An Axiomatic View of the Parimutuel Consensus Wagering Mechanism *

Rupert Freeman
Duke University
rupert@cs.duke.edu

David M. Pennock
Microsoft Research
dpennock@microsoft.com

Abstract

We consider an axiomatic view of the *Parimutuel Consensus Mechanism* defined by Eisenberg and Gale [1959]. The parimutuel consensus mechanism can be interpreted as a parimutuel market for wagering with a proxy that bets optimally on behalf of the agents, depending on the bets of the other agents. We show that the parimutuel consensus mechanism uniquely satisfies the desirable properties of Pareto optimality, individual rationality, budget balance, anonymity, sybilproofness and envy-freeness. While the parimutuel consensus mechanism does violate the key property of incentive compatibility, it is incentive compatible in the limit as the number of agents becomes large. Via simulations on real contest data, we show that violations of incentive compatibility are both rare and only minimally beneficial for the participants. This suggests that the parimutuel consensus mechanism is a reasonable mechanism for eliciting information in practice.

1 Introduction

In 1867, Spanish entrepreneur Joseph Oller invented *parimutuel betting*, a form of wagering still popular today, handling billions of dollars annually on horse races and jai alai games. Each bettor places money on one of several future outcomes—say, horse #1 to win a race. She is allowed to cancel her bet or move her money to a different outcome at any time, even at the last second before the race begins. After the outcome resolves—say, horse #1 wins—agents who picked the wrong outcome lose their wagers to the agents who picked correctly. Winning agents split the pot in proportion to the size of their wagers.

Eisenberg and Gale [1959] analyzed the equilibrium of parimutuel betting, defining the *parimutuel consensus mechanism (PCM)*. The PCM is equivalent to parimutuel betting where each agent has a proxy. Each agent's proxy knows her true probabilities for all outcomes. As bets come in, and the prospective payoff per dollar, or *odds*, for each outcome converge, the proxy automatically switches its agent's money to the outcome yielding the highest expected payoff for that agent. In equilibrium, all the proxies are optimizing and none want to switch outcomes.

*A preliminary version of this paper appears in IJCAI'18.

At any point in time, the odds can be interpreted as probabilities, providing a prediction of the outcome of the event. Thus, facilitating wagering can serve as a source of free information for a principal seeking to forecast a future event.

Eisenberg and Gale discuss one undesirable feature of the equilibrium: it produces odds that sometimes ignore some agents. Manski [2006] further explores in detail how the equilibrium of risk-neutral, budget-constrained agents may fail to aggregate beliefs in a sensible way. Additionally, the PCM is not *incentive compatible*, or truthful. An agent may strategically improve her payoff by taking into account what other agents know or what they may do. In the end, her best action may be to report false probabilities to her proxy that differ from her true subjective probabilities. For a principal whose primary goal is information elicitation, this is problematic because some of the reported probabilities may not faithfully reflect the bettors' private information.

Given the potential for bad equilibria and the lack of incentive compatibility, why is the PCM still prevalent? One answer is that, in practice, it often works fine. Parimutuel betting does consistently induce a wisdom-of-crowds effect, producing odds that encode well calibrated and accurate probabilistic forecasts of the outcomes [Ali, 1977, Snyder, 1978, Thaler and Ziemba, 1988], like many prediction markets do [Arrow et al., 2008]. Plott et al. [2003] tested parimutuel betting in a laboratory experiment, showing that the mechanism is an effective vehicle for information aggregation regardless of why it might go wrong in theory. If agents have concave or risk-averse utility for money, the equilibrium of similar mechanisms is stable and induces sensible belief aggregation [Wolfers and Zitzewitz, 2006, Beygelzimer et al., 2012]. In particular, an agent with logarithmic utility does best by betting an amount on each outcome proportional to her probability [Cover and Thomas, 2006].

We examine another plausible reason why the PCM continues to enjoy usage: the mechanism satisfies a number of desirable axioms for wagering systems. We prove that the PCM is the unique wagering mechanism that is Pareto optimal, individually rational, budget balanced, sybilproof, anonymous, and envy-free, subject to a mild condition on the reports.

Freeman et al. [2017] show that, to gain the key property of incentive compatibility, one of the core properties must be relaxed. Yet we can show that the PCM is near incentive compatible in some cases. Yes, there are scenarios where agents can gain from lying, but we prove that the PCM is incentive compatible *in the large*, as the number of agents grows. In extensive simulations using real forecasts from an online contest, we show that opportunities for agents to profit from untruthful play are rare, mostly vanishing as the number of agents grows. Our results shed light on the practical success of the PCM. Despite its flaws, identified as early as 1959, it does satisfy six natural and desirable properties of wagering mechanisms and it comes close both theoretically and empirically to obtaining a crucial seventh: incentive compatibility.

2 Preliminaries

Consider a random variable (*event*) X which takes a value (*outcome*) in $\{0, 1\}$. There is a set of agents (or bettors) $\mathcal{N} = \mathbb{N}$,[1] each with a private, subjective, immutable belief p_i which is the probability that $X = 1$, and a budget w_i, which is the maximum amount that they are prepared to lose.

A *wagering mechanism* is used to elicit beliefs from the agents. Each agent submits a report $\hat{p}_i \in [0, 1]$ and wager $w_i \in \mathbb{Q}_{\geq 0}$ to the mechanism, where $\hat{p}_i$ captures her belief and w_i captures her budget.[2] We denote the vector of agents' reports by $\hat{\mathbf{p}}$, and the vector of agents' wagers by $\mathbf{w}$. The wagering mechanism defines a payoff $\Pi_i(\hat{\mathbf{p}}; \mathbf{w}; x)$ to each agent that depends on the reports and wagers of the agents and x, the observed value of X. To be a valid wagering mechanism, it must be the case that no agent loses more than her wager (i.e., $\Pi_i(\hat{\mathbf{p}}; \mathbf{w}; x) \geq -w_i$) and that an agent with zero wager does not participate in the mechanism (i.e., $\Pi_i(\hat{\mathbf{p}}; \mathbf{w}; x) = 0$ if $w_i = 0$). We denote the reports and wagers of all agents other than i by $\hat{\mathbf{p}}_{-\mathbf{i}}$ and $\mathbf{w}_{-\mathbf{i}}$ respectively.

2.1 Security Interpretation of Wagering Mechanisms

Freeman et al. [2017] observed that the output of a wagering mechanism can be interpreted as an allocation of Arrow-Debreu securities with payoffs that depend on the realization of X. A `yes` security is a contract that pays off \$1 in the outcome $X = 1$ and \$0 if $X = 0$. Similarly, a `no` security pays off \$0 if $X = 1$ and \$1 if $X = 0$. A *risk neutral* agent with belief p about the likelihood that $X = 1$ would be willing to buy a `yes` security at any price up to p or a `no` security at any price up to $1 - p$.

Suppose a wagering mechanism would yield a net payoff to agent i of $\rho_1 = \Pi_i(\hat{\mathbf{p}}; \mathbf{w}; 1)$ when $X = 1$ and $\rho_0 = \Pi_i(\hat{\mathbf{p}}; \mathbf{w}; 0)$ when $X = 0$. This is equivalent to the payoff that i would receive if she were sold $y_i = \max\{\rho_1 - \rho_0, 0\}$ `yes` securities and $n_i = \max\{\rho_0 - \rho_1, 0\}$ `no` securities for a total cost of $\sigma_i = \max\{-\rho_0, -\rho_1\}$. For example, if $\rho_0 < \rho_1$, then agent i's participation in the wagering mechanism is equivalent to agent i paying the principal $\sigma_i = -\rho_0$ before X is realized and then receiving $y_i = \rho_1 - \rho_0$ from the principal in the outcome $X = 1$.

Therefore, the output of a wagering mechanism can be completely specified by a triple $(\mathbf{y}, \mathbf{n}, \boldsymbol{\sigma})$, where for each agent i, y_i (resp. n_i) is the number of `yes` (`no`) securities allocated to i, and σ_i is the cost paid by i for these securities. The requirement that no agent can lose more than her wager is captured by the constraint $\sigma_i \leq w_i$. Without loss of generality, we assume that for all i, either $y_i = 0$ or $n_i = 0$ (or both), since any (fraction of a) pair of `yes` and `no` securities can be precisely converted into (a fraction of) \$1.

2.2 Properties of Wagering Mechanisms

Lambert et al. [2008] introduced several desirable properties for wagering mechanisms.

[1]Following Lambert et al. [2008], this is without loss of generality since non-participation can be seen as making a zero wager.

[2]The assumption of rational wagers is required in Section 4. Rational wagers can approximate real-valued budgets arbitratily well.

We focus on five of these properties in our analysis: anonymity, individual rationality, incentive compatibility, budget balance, and sybilproofness.

First, *anonymity* says that the payouts do not depend on the identities of the agents. This is a basic property that all wagering mechanisms proposed in the literature, including the PCM, satisfy.

Individual rationality says that agents do not lose money (in expectation) by participating truthfully. We note that this is slightly stronger than the usual definition, which only requires that there exists *some* (not necessarily truthful) report for which agent i does not make an expected loss. Our definition is implied by the combination of the usual definition of individual rationality, and incentive compatibility. We require the stronger version in the proof of Theorem 2.

Definition 1. *A wagering mechanism is* individually rational *if, for any agent* i *and any belief* p_i*, for all* $\hat{\mathbf{p}}_{-\mathbf{i}}, \mathbf{w}$ *her expected payoff is greater than or equal to her cost; that is,*

$$p_i y_i(p_i, \hat{\mathbf{p}}_{-i}; \mathbf{w}) + (1 - p_i) n_i(p_i, \hat{\mathbf{p}}_{-i}; \mathbf{w}) \geq \sigma_i(p_i, \hat{\mathbf{p}}_{-i}; \mathbf{w})$$

Incentive compatibility requires that each agent maximizes her expected payoff by reporting truthfully, regardless of the reports and wagers of other agents.

Definition 2. *A wagering mechanism is* (weakly) incentive compatible *if, for every agent* i *with belief* p_i *and all reports* $\hat{\mathbf{p}}$ *and wagers* $\mathbf{w}$*,*

$$\begin{aligned} p_i y_i(p_i, \hat{\mathbf{p}}_{-i}; \mathbf{w}) + (1 - p_i) n_i(p_i, \hat{\mathbf{p}}_{-i}; \mathbf{w}) - \sigma_i(p_i, \hat{\mathbf{p}}_{-i}; \mathbf{w}) \\ \geq p_i y_i(\hat{\mathbf{p}}; \mathbf{w}) + (1 - p_i) n_i(\hat{\mathbf{p}}; \mathbf{w}) - \sigma_i(\hat{\mathbf{p}}; \mathbf{w}). \end{aligned}$$

A wagering mechanism is budget balanced if the principal never makes a profit or a loss.

Definition 3. *A wagering mechanism is* budget balanced *if, for all* $\hat{\mathbf{p}}$ *and* $\mathbf{w}$*,*

$$\sum_{i \in \mathcal{N}} y_i(\hat{\mathbf{p}}; \mathbf{w}) = \sum_{i \in \mathcal{N}} n_i(\hat{\mathbf{p}}; \mathbf{w}) = \sum_{i \in \mathcal{N}} \sigma_i(\hat{\mathbf{p}}; \mathbf{w})$$

If the mechanism sometimes makes a profit, but never a loss, then we say that it is *weakly budget balanced*.

A mechanism is sybilproof if it is not beneficial for agents to participate under multiple fake identities, or for agents reporting the same probability to merge.

Definition 4. *A wagering mechanism is* sybilproof *if for any subset of players* S*, for any* $\hat{\mathbf{p}}$ *with* $\hat{p}_i = \hat{p}_j$ *for* $i, j \in S$*, for any vectors of wagers* $\mathbf{w}, \mathbf{w}'$ *with* $w_i = w'_i$ *for* $i \notin S$ *and* $\sum_{i \in S} w_i = \sum_{i \in S} w'_i$*, it is the case that:*

$$\sum_{i \in S} (y_i(\hat{\mathbf{p}}; \mathbf{w}), n_i(\hat{\mathbf{p}}; \mathbf{w}), \sigma_i(\hat{\mathbf{p}}; \mathbf{w})) = \sum_{i \in S} (y_i(\hat{\mathbf{p}}; \mathbf{w}'), n_i(\hat{\mathbf{p}}; \mathbf{w}'), \sigma_i(\hat{\mathbf{p}}; \mathbf{w}'))$$

and for all $i \notin S$*,*

$$(y_i(\hat{\mathbf{p}}; \mathbf{w}), n_i(\hat{\mathbf{p}}; \mathbf{w}), \sigma_i(\hat{\mathbf{p}}; \mathbf{w})) = (y_i(\hat{\mathbf{p}}; \mathbf{w}'), n_i(\hat{\mathbf{p}}; \mathbf{w}'), \sigma_i(\hat{\mathbf{p}}; \mathbf{w}')).$$

Freeman et al. [2017] introduce the notion of *Pareto optimality* for wagering mechanisms, which is satisfied if there is no side bet that can be made by the agents on top of that already facilitated by the wagering mechanism, such that some agent benefits without harming others. Freeman et al. show that the following condition is equivalent.

Definition 5. *A wagering mechanism is Pareto optimal if, for all reports* $\hat{\mathbf{p}}$ *and* $\mathbf{w}$*, there exists a* $p \in [0, 1]$ *such that*

$$\forall i : \hat{p}_i < p, \qquad \sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i \quad \textit{and} \quad y_i(\hat{\mathbf{p}}; \mathbf{w}) = 0,$$
$$\forall i : \hat{p}_i > p, \qquad \sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i \quad \textit{and} \quad n_i(\hat{\mathbf{p}}; \mathbf{w}) = 0.$$

Freeman et al. [2017] show that four key properties are incompatible.

Theorem 1 (Freeman et al. [2017])**.** *No wagering mechanism simultaneously satisfies individual rationality, weak incentive compatibility, weak budget balance, and Pareto optimality.*

Lastly, we consider the property of envy-freeness [Foley, 1967]. Envy-freeness is a basic fairness property which says that no player should envy the allocation of securities to any other agent.

Definition 6. *Say that agent* i envies *another agent* j *if* $\sigma_j(\hat{\mathbf{p}}, \mathbf{w}) \leq w_i$ *and*

$$\hat{p}_i y_i(\hat{\mathbf{p}}; \mathbf{w}) + (1 - \hat{p}_i) n_i(\hat{\mathbf{p}}; \mathbf{w}) - \sigma_i(\hat{\mathbf{p}}; \mathbf{w}) < \hat{p}_i y_j(\hat{\mathbf{p}}; \mathbf{w}) + (1 - \hat{p}_i) n_j(\hat{\mathbf{p}}; \mathbf{w}) - \sigma_j(\hat{\mathbf{p}}; \mathbf{w})$$

A wagering mechanism is envy-free if there is no pair of agents (i, j) *such that* i *envies* j*.*

3 The Parimutuel Consensus Mechanism

The Parimutuel Consensus Mechanism (PCM) can be thought of as a direct implementation of the equilibrium of parimutuel betting. The PCM includes the rules of parimutuel betting plus, conceptually, a proxy agent that automatically switches its agent's bet to the outcome with highest expected profit per security. The output of the mechanism is the equilibrium where all proxies are stable. For the binary case of `yes` and `no` outcomes that we consider, the PCM is defined by a price π such that an agent with report less than π is allocated `no` securities at a price of $1 - \pi$ per security, and an agent with report more than π is allocated `yes` securities at a price of π per security. The equilibrium condition is

$$\sum_{i:\hat{p}_i < \pi} \frac{w_i}{1 - \pi} + \sum_{i:\hat{p}_i = \pi} \frac{c_1 w_i}{1 - \pi} = \sum_{i:\hat{p}_i > \pi} \frac{w_i}{\pi} + \sum_{i:\hat{p}_i = \pi} \frac{c_2 w_i}{\pi}, \tag{1}$$

where c_1 and c_2 lie in the interval $[0, 1]$ and $\min\{c_1, c_2\} = 0$. These represent the fractions of budgets spent by agents with $\hat{p}_i = \pi$ who bet (some of) their wager to correctly balance the market prices and allow the market to reach equilibrium, even though they get zero expected profit. At most one of c_1 and c_2 is greater than 0, since it would be redundant to have agents with $\hat{p}_i = \pi$ betting on both `yes` and `no`. Note that the left hand side of Equation 1 is the total number of `no` securities allocated, and the right hand side is the total number of `yes` securities allocated. Eisenberg and Gale [1959] show as their main contribution that such a price is both unique and guaranteed to exist. The output of the PCM is defined by

$$y_i(\hat{\mathbf{p}}; \mathbf{w}) = \begin{cases} 0 & \hat{p}_i < \pi \\ \frac{c_2 w_i}{\pi} & \hat{p}_i = \pi \\ \frac{w_i}{\pi} & \hat{p}_i > \pi \end{cases}, \qquad n_i(\hat{\mathbf{p}}; \mathbf{w}) = \begin{cases} \frac{w_i}{1-\pi} & \hat{p}_i < \pi \\ \frac{c_1 w_i}{1-\pi} & \hat{p}_i = \pi \\ 0 & \hat{p}_i > \pi \end{cases}$$

and

$$\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = \begin{cases} w_i & \hat{p}_i < \pi \\ \max\{c_1, c_2\} w_i & \hat{p}_i = \pi \\ w_i & \hat{p}_i > \pi \end{cases}$$

Example 1. *Suppose that there are four agents, with reports $\hat{\boldsymbol{p}} = (0.3, 0.5, 0.6, 0.8)$ and wagers $\boldsymbol{w} = (1, 1, 3, 6)$. Observe that setting $\pi = 0.6$ and $c_1 = \frac{2}{3}, c_2 = 0$ satisfies Equation 1: each side of the equation has value 10. Further, setting $\pi < 0.6$ results in the right hand side of Equation 1 being greater than the left hand side, for any allowed values of c_1 and c_2, and the opposite is true for any $\pi > 0.6$.*

We can now compute the output of the PCM on this instance, according to the formulae above. Agents 1 and 2 are allocated 2.5 `no` securities each, for a price of 1, agent 3 is allocated 5 `no` securities for a price of 2 (note that this is a $c_1 = \frac{2}{3}$ fraction of agent 3's wager), and agent 4 is allocated 10 `yes` securities for a price of 6.

Recall that, by Theorem 1, no wagering mechanism can simultaneously satisfy individual rationality, incentive compatibility, weak budget balance, and Pareto optimality. Theoretical papers on wagering mechanisms are generally reluctant to give up any of the first three properties, sacrificing Pareto optimality [Chen et al., 2014, Freeman et al., 2017, Lambert et al., 2008, 2015]. However, in practice, Pareto optimality is an important consideration and virtually all real-world wagering mechanisms, including parimutuels, bookmakers, and double auctions, satisfy it. This is because trade drives participation; a mechanism that facilitates little trade is of little use or interest to agents.

Individual rationality seems hard to give up. We cannot force agents to play a game that they expect to lose and, even if we did, they could just wager $w_i = 0$. The center may be willing to pay for the information inherent in the agents' beliefs, subsidizing the mechanism and relaxing budget balance. Market scoring rules [Chen and Pennock, 2007, Hanson, 2003], for example, do just that, losing a strictly bounded amount of money in service of gaining information. However, a patron will only subsidize events that bear on valuable decisions. Nearly all fielded wagering mechanisms have taxes, not subsidies, yielding profits, not losses.

If we want Pareto optimality, individual rationality, and (weak) budget balance, we are forced to give up on incentive compatibility. That's exactly what the PCM does. In the remainder of this paper, we show that the PCM is the unique wagering mechanism that simultaneously satisfies the other six properties from Section 2, subject to a condition on the reports. We then show that, despite not satisfying incentive compatibility, the PCM is approximately incentive compatibile in two senses. First, we prove that, as the number of agents grows, the mechanism is incentive compatible *in the large*. Second, we show empirically that, across thousands of simulated wagers based on real probability estimates, opportunities to profit from misreports are almost negligible.

4 Properties of the Parimutuel Consensus Mechanism

Despite its theoretical flaws, including the possibility of nonsensical information aggregation, the PCM seems well behaved in practice. In this section, we examine one possible reason for this by providing a theoretical justification for the PCM. We first note that the PCM satisfies six desirable properties for wagering mechanisms. Although incentive compatibility is not one of the six, we know that adding it is impossible without giving something up: no mechanism satisfying even just the first three properties can also be incentive compatible.

Proposition 1. *The parimutuel consensus mechanism satisfies individual rationality, budget balance, Pareto optimality, anonymity, sybilproofness, and envy-freeness.*

That the PCM satisfies the first three properties is noted by Freeman et al. [2017].

Proof. For this proof, we assume that $c_2 = 0$ in the equilibrium condition given by Equation 1. The proof for the case where $c_1 = 0$ follows via symmetric arguments for all properties.

1. Anonymity Anonymity clearly holds because Equation 1 and the allocation of securities do not depend on the identities of the agents.

2. Individual rationality: Consider some agent i. If $p_i < \pi$,

$$p_i y_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w}) + (1 - p_i) n_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w}) = (1 - p_i)\frac{w_i}{1 - \pi} > w_i = \sigma_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w})$$

If $p_i > \pi$,

$$p_i y_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w}) + (1 - p_i) n_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w}) = p_i \frac{w_i}{\pi} > w_i = \sigma_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w})$$

Finally, if $p_i = \pi$,

$$p_i y_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w}) + (1 - p_i) n_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w}) = (1 - p_i) c_1 \frac{w_i}{1 - \pi} = c_1 w_i = \sigma_i(p_i, \hat{\mathbf{p}}_{-\mathbf{i}}; \mathbf{w})$$

3. Budget balance: First, note that

$$\sum_{i \in \mathcal{N}} y_i(\hat{\mathbf{p}}; \mathbf{w}) = \sum_{\hat{p}_i > \pi} \frac{w_i}{\pi} = \sum_{\hat{p}_i < \pi} \frac{w_i}{1 - \pi} + c_1 \sum_{\hat{p}_i = \pi} \frac{w_1}{1 - \pi} = \sum_{i \in \mathcal{N}} n_i(\hat{\mathbf{p}}; \mathbf{w}),$$

where the second transition follows from the definition of π. Next,

$$
\begin{aligned}
\sum_{i\in\mathcal{N}} \sigma_i(\hat{\mathbf{p}};\mathbf{w}) &= \sum_{\hat{p}_i>\pi} w_i + \sum_{\hat{p}_i<\pi} w_i + \sum_{\hat{p}_i=\pi} c_1 w_i \\
&= \pi \sum_{\hat{p}_i>\pi} \frac{w_i}{\pi} + (1-\pi)\left(\sum_{\hat{p}_i<\pi} \frac{w_i}{1-\pi} + c_1 \sum_{\hat{p}_i=\pi} \frac{w_1}{1-\pi}\right) \\
&= \pi \sum_{\hat{p}_i>\pi} \frac{w_i}{\pi} + (1-\pi) \sum_{\hat{p}_i>\pi} \frac{w_i}{\pi} \\
&= \sum_{\hat{p}_i>\pi} \frac{w_i}{\pi} = \sum_{i\in\mathcal{N}} y_i(\hat{\mathbf{p}};\mathbf{w}) = \sum_{i\in\mathcal{N}} n_i(\hat{\mathbf{p}};\mathbf{w}),
\end{aligned}
$$

Where the third transition is obtained via the definition of π (and noting that $c_2 = 0$, by assumption).

4. Pareto optimality We show that the price π satisfies the condition of the Pareto optimality definition. From the definition of the PCM, if $\hat{p}_i > \pi$ then $\sigma_i(\hat{\mathbf{p}};\mathbf{w}) = w_i$ and $n_i(\hat{\mathbf{p}};\mathbf{w}) = 0$, and if $\hat{p}_i < \pi$ then $\sigma_i(\hat{\mathbf{p}};\mathbf{w}) = w_i$ and $y_i(\hat{\mathbf{p}};\mathbf{w}) = 0$.

5. Sybilproofness: Consider a set of sybils S such that $\mathbf{w}$ and $\mathbf{w}'$ satisfy the conditions of Definition 4, with corresponding prices π and π' reached by the PCM. By the definition of sybils, the following three conditions hold:

$$
\sum_{i:\hat{p}_i<\pi} w_i = \sum_{i:\hat{p}_i<\pi} w'_i, \qquad \sum_{i:\hat{p}_i=\pi} w_i = \sum_{i:\hat{p}_i=\pi} w'_i, \qquad \sum_{i:\hat{p}_i>\pi} w_i = \sum_{i:\hat{p}_i>\pi} w'_i
$$

It follows immediately that

$$
\begin{aligned}
& \sum_{i:\hat{p}_i<\pi} \frac{w_i}{1-\pi} + c_1 \sum_{i:\hat{p}_i=\pi} \frac{w_i}{1-\pi} = \sum_{i:\hat{p}_i>\pi} \frac{w_i}{\pi} \\
\implies & \sum_{i:\hat{p}_i<\pi} \frac{w'_i}{1-\pi} + c_1 \sum_{i:\hat{p}_i=\pi} \frac{w'_i}{1-\pi} = \sum_{i:\hat{p}_i>\pi} \frac{w'_i}{\pi},
\end{aligned}
$$

so $\pi = \pi'$, with the same value of c_1 in both cases.

Suppose first that $i \notin S$. If $\hat{p}_i > \pi$ then

$$
\begin{aligned}
&(y_i(\hat{\mathbf{p}};\mathbf{w}), n_i(\hat{\mathbf{p}};\mathbf{w}), \sigma_i(\hat{\mathbf{p}};\mathbf{w})) = (\frac{w_i}{\pi}, 0, w_i) \\
&= (\frac{w'_i}{\pi'}, 0, w'_i) = (y_i(\hat{\mathbf{p}};\mathbf{w}'), n_i(\hat{\mathbf{p}};\mathbf{w}'), \sigma_i(\hat{\mathbf{p}};\mathbf{w}')).
\end{aligned}
$$

If $\hat{p}_i < \pi$ then

$$
\begin{aligned}
&(y_i(\hat{\mathbf{p}};\mathbf{w}), n_i(\hat{\mathbf{p}};\mathbf{w}), \sigma_i(\hat{\mathbf{p}};\mathbf{w})) = (0, \frac{w_i}{1-\pi}, w_i) \\
&= (0, \frac{w'_i}{1-\pi'}, w'_i) = (y_i(\hat{\mathbf{p}};\mathbf{w}'), n_i(\hat{\mathbf{p}};\mathbf{w}'), \sigma_i(\hat{\mathbf{p}};\mathbf{w}')).
\end{aligned}
$$

Finally, if $\hat{p}_i = \pi$ then

$$(y_i(\hat{\mathbf{p}};\mathbf{w}), n_i(\hat{\mathbf{p}};\mathbf{w}), \sigma_i(\hat{\mathbf{p}};\mathbf{w})) = (0, c_1\frac{w_i}{1-\pi}, c_1 w_i)$$
$$= (0, c_1\frac{w_i'}{1-\pi'}, c_1 w_i') = (y_i(\hat{\mathbf{p}};\mathbf{w}'), n_i(\hat{\mathbf{p}};\mathbf{w}'), \sigma_i(\hat{\mathbf{p}};\mathbf{w}')).$$

Next, suppose that $i \in S$. If $\hat{p}_i > \pi = \pi'$, then $\hat{p}_j = \hat{p}_i > \pi = \pi'$ for all $j \in S$. We have

$$\sum_{i\in S}(y_i(\hat{\mathbf{p}};\mathbf{w}), n_i(\hat{\mathbf{p}};\mathbf{w}), \sigma_i(\hat{\mathbf{p}};\mathbf{w})) = \left(\sum_{i\in S}\frac{w_i}{\pi}, 0, \sum_{i\in S} w_i\right)$$
$$= \left(\sum_{i\in S}\frac{w_i'}{\pi'}, 0, \sum_{i\in S} w_i'\right) = \sum_{i\in S}(y_i(\hat{\mathbf{p}};\mathbf{w}'), n_i(\hat{\mathbf{p}};\mathbf{w}'), \sigma_i(\hat{\mathbf{p}};\mathbf{w}')).$$

If $\hat{p}_i < \pi = \pi'$, then

$$\sum_{i\in S}(y_i(\hat{\mathbf{p}};\mathbf{w}), n_i(\hat{\mathbf{p}};\mathbf{w}), \sigma_i(\hat{\mathbf{p}};\mathbf{w})) = \left(0, \sum_{i\in S}\frac{w_i}{1-\pi}, \sum_{i\in S} w_i\right)$$
$$= \left(0, \sum_{i\in S}\frac{w_i'}{1-\pi'}, \sum_{i\in S} w_i'\right) = \sum_{i\in S}(y_i(\hat{\mathbf{p}};\mathbf{w}'), n_i(\hat{\mathbf{p}};\mathbf{w}'), \sigma_i(\hat{\mathbf{p}};\mathbf{w}')).$$

Finally, if $\hat{p}_i = \pi = \pi'$ then

$$\sum_{i\in S}(y_i(\hat{\mathbf{p}};\mathbf{w}), n_i(\hat{\mathbf{p}};\mathbf{w}), \sigma_i(\hat{\mathbf{p}};\mathbf{w})) = \left(0, \sum_{i\in S}\frac{c_1 w_i}{1-\pi}, \sum_{i\in S} c_1 w_i\right)$$
$$= \left(0, \sum_{i\in S}\frac{c_1 w_i'}{1-\pi'}, \sum_{i\in S} c_1 w_i'\right) = \sum_{i\in S}(y_i(\hat{\mathbf{p}};\mathbf{w}'), n_i(\hat{\mathbf{p}};\mathbf{w}'), \sigma_i(\hat{\mathbf{p}};\mathbf{w}')).$$

Therefore, the conditions for sybilproofness are satisfied.

6. Envy-freeness: Consider an agent i with $\hat{p}_i < \pi$. Let $j \neq i$. If $\sigma_j(\hat{\mathbf{p}};\mathbf{w}) > w_i$ then i does not envy j, so suppose that $\sigma_j(\hat{\mathbf{p}};\mathbf{w}) \leq w_i$.

Suppose that $\hat{p}_j > \pi$. Then

$$\begin{aligned}\hat{p}_i y_j(\hat{\mathbf{p}};\mathbf{w}) + (1-\hat{p}_i) n_j(\hat{\mathbf{p}};\mathbf{w}) - \sigma_j(\hat{\mathbf{p}};\mathbf{w}) &= \hat{p}_i\frac{w_j}{\pi} - w_j \\ &< 0 \\ &< \hat{p}_i y_i(\hat{\mathbf{p}};\mathbf{w}) + (1-\hat{p}_i) n_i(\hat{\mathbf{p}};\mathbf{w}) - \sigma_i(\hat{\mathbf{p}};\mathbf{w})\end{aligned}$$

Next, suppose that $\hat{p}_j < \pi$. Then

$$\begin{aligned}\hat{p}_i y_j(\hat{\mathbf{p}};\mathbf{w}) + (1-\hat{p}_i) n_j(\hat{\mathbf{p}};\mathbf{w}) - \sigma_j(\hat{\mathbf{p}};\mathbf{w}) &= (1-\hat{p}_i)\frac{w_j}{1-\pi} - w_j \\ &= w_j(\frac{1-\hat{p}_i}{1-\pi} - 1) \\ &\leq w_i(\frac{1-\hat{p}_i}{1-\pi} - 1) \\ &= \hat{p}_i y_i(\hat{\mathbf{p}};\mathbf{w}) + (1-\hat{p}_i) n_i(\hat{\mathbf{p}};\mathbf{w}) - \sigma_i(\hat{\mathbf{p}};\mathbf{w})\end{aligned}$$

Finally, suppose that $\hat{p}_j = \pi$. Then

$$\begin{aligned}
\hat{p}_i y_j(\hat{\mathbf{p}};\mathbf{w}) + (1-\hat{p}_i) n_j(\hat{\mathbf{p}};\mathbf{w}) - \sigma_j(\hat{\mathbf{p}};\mathbf{w}) &= (1-\hat{p}_i)\frac{c_1 w_j}{1-\pi} - c_1 w_j \\
&= c_1 w_j (\frac{1-\hat{p}_i}{1-\pi} - 1) \\
&\leq w_i (\frac{1-\hat{p}_i}{1-\pi} - 1) \\
&= \hat{p}_i y_i(\hat{\mathbf{p}};\mathbf{w}) + (1-\hat{p}_i) n_i(\hat{\mathbf{p}};\mathbf{w}) - \sigma_i(\hat{\mathbf{p}};\mathbf{w})
\end{aligned}$$

The cases where $\hat{p}_i = \pi$ and $\hat{p}_i > \pi$ can be proven similarly. □

4.1 Axiomatic Characterization

We next show that the PCM is the *unique* wagering mechanism satisfying the six properties from Proposition 1, subject to a condition on the reports. Suppose that some non-zero wager is placed on $N > 3$ distinct reports, denoted by $P_1 < P_2 < \ldots < P_N$, and let $W_k = \sum_{i:\hat{p}_i = P_k} w_i$ be the total wager at probability P_k. We say that the *non-extreme* assumption holds if $P_2 W_1 < (1-P_2)\sum_{i=3}^{N} W_i$ and $(1-P_{N-1})W_N < P_{N-1}\sum_{i=1}^{N-2} W_i$. For the data set used in Section 5 and wagers generated according to a Pareto($\alpha = 1.16$) distribution (see Section 5 for details), the non-extreme assumption held on over 99.97% of instances.

Theorem 2. *Let M be a wagering mechanism satisfying anonymity, individual rationality, budget balance, Pareto optimality, sybilproofness, and envy-freeness. If the non-extreme assumption holds, then payoffs defined by M match those defined by the parimutuel consensus mechanism.*

Proof. We first show that any wagering mechanism that satisfies envy-freeness, sybilproofness, and anonymity is defined by fixed prices p_y and p_n for `yes` and `no` securities. That is, for all agents i with $y_i(\hat{\mathbf{p}};\mathbf{w}) > 0$, we have $p_y = \frac{\sigma_i(\hat{\mathbf{p}};\mathbf{w})}{y_i(\hat{\mathbf{p}};\mathbf{w})}$, and for all agents i with $n_i(\hat{\mathbf{p}};\mathbf{w}) > 0$, we have $p_n = \frac{\sigma_i(\hat{\mathbf{p}};\mathbf{w})}{n_i(\hat{\mathbf{p}};\mathbf{w})}$.

To prove this, suppose otherwise for contradiction. That is, suppose that there exist agents i, j with $y_i(\hat{\mathbf{p}};\mathbf{w}) > 0$ and $y_j(\hat{\mathbf{p}};\mathbf{w}) > 0$ such that $\frac{\sigma_i(\hat{\mathbf{p}};\mathbf{w})}{y_i(\hat{\mathbf{p}};\mathbf{w})} > \frac{\sigma_j(\hat{\mathbf{p}};\mathbf{w})}{y_j(\hat{\mathbf{p}};\mathbf{w})}$. Consider a modified instance $(\hat{\mathbf{p}};\mathbf{w}')$ in which both of i and j participate as sybils, denoted by sets S_i and S_j, instead of their individual identities, such that for all $k, \ell \in S_i \cup S_j$, we have that $\sigma_k = \sigma_\ell$. By sybilproofness and anonymity it must be the case that $\sigma_k(\hat{\mathbf{p}};\mathbf{w}') = \sigma_i(\hat{\mathbf{p}};\mathbf{w})/|S_i|$ and $y_k(\hat{\mathbf{p}};\mathbf{w}') = y_i(\hat{\mathbf{p}};\mathbf{w})/|S_i|$ for all $k \in S_i$, with the equivalent equalities holding for all $\ell \in S_j$ also. Therefore, for all $k \in S_i$ and $\ell \in S_j$,

$$\frac{\sigma_k(\hat{\mathbf{p}};\mathbf{w}')}{y_k(\hat{\mathbf{p}};\mathbf{w}')} = \frac{\sigma_i(\hat{\mathbf{p}};\mathbf{w})}{y_i(\hat{\mathbf{p}};\mathbf{w})} > \frac{\sigma_j(\hat{\mathbf{p}};\mathbf{w})}{y_j(\hat{\mathbf{p}};\mathbf{w})} = \frac{\sigma_\ell(\hat{\mathbf{p}};\mathbf{w}')}{y_\ell(\hat{\mathbf{p}};\mathbf{w}')}.$$

Because $\sigma_k(\hat{\mathbf{p}};\mathbf{w}') = \sigma_\ell(\hat{\mathbf{p}};\mathbf{w}')$, k envies ℓ, violating envy-freeness in the modified instance. An identical argument shows the existence of a fixed price p_n for `no` securities.

By budget balance, the wagering mechanism must sell exactly the same number of yes and no securities, and it must be the case that each yes/no pair sells for exactly \$1. Therefore, $p_y + p_n = 1$. By individual rationality, it must be the case that all agents with $\hat{p}_i < p_y$ have $y_i(\hat{\mathbf{p}}; \mathbf{w}) = 0$, and all agents with $\hat{p}_i > p_y$ have $n_i(\hat{\mathbf{p}}; \mathbf{w}) = 0$.

We now use Pareto optimality, along with sybilproofness, anonymity, and envy-freeness, to show that whenever there exist agents i and j, with $\hat{p}_j > \hat{p}_i > p_y$, it must be the case that $\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i$ and $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) = w_j$. We know by Pareto optimality that at least one of the equalities must hold; say, $\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i$. Suppose for contradiction that $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) < w_j$. Again consider a modified instance $(\hat{\mathbf{p}}; \mathbf{w}')$ in which i and j participate as sybils, denoted by sets S_i and S_j, instead of their individual identities, such that for all $k, \ell \in S_i \cup S_j$, we have that $w'_k = w'_\ell$. By anonymity, we have $\sigma_k(\hat{\mathbf{p}}; \mathbf{w}') = w'_k$ for all $k \in S_i$ and $\sigma_\ell(\hat{\mathbf{p}}; \mathbf{w}') < w'_\ell$ for all $\ell \in S_j$. Now, using that fact that all agents are buying yes securities at price p_y, we have that

$$
\begin{aligned}
&\hat{p}_\ell y_\ell(\hat{\mathbf{p}}; \mathbf{w}') + (1 - \hat{p}_\ell) n_\ell(\hat{\mathbf{p}}; \mathbf{w}') - \sigma_\ell(\hat{\mathbf{p}}; \mathbf{w}') \\
&\qquad = \hat{p}_\ell \frac{\sigma_\ell(\hat{\mathbf{p}}; \mathbf{w}')}{p_y} - \sigma_\ell(\hat{\mathbf{p}}; \mathbf{w}') \\
&\qquad < \hat{p}_\ell \frac{w'_\ell}{p_y} - w'_\ell \\
&\qquad = \hat{p}_\ell \frac{w'_k}{p_y} - w'_k \\
&\qquad = \hat{p}_\ell \frac{\sigma_k(\hat{\mathbf{p}}; \mathbf{w}')}{p_y} - \sigma_k(\hat{\mathbf{p}}; \mathbf{w}') \\
&\qquad = \hat{p}_\ell y_k(\hat{\mathbf{p}}; \mathbf{w}') + (1 - \hat{p}_\ell) n_k(\hat{\mathbf{p}}; \mathbf{w}') - \sigma_k(\hat{\mathbf{p}}; \mathbf{w}')
\end{aligned}
$$

Therefore, agent $\ell \in S_j$ envies agent $k \in S_i$, violating envy-freeness. A similar argument can be used to show that $\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i$ and $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) = w_j$ when $\hat{p}_j < \hat{p}_i < p_y$.

Next, suppose that $\hat{p}_j > \hat{p}_i = p_y$. We show that if $y_i(\hat{\mathbf{p}}; \mathbf{w}) > 0$ then $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) = w_j$. First, note that if $\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) < w_i$, then Pareto optimality implies that $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) = w_j$. So suppose that $\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i$. Then, because we have also assumed that $y_i(\hat{\mathbf{p}}; \mathbf{w}) > 0$, we know that $y_i(\hat{\mathbf{p}}; \mathbf{w}) = \frac{\sigma_i(\hat{\mathbf{p}}; \mathbf{w})}{p_y} = \frac{w_i}{p_y}$. We can now use an argument identical to that used in the previous paragraph to argue that if $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) < w_j$, then we can create the same modified instance $(\hat{\mathbf{p}}; \mathbf{w}')$ so that sybils of j will envy sybils of i. Therefore, $\sigma_j(\hat{\mathbf{p}}; \mathbf{w}) = w_j$.

We now show that, provided the condition on reports in the statement of the theorem holds, $y_i(\hat{\mathbf{p}}; \mathbf{w}) > 0$ for all i with $\hat{p}_i = p_{N-1}$ (note that this, along with individual rationality, implies $p_y \le p_{N-1}$). To see this, suppose otherwise. There are two cases. First, if $p_y < p_{N-1} < p_N$, then by our earlier observation it must be the case that $\sigma_i(\hat{\mathbf{p}}; \mathbf{w}) = w_i$ for all i with $\hat{p}_i = p_{N-1}$ or $\hat{p}_i = p_N$. Therefore, $y_i(\hat{\mathbf{p}}; \mathbf{w}) = \frac{w_i}{p_y} > 0$. Second, if $p_y \ge p_{N-1}$ and $y_i(\hat{\mathbf{p}}; \mathbf{w}) = 0$ for some i with $\hat{p}_i = p_{N-1}$, then we can use sybilproofness and anonymity to argue that $y_i(\hat{\mathbf{p}}; \mathbf{w}) = 0$ for *all* i with $\hat{p}_i = p_{N-1}$. Therefore, the total number of yes securities allocated is strictly less than the total number of no securities allocated:

$$\sum_{i\in\mathcal{N}} y_i(\hat{\mathbf{p}};\mathbf{w}) \le \frac{W_N}{p_y} \le \frac{W_N}{p_{N-1}} < \frac{1}{1-p_{N-1}}\sum_{i=1}^{N-2} W_i \le \frac{1}{p_n}\sum_{i=1}^{N-2} W_i \le \sum_{i\in\mathcal{N}} n_i(\hat{\mathbf{p}};\mathbf{w})$$

This violates budget balance. By a symmetric argument, we can show that $p_n \le 1-p_2$ and $n_i(\hat{\mathbf{p}};\mathbf{w}) > 0$ for all i with $\hat{p}_i = p_2$.

In particular, the previous paragraph says that, subject to the conditions of the theorem holding, at least two bettors with distinct reports are allocated yes securities, and at least two bettors with distinct reports are allocated no securities. By the two earlier paragraphs, this implies that for all i with $\hat{p}_i > p_y$, we have $\sigma_i(\hat{\mathbf{p}};\mathbf{w}) = w_i, y_i(\hat{\mathbf{p}};\mathbf{w}) = \frac{w_i}{p_y}, n_i(\hat{\mathbf{p}};\mathbf{w}) = 0$, and for all i with $1-\hat{p}_i > p_n$, we have $\sigma_i(\hat{\mathbf{p}};\mathbf{w}) = w_i, y_i(\hat{\mathbf{p}};\mathbf{w}) = 0, n_i(\hat{\mathbf{p}};\mathbf{w}) = \frac{w_i}{p_n}$.

Finally, it is easy to see that the only value of p_y/p_n that satisfies this condition while allocating an equal number of yes and no securities and satisfying $p_y = 1-p_n$ is that defined by $p_y = \pi$ and $p_n = 1-\pi$, from Equation 1. To characterize the allocations and payments of agents with $\hat{p}_i = p_y$, we note that these agents are required to exactly make up the difference between yes and no securities, if such a difference exists. By anonymity and sybilproofness, each of these bettors must be sold a number of securities that is proportional to their wager. This exactly matches the allocations and payments defined by the PCM. □

4.2 Incentive Properties of the PCM

As a Corollary of Theorem 1 and Proposition 1, we know that the PCM violates incentive compatibility. Intuitively, this is because agents are able to change the price π by changing their reports.

Example 2. *Let* $\mathbf{p} = (0.4, \frac{2}{3}, 0.8)$ *and* $\mathbf{w} = (1,1,1)$. *Then the outcome of the PCM is* $(\mathbf{y} = (0, 1.5, 1.5), \mathbf{n} = (3,0,0), \boldsymbol{\sigma} = (1,1,1))$. *Note that the price* $\pi = \frac{2}{3}$, *so agent 2's utility is 0. However, if agent 2 misreports* $\hat{p}_2 = 0.6$, *then the outcome becomes* $(\mathbf{y} = (0, \frac{5}{6}, \frac{5}{3}), \mathbf{n} = (2.5, 0, 0), \boldsymbol{\sigma} = (1, 0.5, 1))$. *Now the price* $\pi = 0.6$, *so agent* 2*'s utility is* $\frac{5}{6}\cdot\frac{2}{3} - 0.5 = \frac{1}{18} > 0$.

Example 2 has a particular form common to all profitable misreports. In order to change the price in a profitable way, a manipulating agent must ensure that her misreport exactly matches the new equilibrium price. The intuition is that the only way an agent can affect the price is to report a probability on the opposite side of the current price as her belief. However, such a misreport is only profitable if she does not 'over-shoot' and end up buying the wrong type of security.

Theorem 3. *Let* $\hat{p}_i \ne p_i$ *be a profitable misreport for agent* i. *Let* π_T *denote the* yes *security price when agent i reports truthfully, and* π_M *denote the* yes *security price in the instance when* i *misreports* $\hat{p}_i$. *Then it must be the case that* $\hat{p}_i = \pi_M$, *and either* $\hat{p}_i < \pi_T \le p_i$ *or* $p_i \le \pi_T < \hat{p}_i$.

Before we give the proof, we first state and prove a monotonicity lemma which states that, all else being equal, if the report of a single agent increases then the security price π also (weakly) increases.

Lemma 1. *Let* $\hat{\mathbf{p}}_{-\mathbf{i}} = \hat{\mathbf{p}}'_{-\mathbf{i}}$*. Let* $\hat{p}'_i < \hat{p}_i$*, and denote by* π' *the equilibrium price under vector of reports* $\hat{\mathbf{p}}'$*, and* π *the equilibrium price under vector of reports* $\hat{\mathbf{p}}$*. Then* $\pi' \leq \pi$*.*

Proof. Consider the equilibrium condition, Equation 1:

$$\sum_{j:\hat{p}_j<\pi} \frac{w_j}{1-\pi} + c_1 \sum_{j:\hat{p}_j=\pi} \frac{w_j}{1-\pi} = \sum_{j:\hat{p}_j>\pi} \frac{w_j}{\pi} + c_2 \sum_{j:\hat{p}_j=\pi} \frac{w_j}{\pi}$$

Suppose that $\hat{p}_i > \pi$ (other cases can be handled similarly). Suppose for contradiction that $\pi' > \pi$. Let c_1, c_2 represent the values of the equilibrium constants in the case that i reports $\hat{p}_i$, and c'_1, c'_2 represent those values when i reports $\hat{p}'_i$. Then we have

$$\begin{aligned}
\sum_{j:\hat{p}_j>\pi} \frac{w_j}{\pi} + c_2 \sum_{j:\hat{p}_j=\pi} \frac{w_j}{\pi} \geq \sum_{j:\hat{p}_j>\pi} \frac{w_j}{\pi} &> \sum_{j:\hat{p}_j>\pi} \frac{w_j}{\pi'} \\
&\geq \sum_{j:\hat{p}'_j>\pi'} \frac{w_j}{\pi'} + c'_2 \sum_{j:\hat{p}'_j=\pi'} \frac{w_j}{\pi'} \\
&= \sum_{j:\hat{p}'_j<\pi'} \frac{w_j}{1-\pi'} + c'_1 \sum_{j:\hat{p}'_j=\pi'} \frac{w_j}{1-\pi'} \\
&\geq \sum_{j:\hat{p}'_j<\pi'} \frac{w_j}{1-\pi'} \\
&\geq \sum_{j:\hat{p}_j<\pi} \frac{w_j}{1-\pi'} + c_1 \sum_{j:\hat{p}_j=\pi} \frac{w_j}{1-\pi'} \\
&> \sum_{j:\hat{p}_j<\pi} \frac{w_j}{1-\pi} + c_1 \sum_{j:\hat{p}_j=\pi} \frac{w_j}{1-\pi}
\end{aligned}$$

where the equality holds by Equation 1, and the inequalities all hold due to the assumptions that $\hat{p}_i > \pi$ and that $\pi' > \pi$. Comparing the first and last line contradicts that π is the equilibrium price under reports $\hat{\mathbf{p}}$. Therefore, it must be the case that $\pi' \leq \pi$. □

Proof of Theorem 3. Suppose that $p_i > \pi_T$. The cases $p_i < \pi_T$ and $p_i = \pi_T$ can be proven similarly. Note that if $\pi_M = \pi_T$, then $\hat{p}_i$ cannot be a profitable misreport, because under truthful reporting, i already buys as many `yes` securities as her budget allows, and these are the only securities from which she obtains positive expected profit at the current price π_T. Therefore, to show that any profitable misreport satisfies $\hat{p}_i < \pi_T$, we show that $\pi_M = \pi_T$ whenever $\hat{p}_i \geq \pi_T$.

Consider again Equation 1. For $\hat{p}_i > \pi_T$, if we set $\pi = \pi_T$ then each term in the equation takes the same value under truthful reporting and misreporting. Therefore, equality holds in the misreported case with $\pi_M = \pi_T$. Next, if $\hat{p}_i = \pi_T < p_i$, then we know that $\pi_M \leq \pi_T$, by Lemma 1, since $\hat{p}_i < p_i$. It remains to rule out $\pi_M < \pi_T$. So suppose for contradiction that $\pi_M < \pi_T = \hat{p}_i < p_i$. Let c_1^M, c_2^M denote the equilibrium values of c_1 and c_2 when i misreports $\hat{p}_i$, and c_1^T, c_2^T the equilibrium values when i truthfully reports p_i. Then we have a similar system of inequalities as in the proof of Lemma 1 (some steps are omitted),

$$
\begin{aligned}
\sum_{j:\hat{p}_j>\pi_M} \frac{w_j}{\pi_M} + c_2^M \sum_{j:\hat{p}_j=\pi_M} \frac{w_j}{\pi_M} &\geq \sum_{j:\hat{p}_j>\pi_M} \frac{w_j}{\pi_M} \\
&> \sum_{j:\hat{p}_j>\pi_M} \frac{w_j}{\pi_T} \\
&\geq \sum_{j:\hat{p}_j>\pi_T} \frac{w_j}{\pi_T} + \frac{w_i}{\pi_T} + c_2^T \sum_{j:\hat{p}_j=\pi_T} \frac{w_j}{\pi_T} \\
&= \sum_{j:\hat{p}_j<\pi_T} \frac{w_j}{1-\pi_T} + c_1^T \sum_{j:\hat{p}_j=\pi_T} \frac{w_j}{1-\pi_T} \\
&\geq \sum_{j:\hat{p}_j<\pi_T} \frac{w_j}{1-\pi_T} \\
&\geq \sum_{j:\hat{p}_j<\pi_M} \frac{w_j}{1-\pi_T} + c_1^M \sum_{j:\hat{p}_j=\pi_M} \frac{w_j}{1-\pi_T} \\
&> \sum_{j:\hat{p}_j<\pi_M} \frac{w_j}{1-\pi_M} + c_1^M \sum_{j:\hat{p}_j=\pi_M} \frac{w_j}{1-\pi_M}
\end{aligned}
$$

which contradicts that π_M is the equilibrium price when i reports $\hat{p}_i$. Therefore it is not the case that $\pi_M < \pi_T$, so $\pi_M = \pi_T$ and the misreport $\hat{p}_i \geq \pi_T$ is not profitable.

We have shown that $\hat{p}_i < \pi_T < p_i$ must hold for any profitable misreport $\hat{p}_i$. Therefore, by Lemma 1, we know that $\pi_M \leq \pi_T$. We now show that $\pi_M = \hat{p}_i$. First, suppose that $\hat{p}_i < \pi_M$. Then i is buying `no` securities at a price $1-\pi_M \geq 1-\pi_T > 1-p_i$, where $1-p_i$ is her value for a `no` security. Therefore, she obtains negative expected profit from this purchase, meaning that $\hat{p}_i$ is not a profitable misreport. Second, suppose that $\hat{p}_i > \pi_M$. In this case, we can argue by setting $\pi = \pi_M$ in Equation 1. It is easy to see that at this equilibrium price, strictly more `yes` securities are sold than in the truthful case, and strictly fewer `no` securities. This violates budget balance, since equal numbers of `yes` and `no` securities are sold in the truthful case. Therefore, $\pi_M = \pi_T$. However, we have already established that if $\pi_M = \pi_T$, then $\hat{p}_i$ is not a profitable misreport, a contradiction. □

Incentive Compatibility in the Large. We now show that the PCM satisfies an approximate notion of incentive compatibility known as *incentive compatibility in the large (IC-L)*, introduced by Azevedo and Budish [2017].[3] It relaxes incentive compatibility by requiring only that truthful reporting is optimal as the number of agents grows large, and that truthful reporting is only optimal in expectation over the reports, rather than based on the (ex-post) realization of reports, as our definition of incentive compatibility requires.

Conceptually, this section mirrors the work of Azevedo and Budish. Indeed, in cases where only a finite set of reports are allowed, that the PCM satisfies IC-L follows directly from the fact that the PCM satisfies envy-freeness (Azevedo and Budish show

[3]There is a large body of work focusing on other limiting IC criteria, including ϵ-strategyproofness Roberts and Postlewaite [1976], Ehlers et al. [2004], that we do not focus on here.

that this implies IC-L when the number of possible reports is finite). Since finite sets of reports can provide arbitrary precision, this is usually enough for practical purposes. Most real-world mechanisms only allow reports up to a precision of 1%, and this is also the precision we use in our simulations (see Footnote 5). However, for completeness, we provide an independent proof of IC-L for the case where an infinite number of reports are allowed. The proof is a simple extension of the finite reports case.

Let D denote a probability distribution over $[0, 1]$ with full support. We model each agent as drawing a report $\hat{p}_i$ i.i.d. according to D. So D models the distribution of reports, not necessarily beliefs. We will assume that wagers are drawn i.i.d. from some fixed distribution bounded by the interval $[1, W]$ for some $W \geq 1$. In particular, the ratio of the wagers of any two agents is bounded by W. Denote the expected value of a randomly drawn wager by $\overline{w}$.

For the remainder of this section, let $(y_i(\hat{p}_i, w_i, D, n), n_i(\hat{p}_i, w_i, D, n), \sigma_i(\hat{p}_i, w_i, D, n))$ denote the expected allocation of securities and payment for an agent reporting $\hat{p}_i$ and wagering $w_i \in [1, W]$ when there are n other agents that draw reports according to D and wagers from the fixed wager distribution. Let $(y_i(\hat{p}_i, w_i, D, \infty), n_i(\hat{p}_i, w_i, D, \infty), \sigma_i(\hat{p}_i, w_i, D, \infty)) = \lim_{n\to\infty}(y_i(\hat{p}_i, w_i, D, n), n_i(\hat{p}_i, w_i, D, n), \sigma_i(\hat{p}_i, w_i, D, n))$. We can now formally define incentive compatibility in the large.

Definition 7. *A wagering mechanism is* incentive compatible in the large *if, for any D with full support over $[0, 1]$, and any $\hat{p}_i$ and w_i,*

$$\begin{aligned} &p_i y_i(p_i, w_i, D, \infty) + (1 - p_i) n_i(p_i, w_i, D, \infty) - \sigma_i(p_i, w_i, D, \infty) \\ &\geq p_i y_i(\hat{p}_i, w_i, D, \infty) + (1 - p_i) n_i(\hat{p}_i, w_i, D, \infty) - \sigma_i(\hat{p}_i, w_i, D, \infty). \end{aligned}$$

To show that the PCM satisfies incentive compatibility in the large, we first show that when the number of bettors is large, no single agent can affect the security price π; that is, agents are price-takers in the large market limit. The second step is to show that price takers have no profitable manipulations, which follows immediately from Theorem 3.

Theorem 4. *The parimutuel consensus mechanism satisfies incentive compatibility in the large.*

Proof. Let π^n denote the price defined by the PCM in expectation when there are n agents drawing reports from D, as well as agent i reporting $\hat{p}_i$, and let $\pi^\infty = \lim_{n\to\infty} \pi^n$. We first show that π^∞ exists. For contradiction, suppose otherwise. Fix $\epsilon > 0$. Then there exist arbitrarily large N_1, N_2 such that $|\pi^{N_1} - \pi^{N_2}| > \epsilon$ for some $\epsilon > 0$. Suppose without loss of generality that $\pi^{N_1} > \pi^{N_2} + \epsilon$. Note that we can rewrite the equilibrium condition, Equation 1,

$$\pi = \frac{\sum_{j:\hat{p}_j > \pi} w_j + c_2 \sum_{j:\hat{p}_j = \pi} w_j}{\sum_{j:\hat{p}_j \neq \pi} w_j + (c_1 + c_2) \sum_{j:\hat{p}_j = \pi} w_j}$$

Therefore, π^{N_1} and π^{N_2} are defined by

$$\pi^{N_1} = \frac{\sum_{j\neq i:\hat{p}_j > \pi^{N_1}} \bar{w} + c_2 \sum_{j \neq i:\hat{p}_j = \pi^{N_1}} \bar{w} + y_1 w_i}{\sum_{j\neq i:\hat{p}_j \neq \pi^{N_1}} \bar{w} + (c_1 + c_2) \sum_{j\neq i:\hat{p}_j = \pi^{N_1}} \bar{w} + w_i} \tag{2}$$

$$\pi^{N_2} = \frac{\sum_{j \neq i : \hat{p}_j > \pi^{N_2}} \bar{w} + c_2 \sum_{j \neq i : \hat{p}_j = \pi^{N_2}} \bar{w} + y_2 w_i}{\sum_{j \neq i : \hat{p}_j \neq \pi^{N_2}} \bar{w} + (c_1 + c_2) \sum_{j \neq i : \hat{p}_j = \pi^{N_2}} \bar{w} + w_i} \quad (3)$$

Where $y_1 = 1$ if $\hat{p}_i > \pi_1^N$ and $y = 0$ if $\hat{p}_i < \pi_1^N$, and similarly for y_2 in Equation 3 (for simplicity of notation, we ignore the case where $\hat{p}_i = \pi$, but it can be handled similarly). We can replace wagers w_j by $\bar{w}$ because we are interested in the price in expectation.

Since we can choose N_1 and N_2 to be artbitrarily large, the sum of all wagers w_j becomes large, and the effect of the wager w_i becomes arbitrarily small. Therefore, π^{N_1} and π^{N_2} become arbitrarily close to one another, violating the assumption that they are bounded apart by ϵ. Thus, π^∞ exists.

To see that π^∞ is independent of $\hat{p}_i$, we divide both the numerator and denominator of Equation 2 by N_1 and let $N_1 \to \infty$. The equilibrium condition becomes

$$\pi^\infty = \frac{Pr_{x \sim D}(x > \pi^\infty) + c_2 Pr_{x \sim D}(x = \pi^\infty)}{Pr_{x \sim D}(x \neq \pi^\infty) + (c_1 + c_2) Pr_{x \sim D}(x = \pi^\infty)}$$

Since this equation has no dependence on $\hat{p}_i$ (or w_i), π^∞ is independent of $\hat{p}_i$.

It now follows immediately from Theorem 3 that the PCM satisfies IC-L, since any profitable manipulation must alter the security price. But in the limit as the number of agents goes to ∞, it is impossible for i to affect the price π^∞. □

5 Simulations

We tested the incentive compatibility of the PCM on a data set consisting of probability reports gathered from an online prediction contest called ProbabilitySports Galebach [2004].[4] The data set consists of probabilistic predictions about the outcome of 1643 National Football League matches from the start of the 2000 NFL preseason until the end of the 2004 season. For each match, between 64 and 1574 players reported their subjective probability of a fixed team (say, the home team) winning the match. Each match was scored according to the Brier scoring rule, with points contributing to a season-long scoreboard.

ProbabilitySports users submitted probabilities but not wagers. We generated wagers from a variety of Pareto distributions. Pareto distributions are a natural choice as they approximately model the distribution of wealth in a population. A Pareto distribution is defined by two parameters: a scale parameter k, which has the effect of multiplicatively scaling the distribution, and a shape parameter α, which affects the size of the distribution's tail. To allow for a fair comparison between distributions and instance sizes, we scaled each set of randomly generated wagers so that the average wager is 1. This means that changing the scale parameter has no effect, as the wagers are rescaled anyway. Therefore, we fix the scale parameter to 1 and vary only the shape parameter.

The first Pareto distribution we use for wager generation has $\alpha = 1.16$, which is often described as "20% of the population has 80% of the wealth," and classically

[4]We thank Brian Galebach for providing us with this data.

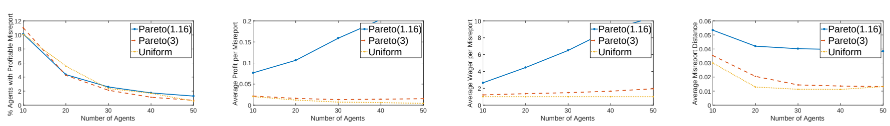
	% Agents With Profitable Misreports	Average Profit	Average Wager per Profitable Misreport	Average Misreport Distance
Pareto($\alpha = 1.16$)	0.07	1.55	118.8	0.044
Pareto($\alpha = 3$)	< 0.01	0.03	5.76	0.015
Uniform	0	n/a	n/a	n/a

Table 1: Profitable misreports under Pareto and uniform wager generation.

viewed as a realistic distribution of wealth. Second, we use $\alpha = 3$, which produces a more equal distribution of wagers in comparison to $\alpha = 1.16$. Finally, we consider a uniform distribution of wagers (that is, $w_i = 1$ for all agents), corresponding to a situation either where all agents are equal, or where they do not have the opportunity to choose their wager (as in the ProbabilitySports competition). Note that the uniform distribution is the limit of the Pareto distribution as $\alpha \to \infty$.

Our first step was to examine the entire dataset. For each of the 1643 matches and each wager distribution, we randomly generated a set of wagers drawn from that distribution. For each set of wagers we chose 50 random agents and simulated 101 reports for them in the range $\{0, 0.01, \dots, 0.99, 1\}$.[5] For each report, we computed the agent's expected utility, taking their true belief to be their original report p_i. If there exists a misreport $\hat{p}_i \neq p_i$ that yields a higher utility than reporting their true belief, then the agent has a profitable misreport.

The results are summarized in Table 1. We report four statistics. The '% Agents With Profitable Misreports' column states the percentage of agents that are able to benefit from misreporting. The 'Average Profit' column gives, out of those agents with a profitable misreport, the average benefit that the agent can gain from misreporting optimally, over and above her utility from reporting truthfully. The 'Average Wager per Profitable Misreport' column gives the average wager of agents with a profitable misreport available. Finally, the 'Average Misreport Distance' gives, for those agents with a profitable misreport, the average distance between the optimal misreport and the true belief.



(a) % Profitable Misreports

(b) Average Profit per Misreport

(c) Average Wager per Misreport

(d) Average Misreport Distance

Figure 1: Profitable misreport characteristics for varying wager distributions and varying numbers of agents, n.

[5] In principle, our setup allows agents to report at a higher precision than this, so there will be some possible misreports that we do not detect. However, we believe that considering reports of only multiples of 0.01 is reasonable, due to limited cognitive capacity of the agents and the practical constraints of many wagering systems.

For wagers generated from a Pareto distribution with $\alpha = 1.16$, we found 55 profitable misreports (out of 82,150 agents that we checked), which means that only around 0.07% of the agents that we checked had a profitable misreport. It is striking to consider the makeup of this small percentage of agents. The average wager of these agents is 118.8 (recall that each set of wagers is scaled so that the average wager is 1). What we are seeing is that only agents with very, very high wagers have sufficient power to change the price π. In contrast, the average profit that these agents obtain by misreporting is only 1.55, suggesting that even these high-wager agents are unable to have too large an effect on the security price π. This average profit is on the order of 1-2% of the misreporting agents' wagers – arguably an insignificant amount. For those agents that do misreport, the optimal misreport only differs from their belief by around 0.04.

As α increases, the number of agents with opportunity to misreport decreases. Indeed, for the uniform wagers, we did not find a single opportunity to profitably misreport. This is not surprising, since when wagers are uniform and there are a large number of agents, no agent will ever be able to significantly affect the price.

So for the full data set, with $64 \leq n \leq 1574$ agents per match, opportunities to profitably misreport are scarce, as we would expect because the PCM satisfies IC-L. But what about instances with fewer agents? To investigate smaller instances, we subsampled smaller values of n from the complete set of reports and ran the same simulation. For each match, each value of $n \in \{10, 20, 30, 40, 50\}$, and each wager distribution, we randomly sampled n reports and generated wagers. For every instance generated in this way, we tested *every* agent to see whether they had a profitable misreport.

Figure 1(a) shows how the percentage of agents that can profitably misreport changes with instance size. Even with only 10 agents per instance, there are relatively few opportunities to profitably misreport, with around 10% of all agents being able to do so. This fraction decreases quickly as n increases – for instances with 50 agents, less than 2% of agents are able to profitably misreport. Interestingly, all wager distributions exhibit approximately the same susceptibility to manipulation, in contrast to the full instances. We speculate that this is because, while high-wager agents are more likely to have profitable manipulations available, their existence also prevents low-wager agents from being able to manipulate, thus rendering the existence of high-wager agents something of a wash for small n. For large n, the latter effect disappears, since low-wager agents are unable to profitably misreport, even in the absence of high-wager agents.

Figure 1(b) shows how the average value of each profitable misreport changes with n, where the value of a profitable misreport is the difference in expected utility between the agent's optimal misreport and their truthful report. Interestingly, we see three very different trends depending on the wager distribution, all of which are consistent with the results on the full dataset. For $\alpha = 1.16$, the average value of a misreport steadily increases with n, as high-wager agents (who have high-value misreports) become more and more frequent, while low-value misreports become less frequent. With uniform wagers, the value of a misreport quickly decreases with n. With only 10 agents, a misreporting agent may be able to affect the price quite significantly, however with increasing n, misreports will consist of only being able to make small adjustments to the security price. For $\alpha = 3$, the value of a misreport remains approximately constant as n increases, suggesting some combination of the two previous effects.

Figure 1(c) shows how the average wager of agents with a profitable misreport changes with n. For uniform wagers this line is flat, since all agents have wager $w_i = 1$. The other two wager distributions display increasing wagers, which is again explained by increasing frequency of high-wager agents (with this frequency increasing faster for $\alpha = 1.16$ than for $\alpha = 3$), and decreasing frequency of low-wager agents that are actually able to profitably misreport.

Finally, in Figure 1(d) we plot the average distance between a profitable misreport and an agent's true belief. In contrast with the other statistics that we consider, this one is actually relatively flat as n increases (with the exception of a significant drop from $n = 10$ to $n = 20$). This tells us that even for small numbers of forecasters, misreporting is limited to agents with beliefs fairly close to the price π and does not significantly affect the equilibrium price.

We note that we have considered an omniscient setting where manipulating agents have precise knowledge of the reports of other agents. In practice, of course, the manipulating agent has uncertainty about her opponents. A misreport is risky, involving some possibility of being forced to buy securities at a price favorable to her misreport but not her true belief. High-budget agents have the most opportunities to misreport but also the most to lose if they miscalculate.

6 Conclusion

We have provided an axiomatic justification of the parimutuel consensus mechnanism. While no wagering mechanism can satisfy anonymity, individual rationality, budget balance, Pareto optimality, sybilproofness, envy-freeness and incentive compatibility, we show that the PCM comes very close in that it satisfies all of the first six properties, and a relaxation of the seventh: incentive compatibility in the large. Subject to a mild condition on the reports, the PCM is the only wagering mechanism that satisfies all six properties. Via comprehensive simulations based on real contest data, we have shown that on large instances, opportunities to profitably manipulate are extremely rare. Even on small instances, the vast majority of agents cannot gain from misreporting.

Acknowledgments

The authors thank Jennifer Wortman Vaughan for helpful discussions. This work was conducted in part while Freeman was an intern at Microsoft Research. He also thanks a Facebook Ph.D. Fellowship for support.

References

Mukhtar M. Ali. Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, 85(4):803–816, 1977.

Kenneth J. Arrow, Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O. Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D. Nelson, George R. Neumann, Marco Ottaviani, Thomas C. Schelling, Robert J. Shiller, Vernon L. Smith, Erik Snowberg, Cass R. Sunstein, Paul C. Tetlock, Philip E. Tetlock, Hal R. Varian, Justin Wolfers, and Eric Zitzewitz. The promise of prediction markets. *Science*, 320(5878):877–878, 2008.

Eduardo M Azevedo and Eric Budish. Strategy-proofness in the large. Technical report, National Bureau of Economic Research, 2017.

Alina Beygelzimer, John Langford, and David M. Pennock. Learning performance of prediction markets with Kelly bettors. In *Proceedings of the 11th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1317–1318, 2012.

Yiling Chen and David Pennock. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence (UAI)*, 2007.

Yiling Chen, Nikhil R. Devanur, David M. Pennock, and Jennifer Wortman Vaughan. Removing arbitrage from wagering mechanisms. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, 2014.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory, 2nd Edition*. Wiley, New York, 2006.

Lars Ehlers, Hans Peters, and Ton Storcken. Threshold strategy-proofness: On manipulability in large voting problems. *Games and Economic Behavior*, 49(1):103–116, 2004.

Edmund Eisenberg and David Gale. Consensus of subjective probabilities: The pari-mutuel method. *The Annals of Mathematical Statistics*, 30(1):165–168, 1959.

D. Foley. Resource allocation and the public sector. *Yale Economics Essays*, 7:45–98, 1967.

Rupert Freeman, David M Pennock, and Jennifer Wortman Vaughan. The double clinching auction for wagering. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC)*, pages 43–60, 2017.

Brian Galebach. Probabilitysports. `http://probabilitysports.com/`, 2004.

Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.

Nicolas S. Lambert, John Langford, Jennifer Wortman, Yiling Chen, Daniel Reeves, Yoav Shoham, and David M. Pennock. Self-financed wagering mechanisms for forecasting. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC)*, 2008.

Nicolas S. Lambert, John Langford, Jennifer Wortman Vaughan, Yiling Chen, Daniel Reeves, Yoav Shoham, and David M. Pennock. An axiomatic characterization of wagering mechanisms. *Journal of Economic Theory*, 156:389–416, 2015.

Charles F. Manski. Interpreting the predictions of prediction markets. *Economics Letters*, 91(3):425–429, June 2006.

Charles R Plott, Jorgen Wit, and Winston C Yang. Parimutuel betting markets as information aggregation devices: Experimental results. *Economic Theory*, 22(2):311–351, 2003.

Donald John Roberts and Andrew Postlewaite. The incentives for price-taking behavior in large exchange economies. *Econometrica: Journal of the Econometric Society*, pages 115–127, 1976.

Wayne W. Snyder. Horse racing: Testing the efficient markets model. *Journal of Finance*, 33(4):1109–1118, 1978.

Richard H. Thaler and William T. Ziemba. Anomalies: Parimutuel betting markets: Racetracks and lotteries. *Journal of Economic Perspectives*, 2(2):161–174, 1988.

Justin Wolfers and Eric Zitzewitz. Interpreting prediction market prices as probabilities. Technical Report Working Paper 12200, National Bureau of Economic Research, 2006.