Eliciting Socially Optimal Rankings from Biased Jurors: Two Juror Case.

Brandon Campbell[1]
Department of Economics,
Texas A&M University
April 15, 2014

Abstract

I extend the results of Amorós (2009) to the two juror case. Amorós looked at the environment where a jury of 3 or more had to report a ranking of contestants. There exists a true ranking which is known to all the jurors, but is not known nor verifiable by the social planner. The social planner's social choice rule is to figure out the true ranking from the jurors. The jurors can be biased over contestants, so I use partially-impartial and partially-indifferent preferences to get implementation. I show that it is impossible to subgame perfect implement in the two juror case with restrictions only on partially-impartial preferences, but I show that with restrictions on partially-impartial and partially-indifferent preferences we can get implementation, and how large is the Universe in which implementation occurs. I also show that the simple two-turn extensive form game, where one juror suggests a ranking, then the second juror suggests a ranking dependent on the previous juror's suggestion, is an optimal mechanism for Subgame perfect implementation in this problem. Finally, Nash implementation sufficient results are characterized.

JEL Classification: D78, C70
Keywords: Implementation Theory, Subgame Perfect Equilibrium, Nash Equilibrium.

[1]Brandon Campbell: Spielmaus@neo.tamu.edu. I thank feedback from Dr. Tian, Dr. Velez, attendants at the Texas A&M Theory and Experimental Student Seminar, and faculty at the Private Enterprise Research Center.

1 Introduction

Consider the situation: You are sick. You go to a doctor, and he tells you you have a 1
stomach ulcer, and you need to take an expensive drug brand. You ask him about other drugs and talks about the generic brand, as well as some competitor, but recommends the expensive drug brand first most. He even gives you a free sample of the branded drug. Often in such a situation, we hear of people getting a "second opinion" by visiting another doctor and seeing what he recommends. Can this work? Can we find the drug that is best for our ulcer, and then the next best, and so on? In this paper, I say yes, even with just two advisers.

Past papers on similar topics have generally garnered impossibility results. In Wolinsky 2
(2002) they look at a typical problem where they have two biased advisers, and are trying to find the correct θ in the range of $[0, 1]$. This is a problem relevant to finding a price for a drug brand, and they find they need advisers biased in opposite direction to get closer to the true value of θ, but find its impossible to guaranteed discovery of true θ with any observable features of the advisers. My paper is different. It is looking for a ranking of outcome, and will use observable features of the advisers to get positive results. In the example of the doctor prescribing medication for ulcers, we would say a doctor is biased between recommending a drug from a company that is sponsoring him, and one that is not (he will have free samples from the sponsoring company, or you can look up the drug companies he receives gifts from in excess of 10$ due to the Physician Payments Sunshine Act). However, if given a choice between two rankings where only two sponsored, or two non-sponsored drugs are the only drugs that change their relative ranking, then we say he is impartial between those two rankings and order the drugs by their effectiveness. Thus, if the doctor is sponsored by by both drug A and drug B, we learn about the relative effectiveness between the two drugs when he chooses between rankings ABC or BAC.

This paper is similar to Amorós (2009). He looked at implementation in a environment 3
where he had 3 or more jurors. This is more common for institutional juries, such as the judges for a figure skating competition. This is a subjective criteria, that only becomes more objective to trained judges, such that the social planner has difficulty judging the skaters themselves and must rely on the judges to report the true ranking of skaters from best performance to worst. But judges can be biased, and can seek to improve the chances of their countries candidates at the expense of others, but be impartial when ranking their own countries skaters against each other. He ultimately finds that the largest environment the social planner can get the judges to reveal the true ranking in, is when he knows a judge impartial over each contestant pair, much like the doctor above being impartial over Drug A and Drug B, and the social planner also has that all contestant pairs have a judge impartial over them. If everyone prefers skater A to B, then we can't expect them to rank B before A, even if B out skated A.

For my results, I will require the same two assumptions. However, I will also need to make 4
some restrictions on the distribution of the jurors impartiality over the contestant pairs. I will also need to introduce indifference over contestant pairs to get implementation. I also show that in such an environment, the optimal mechanism is the Stackleberg solution using Subgame Perfect Nash Implementation (a two turn mechanism). Under stronger restrictions, I show that Nash implementation is possible. Ultimately, while I get positive results, they are quite particular, and ultimately show the value of a second opinion, and how honest people are in general (rarely strategic, which makes implementation easy).

The paper is organized as follows: Section 2 introduces the model and the environment, 5
and formal definitions. In section 3 I introduce important sets used in the proofs and the concept by which I determine jurors incentive compatibility. Section 4 outlays the results. It consists of subsections for the necessity and sufficient characterization and some more

intuitive assumptions to satisfy the necessity results. Section 5 is the conclusion.

2 Model

The structure of the problem looks similar to social choice problems, but is usually called 1
a Juror problem. In this paper, I have only two jurors $\{i, j\} = J$. Let N be a set of $n \geq 3$ contestants, and N^2 the set of all contestant pairs. An outcome, π, is a ranking of all contestants in N, where Π denotes the set of all rankings of N. Define p_A^π as the position of contestant A in ranking π. For example, if $\pi = (A, B, C)$, then $p_A^\pi = 1$. Accepting some sloppy notation, rankings will often be written ABC and subrankings $...ABC$, when context clear.

We need the jurors because they observe the **true ranking**, π_t, for all $j \in J$, which is 2
unknown by the social planner. The **socially optimal rule** is that the ranking we elicit from the jurors is the same as the true ranking.

Jurors preferences are dependent on the true ranking. This gives a state-dependent 3
preference function $R : \Pi \to \mathcal{R}$, where $\mathcal{R}$ is the set of all possible preferences over rankings, such that for any true ranking, $\pi_t \in \Pi$, we get a preference relation $R_j(\pi_t) \in \mathcal{R}_j$ for juror j. Thus, with the same preference function, if we change the true ranking, we get different preference relations. I will define $P_j(\pi)$ as the strict part of juror j's preferences, and $I_j(\pi)$ as the equivalence. This will allow us to start talking about impartialness on preferences.

I will say a juror has the **impartial pair** over (A, B), if given a choice between two 4
rankings where A and B are adjacent, and the only difference is A and B have swapped their positions with each other, the juror strictly prefers the ranking where A and B are in the same order as in the true ranking. I denote the set of juror j's impartial pairs as $U_j \subset N^2$.

Definition 1 (Partially Impartial Preferences)**.** *Preferences R_j are partially-impartial for contestants $\{A, B\} \in U_j$ if: for any rankings $\{\pi, \hat{\pi}\} \in \Pi$ for juror j, then $\pi P_j(\pi_t)\hat{\pi}$ if:*

- $p_A^{\pi} = p_B^{\pi} - 1$
- $p_A^{\hat{\pi}} = p_B^{\hat{\pi}} + 1$
- $p_C^{\pi} = p_C^{\hat{\pi}}$ *for all* $C \in N \setminus \{A, B\}$, *and*
- $p_A^{\pi_t} < p_B^{\pi_t}$

Denote the set of Partially Impartial Preferences as $\mathcal{U}_j \subset \mathcal{R}_j$ *for any* $j \in J$.

For an example, consider a juror ranking three contestants **A**lice, **B**arbara, and **C**hristina 5
in a gymnastics competition. Consider the case where the juror is impartial over Alice and Christina, because they are her friends, $\{A, C\} \in I$, but she is biased between Barbara and Alice because only Alice is a friend, $\{A, B\} \notin I$. Then if the juror was given a choice over the rankings (Barbara, Alice, Christina) and (Barbara, Christina, Alice), hereafter written BAC and BCA, he would choose the ranking where Alice and Christina are in the same order as the true ranking. Suppose the true ranking is ABC, then the juror would prefer to choose BAC. If he had chosen BCA, then Christina is ranked before Alice. This is the opposite ranking from the true ranking, ABC, therefore $BAC\ P_j(ABC)\ BCA$. So, even though the juror is biased, we could get some meaningful information from them by just knowing Alice and Christina are his friends.

An advantage of using impartial pairs is that they are about the weakest assumption you 6
can make about how someone's honesty translates into preferences. The fact the contestants are adjacent in the rankings when swapped is important. When we look at BAC and BCA, the A and C are adjacent and swapped between the two choices. That means the relative order of A and C changes between the two rankings, but B ranks before C and A still, such that their orders don't change, and any bias in preferences related to the pairings $\{A, B\}$ or $\{B, C\}$ doesn't matter. As a result, the only motivations that matter in choosing between BAC and BCA that matter are the ones related to the relative order of A and C, and if we

know they are impartial, $\{A, C\} \in I$, we know they will report honestly in a simple choice between those two rankings.

A similar concept is indifference pairs, and partially indifferent preferences. The definition 7
is almost identical to impartial pairs, but instead of creating a strong preference to reveal the ranking, they are indifferent. This could be considered a true indifference, or you could consider it as a bias weaker than the jurors preference towards impartiality. Thus, when forced trade off being impartial or being weakly biased towards a pair in the indifference pairs, they will choose to be impartial. I denote the set of indifference pairs for juror j as V_j.

Definition 2 (Partially Weakly Indifferent Preferences)**.** *Preferences R_j are partially-indifferent for all alternatives $\{A, B\} \in V_j$ if: for any rankings $\{\pi, \hat{\pi}\} \in \Pi$ for juror j, then $\pi I_j(\pi_t)\hat{\pi}$ if:*

- $p_A^{\pi} = p_B^{\pi} - 1$
- $p_A^{\hat{\pi}} = p_B^{\hat{\pi}} + 1$
- $p_C^{\pi} = p_C^{\hat{\pi}}$ *for all* $C \in N \setminus \{A, B\}$

Denote the set of Partially Indifferent Preferences as $\mathcal{V}_j \subset \mathcal{R}_j$ for any $j \in J$.

8

The above definition does not allow learning like the impartial pairs. If $\{A, B\} \in V_j$, then 9
$ABC \ \ I_j(\pi_t) \ \ BAC$ but the ranking in the true ranking can be either $p_A^{\pi_t} < p_B^{\pi_t}$ or $p_A^{\pi_t} > p_B^{\pi_t}$. For example, let the true ranking be ABC, and the juror to be indifferent on the contestant pair of $\{A, B\} \in V$. Then if the juror is given a choice over $\{ABC, BAC\}$, he would be indifferent between the two. He could choose ABC or BAC just as easily. Likewise, suppose he chooses ABC, that doesn't tell us the true ranking is ABC, it could just as easily be any other ranking, such as CBA, which is completely different. Thus, the social planner can't use choices using only indifferent pairs to learn anything, and for this paper we treat indifferent pairs as bias weaker than impartiality, since our goal is to guarantee elicitation of the true ranking.

The solution to the problem is two fold, first, the designer needs restrictions on jurors 10
preferences, second, he needs to have an optimal mechanism for elicitation. An optimal mechanism for two juror elicitation does not exist previously for this problem, Amorós (2009) found an optimal mechanism for the case of $|J| \geq 3$, but he also shows by proposition 3 in appendix A that his mechanism is insufficient for $|J| = 2$. The next section will consist of showing the design of the mechanism used in this paper, and how it is optimal.

The **State of the World** is the vector of true ranking and preference functions given 11
by $e = (\pi_t, R_1, R_2) \in \Pi \times \mathcal{R}_1 \times \mathcal{R}_2$. Let $\mathcal{S}$ be the restricted universe on which we get implementation, formally $\mathcal{S} \subset \Pi \times \mathcal{F}_1 \times \mathcal{F}_2$. Here, $\mathcal{F}_i$ is the set of preferences that satisfies both juror i's impartiality and indifference restrictions, formally $\mathcal{F}_i = \mathcal{U}_i \cap \mathcal{V}_i$.

A **mechanism** is the vector $\Gamma = (T, >, D, \delta, \omega)$. A node in the extensive form game is 12
$t \in T$. I will assume there is only 1 root node, and label it t_0. The order of the nodes is captured by $>$, such that if $t_4 > t_2$, that means t_4 is on a lower tier of the tree than node 2. The set of possible choices is D, which will be $D = \Pi$ since all rankings could be the true ranking. As a function $\delta : \{T - t_0\} \to T$, delta takes each node and maps to the node from which it departed. For example, $\delta(t_1) = t_0$ means that node t_1 was chosen at node t_0. This allows us to construct the choices at each node by $\delta^{-1}(t_0) = \{t \in T | \delta(t) = t_0\}$, where $\delta^{-1}(t_0)$ indicates the choices available at node t_0, $\delta^{-1}(t_1)$ would be choices at t_1, et cetera. To clean up notation, let $D(t_n) \equiv \delta^{-1}(t_n)$ be my conditional choice set at node n. Finally, ω maps the terminal nodes into outcomes. This is a direct mechanism such that the terminal nodes are rankings, which are the outcomes, such that ω is simply an identity map. For ease, the first to move in the mechanism will be juror 1, and the second, and last, to move will be juror 2.

Let $S_i(t)$ be juror i's strategies available at node t. If we observe $|S_i(t)| > 1$ and $|S_j(t)| >$ 13
1, then jurors i and j are moving simultaneously at node t, whereas if $|S_i(t)| > 1$ but $|S_j(t)| = 1$ for any juror j not i, then only juror i moves at node t. For the most part, all moves will sequential. Let the space of juror i's strategies be $S_i = \times_{t \in T} S_i(t)$, the space of all

strategies from each node. Then, the strategy space for all jurors would be $S = S_1 \times S_2$. For any $s \in S$ and $t \in T$, let $\omega(s|t)$ be the outcome be achieved using strategy s starting from node t.

A strategy $s \in S$ is a **subgame perfect equilibrium** for the mechanism Γ and state 14
of the world $e \in \mathcal{S}$ if $\forall i \in J, \forall t \in T, \forall \hat{s}_i \in S_i, \omega(s_i, s_{-i}|t) R_i(\pi_t) \omega(\hat{s}_i, s_{-i}|t)$. Where its worth keeping in mind that s_{-i} is all other jurors strategy, which is determined by choices at each node. Thus, s_i includes responses at nodes off the equilibrium path. Denote the space of SPNE strategies as $SPE(\Gamma, e)$ and the outcomes as $\omega(SPE(\Gamma, e)|t_0)$.

An extensive form game Γ subgame **implements** the true ranking π_t in state $e \in \mathcal{S}$ if 15
$\omega(SPE(\Gamma, e)|t_0) = \pi_t$, for all $e \in \mathcal{S}$.

3 Mechanism

With two jurors, there are few choices for optimal mechanisms. The choice of a dicta- 1
torship would not work in this setting, because we don't want the jurors to reveal their true preferences, but the true ranking, which won't occur under a dictatorship unless some juror is fully impartial. I exclude the possible case of a fully impartial juror, because it makes the second juror redundant. Therefore, I will instead use a two-turn extensive form game, where juror 1 chooses a message, and juror 2 observes juror 1's message and chooses a message in response. This is known as a Stackleberg solution, and since I have complete information, I can guarantee that a sequential equilibrium exists. Using longer extensive form games can be useful for implementation sometimes, but that involves a social choice correspondence that is attempting to enforce compliance, in my case, I could not do so because the designer is ignorant of R_i. Further more, since the game is perfect information for the jurors, they could backwards induct the outcome of their choices, and the designer has nothing to ask them about besides the true ranking, such that having each juror make more than 1 choice

is unnecessary. Proof of this to be shown later.

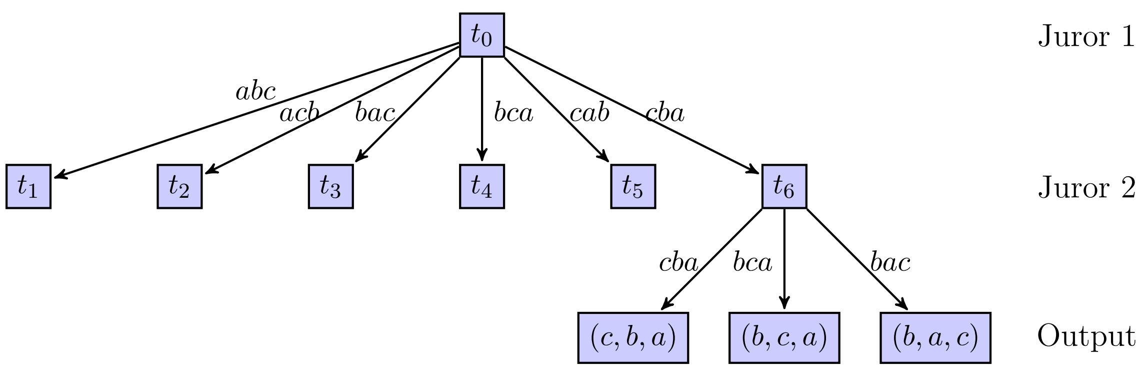
Since the "optimality" of a mechanism depends on the scope of the environment, let me 2
introduce some initial assumptions that restrict our choice of mechanism. First, no juror has all possible impartial pairs, or stated differently, no juror is fully impartial **(A1)**. If you had a juror who was fully impartial, this satisfies the requirements for implementability with only one juror, such that any additional juror is unnecessary to consult. Second, any impartial pair is assigned to a juror **(A2)**. If this is not the case, that means there are two contestants in a ranking no judge is willing to place in the correct order, which makes implementing the true ranking impossible. Third, the Designer needs to know for each pair of contestants at least 1 juror who has them as an impartial pair **(A3)**. If the juror does not know which juror is impartial over a pair of contestants, he does not know whom he can trust to elicit that information from. All three assumptions are discussed in Amorós (2009), where the last two are shown to be necessary to get implementation, but not sufficient in the two juror case.

Now if we attempt to look at other mechanisms, it can be shown that they will violate at 3
least one of the three base line assumptions. If we consider the Dictatorship, we suppose you randomize and juror 2 is chosen such that his choice of ranking is the mechanism's choice. The choice of the mechanism can be guaranteed to be the true ranking only if juror 2 is fully impartial. Only if he is fully impartial can you ensure that the top of $R_2(\pi_t)$ is the true ranking. In the deferred acceptance algorithm the problem is similar, and it stems from the fact in this problem we want the true ranking, and the relation of jurors preference over true ranking is independent of the SCC. Or in other words, the true ranking does not cease to be the true ranking if it becomes less popular with the jurors; the jurors opinions of the truth does not change your objective from figuring out the truth.

The next mechanism would be the Maskin mechanism, or a slight variation on it. It is 4
much more subtle to show, but it is also not the optimal mechanism choice in this environment. The Maskin mechanism handles multiple intractable messages by having the jurors

play the integer game, a game with no equilibrium. With two jurors, anytime the jurors do not send the exact same ranking as a message, then we would be playing the case with no equilibrium. As a result, unilateral deviations are not equilibrium play. Strictly speaking, that means any ranking could be equilibrium. However, in this setting with both jurors almost completely impartial, each just missing one impartial pair, it is still possible for there to be rankings in common in both jurors upper contour sets of the true ranking. Thus, if the true ranking was an equilibrium, these rankings in common in the upper contour set would also be equilibrium, violating SCC implementability.

Thus, the mechanism I use will be structured as a simple Stackleberg Solution, also 5
known as a two-turn extensive form game. Juror 1 will move first, choosing a message at node t_0 from his message space $D(t_0)$, then juror 2 observes juror 1's message, and chooses a message from the resulting node t_n. Only juror 1 has to consider a implications of message choice, since juror 2 goes last after observing juror 1's choice. The result of the mechanism is juror's 2 choice.



The above image shows an example of how the mechanism works. Rather than draw out 6
all of the terminal nodes, I simply showed one path of play. In this instance, the first juror chose the ranking (c, b, a) as their message. This forces juror 2 to make a choice amongst the selection at t_6, which can be $(c, b, a), (b, c, a)$, or (b, a, c). Whichever ranking juror 2 chooses will be the output of the mechanism. Given juror 2's impartial pairs, we can ensure he reports the true ranking if it is among his choices by restricting his set of choices, just as we did above

by restricting him to choosing among just three rankings. A jurors **choice** of rankings in set β will be defined as the correspondence $C_j(\beta) \equiv \{\pi \in \beta | \nexists \pi' \in \beta, \text{ such that } \pi' P_j(\pi_t)\pi\}$.

Ultimately, since whichever juror 2 chooses is the outcome, we want his choice to be the 7
true ranking, which means he chooses the true ranking whenever he is able. I will define a set $\beta(\pi, U_j)$ as the set from which we can assure that the true ranking will be chosen if it is present. Since the information available to the Designer is the impartial sets, we design this set based on the impartial sets, and the message juror 1 sends.

Definition 3. *Let $\beta(U,V) = \{\alpha \subset \Pi | \forall e \in \mathcal{S}, \pi_t \in \alpha \rightarrow C(\alpha | U \cup V) = \pi_t\}$ is a collection of sets of rankings. Each element $\alpha \in \beta(U,V)$ is a set of rankings, such that if the true ranking is in the set α, then juror 2 will choose the true ranking when his choice is restricted to set α.*

The sets in β will be important, because if we want the resulting terminal node of 8
the mechanism to be a true ranking, then the last juror has to be making a choice from $D(t) \subset \alpha \in \beta$. As it turns out, by using impartial pairs we can always create sets in β, however, without restrictions some $\alpha \in \beta$ can be singletons - which would be a problem. However, while the sets in β will be an important set to look at for implementation, it is not sufficient. The following shows that for any proper subset of contestant pairs, there will always be an $\alpha \in \beta$, such that there is a ranking not in that α. That ranking could always be the true ranking, π_t. As a result, juror 1's message becomes necessary to choose a set α from which juror 2 is to choose, and juror 1 can always choose an α where the true ranking is not available to juror 2.

Proposition 1 (Manipulation of juror 2's Choices)**.** *For any $U_j \neq N^2$, then for any $\alpha \in \beta(U_j \cup V_j)$, there is a ranking $\pi \in \Pi$ such that $\pi \notin \alpha$.*

However, some good news comes from the fact that the a two turn extensive form game 9
is optimal. So long as a juror isn't what I call *mostly honest*, which occurs if $U_i \cup V_i = N^2$,

then adding more turns simply makes it easier for jurors to manipulate one another - making it harder to get the true ranking. What $U_i \cup V_i = N^2$ means, is that if a juror is not impartial over a pair of contestants $\{A, B\}$, then they are indifferent[2]. This is a restrictive case, one in which I can get Nash Implementation, thus, for all practical purposes, a two - turn extensive form game is optimal.

Proposition 2 (Two-turn Optimality)**.** *If for both $j \in J$, $U_j \cup V_j \neq N^2$, then the two turn mechanism used in this paper is optimal compared to any mechanism using more turns where any juror goes twice or more.*

The proofs for above propositions are in the appendix. The following are three functions 10
useful in the proofs. The first is the set of swapped pairs, accounting for all possible rankings achieved from π by stringing together swaps of adjacent pairs over which juror j is impartial or indifferent. This set is important because the set $\alpha \in \beta(U_j, V_j)$ always contains it, and is often equal to it.

Definition 4 (Set of Swapped Pairs)**.** *Let $X(\pi, U_i) \subset \Pi$ be the set of all rankings achieved by swapping any or all adjacent pairs in U_i, in any order.*

The set of Misranked Pairs satisfies a lot of properties. It is injective, well defined, but 11
not surjective. It finds all the contestant pairs that are misranked, and for any two rankings, this collection of contestant pairs is unique. This is useful in determining the necessary distribution of impartial pairs, or the above two propositions.

Definition 5 (Set of Misranked Pairs)**.** *Let $Z(\pi, \pi_t) \subset N^2$ be the set of all contestant pairs misranked when comparing π to π_t. Formally this is written as $Z(\pi, \pi_t) \equiv \{\{A, B\} \in N^2 |$ if $p_A^{\pi_t} < p_B^{\pi_t}$ then $p_A^{\pi} > p_B^{\pi}\}$.*

The Set of Partially-Honest Rankings is probably the most important set. It is a set of 12
all rankings, such that juror j has all the contestant pairs in his impartial pairs U_2 ranked

[2] With these restricted preferences and mechanism, this is the same as the juror is not a strategic liar.

similarly to the true ranking, π_t. The set can range from the size of half of all the rankings, to only the true ranking if juror fully impartial (which we exclude by assumption). Later results will show that if we can get just one juror to report a ranking from the partially-honest set, then the other juror will want to as well. When both jurors report from their partially-honest sets respectively, then the only outcome is the true ranking. So the partially-honest set is integral, but it is not incentive compatible, wherein lies the problem.

Definition 6 (Set of Partially Honest Rankings)**.** *Formally defined as* $H(\pi_t, U) = \{\pi \in \Pi | \forall \{a, b\} \in U, p^a_{\pi_t} < p^b_{\pi_t} \rightarrow p^a_\pi < p^b_\pi\}$*, the set* H *is the set of all rankings where contestants in the juror's impartial pairs are in the same order as the true ranking.*

4 Results

I will continue with information requirements established as necessary for all equilibrium 1
concepts by Amorós (2009). He required all impartial pairs to be assigned to a set of jurors, and the designer must know which judge has which impartial pair.[3] I will also look at the strict case when a single juror does not have all the impartial pairs - since the solution in such a case would simply be to implement the fully impartial juror's most preferred ranking. Since I have only two jurors, then any alternative not in juror i's impartial collection must be in the other juror's collection.

Assumption 4 (Transitive over sets in β)**.** $\forall \pi, \pi', \pi'' \in \alpha \in \beta(U_j \cup V_j)$*, if* $\pi R_i(\pi_t)\pi'$ *and* $\pi' R_i(\pi_t)\pi''$ *then* $\pi R_i(\pi_t)\pi''$*.*

The assumptions that I will look at to characterize the implementation in this environ- 2
ment, in addition to Amorós results, include partial transitivity such that jurors are partially rational. This is a fairly standard assumption, which will be useful to establish consistency

[3] Notice that maskin monotonicity, a necessary condition for implementation in nash equilibrium, is satisfied by all impartial pairs being assigned to a set of judges.

of choices with preferences.

Assumption 5 (Non-Partial over Alternative)**.** *If* $\{A, B\} \in U_i$*, then* $\forall X \in N_{-A}, \{A, X\} \in U_i \cup V_i$*.*

To prevent the first juror from manipulating the last juror requires some restrictions on 3
the distribution of the last juror's impartial pairs, as well as partially-indifferent preferences. The last juror being non-partial over an alternative is a fairly weak restriction, that boils down to a juror not being partial over the ranking of an alternative and any other one.

Assumption 6 (Condition-λ)**.** *If* $\{A, B\} \in U_j$ *then* $\{A, X\} \in V_2$*, for any* $X \in N_{-B}$*.*

An intuitive version of above is that if juror j is impartial between contestants A and 4
B, then he is more impartial over A and B than he is over impartial or bias over any other contestant pair containing A. This precludes juror j from having $\{A, C\}$ and $\{B, C\}$ as an impartial pair, forcing them to be indifferent pairs. This looks like its weakening juror j, but in fact makes him harder for juror 1 to manipulate.

4.1 Necessary and Sufficient Characterization

For the necessity results I will not be able to take advantage of many features in other 1
works, such as Moore and Repullo (1990). In their work they have a known ranking π that is less preferable than π_t for all environments for both jurors. However, my results resonate with the necessity that deviating choices have to lead to subgame perfect equilibrium outcomes that is less favorable for the deviator than the true ranking. By satisfying this condition, I will satisfy Condition C from Moore and Repullo (1988), which is the necessary condition that must be satisfied for Subgame perfect implementation in a general environment. Further, in a two juror case there is no way to tell who is deviating. In a three juror case this is handled by the fact the one juror with a different message is ignored. Thus, implementation

for three juror implementation needs to find coalition proofness restrictions, while two juror case must also find preference restrictions to prevent unilateral deviation in addition. These restrictions will need to force consequences for deviating, since the deviators are unknown to the social planner.

To avoid the difficulty to ensure truth telling of π_t as a message space for both jurors, my 2
mechanism tries to get a specific one of them to report π_t, the last guy. We will see in the necessity proof this is necessary. In what follows I will show that assumptions A5 and A6 are necessary, as well as restricting the last juror's choice set to $D(t) = \alpha$ for some $\alpha \in \beta$, $\forall t \in T_{-\{t_0\}}$. I will also show an irreducible condition I call condition-λ, which cannot be satisfied with restrictions only on impartial pairs.

Let us define the set of admissible states of the world $\mathcal{S}$ as the set of all states of the world 3
where jurors 1 and 2's preference satisfy assumptions A1 through A4, and juror 2's preference also satisfy A5 and A6. Further, if we restrict the mechanism such that $\forall t \in T_{-\{t_0\}}$, there exists some $\alpha \in \beta$ such that $D(t) = \alpha$, then we have our necessary results. Thus if $e \notin \mathcal{S}$ we can not guarantee SP implementation. The proof below proceeds more cautiously to show the necessity of each part.

To begin, I will show a rather innocent looking lemma. The lemma's results show that 4
the sets in β and the set of partially-honest rankings can only have one ranking in common. The problem with the partially-honest set is that the most preferred ranking in the set is most likely not the true ranking. As a result, if juror 2 had a choice over multiple partially-honest answers, with one being the true ranking, there are states of the world where he would deviate. With the sets in β we can avoid that problem for juror 2, but not juror 1.

Lemma 1. *If juror 2 is choice set is restricted to $D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U)$, then for any ranking juror 1 chooses, juror 2 can have at most one ranking in $D(T_{-\{t_0\}})$ that is a partially-honest message.* $\forall U \subset N^2, \forall t_n \in \Pi, |D(t_n|U) \cap H(\pi_t, U)| \leq 1.$

Now, I will proceed to solve the mechanism backwards, and show what it takes to get 5 juror 2 to be incentive compatible. It turns out we can get juror 2's cooperation practically for free. We just need to restrict his choices to sets in β and require him to be transitive over the sets in β so he makes rational choices on the set. By forcing juror 2 to makes decisions on β, juror 2 will always report a partially-honest message when he is able. This is useful, because if juror 2 reports a partially-honest message, but juror 1 does not, than juror 1 will prefer the true ranking to the ranking that resulted from their choice. The following proof shows that if one juror chooses a partially-honest message for themselves, then the other one will choose the true ranking, if we restrict the responders choice set to sets in β.

Proposition 3. *If both jurors satisfy A4, and one juror chooses a partially honest message, then* $X(\pi, U_j) \subset D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U_j)$ *iff* $C_j(D(T_{-\{t_0\}})) = \pi_t$*, where* $D(T_{-\{t_0\}})$ *is juror j's choices given juror i chose* π*.*

It is important for the success of the mechanism for both agents to report partially-honest 6
messages. Because of A1 and A2, the true ranking is where both juror's set of Partially-honest message overlap. However, while it is easy to get juror 2 to be cooperative, it is not easy to get juror 1 to be cooperative. However, If we can ensure juror two can always send a partially-honest ranking, then juror 1 would prefer the result if he had sent one a message where he was being partially-honest, i.e. he would prefer the true ranking to the ranking outcome where he lied, and the juror 2 was honest.

Lemma 2. *If the jurors satisfy A4, Juror 1 reports a message* t_n *such that* $\pi_t \in D(t_n|U)$ *iff juror 2 can always choose a partially-honest ranking:* $\forall e \in \mathcal{S}, \forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, R_1(\pi_t)) \neq \emptyset \iff C_2(D(t_n|U_2)) \cap H(\pi_t, U_2) \neq \emptyset$.

A simple illustration. Let $\pi = CBAD$. Suppose $\pi_t = ABCD$. then A is the first mis- 7
ranked contestant. The pairs between A and C that are misranked are $\{A, B\}, \{A, C\}, \{B, C\}$, whereas all of the impartial pairs including contestant D are ranked properly. Since we know juror 2 was partially-honest, that means all the misordered pairs are in U_1. Which means

that juror 1 can swap C and B to get $BCAD$, then swap A and C to get $BACD$, and finally swap A and B to get $ABCD$. If we had different true ranking, we would get different ownership of the impartial pairs. Suppose that $\{A, B\}, \{A, C\} \in U_1$, then the others are in U_2 and are ranked correctly, which means the true ranking must be either $ACBD$, $CABD$, or $CBAD$, depending on which pairs in U_1 are misranked. In sum, if juror 2 is honest, juror 1 could always swap his pairs to be honest.

Finally, to get juror 2 to always be able to send an partially-honest ranking to the 8
mechanism for output, takes a bit of restriction on the distribution on juror 2's impartiality. In addition to the preference restrictions the two agents have in common with A1 through A4, juror 2 will also need to satisfy assumptions A5 and A6. These two restrictions combine will prevent juror 1 from being able to manipulate juror 2's choice set from sets in β, where juror 2 has no partially-honest ranking to choose. By A5, juror 1 will not be able to lie and place a contestant in juror 2's way so that contestant pairs in $U_2 \cup V_2$ are never adjacent, and thus can never be swapped by him. By A6, juror 1 can never force juror 2 to face a tradeoff in satisfying one impartial pair at the expense of another. Taken together, A5 and A6 make it impossible for juror 1 to choose a choice set for juror 2 that doesn't contain any ranking that is partially-honest for him.

Proposition 4. *Juror 2 can always choose a partially-honest message iff his preferences satisfy A2 through A6.*

Theorem 1 (Implementation)**.** *The mechanism Γ subgame implements the true ranking iff Γ is the Stackleberg Solution and $D(T_{-\{t_0\}}) \subseteq \alpha \in \beta$, jurors 1 and 2's preferences satisfy A1 through A4, and juror 2 satisfies A5 and A6 in addition or some juror j is non-strategic.*

Proof. $(\rightarrow)$

If we subgame implement the true ranking all the time, that means the jurors are always being at least partially-honest every time. By lemma 2, juror 1 only reports a partially-honest message if juror 2 always can. By proposition 4, juror 2 can always report a partially-honest message if he satisfies assumptions 1 through 6. By proposition 3, juror 1 must satisfy A4, and juror 2 choices must be restricted to sets $\alpha \in \beta(U_2, V_2)$ if juror 1 to prefer sending a partially-honest ranking. As result, if we always get implementation, than we must satisfy restricting juror 2's choice set to $\alpha \in \beta(U_2, V_2)$, and jurors 1 and 2 satisfy A1-A4, and juror 2 satisfy A5 and A6 as well.

$(\leftarrow)$

By proposition 3, we restrict juror 2's choice set to $\alpha \in \beta(U_2, V_2)$, and by proposition 4, with A1-A6 restrictions on juror 2's preference, we can ensure he always can choose a partially-honest message. By lemma 2, this means juror 1 will also choose a partially-honest message. As a result, when both be partially-honest messages respectively, than the resulting outcome is the true ranking. □

Finally, unlike Amorós (2009), this paper takes advantage of partially-indifferent prefer- 9
ences in addition to partially-impartial preferences. This turns out to be necessary, because A6 explicitly creates atleast 1 indifference pair for juror 2 in conjunction with A5. Therefore, because these two assumptions are necessary, it is impossible to garuantee implemententation, even subgame implementation, with only impartial pairs.

Corrolary 1 (Impossibility). *A6 and A5 is necessary, therefore it is impossible to implement using impartial pairs alone.*

4.2 Nash Equilibrium Implementation

The following are sufficiency results to get nash implementation. In the context of this 1
environment, an agent is non-strategic if over any contestant pair he is either impartial or indifferent. Intuitively, this means he prefers being partially-honest over being a strategic

liar at any point, but does not mean he is unbiased.

Definition 7 (Non-Strategic)**.** *Juror j is non-strategic if $U_2 \cup V_2 = N^2$.*

It is straightforward to see that an agent that is non-strategic satisfies A5, but it looks 2
to breach A6. That is because A6 is necessary for $\alpha \in \beta(U_j, V_j)$ to be non-manipulatable. But if a juror is non-strategic, that means they will report partially-honest, where they don't manipulate, and thus we can let them go first. In effect, we can use a different mechanism. In fact, they could move simultaneously and we can still get implementation. Implementation under bias, but not strategic liars anymore.

Remark 1. *Its clear that A5 and A6 are weaker than some juror j being non-strategic, but does being non-strategic satisfy A5 and A6? The answer is no, but it could. Consider the example of three contestants ABC, then if the juror j has $\{A, B\}, \{B, C\} \in U_2$ and $\{A, C\} \in V_2$, then this is non-strategic, but it doesn't satisfy A6 since $\{A, B\} \in U_2$ means $\{B, C\}$ should not be impartial, but indifferent, and its not. However, if we set $\{B, C\} \in V_2$, then it is still non-strategic, but it also now satisfies A6.*

Theorem 2 (Nash Implementation)**.** *If jurors 1 and 2 satisfy A1 through A4, and one juror is non-strategic, we can get Nash implementation.*

Proof. Forthcoming. Must show well-defined outcome function. □

I do not know if this condition is necessary. It likely is, which would make this a negative 3
result.

5 Conclusion

So, with only two jurors, can we get implementation of the true ranking? Yes, using 1
subgame implementation. But to do so, requires weak restrictions on juror 1, forcing him to

satisfy the same assumptions shown to be necessary in Amorós (2009), plus transitivity. Juror 2 would need to satisfy restriction on the distribution of his impartialness and indifference in addition. He would have to contain all of the indifference pairs for a contestant, if he is impartial over the contestant. How does this connect to the example of the two doctors and their ranking of drugs for an ulcer?

Well, we can get information on the the doctors sponsorship from Physician Payment 2
Sunshine Act. Thus, we can ensure that the doctors satisfy A1 through A4, considering drugs that both have sponsors to be an impartial pair, and two drugs that don't have any sponsor to be an impartial pair as well. But to satisfy A5 and A6, means we know there is one drug that the doctor wants to make sure is ranked correctly. That is unlikely. But at the end of the day, most of us probably get very close to the true ranking of the best drugs for our condition.

How so? The truth is, while more restrictive, the non-strategic requirement for Nash 3
implementation is more universally applicable. If we consider a difference between telling a lie, and abusing someone's trust, then we can make the difference between being bias (lie) and being strategic. In the case of doctors, they got in the business to help people, so are unlikely to be strategic at their patients expense, but can still be biased because drug sponsorship also provides more information that might make the drug more appealing to the doctor. Thus, while the results are negative in general, much as Amorós (2009), or any paper looking eliciting information from biased jurors or advisers, there is some good news when one juror is non-strategic.

References

Amorós, Pablo, "Eliciting socially optimal rankings from unfair jurors," *Journal of Economic Theory*, May 2009, *144* (3), 1211–1226.

Jackson, Matthew O. and Sanjay Srivastava, "On two-person Nash implementable choice functions," *Social Choice and Welfare*, 1992, *9*, 263–264. 10.1007/BF00192882.

Krishna, Vijay and John Morgan, "A Model Of Expertise," *The Quarterly Journal of Economics*, May 2001, *116* (2), 747–775.

Maskin, Eric, "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, January 1999, *66* (1), 23–38.

Moore, John and Rafael Repullo, "Subgame Perfect Implementation," *Econometrica*, September 1988, *56* (5), 1191–1220.

_ and _ , "Nash Implementation: A Full Characterization," *Econometrica*, September 1990, *58* (5), 1083–99.

Wolinsky, Asher, "Eliciting information from multiple experts," *Games and Economic Behavior*, October 2002, *41* (1), 141–160.

6 Appendix

The Following are results with their proofs.

Proposition 1 (Manipulation of juror 2's Choices). $\forall U \subset N^2, \exists \alpha \in \beta(U), \exists \pi \in \Pi,$ *such that* $\pi \notin \alpha$.

Proof. By $U \subset N^2$ this implies $\exists \{A, B\} \notin N^2 \setminus U$. Then taken, we get $\exists \alpha, \alpha' \in \beta(U)$ such that the difference between α and α' is that their exists $\pi \in \alpha \setminus \alpha'$ and $\pi' \in \alpha' \setminus \alpha$, where $p_A^\pi < p_B^\pi$ and $p_A^{\pi'} > p_B^{\pi'}$.

Since, if $\pi, \pi' \in \alpha$, then to get a contradiction, just choose $e \in \mathcal{S}$ such that $\pi' R_j(\pi)\pi$ and $\pi = \pi_t$. Then $C_j(\alpha) \neq \pi_t$ which contradicts $\alpha \in \beta(U)$ because $\alpha \in \beta(U)$ is defined such that $C_j(\alpha) = \pi_t$.

Therefore, $\alpha \neq \alpha'$ which implies $\exists \pi \notin \alpha'$ and $\exists \pi' \notin \alpha$ which is what we wanted to prove. □

Proposition 2 (Two-turn Optimality). *If for any $j \in J$, $U_j \cup V_j \neq N^2$, then the two turn mechanism used in this paper is optimal over any mechanism using more turns where any juror goes twice or more.*

Proof. Let juror 2 go last. Since juror 2 goes last, his last choice will be from some $\alpha \in \beta(U_2 \cup V_2)$ and so if he goes before juror 1, and after, his first choice is among two different $\alpha, \alpha' \in \beta(U_2 \cup V_2)$. First, 2 cases:

Case 1: $U_j \cup V_j \neq N^2$, for all $j \in J$.

Take $\alpha, \alpha' \in \beta(U_2 \cup V_2)$ such that $\alpha \neq \alpha'$, then there exists $\pi' \neq \pi_t$ such that $\pi_t \in \alpha$ and $\pi' \in \alpha'$ and $\pi_t, \pi' \notin \alpha \cap \alpha'$. Since the state of the world is suppose to be arbitrary, we can choose $e \in \mathcal{S}$ such that $C_2(\alpha) = \pi_t$ and $C_2(\alpha') = \pi'$ and $\pi' P_2(\pi_t)\pi_t$ such that:$C_2(C_2(\alpha), C_2(\alpha')) = C_2(\pi_t, \pi') = \pi'$.

Thus, we do not implement. We could fix this by forcing all possible resulting terminal nodes from juror 2's first choice to belong to either α or α' exclusively. But if juror 2 goes first, then we can not choose between the two sets otherwise than randomizing. Thus, we let juror 1 go before juror 2, then we can use juror 1's choice to choose between α or α', which is the same as the two-turn mechanism except with more turns. As a result, because of the juror's being perfectly informed, it is their very first choices each make, that matters to the resulting outcome. The restrictions to get implementation on each jurors first choice, is the same in Stackleberg Solution mechanism, where each only goes once.

Case 2: $U_2 \cup V_2 = N^2$

Under this restriction we can get implementation with agent 2 going first and last, thus more than two turns, however, this result is also sufficient to get Nash Equilibrium Implementation, which is shown elsewhere. Thus, because we get the stronger NE implementation, we wouldn't worry about a many turn mechanism and its SPN Implementation. □

Lemma 1. *If juror 2 is choice set is restricted to $D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U)$, then for any ranking juror 1 chooses, juror 2 can have at most one ranking in $D(T_{-\{t_0\}})$ that is a partially-honest message.* $\forall U \subset N^2, \forall t_n \in \Pi, |D(t_n|U) \cap H(\pi_t, U)| \leq 1$.

Proof. Proceed by way of contradiction: let $|D(t_n|U) \cap H(\pi_t, U)| \geq 2$. Then by $D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U)$ this implies that there exists $\alpha \in \beta(U)$ such that there is two rankings $\pi, \pi' \in \alpha$ which are ranked as is the true ranking, i.e. if $p_A^{\pi_t} < p_B^{\pi_t}$ then $p_A^{\pi} < p_B^{\pi}$ and $p_A^{\pi'} < p_B^{\pi'}$. Then as a result, it becomes very easy to get a contradiction. Choose $e \in \mathcal{S}$ such that $\pi_t = \pi$ and $\pi' P_j(\pi_t)\pi_t$, which is expected with bias advisors, and you get that $C_j(\alpha) \neq \pi_t$, a contradiction of how β defined. Therefore, $|D(t_n|U) \cap H(\pi_t, U)| \leq 1$.

□

Lemma 2. *The function Z is well-defined, injective, but not surjective.*

Proof. To see that Z is well defined: Take two rankings $\pi, \pi' \in \Pi$, where $\pi \neq \pi'$. Wlog, let A be the first contestant that is in a different position for the two rankings, such that $P_A^{\pi} \neq P_A^{\pi'}$. Let $P_A^{\pi} < P_A^{\pi'}$, then this implies there is some $B \in N$ such that $P_A^{\pi} = P_B^{\pi'}$, which therefore implies that $P^{\pi}_A < P^{\pi}_B$ and $P^{\pi'}_B < P^{\pi'}_A$. Thus, $\{A, B\} \in n$, where n is $Z(\pi, \pi') = n$. Thus, if the two rankings are different, then the function has a nonempty solution. Further, for every two rankings it has only one solution.

Let there be two rankings such that $Z(\pi, \pi') = \{n, n'\}, n \neq n'$. Being different, this implies $\exists\{A, B\} \in N^2$, such that $\{A, B\} \in n' \setminus n$, which means that if $P^{\pi'}_A < P^{\pi'}_B$ and $P^{\pi}_A < P^{\pi}_B$ by n, then by n', $P^{\pi}_B < P^{\pi}_A$. But π can not have both that $P^{\pi}_A < P^{\pi}_B$ and $P^{\pi}_B < P^{\pi}_A$ by n and n', therefore a contradiction. Thus, $n = n'$, and Z is in fact, a well defined function.

To show that Z is injective, I will do so by way of contradiction. Suppose $\exists n \subset N^2$ such that $\forall \pi \neq \pi' \in \Pi$, we get $Z(\pi, \pi_t) = n$ and $Z(\pi', \pi_t) = n$. But this would imply that all contestant pairs not in n are ranked the same as the true ranking: i.e. $\forall\{A, B\} \in N^2_{-\{n\}}$, if $p^{\pi_t}_A < p^{\pi_t}_B$ then $p^X_A < p^X_B$ for $X = \{\pi, \pi'\}$.

Likewise, for any contestant pair in $\{C, D\} \in n$, $p^{\pi_t}_A < p^{\pi_t}_B$ then $p^X_A > p^X_B$ for $X = \{\pi, \pi'\}$. Therefore, if we take π there will be a contestant $A \in N$ such that its first: $\forall B \in N, p^{\pi}_A < p^{\pi}_B$, and if $\pi \neq \pi'$ there has to be a difference between the two rankings, wlog let it be what is first: Then this implies that there is some $B \in N, B \neq A$ such that B is first in π'. But this implies $p^{\pi}_A < p^{\pi}_B$, but $p^{\pi'}_B < p^{\pi'}_A$. Thus, if $p^{\pi_t}_A < p^{\pi_t}_B$, then $Z(\pi, \pi_t) = n, \{A, B\} \notin n$, whereas $Z(\pi', \pi_t) = n', \{A, B\} \in n'$ such that $n \neq n'$ a contradiction. Thus the map Z is injective.

Counter example to show not Surjective:

Consider a ranking $\pi = ABC$, and $n = \{A, C\} \in N^2$. Then there doesn't exist a ranking $\pi' \in \Pi$ such that $Z(\pi', ABC) = \{\{A, C\}\}$. Because if A and C are misranked, then some other contestant pairs must also be misranked. As a result, the lone pair does not have a pre-image in the domain. And this problem is not consistent as we swap π for other rankings. Therefore, the function is not surjective.

$\square$

Proposition 3. *If one juror chooses a partially honest message, then* $X(\pi, U_j) \subset D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U_j)$ *iff* $C_j(D(T_{-\{t_0\}})) = \pi_t$*, where* $D(T_{-\{t_0\}})$ *is juror j's choices given juror i chose* π*.*

Proof. I will show ($\rightarrow$) first.

Suppose $\pi \in H(\pi_t, U_i)$ and $X(\pi, U_i) \subset D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U_j, V_j)$. I want to show that $C_j(D(T_{-\{t_0\}})) = \pi_t$.

Notice, that for each $\pi \in \Pi$, $Z(\pi, \pi_t)$ is unique. Thus, by $\pi \in H(\pi_t, U_i)$, and by $X(\pi, U_i) \subset D(T_{-\{t_0\}})$ this implies that $H(\pi_t, U_i) \subset D(T_{-\{t_0\}})$, which means that for any $\pi \in D(T_{-\{t_0\}})$, that $Z(\pi, \pi_t) \subset U_j$.

Thus, if $C_j(D(T_{-\{t_0\}})) = \hat{\pi} \neq \pi_t$ then $\exists n \subset U_2, Z(\hat{\pi}, \pi_t) = n$. Now because agent ones message was honest, that means $\pi_t \in D(T_{-\{t_0\}})$, which means there exists $\pi^1 \in X(\pi, U_j), Z(\pi^1, \pi_t) = n_{-\{A,B\}}$ such that $\pi^1 P_(\pi_t)\hat{\pi}$.

And we can repeat this step: $\exists \pi^2 \in X(), Z(\pi^2, \pi_t) = n_{-\{\{A,B\},\{D,C\}\}}$ such that $\pi^2 P_J(\pi_t)\pi^1$. Continue this until we have a $\pi^n \in X()$ such that $Z(\pi^n, \pi_t) = \emptyset$, then $\pi^n = \pi_t$ and we can construct a change of the rankings as $\pi^n P_j(\pi_t)\pi^{n-1} P_j(\pi_t)...P_j(\pi_t)\pi^1 P_j(\pi_t)\hat{\pi}$ and conclude by transitivity that $\pi_t P_j(\pi_t)\hat{\pi}$. Therefore, $C_j(D(T_{-\{t_0\}})) \neq \hat{\pi}$, which thus for an arbitrary ranking that is not the true ranking means he chooses the true ranking.

Now to show ($\leftarrow$).

Suppose $\pi \in H(\pi_t, U_i)$ and $C_j(D(T_{-\{t_0\}})) = \pi_t$ and $D(T_{-\{t_0\}}) \subset \alpha \in \beta(U_j)$ from lemma 1. I want to show that $X(\pi, U_j) \subset D(T_{-\{t_0\}}) \subseteq \alpha \in \beta(U_j)$.

From lemma 1, we already know that $|\alpha \cap H(\pi_t, U_j)| \leq 1$ such that $|D(T_{-\{t_0\}}) \cap H(\pi_t, U_j)| \leq 1$. Since we know that $C_j(D(T_{-\{t_0\}})) = \pi_t$, we know that $D(T_{-\{t_0\}}) \cap H(\pi_t, U_j) = \pi_t$.

Now to show that $\exists \alpha \in \beta(U_j), X(\pi, U_j) \subseteq \alpha$, I will do so by way of contradiction. Suppose otherwise, then this implies there exists $\pi' \in X(\pi, U_j)$ such that for any α which contains the rest of $X(\pi, U_j)_{-\pi'}$, $\pi' \notin \alpha$. Then simply choose $e \in \mathcal{S}$ such that $\pi \in H(\pi', U_i)$ so that the other juror would still choose π, then $C_j(\alpha) \neq \pi'$ a contradiction of how α defined. Thus, for the same reason, $X(\pi, U_j) \in D(T_{-\{t_0\}})$, otherwise it would contradict the fact that $C_j(D(T_{-\{t_0\}})) = f(e)$.

This concludes the proof.

□

Lemma 3. *Juror 1 reports a message t_n such that $\pi_t \in D(t_n|U)$ iff juror 2 can always choose a partially-honest ranking: $\forall e \in \mathcal{S}, \forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, R_1(\pi_t)) \neq \emptyset \iff C_2(D(t_n|U_2)) \cap H(\pi_t, U_2) \neq \emptyset$.*

Proof. To ensure the first juror is incentive compatible, I require that for $\forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, R_1(\pi_t)) \neq \emptyset$ when $C_2(D(t_n|U)) \neq \pi_t$. I will show this is equivalent to $\forall e \in \mathcal{S}, \forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, R_1(\pi_t)) \neq \emptyset \iff C_2(D(t_n|U_2)) \cap H(\pi_t, U_2) \neq \emptyset$.

I Will show $(\rightarrow)$ first.

To show $\forall e \in \mathcal{S}, \forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, R_1(\pi_t)) \neq \emptyset \rightarrow C_2(D(t_n|U_2)) \cap H(\pi_t, U_2) \neq \emptyset$. I will take the contradiction of the contrapositive, such that $\forall e \in \mathcal{S}, \forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, U_2) \neq \emptyset$ is rewritten as $C_2(D(t_n|U_2))^c \cup L_1(\pi_t, R_1(\pi_t))^c = \emptyset$. Now to determine the contradiction:

First: $C_2(D(t_n|U_2))^c = \emptyset$ only possible if $\exists t_n \in \Pi, D(t_n|U_2) = \Pi$ which is only possible if $U_2 = N^2$ since $D(t_n|U) \subset \alpha \in \beta$. However, this would imply the last juror is fully impartial, a contradiction of A1.

Second: $L_1(\pi_t)^c = \emptyset$ equivalent to saying $\forall \pi_t \in \Pi, \pi_t P_1(\pi_t)\pi$. However, requiring the true ranking to be the preference maximizing for the first juror makes it unnecessary to consult two judges, therefore also a contradiction of A1. Therefore, by way of contradiction, $\forall e \in \mathcal{S}, \forall t_n \in D(t_0), C_2(D(t_n|U_2)) \cap L_1(\pi_t, R_1(\pi_t)) \neq \emptyset \rightarrow C_2(D(t_n|U_2)) \cap H(\pi_t, U_2) \neq \emptyset$.

Now to show $(\leftarrow)$.

First, we know that by the construction of the sets $\alpha \in \beta(U_2, V_2)$, that if juror 2 has a choice of a partially-honest message, he picks it. Since we are looking at the reverse, we can assume that juror 2's choice always contains a partially-honest message.

Therefore, let us assume juror 1 chose t_n, such that juror 2 chose the ranking $\pi \neq \pi_t$, but $\pi \in H(\pi_t, U_2)$. Then, $\pi \neq \pi_t$ implies $\pi \notin H(\pi_t, U_1)$. To show that juror 1 prefers the true ranking to pi, we just need to show that $\exists \alpha \in \beta(U_2, V_2)$, such that $\pi, \pi_t \in \alpha$. First, notice that $Z(\pi, \pi_t) \subset U_1$, since juror is partially-honest, but we did not get the true ranking. As a direct result, $\pi_t \in X(\pi, U_1 \cup V_1) \subset \alpha \in \beta(U_1, V_1)$, where the true ranking is in the set of swapped pairs of π.

Now, such a set is always possible because we know juror 2 is being partially-honest. I will show recursively how. Take a ranking in juror 2's honest set, but not juror 1's, call it π. Now, take the first contestant $B \in N$ that is not in the proper order. It can't be out of order with something before it, because its the first, so their is a contestant later in the ranking that is suppose to be before B, let it be A. For all the contestants between B and A in π there are the impartial pairs. By assumption, all the impartial pairs in U_2 are honestly-ordered. Therefore, all the impartial pairs between A and B misordered are in juror 2's U_2 and thus in his ability to swap. So that fixes A and B. Move on to the next misordered and it fixes similarly. Since π finite, this process will end.

A simple illustration. Let $\pi = CBAD$. Suppose $\pi_t = ABCD$. then A is the first misranked contestant. The pairs between A and C that are misranked are $\{A, B\}, \{A, C\}, \{B, C\}$, whereas all of the impartial pairs including contestant D are ranked properly. Since we know juror 2 was partially-honest, that means all the misordered pairs are in U_1. Which means that juror 1 can swap C and B to get $BCAD$, then swap A and C to get $BACD$, and finally swap A and B to get $ABCD$. If we had different true ranking, we would get different ownership of the impartial pairs. Suppose that $\{A, B\}, \{A, C\} \in U_1$, then the others are in U_2 and are ranked correctly, which means the true ranking must be either $ACBD$, $CABD$, or $CBAD$, depending on which pairs in U_1 are misranked. In sum, if juror 2 is honest, juror 1 could always swap his pairs to be honest. □

Proposition 4. *Juror 2 is always able to choose a partially-honest message iff his preferences satisfy A6, and a condition*$-\lambda$.

Proof. First, check $(\rightarrow)$.

Case 1.

BWOC, suppose juror 2 is always able to choose a partially-honest message, but doesn't satisfy A6. Let the true ranking be ABC, and $\{A, C\} = U_2 \cup V_2$. Then if juror 1 is not honest and choose node $t_n = CBA$, then $\{CBA\} = X(CBA, \{A, C\}) = \alpha \in \beta(U_2, V_2)$, is juror 1's only choice and is not a partially-honest message for juror 2. A contradiction[4], therefore, we must have at least that $\{A, B\} \in V_2$, which makes $\{CBA, CAB, ACB\} = X(CBA, U_2, V_2)$, and $ACB \in H(\pi_t, U_2)$.

Case 2.

BWOC, suppose juror 2 is always able to choose a partially-honest message, satisfies A6, but doesn't satisfy condition$-\lambda$. Let the true ranking be ABC and juror 2 has the impartial pairs $\{A, B\}, \{B, C\} \in U_2$, but $V_2 = \emptyset$. Then suppose juror 1 was not honest and chose the node $t_n = CBA$. Then $\alpha = X(CBA, U_2) = \{BCA, CBA, CAB\}$, all of which are not in $H(\pi_t, U_2)$, since none of the ranking ranks both B before C and A before B. This is a contradiction of juror 2 always able to pick a partially-honest ranking. However, if $\{A, B\} \in V_2$, which satisfies condition$-\lambda$, then $BCA \in H(ABC, U_2)$. Therefore, if juror 2 is always able to be partially-honest, he satisfies A6 and condition$-\lambda$.

Now for the reverse case: $(\rightarrow)$

Suppose juror 2 satisfies A6 and condition$-\lambda$. Let juror 1 choose a node t_n. Then $D(t_n) = \alpha \in \beta(U_2, V_2)$ is such that $D(t_n) \cap H(\pi_t, U_2) \neq \emptyset$. Let $t_n = \pi \neq \pi_t$. Then for $Z(\pi, \pi_t)$, let $\{A, B\} \in U_2$ be be misranked. By A6 and condition$-\lambda$, wlog, $\{A, X\} \in V_2, \forall X \in N_{-B}$, which means there exists $\pi' \in X(\pi, U_2, V_2)$ where A and B properly ranked. Since this was done for an arbitrary pair in U_2, it can be done for all of them, such that $\pi^r \in X(\pi, U_2, V_2)$, and thus $\pi^r \in H(\pi_t, U_2)$. Therefore, since this was done for an arbitrary true ranking, and

[4]Notice, this proves we need more than impartial pairs to get implementation with two jurors.

choice by juror 1, juror 2 can always choose a partially-honest message.

□

Lemma 4. *If the mechanism subgame implements the true ranking, then $\forall t \in T_{-\{t_0\}}$, there exists some $\alpha \in \beta$ such that $D(t) = \alpha$.*

Proof. □

Proposition 5. *Given that $D(T_{-\{t_0\}}) \subseteq \alpha \in \beta$, juror 2 is transitive over rankings in δ, and juror 2 satisfies Non-Partial over Alternatives, then Juror 1 will send a partially-honest message iff condition-λ is met.*

Proof. □

Theorem 1 (Implementation)**.** *The mechanism Γ subgame implements the true ranking iff $D(T_{-\{t_0\}}) \subseteq \alpha \in \beta$, juror 2's preferences are transitive over $D(T_{-\{t_0\}})$, juror 2 satisfies Non-Partial over Alternatives, and both jurors satisfy condition-λ.*

Proof. □