Simulated Climate Prediction Markets and Convergence of Predictive Models

August 28, 2015

Abstract

We present an agent-based model of a climate prediction market in which traders adapt their belief about the climate based on the monetary performance of their neighboring traders in a social network. We analyze the model to determine how a variety of factors affect whether climate prediction markets foster a convergence of market participants' beliefs regarding a true model of the earth's climate. We find that only the imposed "true model" itself significantly affects the convergence of beliefs. If the true model from which the temperature data is generated is auto-regressive, with no human-induced effects, participation in the prediction market, on average, does not cause convergence of traders' beliefs. On other hand, if anthropogenic climate change is real, traders' beliefs are more likely than not to converge toward the correct climate model due to participation in the market.

Introduction

The previous two decades saw a strong polarization of the climate change debate. Although the scientific consensus on the anthropogenic nature of climate change strongly increased, the average belief about anthropogenic climate change did not evolve much within the public [18]. In addition, the divide on anthropogenic climate change between liberals and conservatives has grown steadily, indicating that the question is becoming increasingly politicized and potentially disconnected from scientific evidence. This is troubling given the costs of a misinformed climate policy. If anthropocentric climate change is a myth but is taken to be true, a tremendous amount of public resources will be wasted on adaptation and mitigation efforts. On the other hand, if climate change is human induced but not recognized as so, the costs of inaction would be huge. Given that effective climate policies require long-lasting measures to be taken, the importance of an accurate consensus on the issue is manifest.

Attempts to foster such consensus face many social and psychological challenges. Recently, people have argued that such challenges could be met by setting up climate prediction markets [9, 18]. The idea of using prediction markets to efficiently aggregate information about uncertain events has been discussed for decades [8]. Prediction markets have been shown to have interesting theoretical properties [16, 5] and to perform well in terms of prediction accuracy in experiments [6, 7], agent-based models (ABM) [11, 10], and in the real-world [19].

However, to the best of our knowledge, the idea that prediction markets can generate consensus *on the factors* affecting uncertain events has never been quantitatively explored. ABMs of prediction markets have been studied [11, 17, 10], some of which feature communication between agents. In these models, however, beliefs about the uncertain outcomes are constructed in rather abstract ways. In particular, beliefs are not based on structural models from which agents could derive causal implications, and therefore these models are not suitable for investigating the convergence of the underlying explanatory models agents employ for prediction.

From a public policy perspective, changing the explanatory models of market participants is one of the most important roles that prediction markets might play.[1] Effective climate policy not only requires an accurate consensus on future climate *outcomes*. It also requires an accurate consensus on the causal *mechanisms* influencing such outcomes. If people agree that temperature will rise, but some believe it will be due to greenhouse gases, while others believe that it will be caused by increased solar activity, inconsistent and ineffective policies may be implemented. Our model is designed to investigate whether, and under what conditions, prediction markets can foster such convergence in predictive models.

Model Design

TODO: Martin use this text below (it is commented out) from the original document and flesh it out for the Supplementary Material and then write a couple paragraph summary and the key equations to put in this main article right here.

In particular, lets go over main design and then define what each of these mean in our model:

- `ideo` $\sim Unif(0, 1)$
- `market.complet` $\sim Unif(0, 1000)$ (mapped into integer)
- `n.edge` $\sim Unif(100, 200)$ (mapped into integer)
- `n.traders` $\sim Unif(50, 250)$ (mapped into integer)
- `risk.tak` $\sim Unif(0, 1)$

[1] And perhaps the most important potential social benefit of prediction markets given the comparable predictive power of statistical models, which are less expensive than setting up and maintaining prediction markets [4].

- `seg` $\sim Unif(0, 1)$
- `true.model` $\sim Binom(0.5)$

Model Analysis

We estimated the effects of the parameters of our model on the difference between the fraction of traders who believe in the true model at the beginning of the experiment and the fraction who believe in the true model at the end of all trading sequences. The sensitivity analysis is based on a Latin Hypercube Sampling of 10,000 parameter sets from the following distributions [1, 2].

- `ideo` $\sim Unif(0, 1)$
- `market.complet` $\sim Unif(0, 1000)$ (mapped into integer)
- `n.edge` $\sim Unif(100, 200)$ (mapped into integer)
- `n.traders` $\sim Unif(50, 250)$ (mapped into integer)
- `risk.tak` $\sim Unif(0, 1)$
- `seg` $\sim Unif(0, 1)$
- `true.model` $\sim Binom(0.5)$

We use the model to simulate 10,000 outcomes based on the input paramter sets and then conduct a partial rank correlation coefficient analysis on the relationship between the input matrix, X, resulting simulated vector, y [12, 13, 15]. Partial correlation computes the linear relationship between the part of the variation of X_i and y that are linearly independent of other X_j ($j \neq i$). The only difference with the partial correlation and partial *rank* correlation (which we use here), is that the outcome variable y is first ranked-transformed in order to allow us capture potentially non-linear relationships.[2]

Our sensitivity analysis, illustrated by Fig. 1, strongly suggests that the magnitude of convergence to the true model is conditional on what the true model actually is: if anthropogenic climate change is true, a prediction market is likely to cause convergence to the true model; while if anthropogenic climate change is false, and the model of the climate is in fact auto-regressive, convergence to this true model is unlikely (Fig. 2).

[2]That is every value, y_s, in the data set is replace by a number $f(y_s) \equiv |\{y_r \mid y_s > y_r\}|$, where y_r are the other observed values of y in the data set.

Discussion

Proponents of climate prediction markets argue that climate prediction markets will foster convergence to the best approximate model, whatever the actual true model is. It is framed as an apartisan proposal: if climate change is not anthropogenic, traders will converge to non-anthropogenic beliefs similar to how they would converge to anthropogenic beliefs if the climate is truly influenced by human activities. Our findings contradict this assumption.

TODO: Jonathan, perhaps talk about the policy implications of this finding...

Supplemental material

0.1 Detailed description of the model

0.1.1 Elements of the model

The ABM is made of the following elements.

A simple "true" model of world temperatures. The true model is determined by parameter `true.model`.

If `true.model` = 0, then anthropogenic climate change is a myth and the temperatures follow the autoregressive process

$$T_t \;=\; \lambda T_{t-1} + \epsilon_t, \tag{1}$$

where T_t is the temperature at time t, and ϵ_t^T is a random error term.

The value of λ is obtained by calibration on actual data. Specifically, λ is the ordinary least square estimator of (1), where the temperature time series are US yearly averages based from 1895 to 2010, based on US monthly averages from http://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp (TAVG in the data set, retrieved on May 5th 2015).

The distribution of ϵ_t is the empirical distribution of the residual from this linear regression.

If `true.model` = 1, then anthropogenic climate is true and the temperatures follow the process

$$T_t \;=\; \beta GHG_{t-1} + \mu_t, \tag{2}$$

where GHG_t is the level of greenhouse gas emissions at time t, and ϵ_t^{GHG} is a random error term.

Again, the value of β is obtained by calibration on actual data. Specifically, β is the ordinary least square estimator of (2), where the temperature time series are as in `true.model` = 0, and the

GHG time series are Global Mean CO2 Mixing Ratios (ppm) on the same time period from http://data.giss.nasa.gov/modelforce/ghgases/Fig1A.ext.txt (MixR in the data set, retrieved on May 5th 2015).

The distribution of μ_t is the empirical distribution of the residual from this linear regression.

In either cases (`true.model` = 0 and `true.model` = 1), the chosen true model is used to generate and artificial temperature time series $\{\hat{T}_t\}_{t=1895}^{t=2010}$.

We emphasize the artificial time series is in general different from the temperature in the empirical data set.

- When `true.model` $= 0$, $\hat{T}_{1895}$ is chosen to be the temperature in 1895 from the empirical data set, and $\{\hat{T}_t\}_{t=1896}^{t=2010}$ are computed recursively using (1).
- When `true.model` $= 1$, the value of GHG_t for all $t \in \{1895, \ldots, 2010\}$ is taken from the empirical data set and $\{\hat{T}_t\}_{t=1895}^{t=2010}$ are computed using (2).

Beliefs. Traders use either (1) or (2) to forecast future temperatures in order to trade on the prediction market. These 2 models are interpreted as the trader's beliefs about the true climate model. They are meant to represent pervasive positions on climate change in the public debate. In order to approximatively match the current configuration of beliefs in climate change in the US, model (1) is randomly assigned to half of the traders, while (2) is assigned to the other half of the traders.

Thus, both when `true.model` $= 0$ and when `true.model` $= 1$, some traders use the true model to make predictions. However, this does not mean that traders using the true model make perfectly accurate predictions. Although these traders believe in the correct *functional form* of the model, they still need to calibrate it based on limited noisy data. Therefore, the values they use as the *parameters* of the model will typically still be different from λ and β (see description of the model's timing below).

Traders. Traders are initial endowed with one single unit Experimental Currency Unit (ECU). Traders use their believed model to forecast the distribution of future temperatures and determine their reservation price for different securities. Securities pay 1 ECU if the actual temperature at some time t^* falls between a certain range. Traders are assumed to be *risk-neutral expected utility maximizers*. Therefore, their reservation price for a security $\tilde{s}$ is simply their assessment of the probability that the temperature at time t^* falls in the range covered by $\tilde{s}$.

Based on their reservation price, traders behave as "*zero-intelligence*" agents ([3]), that is they sell at a random price above their reservation, and buy at a random price below their reservation. The distribution from which buy and sell prices are drawn is determined by the traders risk attitude, characterized by parameter $\texttt{risk.tak}_i$ (more details below). These strategies are simple but have been shown to provide good approximations of the behavior on prediction markets (see for instance [11]), and in financial markets more broadly.

Continuous-Double Auction (CDA) market structure. Continuous-Double auctions (or some variants thereof) are common *market mechanisms*, i.e. procedures to match buy and sell orders. CDA are notably used on large stock markets (see for instance [17]). Our model features a particular version of CDA through which traders exchange securities (more in section **??**).

A social network. Traders are part of a social network. Every time a security is realized, each trader i looks at the performance of their neighbors in the network. If one of i's neighbors, say j, is richer than i, i interprets this has an indication that j has a better approximate model. Then i considers adopting j's approximate model.[3] For each trader, the willingness to revise her belief is determined by how ideologically loaded her belief is, which is characterized by parameter $\texttt{ideo}_i$.

An example of such social network is depicted in Figure 4. We assume that the initial network is segregated, that is traders who believe in (1) (resp. (2)) are more likely to be linked with other traders who believe in (1) (resp. (2)). The degree of segregation is determined by parameter `seg` (more details below). Although traders can change their approximate model as time passes, the *connections* between traders do not change as the market unfolds (i.e. the edges are fixed).

0.1.2 Timing

The time periods t are grouped in trading *sequences*. In a given sequence, the potential payments associated with traded securities are all based on the temperature at the end of the sequence. For instance, the third trading sequence might start in period $t = 1964$ and end in period $t = 1970$. In this case, a security traded in the third sequence pay 1 ECU if the temperature at $t = 1970$ falls into the range of temperatures covered by the security.

At each time t, traders are assumed to know the past value of the temperature $\hat{T}_{0:t}$ and greenhouse gas emissions $GHG_{0:t}$. In a sequence finishing at time t^*, traders also have common knowledge of $GHG_{t:t^*}$, the future values of greenhouse gas emissions up to t^*. However, at any t, traders do not know the value of any future *temperatures*. In particular, in a given sequence, traders do not know the value of $\hat{T}_{t^*}$. Traders can only predict $\hat{T}_{t^*}$ using their approximate model and their knowledge of $\hat{T}_{0:t}$ and $GHG_{0:t^*}$. Notice that because $GHG_{0:t^*}$

[3] Traders start with the same initial amount of money, so differences in money among traders can only come from traders' interactions through the market.

is common knowledge, in each period t every trader i with the same approximate model forms the same *pseudo*-expectation (see footnote 4)

$$\mu_{ti} \equiv E_t(\hat{T}_{t^*} \mid \hat{T}_{0:t}, GHG_{0:t^*}, i's \text{ approximate model of at } t). \qquad (3)$$

We assume that the believed probability distribution of $\hat{T}_{t^*}$ for trader i at any time t is

$$\mu_{ti} + \eta_t,$$

where the distribution η_t is obtained by trader i by bootstrapping the residuals from the regression of her believed true model over the past periods.[4]

At each time t, traders

- recalibrate their approximate model based on the new set of past data available at t (via pooled ordinary least squares),
- compute their belief-density for $\hat{T}_{t^*}$ and use it to determine the expected value they attached to each security,
- and trade on the CDA market based on their zero-intelligence decision rule as follows.

  1. Every trader i choses at random a security s_i^B she will try to buy.
  2. Every trader i also chooses at random among the securities she owns a positive amount of (if any) a security s_i^S she will try to sell.
  3. Traders then decide of their selling p_i^S and buying price p_i^S. To do so, they compute the expected value of the securities s_i^B and s_i^S according to (3). Then they set a selling and a buying price following the zero-intelligence rule described above.
  4. Traders go to the market one at the time, in a random order.
  5. When i comes to the market, she place *limit orders* in the order book. These orders specify that i is willing to buy s_i^B at any price below p_i^B, and to sell s_i^S at any price above p_i^S.
  6. The market maker then tries to match i's orders with some order which was put in the order book *before* i came to the market.
  7. If there are outstanding sell offers for s_i^B at price p lower than p_i^B, then a trade is concluded. Trader i buys one unit from the sellers who sells at the *highest* price below p_i^B, and the sell and buy offers are removed from the order book.
  8. If there are outstanding buy offers for s_i^S at price p higher than p_i^S, then a trade is concluded. Trader i sells one unit to the buyer who buys at the *highest* price above p_i^S, and the sell and buy offers are removed from the order book.

[4] This explains the use of the term *pseudo*-expectation. Because η_t is not bound to have expected value zero, a traders' effective expected value of $\hat{T}_{t^*}$ may be different from μ_{ti}.

9. Whenever all traders have come to the market, any remaining outstanding offer is removed from the order book, and the trading period is concluded.

At t^*, when the sequence ends, there is only one security s^* associated with a range of temperatures including the actual $\hat{T}_{t^*}$. Thus,

- traders receive 1 ECU per unit of s^* they own, and
- consider adapting their neighbors' approximate model as described above.

0.1.3 The models' parameters and details about their effects

The model depends on the following parameters.

Network parameters.

`n.traders`**:** the number of traders.

`n.edg` **:** the number of edges in the social network. This number is fixed throughout the experiments.

`seg` **:** determines the initial degree of homophily in the network. The higher `seg`, the higher the initial homophily.

When constructing the `n.edg` edges of the network, the probability that a link between two traders be formed depends on whether the traders share the same approximate model in the following way

$$\begin{cases} \frac{(1-\mathtt{seg})}{\mathtt{n.edg}}, & \text{if the traders have different approximate models} \\ \frac{1}{\mathtt{n.edg}}, & \text{otherwise} \end{cases}$$

Market structure parameter.

`market.complet`**.** Determines market's completeness, i.e. the number of securities which can be traded. With more securities, the interval of temperatures corresponding to each security is smaller. Traders can then trade on more precise temperature intervals. Higher values of `market.complet` may, however, reduce the number of exchanges. Because traders pick the securities they buy and sell at random, the probability that a match between sellers and buyers is found is lower for higher values of `market.complet`.

Behavioral parameters.

`risk.tak`**.** Determines the distribution of risk taking behavior. The higher `risk.tak`, the more traders will try to buy (resp. sell) lower (resp. higher) than their reservation price. Formally, at time t, trader i picks her buying (resp. selling) prices randomly in the interval $[\mathtt{reserv}_{it}, \mathtt{reserv}_{it} * (1 - \mathtt{risk.tak}_i)]$ (resp. $[\mathtt{reserv}_{it}, \mathtt{reserv}_{it} *$

$(1+\texttt{risk.tak}_i)]])$, where $\texttt{reserv}_{it}$ is i's reservation price at time t for the security i picked to buy (resp. sell).

For each trader i, the level of $\texttt{risk.tak}_i$ is drawn uniformly at random from [0,`risk.tak`]

$\texttt{ideo}_i$. Determines the degree of "ideology" of traders. If `ideo` is high, traders will not revise their approximate models easily, even when faced with strong evidence that their neighbors are doing better than them.

For each trader i and each sequence, a parameter d_i is drawn from $[0, \texttt{ideo}_i]$. The value of d_i is the probability that i adopts one of her neighbors' approximate model if this neighbor is doing better than i at the end of the sequence (in monetary terms).

For each trader i, the level of $\texttt{ideo}_i$ is drawn uniformly at random from [0,`ideo`]

Timing parameters.

`burn.in`. The number of burn-in periods in which no securities are traded (necessary to allow believed models to be estimated in the first period).

`n.seq`. Number of trading sequences.

`horizon`. Number of trading periods in each trading sequence.

Acknowledgements

We gratefully acknowledge the authors of R [14]. This manuscript was prepared using knitr [20]. We would like to thank Yevgeniy Vorobeychik for feedback on this project.

References

[1] Brian Beachkofski and Ramana Grandhi. Improved Distributed Hypercube Sampling. In *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, 2002.

[2] Rob Carnell. lhs Latin Hypercube Samples, July 2012.

[3] Dhananjay K. Gode and Shyam Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, February 1993.

[4] Sharad Goel, Daniel M. Reeves, Duncan J. Watts, and David M. Pennock. Prediction Without Markets. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, EC '10, pages 357–366, New York, NY, USA, 2010. ACM.

[5] Robin Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *The Journal of Prediction Markets*, pages 3–15, 2012.

[6] Robin Hanson, Ryan Oprea, and David Porter. Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior & Organization*, 60(4):449–459, August 2006.

[7] Paul J. Healy, Sera Linardi, J. Richard Lowery, and John O. Ledyard. Prediction markets: Alternative mechanisms for complex environments with few traders. *Management Science*, 56(11):1977–1996, November 2010.

[8] Christian Franz Horn, Bjoern Sven Ivens, Michael Ohneberg, and Alexander Brem. Prediction markets – a literature review 2014. *The Journal of Prediction Markets*, 8(2):89–126, September 2014.

[9] Shi-Ling Hsu. Prediction market for climate outcomes, a. *University of Colorado Law Review*, 83:179, 2011.

[10] Janyl Jumadinova and Prithviraj Dasgupta. A multi-agent system for analyzing the effect of information on prediction markets. *International Journal of Intelligent Systems*, 26(5):383–409, 2011.

[11] Frank MA Klingert and Matthias Meyer. Comparing prediction market mechanisms using an experiment-based multi-agent simulation. In *ECMS*, pages 654–661, 2012.

[12] Simeone Marino, Ian B. Hogue, Christian J. Ray, and Denise E. Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, 254(1):178–196, September 2008.

[13] Gilles Pujol, Bertrand Iooss, Alexandre Janon with contributions from Paul Lemaitre, Laurent Gilquin, Loic Le Gratiet, Taieb Touati, Bernardo Ramos, and Jana Fruth and Sebastien Da Veiga. sensitivity: Sensitivity Analysis, August 2014.

[14] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.

[15] A. Saltelli, K. Chan, and E. M. Scott. *Sensitivity Analysis*. Wiley, Chichester, 1 edition edition, March 2009.

[16] Reinhard Set and Reinhard Selten. Axiomatic characterization of the quadratic scoring rule. *Experimental Economics*, 1(1):43–62, June 1998.

[17] Jie-Jun Tseng, Chih-Hao Lin, Chih-Ting Lin, Sun-Chong Wang, and Sai-Ping Li. Statistical properties of agent-based models in markets with continuous double auction mechanism. *Physica A: Statistical Mechanics and its Applications*, 389(8):1699–1707, April 2010.

[18] Michael P. Vandenbergh, Kaitlin E. Toner, and Jonathan M. Gilligan. Energy and climate change: A climate prediction market. SSRN Scholarly Paper ID 2372321, Social Science Research Network, Rochester, NY, December 2013.

[19] Justin Wolfers and Eric Zitzewitz. Prediction markets in theory and practice. 2006.

[20] Yihui Xie. *Dynamic Documents with R and knitr*. Chapman & Hall/CRC, Boca Raton, 2014.

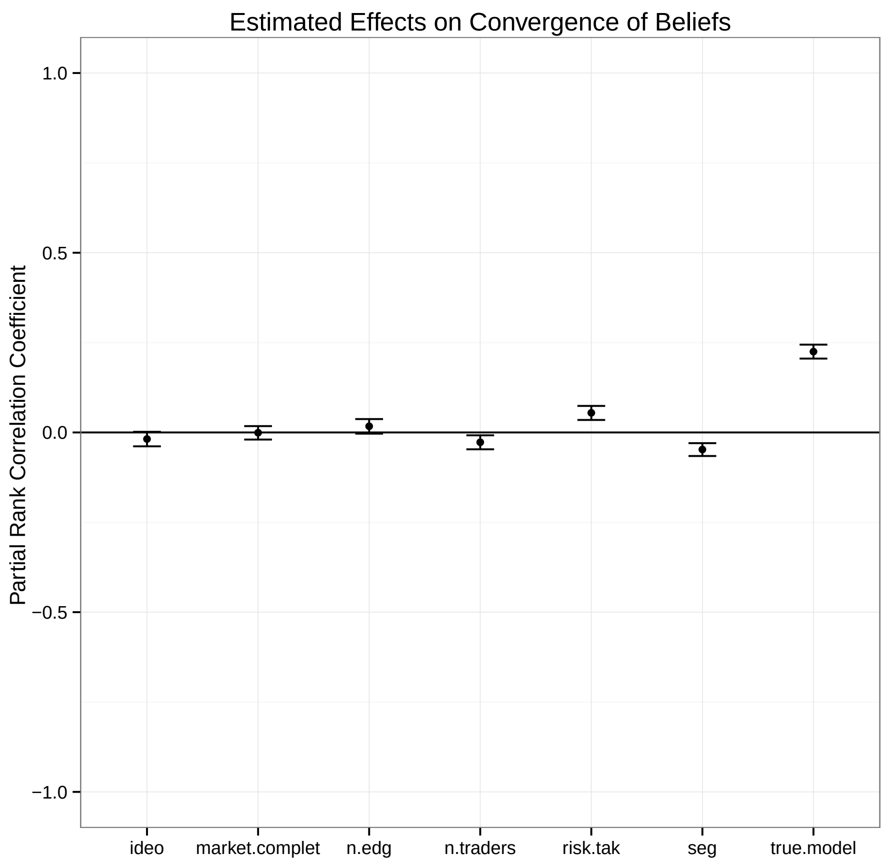

Figure 1: Partial rank correlation coefficient analysis based on 10,000 simulated parameter sets [13, 15] of the effects of ideology, market completeness, number of edges in social network, number of traders in market, risk taking propensities, segregation measure of the social network, and the true climate model on the convergence of traders' beliefs. Lines are bootstrapped 95% confidence intervals.

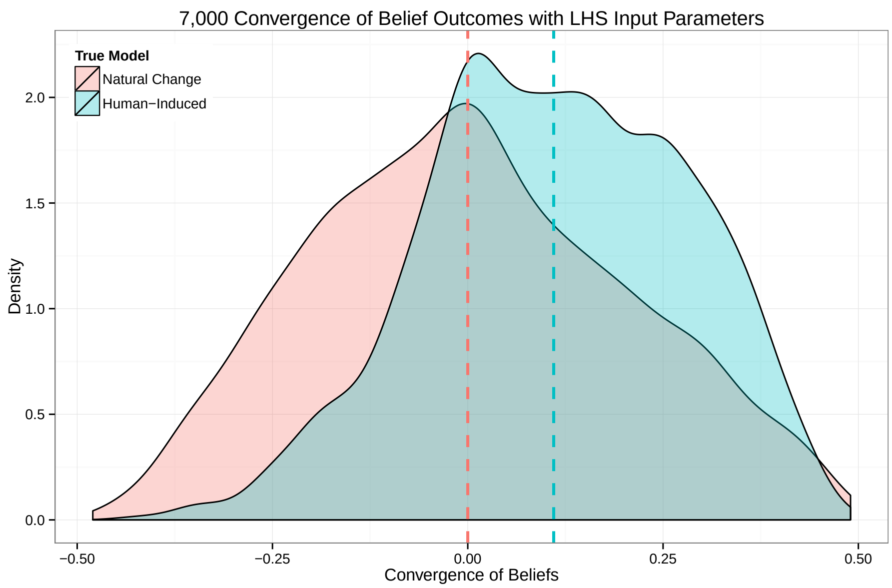

Figure 2: Density of 7,000 convergence of belief model outcomes with 7,000 Latin Hypercube Sampled input parameter sets. 3,500 of the Latin Hypercube Sampled input parameter sets have "human-induced" climate change as the true model and 3,500 have "natural change" climate change as the true model. Dashed lines are the medians of the two densities (0, 0.11). Possible values of convergence of beliefs range from -1 to 1.

```
## Error in loadNamespace(name): there is no package called
'gtools'
```

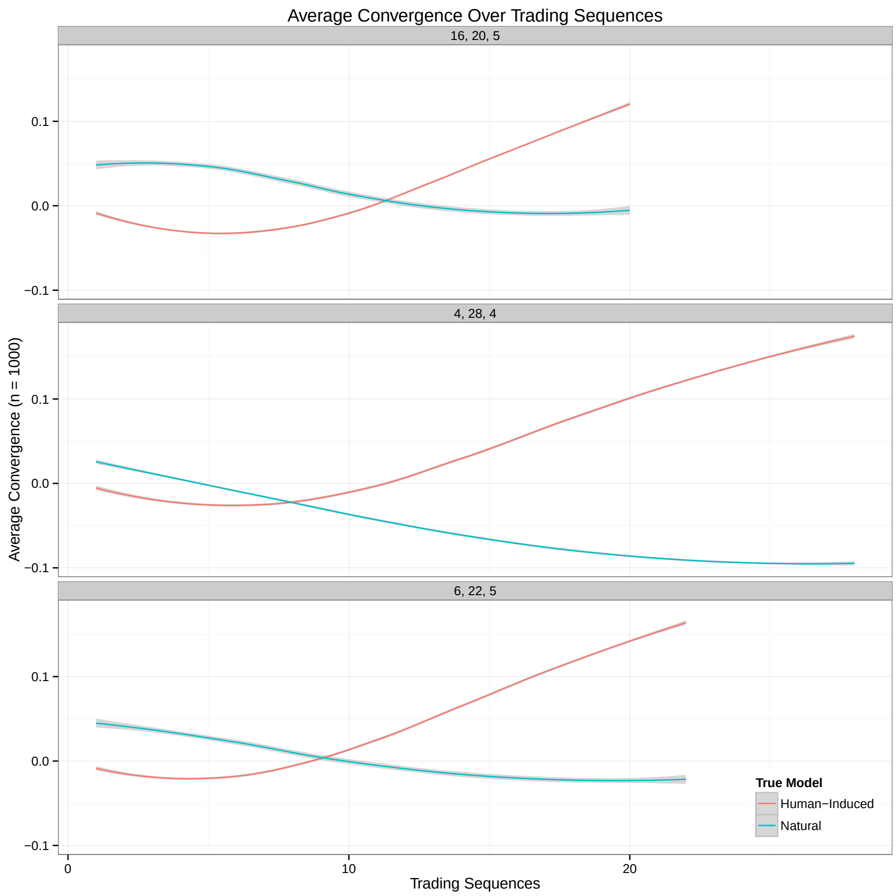

Figure 3: Temporal evolution of convergence of beliefs for the ABM with 1,000 model runs with "True Model" parameter set to natural climate change and 1,000 model runs with "True Model" parameter set to human-induced climate change. All other parameters besides burn.in, n.seq, and horizon are sampled from their full support with Latin Hypercube Sampling; therefore, the plot represents, in effect, a marginalizing out of the effects of all other parameters to show the moderating effect that the true model has on the effect of the prediction market on the convergence of traders' beliefs about what the true model is. The prediction market causes opposite convergence effects for the two true climate models. The top figure used the timing settings: burn.in = 4, n.seq = 28, horizon = 4. Middle figure: burn.in = 6, n.seq = 22, horizon = 5. Bottom figure: burn.in = 16, n.seq = 20, horizon = 5.

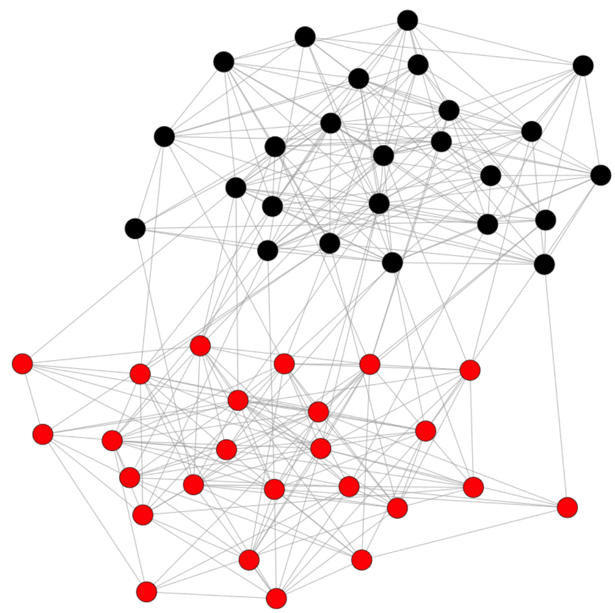

Figure 4: An example of the initial state of a belief network with 50 traders. Red nodes are traders who initially believe in (1) and black nodes are traders who initially believe in (2). Notice that the