def bidPriceAlpha(self, model):
        return self.belief - 0.05

def bidPriceBeta(self, model):
        return (1 - self.belief) - 0.05
```

Listing 4: Strategic Agent

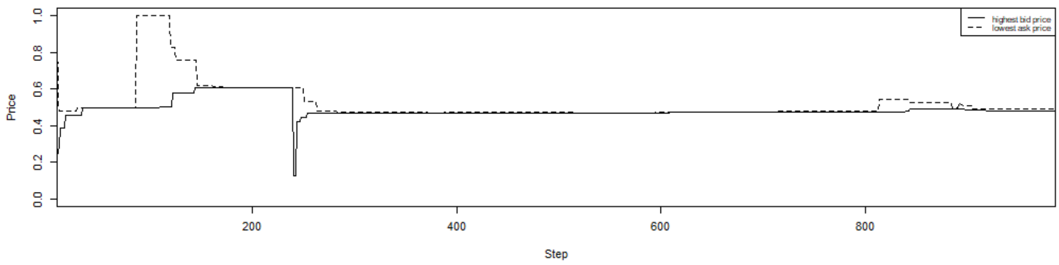
```
def askPriceAlpha(self, model):
        return model.highestAskAlpha().price -
        ((model.highestAskAlpha().price - self.belief -
        0.01)*((len(model.bidQueueAlpha)+1)/(len(model.bidQueueAlpha)
         + len(model.askQueueAlpha)+1)))
```

```
def askPriceBeta(self, model):
        return model.highestAskBeta().price - ((model.highestAskBeta().price
         - (1-self.belief -0.01))*((len(model.bidQueueBeta)+1)
         /(len(model.bidQueueBeta) + len(model.askQueueBeta)+1)))


def bidPriceAlpha(self, model):
        return model.lowestBidAlpha().price + ((self.belief - 0.01 -
        model.lowestBidAlpha().price)*((len(model.askQueueAlpha)+1)
        /(len(model.askQueueAlpha) + len(model.bidQueueAlpha)+1)))

def bidPriceBeta(self, model):
        return model.lowestBidBeta().price + (((1-self.belief - 0.01)
         - model.lowestBidBeta().price)*((len(model.askQueueBeta)+1)
         /(len(model.askQueueBeta) + len(model.bidQueueBeta) +1)))
```

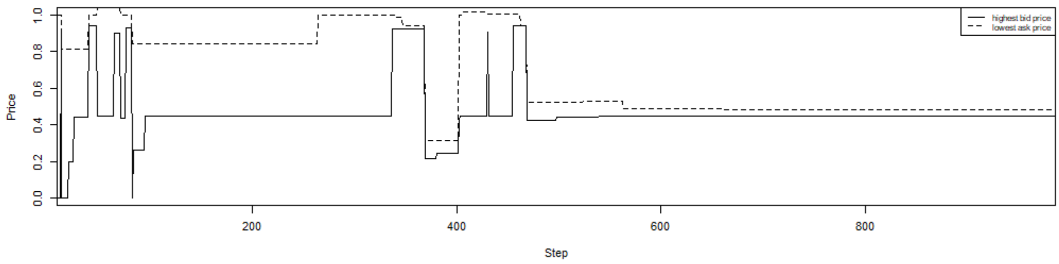
C Appendix

Sample Characteristics	Frequency Alpha	Frequency Beta	Average
Dependent Variables			
Average Price Alpha	134012		0.5268 (sd=0.1661)
Average Price Beta		134011	0.5258 (sd=0.1658)
Independent Variables			
Mean Belief	134012	134011	0.5000 (sd=0.1649)
Median Belief	134012	134011	0.5001 (sd=0.1738)
Control Variables			**Average Price**
Agent Strategy			
Strategic	33839	33839	α: 0.55, β: 0.54
Clearing Price	34990	34992	α: 0.51, β: 0.51
Random Markup	35000	35000	α: 0.51, β: 0.51
Fixed Markup	30183	30180	α: 0.55, β: 0.55
Wealth Distribution			
Equal	26830	26832	α: 0.52, β: 0.52
Random	26836	26835	α: 0.52, β: 0.52
Pareto (high imbalance)	26715	26718	α: 0.53, β: 0.53
Pareto (moderate imbalance)	26772	26768	α: 0.53, β: 0.53
Pareto (low imbalance)	26859	26858	α: 0.53, β: 0.53
Belief Distribution			
Random	19812	19814	α: 0.54, β: 0.54
Normal - High Mean & Low Standard Deviation	19896	19891	α: 0.76, β: 0.31
Normal - High Mean & High Standard Deviation	19711	19708	α: 0.69, β: 0.37
Normal - Medium Mean & Low Standard Deviation	15170	15167	α: 0.50, β: 0.50
Normal - Medium Mean & High Standard Deviation	19848	19848	α: 0.52, β: 0.52
Normal - Low Mean & Low Standard Deviation	19886	19894	α: 0.31, β: 0.75
Normal - Low Mean & High Standard Deviation	19689	19689	α: 0.37, β: 0.69

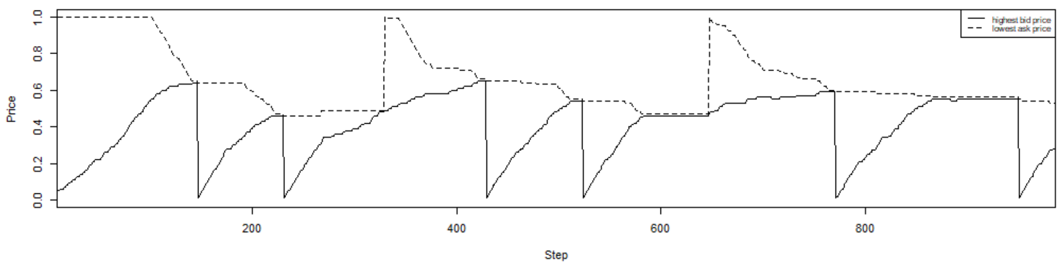
D Appendix



highest bid price
lowest ask price
Price
Step



highest bid price
lowest ask price
Price
Step



highest bid price
lowest ask price
Price
Step

E Appendix

Regression Results - Multiple Linear Regression	Model 1α		Model 2α		Model 3α		Model 4α		Model 5α	
	Estimate	**Std. Error**	**Estimate**	**Std. Error**	**Estimate**	**Std. Error**	**Estimate**	**Std. Error**	**Estimate**	**Std. Error**
Control Variables										
Agent Strategy [contrast = Clearing price]										
Strategic	0.0378902	0.0005354***	0.0377832	0.0005279***	0.0377839	0.0005308***			0.0377893	0.0005278***
Random Markup	-0.0045333	0.0005309***	-0.0046818	0.0005234***	-0.0046507	0.0005263***			-0.0046784	0.0005234***
Fixed Markup	0.0344259	0.0005579***	0.0343654	0.0005500***	0.0343523	0.0005530***			0.0343711	0.0005500***
Wealth Distribution [contrast = equal]										
Random	0.0033245	0.0006063***	0.0033491	0.0005977***	0.0034704	0.0006010***			0.0033275	0.0005977***
Pareto (high imbalance)	0.0108093	0.0006070***	0.0110950	0.0005984***	0.0110661	0.0006017***			0.0110832	0.0005984***
Pareto (moderate imbalance)	0.0093784	0.0006066***	0.0093693	0.0005981***	0.0094552	0.0006013***			0.0093555	0.0005980***
Pareto (low imbalance)	0.0044242	0.0006061***	0.0045772	0.0005976***	0.0046292	0.0006009***			0.0045597	0.0005976***
Belief Distribution [contrast = Normal - High Mean & High Standard Deviation]										
Random	-0.1526367	0.0007065***	0.0030999	0.0025987	-0.0604198	0.0020177***			0.0046745	0.0026213
Normal - High Mean & Low Standard Deviation	0.0676379	0.0007057***	0.0123954	0.0011282***	0.0465173	0.0008229***			0.0099065	0.0012528***
Normal - Medium Mean & Low Standard Deviation	-0.1778895	0.0007666***	-0.0219580	0.0026182***	-0.0855507	0.0020415***			-0.0203826	0.0026407***
Normal - Medium Mean & High Standard Deviation	-0.1700105	0.0007062***	-0.0141852	0.0026000***	-0.0777107	0.0020192***			-0.0126148	0.0026224***
Normal - Low Mean & Low Standard Deviation	-0.3770524	0.0007058***	-0.0099656	0.0059422	-0.1712906	0.0042799***			-0.0043217	0.0060689
Normal - Low Mean & High Standard Deviation	-0.3198379	0.0007076***	-0.0084006	0.0050550	-0.1354949	0.0038472***			-0.0052402	0.0051018
Independent Variables										
Mean Belief			0.8686079	0.0139638***			1.0375590	0.0148490***	0.9630527	0.0249528***
Median Belief					0.4603941	0.0094474***	-0.1247335	0.0140834***	-0.0766792	0.0167904***
R-squared	0.8214		0.8264		0.8245		0.8091		0.8264	
F-statistic	47400***		45560***		44960***		283900***		42530***	

* p < 0.05
** p < 0.01
*** p < 0.001

Regression Results - Multiple Linear Regression	Model 1β		Model 2β		Model 3β		Model 4β		Model 5β	
	Estimate	**Std. Error**	**Estimate**	**Std. Error**	**Estimate**	**Std. Error**	**Estimate**	**Std. Error**	**Estimate**	**Std. Error**
Control Variables										
Agent Strategy [contrast = Clearing price]										
Strategic	0.0276575	0.0005302***	0.0405109	0.0005441***	0.0277653	0.0005254***			0.0277603	0.0005221***
Random Markup	-0.0043363	0.0005257***	-0.0041831	0.0005177***	-0.0042171	0.0005209***			-0.0041876	0.0005176***
Fixed Markup	0.0404471	0.0005525***	0.0277682	0.0005221***	0.0405219	0.0005474***			0.0405036	0.0005440***
Wealth Distribution [contrast = equal]										
Random	0.0035111	0.0006004***	0.0034858	0.0005912***	0.0033634	0.0005949***			0.0035143	0.0005911***
Pareto (high imbalance)	0.0119010	0.0006010***	0.0116058	0.0005919***	0.0116399	0.0005955***			0.0116212	0.0005918***
Pareto (moderate imbalance)	0.0088078	0.0006007***	0.0088159	0.0005916***	0.0087292	0.0005952***			0.0088341	0.0005915***
Pareto (low imbalance)	0.0055573	0.0006002***	0.0054002	0.0005911***	0.0053497	0.0005948***			0.0054233	0.0005910***
Belief Distribution [contrast = Normal - High Mean & High Standard Deviation]										
Random	0.1771783	0.0006996***	0.0168743	0.0025705***	0.0836602	0.0019973***			0.0147957	0.0025926***
Normal - High Mean & Low Standard Deviation	-0.0561224	0.0006989***	0.0007412	0.0011160	-0.0347022	0.0008147***			0.0040276	0.0012392**
Normal - Medium Mean & Low Standard Deviation	0.1441759	0.0007592***	-0.0163296	0.0025898***	0.0505340	0.0020209***			-0.0184095	0.0026118***
Normal - Medium Mean & High Standard Deviation	0.1512640	0.0006993***	-0.0091316	0.0025718***	0.0576622	0.0019988***			-0.0112048	0.0025938***
Normal - Low Mean & Low Standard Deviation	0.3881470	0.0006989***	0.0102937	0.0058777	0.1794803	0.0042365***			0.0028425	0.0060026
Normal - Low Mean & High Standard Deviation	0.3202401	0.0007007***	-0.0003355	0.0050002	0.1332924	0.0038083***			-0.0045073	0.0050461
Independent Variables										
Mean Belief			-0.8940857	0.0138122***			-1.0481454	0.0147061***	-1.0188062	0.0246789***
Median Belief					-0.4668959	0.0093518***	0.1348354	0.0139480***	0.1012644	0.0166064***
R-squared	0.8242		0.8295		0.8274		0.8120		0.8296	
F-statistic	48320***		46570***		45880***		289500***		43480***	

* p < 0.05
** p < 0.01
*** p < 0.001