university of groningen / faculty of science and engineering

Belief Elicitation In Prediction Markets

Bachelor's Project Thesis

Marnix Jansma
Supervisor: Prof. D. Grossi

January 30, 2020

Abstract: Prediction markets have increasingly received interest as a tool for information aggregation from both researchers and corporations due to their utility in forecasting of various types of (binary) events. Some models have been proposed and put under scrutiny. This paper considers a new model and looks to give insight on the ability of prediction markets to stabilise and elicit information on the underlying beliefs of participants, as well as the robustness of this information given various market parameters such as market size. It uses a software implementation of the model to generate simulated data. The results indicate that price stabilisation is almost certain and that if some conditions are met, the markets are able to provide very accurate information on their participants.

1 Introduction

Suppose you have the night off and are getting ready to watch a sports match, the analysts have finished talking about the match expectations and the match is formally introduced. Generally this means you get to see the names of the teams, players, the venue and possibly the audience count and weather conditions. That's nice enough. Occasionally however, you also see an estimation of the outcome of the match: "60% chance of victory for Team Local Heroes, brought to you by your local betting outlet". It makes sense, if it can be assumed that the people that are betting are smart about it, they will bet on the team most likely to win. What a nice use of data!

Prediction markets like the Iowa Electronic Markets (Arrow et al., 2008) are similar to sports betting sites, in the sense that they allow participants to bet their money on the (often binary, but not always) outcome of an event like for example an election. Unlike betting sites, these markets are generally set up somewhat similar to stock markets (Wolfers & Zitzewitz, 2004). Where instead of just putting in your money on a specific outcome (e.g. candidate A will win the election), you can both buy and sell contracts that will pay out when that specific outcome occurs. However, if the alternative is true you get nothing. Bids are submitted to the market and then matched by the operator of the market. This continues until the market ends, usually shortly before the outcome is definitively determined. (For a more in-depth review of prediction markets Wolfers & Zitzewitz (2004) is recommended.)

The idea behind prediction markets is specifically to elicit the beliefs of participants about the event the market is based on. The idea being that if prices of contracts for a certain outcome are high, that might mean that the actual chance of that outcome occurring is high or at least that the participants in the market believe it to be so. If it can be assumed that these participants are well informed on the topic at hand, this information might well be very useful.

Research into prediction markets has generally confirmed this idea, presenting them as an efficient method to aggregate information from a large pool of experts into a single reliable statistic (Hearst, Hunson, & Stork (1999), Wolfers & Zitzewitz (2004), Borison & Hamm (2010)) while giving them an incentive to remain honest. Over time, they have been tested for accuracy and have been found to be quite good (Wilson (2012), Manski (2006), Arnesen & Bergfjord (2014)) at predicting public matters but also when used within corporations (Cowgill & Zitzewitz, 2015).

There has been some interest in modelling prediction markets and efforts have been made in creating models (e.g. Jumadinova, Matache, & Dasgupta (2011)) to demonstrate and explain various effects that are in prediction markets. Examples include the favourite-longshot bias (Page & Clemen, 2013), the effect of differing information (Jumadinova & Dasgupta, 2011) and risk-averse agent populations (He & Treich, 2017). It has been argued that these models need to be of sufficient complexity in order to capture all the dynamics that can be found in prediction markets well enough (Restocchi, McGroarty, Gerding, & Johnson, 2018).

Many things still remain to be researched when it comes to the topic of prediction markets. For example: What can be learned about the participants of prediction markets from the price and does it matter how many traders participate? What are the effects when there are differences in the amount that participants invest? Does the distribution of the beliefs of the participants matter at all? In this paper these topics will be explored and a simulation is developed in order to provide a base to possibly explore even more questions like these in the future.

Research question A model of prediction markets as provided in Grossi (unpublished) will be considered, in an attempt to answer the following question: *Do prediction markets elicit their participants' beliefs?*. To do so the following questions with regard to the model are answered: *When modelling prediction markets do market prices stabilise?* and *When modelling prediction markets do the market end prices correlate with the beliefs of the participants?*

2 Method

Data is generated by running simulations of an implementation based on the model proposed in (Grossi, unpublished). For an in-depth discussion of the model and the changes with regards to the model in Grossi (unpublished) please refer to section 3.

The model is an abstraction of prediction markets in that there is no trading/interaction between agents themselves and there is no limit to the supply of contracts.

The model is implemented as a simulation that allows for testing of parameters individually and aims to maximise reproducibility. Any data set generated using the simulation (including the ones used in the analyses for this paper) using whatever parameters should always be reproducible.

The simulation is written as a C++ program, as the possibility for object-oriented programming lends itself well for implementing multi-agent systems such as the one in the model. Additionally the potential efficiency and speed may possibly allow for easy generation of large sets of data points. The program will be able to write data to file in a `.csv` format which is then further processed and visualised using the R and RStudio software programs.

To answer the question the main variable of interest will be the market end price, i.e. the price when the market has reached either the maximum number of iterations or has stabilised. Especially if it is found that the markets do stabilise under the parameters it will be interesting to see how and if the market end price correlates with the belief distribution of the agents in the simulation.

Data generation A baseline data set is generated based on a specific parameter set that will suit further exploration by altering parameters of interest one at a time and comparing to the baseline, looking at the stabilisation rate of that set of parameters and the end price correlation with the average belief.

A random seed is used to generate the values for the agent beliefs (and in some cases endowments), such that the same seed should result in the same values and thus the same data set.

To make sure that any difference between the data sets can be easily attributed to the parameter change, only the baseline set will be generated using a random seed. All other sets will use the seed that was generated in the baseline set. This will make it so that the differences between the sets should come exclusively from the change in parameters.

This can be considered feasible as it is possible to readily generate such a large amount of data points, there should be little doubt about the (im)possibility of these results being specific to one seed.

3 Model

The used model is described in Grossi (unpublished) but also explained here as it has been adapted in a few areas. The model is a market game in which the agents can apply a mixed investment strategy. The model is simplified in that the signals (s) from the original model have been replaced by a fixed belief (b).

The model contains a set of agents $N = \{1, \ldots, n\}$, that have to invest their endowment into a prediction market that trades contracts on the odds of the outcome of event θ. It is a binary event so $\theta \in \{A, B\}$ i.e. when the contracts end either $\theta = A$ or $\theta = B$.

Each agent has its own belief $b \in [0, 1]$. This belief can be interpreted as their confidence that $\theta = A$ will be true, i.e. if $b = 0.5$ they belief that there is a 50% chance of A being the outcome of the event. Since binary events are under consideration here $1 - b$ is the belief of an agent that $\theta = B$ (the alternative) will be the outcome. $\boldsymbol{b}$ is the belief profile, which contains the set of beliefs $b_i \mid i \in N$ where i is an agent. Additionally each agent has an endowment $w_i \in \boldsymbol{w}$, a positive numerical value representing the wealth they can invest.

In the game agents submit a fraction of their endowment (σ_i, their investment strategy) in contracts for α according to some strategy update method. A contract for α pays out 1 if $\theta = A$ and 0 otherwise, and the same goes for β with respect to B. Agents always spend their complete endowment and so $1 - \sigma_i(b)$ is the fraction of wealth invested in contracts of type β. When all agents have placed their bets the investment strategy profile is obtained $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$.

Given the investment strategy profile the assets' market clearing prices can be calculated as:

$$p^{\alpha}(\boldsymbol{\sigma}) = \sum_{i \in N} \sigma_i \cdot w_i \tag{3.1}$$

$$p^{\beta}(\boldsymbol{\sigma}) = \sum_{i \in N} (1 - \sigma_i) \cdot w_i \tag{3.2}$$

Given that an agents endowment w_i sums to 1 for all agents:

$$\sum_{i \in N} w_i = 1 \tag{3.3}$$

Therefore:

$$p^{\alpha}(\boldsymbol{\sigma}) + p^{\beta}(\boldsymbol{\sigma}) = 1 \tag{3.4}$$

Now with the prices set it can be determined how many contracts an agent would be able to buy given its endowment.

$$x_i^{\alpha}(\boldsymbol{\sigma}) = \begin{cases} \frac{\sigma_i \cdot w_i}{p^{\alpha}(\boldsymbol{\sigma})} & \text{if } \sigma_i > 0 \\ 0 & otherwise \end{cases} \tag{3.5}$$

$$x_i^{\beta}(\boldsymbol{\sigma}) = \begin{cases} \frac{(1-\sigma_i) \cdot w_i}{p^{\beta}(\boldsymbol{\sigma})} & \text{if } \sigma_i < 1 \\ 0 & otherwise \end{cases} \tag{3.6}$$

Which, given our agents' endowments will add up to:

$$\sum_{i \in N} x_i^{\alpha}(\boldsymbol{\sigma}) = 1 \tag{3.7}$$

and

$$\sum_{i \in N} x_i^{\beta}(\boldsymbol{\sigma}) = 1 \tag{3.8}$$

Equation 3.5 holds as long as some agent has $\sigma_i > 0$, i.e. there is some agent that believes that $\theta = A$ with some probability and vice versa for $\theta = B$. This is a sensible assumption, since if agents agree unanimously on a subject the outcome of the event θ must be clear anyway. So in this model the supply of contracts is fixed and normalized to one unit, and perfectly matched to the demand.

To clarify, these assertions mean that in the model both the summed wealth of the agents as well as the amount of contracts available to them is normalised to 1. This means that an agent's endowment (w_i) can be interpreted as their share of the total wealth in the system. The price of a contract (e.g. p^{α}) is the proportion of the total wealth invested in the asset α. That also means that an agent does not (unless he is the only one interested) buy whole contracts but buys a fraction of the one contract available to all agents.

The complete market game now becomes $M = (N, \Theta, (b_i, w_i, \sigma_i, u_i)_{i \in N})$ with $\theta = \{A, B\}$ the binary set of outcomes of the event the market models, $b_i \in [0, 1]$ the possible belief that an agent may receive (essentially defining an agent as this is their only differing factor), $w_i \in (0, 1]$ an agents' endowment, $\sigma_i \in [0, 1]$ signifying the fraction of its endowment invested in α contracts.

The symbol u_i represents an agents' expected utility function:

$$u_i(b_i, \boldsymbol{\sigma}) = b_i \cdot x_i^{\alpha}(\boldsymbol{\sigma}) + (1 - b_i) \cdot x_i^{\beta}(\boldsymbol{\sigma}) \tag{3.9}$$

Which is equal to an agents' expected returns from the market game.

3.1 Agent Strategy Update

In the implementation, an agent's initial strategy is simply dependent on their belief, as that is all the information they have available at the start. Buying their belief (i.e. $\sigma_i = b_i$) is a reasonable strategy in this case. However, when the agent has knowledge of the price it can use that information to adjust their strategy in order to increase their expected utility. This is implemented by means of a strategy update function.

Before continuing equation 3.5 and 3.6 is expanded in equation 3.9 in order to provide a basis to develop some strategy update methods. This results in:

$$u_i(b, \boldsymbol{\sigma}) = \frac{b_i \cdot \sigma_i \cdot w_i}{p^{\alpha}(\boldsymbol{\sigma})} + \frac{(1 - b_i) \cdot (1 - \sigma_i) \cdot w_i}{p^{\beta}(\boldsymbol{\sigma})} \tag{3.10}$$

Continuing on this equation the endowment (w_i) is moved to the left-hand side and 3.4 is used to get:

$$\frac{u_i(b, \boldsymbol{\sigma})}{w_i} = \frac{b_i \cdot \sigma_i}{p^{\alpha}(\boldsymbol{\sigma})} + \frac{(1 - b_i) \cdot (1 - \sigma_i)}{1 - p^{\alpha}(\boldsymbol{\sigma})} \tag{3.11}$$

Equation 3.11 thus gives the expected utility of an agent per unit of endowment it possesses, although it should be noted that p^{α} is calculated using the agent's (previous) strategy and its endowment. The implication of this fact is considered in a strategy update presented in section 3.1.2.

It is also important at this point to remember that b_i and $1 - b_i$ signify an agent's belief in respectively outcome A and B and that σ_i and $1 - \sigma_i$ are dependent on each other and signify the portion of their endowment invested in contracts of either type. Here $\sigma_i \in [0, 1]$) is the only variable in this equation as all other parameters can be considered a constant.

The following sections propose a number of possible strategy update functions.

3.1.1 Simple all-in

Define:

$$ROI_{\alpha}(p^{\alpha}, b_i) = \frac{b_i}{p^{\alpha}} \tag{3.12}$$

and

$$ROI_{\beta}(p^{\alpha}, b_i) = \frac{1 - b_i}{1 - p^{\alpha}} \tag{3.13}$$

realising both are constants (given an asset price) where these parameters are understood to correspond to the return on investment of choosing a strategy that invests everything in contracts α and β respectively (i.e. $\sigma_i = 1 \mid \sigma_i = 0$). When put into equation 3.11 one possible strategy an agent could adopt becomes apparent.

$$\frac{u_i(b, \sigma_i)}{w_i} = \sigma_i \cdot ROI_{\alpha}(p^{\alpha}, b_i) + (1 - \sigma_i) \cdot ROI_{\beta}(p^{\alpha}, b_i) \tag{3.14}$$

A good strategy could be to select $\sigma_i = 1$ if $ROI_{\alpha} > ROI_{\beta}$ and $\sigma_i = 0$ if $ROI_{\alpha} < ROI_{\beta}$. If the expected returns are equal, all strategies have the same expected utility, and an agent might want to attempt to simply minimise losses by setting $\sigma = 0.5$.

However as it is right now the strategy is not sound, which is why it is adapted in the next section.

3.1.2 Smart all-in

A caveat in the Simple all-in strategy update is that if an agent changes its strategy, then the asset price will adapt to that change. Given this information an agent might want to amend the new strategy somewhat to account for this. To do this the strategy update must be adapted. A way of doing this is to use a variable $p^{\alpha\prime}$ instead of the constant p^{α} when computing the utility.

$$p^{\alpha\prime}(\sigma_i) = p^{\alpha}(\boldsymbol{\sigma}_{prev}) + (\sigma_i - \sigma_{iprev}) \cdot w_i \tag{3.15}$$

This equation uses what is known from equation 3.1, which is that part of the market price is directly influenced by each agent(by an amount of $\sigma_i \cdot w_i$), so when subtracting an agent's previous strategy from it (σ_{prev}, a constant) the price if this agent were to change to that different strategy is obtained.

A different strategy could then be one based on the previous strategy but substituting ${p^{\alpha}}'(\sigma_i)$ for p^{α} when computing the ROI's. This gives us the following:

$$\sigma_i = \begin{cases} 1 \text{ if } ROI_{\beta}({p^{\alpha}}'(0), b_i) < ROI_{\alpha}({p^{\alpha}}'(1), b_i) \\ 0 \text{ if } ROI_{\beta}({p^{\alpha}}'(0), b_i) > ROI_{\alpha}({p^{\alpha}}'(1), b_i) \\ \sigma_{prev} \text{ otherwise} \end{cases} \tag{3.16}$$

I.e. if it is profitable to invest maximally into either asset, the agent will choose to do so. If both assets are equally attractive, the previous strategy is kept.

3.1.3 Brute Force Maximisation

As discussed the utility function (3.11) should also take into consideration the price change. Substituting 3.15 into 3.11 the following is obtained:

$$\frac{u_i(b, \sigma_i)}{w_i} = \frac{b_i \cdot \sigma_i}{p^{\alpha} + (\sigma_i - \sigma_{prev}) \cdot w_i} + \frac{(1 - b_i) \cdot (1 - \sigma_i)}{1 - p^{\alpha} + (\sigma_i - \sigma_{prev}) \cdot w_i} \tag{3.17}$$

This change to the utility calculation can be significant. In figure 3.1 it can be observed that the utility function represents (for these cases) a curve, but it is bounded by the values the strategy can actually assume ($[0.0, 1.0]$). In order to maximise the utility, the maximum of the function needs to be computed. It should be noted however, that most of the time this reduces to the simple all-in strategy given previously because the gradient of the curve scales with the endowment of the agents which in turn scales with the number of agents. The consequence is that for large N, generally there is no proper maximum within the bounds of the function so the optimal strategy will still be to invest maximally in either asset.

This can be seen in figure 3.1 in the solid blue line, here the agent has only a fraction 0.1 of the total wealth and his expected utility given his strategy is almost a straight line due to the fact that the price is not influenced much by him. If the agent does posses a larger portion of the total wealth (as with the dotted curves) there are clearly visible local maxima.

There are possibly situations in which there might well be a local maximum available for some agent. For example, if a small number of agents possess a large portion of the endowment and a majority a small amount (i.e. if their endowments are distributed unevenly) some agents might very well have an optimal strategy that is not 1 or 0. To find this maximum one could try to solve the derivative for zero, but in practice here brute force computation is used.

A note on uneven distributions of wealth in prediction markets It is interesting to explore different distributions when it comes to endowments (and doing so is available through a parameter in the simulation) but it is important to remind oneself that prediction markets are often intended as scientific tools to aggregate information and so in practice this type of distribution might not actually occur due to the limit to how much can be invested that is often imposed (as in Iowa Electronic Market (2020)). This makes it so that in practice most participants in prediction markets invest the same or a similar amount.

3.1.4 Prudent

A final strategy that is proposed here is as follows:

$$\sigma_i = \begin{cases} \sigma_{prev} + \delta_i & \text{if } u_i({p^{\alpha}}', \sigma_{prev} + \delta_i) > u_i(p^{\alpha}, \sigma_{prev}) \\ \sigma_{prev} & \text{otherwise} \end{cases} \tag{3.18}$$

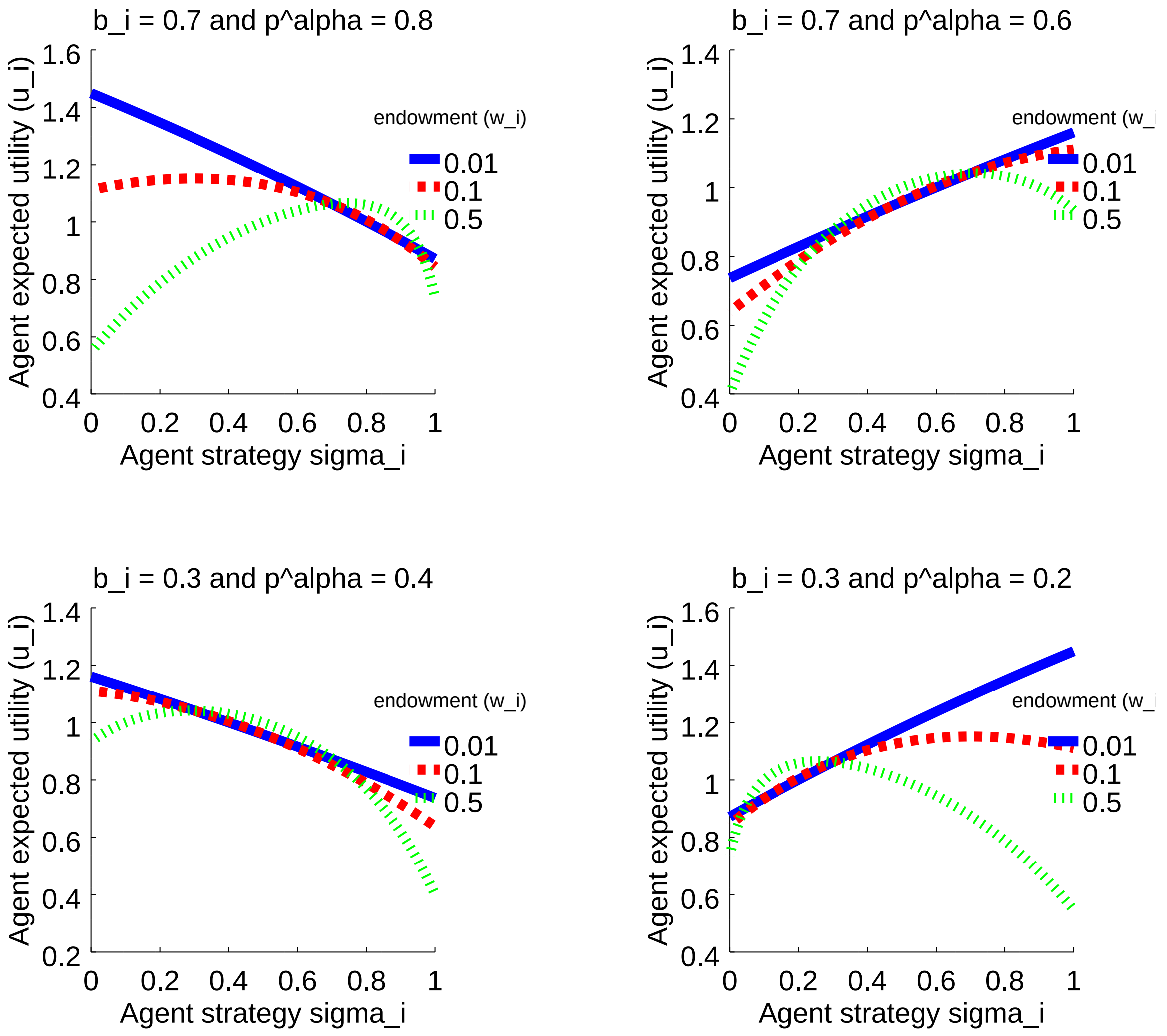
$$\delta_i = \begin{cases} 0.1 & \text{if } ROI_{\alpha}(p^{\alpha}, b_i)) > ROI_{\beta}(p^{\alpha}, b_i)) \\ -0.1 & \text{if } ROI_{\beta}(p^{\alpha}, b_i)) > ROI_{\alpha}(p^{\alpha}, b_i)) \end{cases} \tag{3.19}$$

That is to say, an agent will update its strategy (by a fixed amount) if the agent believes that it might gain a benefit from this change and the strategy remains within the $[0, 1]$ bounds. This strategy is considered because the small increments in strategy that are performed here might lead to different stable points than for example the maximisation strategies.

3.2 Update order

The order in which agents update their strategy can be viewed as a parameter in itself, and is avail-

Figure 3.1: Some sample utility curves with $w_i \in \{0.01, 0.1, 0.5\}$ and varying b_i, p^α



able in the simulation. Both methods will be briefly explained, of which the first seems the most reasonable and will be used as a baseline. For both update methods the first iteration will be computed using the aforementioned buy-your-belief scheme in order to establish a base price.

Random order Every iteration a (seeded) random order will be generated and the agents will be given the chance to revise their strategy in that order. Every time an agent submits its bid, the price will be updated according to that bid, and this price is given to the next agent in line.

Simultaneous update In this update scheme all agents will generate their bid based on the same price, which was generated on the previous iteration. This causes greater fluctuations in price and could prove to be more resilient to stabilisation.

4 Program Specification

The requirements and specification for the simulation are fairly simple. In the following sections first the practical requirements of the simulations are explained, and after that a more formal specification of what the elements of the program will look like and what they can do is given.

4.1 Program Outline

The simulation can do the following:

1. Simulate a prediction market-like game of a variable number (N) of agents as specified in the section 3.
2. Each agent has a fixed belief (b) which can be pulled from a specifiable distribution. The following distributions are provided in the program.
   (a) Uniform random distribution.
   (b) Normal distribution with given mean and variance.
   (c) Pareto distribution with given shape.
3. The agent endowment may also be pulled from a specifiable distribution, but the values are normalised such that they sum to one.
4. In the first iteration of the simulation the agents use a basic buy-your-belief strategy to decide how to invest their endowment.
5. In all other iterations the agents use one of the strategies discussed, selectable at the initialisation of the simulation as a parameter.
6. The simulation continues iterating until a given number (one hundred) of iterations is reached or the game reaches a stable state.
7. The game is stable if the asset price stays the same for ten iterations.
8. Market data can be saved to file.
9. The random elements of the simulation (such as order and beliefs) can be seeded to allow for reproducibility.
10. All agents will have the same investment strategy update and can compute the expected utility of any given investment (given an asset price and a strategy).

Distribution Implementation The way that the distributions are implemented internally is by generating numbers from bounded normal and Pareto distributions (based on the parameters given to the program) in the range $[0.0, 1.0]$ in order to fit the model requirements. The same generator is used for the endowments, which therefore are normalised in order to sum to 1.

4.2 The Agent and Market objects

The agent objects only maintain absolutely necessary state variables:

- Their belief.
- Their endowment.
- Their preferred strategy update.
- Their most recently submitted bid/strategy.

No other information is required. The agent has the following abilities:

- The ability to compute utilities given a price.

Table 4.1: Baseline simulation parameters

N	100
Endowment distribution	Equal
Belief distribution	Uniform random
Update sequence	Random order
Number of resets	100,000
Agent strategy	Smart all-in

- The ability to select a strategy based on a price.

All other information required for the simulation is held by the Market object:

- The current price.

The market has the following routines:

- The ability to request bids from agents.
- The ability to iterate, requesting bids from all agents and generating a new price.

In practice, more values and routines may be available (such as the minimum and maximum price), but what is presented here are the bare essentials.

4.3 Data sets

To explore the model several data sets will be generated using various sets of parameters. The first data set will be used as a baseline to compare the other results, which are generated with a set of parameters based on the baseline but altered on a parameter of interest. The baseline parameters are outlined in table 4.1. These parameters represent an ideal situation in which the market is expected to be the most informative.

N is chosen such that there are enough agents in the system that a single agent is unlikely to affect the price too much, but is still relatively small. Distribution of endowment is set to equal, as this will make it so that each agent is able to influence the market price equally. The beliefs are set to be random, to make sure that agents are different enough that you can assume their expected utilities for any price state to be quite different (meaning they make different decisions). The random order update sequence is used in order to prevent the order of the agents from becoming a confounding factor to the price. Smart all-in is the base strategy as it seems the most simple plausible strategy while still being rational.

Additionally the number of simulated markets to generate is set to 100,000 and they are run for up to 100 iterations or until stabilisation is achieved. A market is considered stable if the price does not change for ten consecutive iterations. In theory a market can be considered stable if it has the same price for two consecutive iterations but to reduce the chances of errors ten is set as a safe amount.

Parameters of interest The following parameters will be altered from 4.1 in order to generate data sets:

1. N: tiny (10), small (30), large (1000).
2. **Endowment distribution**: random, Pareto and normal.
3. **Belief distribution**: random, Pareto and normal.
4. **Update sequence**: simultaneous and random order.
5. **Strategy**: Smart all-in, prudent and brute force maximisation.

Each parameter will be changed one-by-one to one of the given alternatives, with all the other parameters remaining the same.

5 Results

As mentioned the baseline set will be generated according to the parameters in table 4.1 with the resulting spread of the equilibrium prices visible in the plot in figure 5.1.

Table 5.1 shows the correlation values for the averaged belief of the agents in a market to the end price. This is done for all the different parameters of interest mentioned in section 4.3, with the baseline results in the top row. What becomes immediately apparent is that all strategy update variations seem to result in the same correlation values. For all the results in the table it is true that every market stabilised (well) within 100 iterations.

Table 5.1: End price and average belief correlations for the different parameter tweaks

Strategy	Smart all-in	Brute force	Prudent
Altered parameter	Correlation value		
Baseline	0.99	0.99	0.99
N = 1000	>0.99	1.00	1.00
N = 30	0.97	0.97	0.97
N = 10	0.90	0.91	0.91
Normal belief distribution ($\mu = 0.5, \sigma = 0.15$)	0.96	0.96	0.96
Pareto belief distribution ($shape = 0.4$)	0.98	0.98	0.98
Normal endowment distribution ($\mu = 0.5, \sigma = 0.15$)	0.90	0.90	0.90
Pareto endowments	0.46	0.46	0.46
Randomised endowments	0.76	0.76	0.76

5.1 Endowment Distributions

To further explore the results when altering the endowment distribution, a data set was generated with various numbers of agents (N). The results of that are in Table 5.2.

Table 5.2: Results for unequal endowments distribution and varying N

N	Pareto endowments	Normal endowments
Large N (1000)	0.46	0.91
Baseline N (100)	0.46	0.90
Small N (30)	0.46	0.87
Tiny N (10)	0.49 *	0.78

For the normal distribution this results in a slight positive trend for larger N. For the Pareto distribution the results are mostly the same as the results from the initial data set, except for tiny N where the correlation is slightly stronger.

5.2 Simultaneous update

Tables 5.3, 5.4 and 5.5 were generated using the baseline parameters except for the strategy updates and the price update method. All results came from using the simultaneous update method.

*Two of the runs from the simulation did not stabilise, and the data showed that they had reached invalid prices. ($p^\alpha < 0$ and $p^\alpha > 1$). These data points have not been counted towards the correlation value.

Table 5.3: Results for simultaneous price update

Strategy update	Correlation value	Stabilised prices only (stable %)
Brute force	-0.02	0.17 (32%)
Prudent	0.87	0.86 (89%)
Smart all-in	0.11	0.88 (64%)

Table 5.4: Results for simultaneous price update (more iterations)

Strategy update	Correlation value	Stabilised prices only (stable %)
Brute force	-0.02	0.17 (32%)
Prudent	0.87	0.87 (>99%)
Smart all-in	0.11	0.88 (64%)

When looking at the results (5.3) under the simultaneous update condition it is immediately apparent that the correlation is much lower and for two of the strategy updates there is seemingly no correlation at all. Under this condition the systems also no longer always stabilise, but when only considering those that did do so the results get closer to the baseline for the Smart all-in strategy.

Because of this fact the data set was generated another time, this time with the maximum number of iterations set to 300 (three times the default).

What can be seen in table 5.4 is that more than

Table 5.5: Results for simultaneous price update (counting oscillations)

Strategy update	Correlation value (all stable or oscillating)
Brute force	-0.02
Prudent	0.87
Smart all-in	0.11

99% of the prudent markets have now stabilised, but otherwise the results seem to be basically identical.

Due to the volatility of the market prices when all agents submit their bids at the same time, the possibility of a sort of pseudo-equilibrium in which the market price oscillates between two points is also considered. The end price in this case is defined as the average of the two points.

Table 5.5 shows what happens when counting oscillating markets as equilibria. All markets eventually reach a stable or oscillating state but the correlation values remain the same.

Spread To support an understanding of what these correlation values mean a plot of the data spread of these sets is provided. Most of the plots are very similar to the baseline so only the plots of the data sets generated with non-normal endowments are included (figure 5.1).

What can be seen is that in general the end price seems to have a linear (one-to-one) relation to the average belief, and that even in the cases where inequality of endowment is introduced the linearity still seems to hold to some extent.

6 Discussion

What immediately becomes apparent when looking at the results in table 5.1 is that in the baseline data there is an almost one-to-one correlation between the average belief and the end price of the assets. Varying the number of agents seems to slightly influence this, which would be expected as individual agents have a greater influence for small N and so agents with a belief that deviates much from the average will have a stronger impact on the price. However, even with $N = 10$ the correlation stands strong.

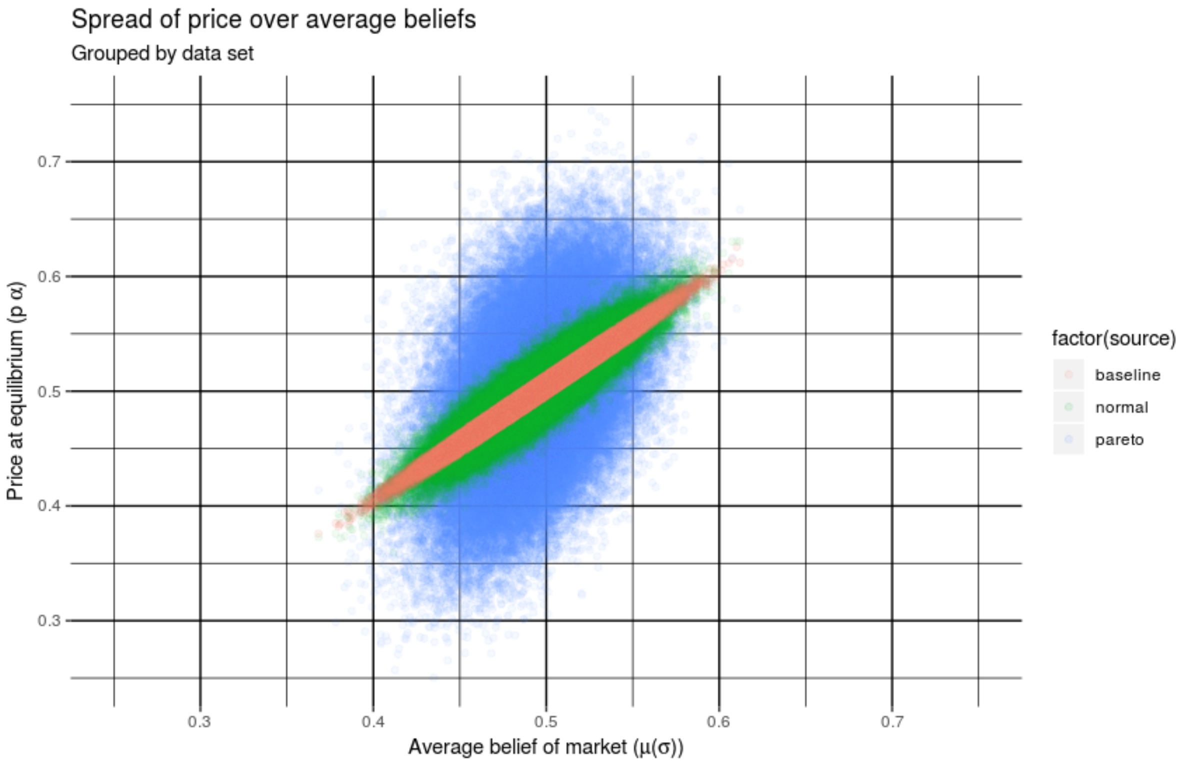
Not much of a difference can be seen when it comes to the various methods of distributing belief but when altering the endowment distributions the correlation becomes a lot weaker, especially for the Pareto distribution. Again, this is in line with what would be expected as a few 'rich' agents get to influence the price more strongly. Even so the correlation is still quite strong for the normally distributed endowments, but it is evident that a key component of an informative prediction market is that participants invest an equal or similar amount of funds. This is quite intuitive and for example the Iowa Electronic Markets' prospectus (Iowa Electronic Market, 2020) already contains a clause that limits a trader's investment to up to 500 US Dollars. On top of that there might possibly be other reasons to limit a person's investment in a market especially if the goal is to use it as an informative tool rather than to provide income to the operator.

The assumption that the smart all-in strategy update should perform similarly to the brute force strategy update seems to be correct. This is good as simple strategies are in general intuitively more plausible.

Even when using the simultaneous update parameter it seems that equilibria can still be reached, however due to the volatility of the price it is less likely for the market to settle. The reason that the market stabilises more often for the prudent agent strategy is because the volatility is kept in check due to the fact that the agent only changes their strategy by a value of δ at a time the price can only change by at most $N \cdot \delta \cdot \boldsymbol{w}$. This modulation makes it so that the market price fluctuates less and is thus more likely to reach a stable point.

As mentioned, the initial price (iteration one) is calculated by soliciting a buy-your-belief bid from all agents. In the case of equal endowment distribution, this will mean that the initial price will be exactly the average belief of the agents in the market and this is also true to an extent for some of the other parameter sets. To be able to assure ourself that the end prices are not simply a consequence of this fact it would also be interesting to compare the data sets to other sets that start with a random/adjusted price. This will also give some insight in the possibility of price convergence towards the average belief, because if it is found that the correlation holds under these circumstances the average belief value might be an attractor for the market

Figure 5.1: Spreads using different endowments distribution (baseline/equal, normal distribution, Pareto)



price.

Possible implementation improvements As mentioned in section 4, the way distributions are currently implemented means that the beliefs and endowments in the simulation are not generated using a regular normal distribution or pareto distribution(since they are bounded). That being said, the resulting belief and endowment profiles in the systems are expected to fit well enough for the purposes of this research.

An updated implementation could generate all values required in a simulation before hand, after which they can then be normalised for use as beliefs or endowments. In this case the benefit would be that any parameters for the distributions could be used without values for the beliefs having to be rejected, but normalisation also makes it so that the parameters of the distribution are no longer directly related to the beliefs and endowments in the system due to the extra step.

The model as used in this research also does not use an explicit source of information. Beliefs are generated from a distribution, which could be interpreted as modelling the true probability but this is not the intention. An expanded model could generate a true probability π which is either shared with the agents with a layer of noise or some function that represents uncertainty and interpretation (like the signals in Grossi (unpublished)).

7 Conclusion

To answer the question: *When modelling prediction markets do market prices stabilise?* The answer seems to be yes, even markets with small numbers of agents should be able to arrive at a stable price. If it can be assumed that the brute force strategy is optimal and that the smart all-in strategy is a close approximation of that strategy, it is even reasonable to assume that these are Nash equilibria. However, this is not proven here. It does provide a solid incentive however to find a formal proof as it seems plausible there should be one.

With regards to the question: *When modelling prediction markets do the market end prices correlate with the beliefs of the participants?* The answer seems also to be a definite yes, given that endowments are distributed evenly and the volatility of the market price is kept in check, the market end price shows a strong correlation to the population belief.

The answer to the main question: *Do prediction markets elicit their participants' beliefs?* seems to be yes, if certain conditions are known to hold true. This combined with existing data suggesting that prediction markets are a good and accurate tool for forecasting and prediction (Wolfers & Zitzewitz (2004), Wilson (2012)) might be an incentive to use them even on small scales, if the facilities for creating such a market are available. Perhaps it might even be interesting to create such facilities to spread their usage.

It should be considered however that compared to some other models of prediction markets (Jumadinova et al. (2011), He & Treich (2017) and Restocchi et al. (2018)) the model used in this paper is quite simple. This can be considered a strength of the model as simple models are often preferable but it also means that it can not capture certain complexities. An example of such complexities could be how a market reacts to the introduction of new information. Specifically in Restocchi et al. (2018) the importance of complexity in order to capture characteristics observed in real prediction markets is emphasised, and Restocchi, McGroarty, & Gerding (2019) proposes some validation methods to be used on models of prediction markets. It would be interesting to see if the current model could be expanded and tested using such validation measures, and an improved implementation could perhaps even be tested against real prediction market data as in Wilson (2012).

Further research The model used in this research shows that given a set of rational agents in a market game, the end price will almost directly correlate with the average belief (aggregated knowledge) of the agents. It also seems from the results that this correlation is very strong with regards to the parameters that have been considered, showing that (perhaps surprisingly) not a lot of individuals are needed for the end price to be indicative of the underlying beliefs. Of course the fact that in this model the agents buy their contracts from an unlimited source, which in the real life is implausible, can not be disregarded. If supply and demand were modelled in the system, it is quite likely problems would arise that are akin to the thin market problem (Hanson, 2003).

Some further expansions to what is provided in this research could be to introduce agents into the system that are not rational, to see to what extent they could sabotage the reliability of the information. This could be useful as in practice traders could have an incentive to influence the market price in a certain way, and market holders must consider their options when they detect this kind of interference or at least account for the possibility.

In general the robustness of the results presented here could be further investigated (e.g. by altering multiple parameters at a time), which might then allow for provision of a recommendation of how a prediction market should be set up to ensure that it fulfils its goal of providing honest, unbiased and valuable information.

References

Arnesen, S., & Bergfjord, O. (2014). Prediction markets vs polls-an examination of accuracy for the 2008 and 2012 elections. *Journal of Prediction Markets*, *8*(3).

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., ... Zitzewitz, E. (2008). The promise of prediction markets. *Science*, *320*(5878), 877-878. Retrieved from `http://science.sciencemag.org/content/320/5878/877.abstract` doi: 10.1126/science.1157679

Borison, A., & Hamm, G. (2010). Prediction markets: A new tool for strategic decision making. *California management review*, *52*(4), 125-141.

Cowgill, B., & Zitzewitz, E. (2015, 04). Corporate Prediction Markets: Evidence from Google, Ford, and Firm X *. *The Review of Economic Studies*, *82*(4), 1309-1341. Retrieved from `https://doi.org/10.1093/restud/rdv014` doi: 10.1093/restud/rdv014

Grossi, D. (unpublished). *Condorcet markets (unpublished).*

Hanson, R. (2003, Jan). Combinatorial information market design. *Information Systems Frontiers*, *5*(1), 107-119. Retrieved from `https://search.proquest.com/docview/232060218` doi: 1022058209073

He, X.-Z., & Treich, N. (2017). Prediction market prices under risk aversion and heterogeneous beliefs. *Journal of Mathematical Economics*, *70*, 105-114.

Hearst, M. A., Hunson, R. D., & Stork, D. G. (1999). *Building intelligent systems one e-citizen at a time* (Vol. 14) (No. 3). doi: ID:110.1109/5254.769877

Iowa Electronic Market. (2020). *Iem prospectus: Pres20 vs 2020 us presidential vote share market.* Retrieved from `https://iemweb.biz.uiowa.edu/markets/pr_pres20_VS.html`

Jumadinova, J., & Dasgupta, P. (2011, May). A multi-agent system for analyzing the effect of information on prediction markets. *International Journal of Intelligent Systems*, *26*(5), 383-409. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1002/int.20471` doi: 10.1002/int.20471

Jumadinova, J., Matache, M. T., & Dasgupta, P. (2011). A multi-agent prediction market based on boolean network evolution. In *Proceedings of the 2011 ieee/wic/acm international conferences on web intelligence and intelligent agent technology-volume 02* (p. 171-179). IEEE Computer Society.

Manski, C. F. (2006). *Interpreting the predictions of prediction markets* (Vol. 91) (No. 3). Retrieved from `http://www.sciencedirect.com/science/article/pii/S016517650600022X` (ID: 271682) doi: //doi.org/10.1016/j.econlet.2006.01.004

Page, L., & Clemen, R. T. (2013). Do prediction markets produce well-calibrated probability forecasts?*. *The Economic Journal*, *123*(568), 491-513. Retrieved from `https://doi.org/10.1111/j.1468-0297.2012.02561.x` (doi: 10.1111/j.1468-0297.2012.02561.x; 26) doi: 10.1111/j.1468-0297.2012.02561.x

Restocchi, V., McGroarty, F., & Gerding, E. (2019). *The stylized facts of prediction markets: Analysis of price changes* (Vol. 515). Retrieved from `http://www.sciencedirect.com/`

science/article/pii/S037843711831313X (ID: 271529) doi: //doi.org/10.1016/j.physa.2018.09.183

Restocchi, V., McGroarty, F., Gerding, E., & Johnson, J. E. V. (2018). *It takes all sorts: A heterogeneous agent explanation for prediction market mispricing* (Vol. 270) (No. 2). Retrieved from http://www.sciencedirect.com/science/article/pii/S037722171830314X (ID: 271700) doi: //doi.org/10.1016/j.ejor.2018.04.011

Wilson, J. L. (2012, Nov 16,). Prediction markets: How accurate are they? *CHANCE*, *25*(4), 32-38. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09332480.2012.752284 doi: 10.1080/09332480.2012.752284

Wolfers, J., & Zitzewitz, E. (2004, May 1,). Prediction markets. *The Journal of Economic Perspectives*, *18*(2), 107-126. Retrieved from http://www.ingentaconnect.com/content/aea/jep/2004/00000018/00000002/art00006 doi: 10.1257/0895330041371321