Arrow's, May's, and the Gibbard-Satterthwaite Theorems*

Adithya Bhaskara

July 25, 2024

1 Motivation & Preliminaries

In this set of notes, we first examine Arrow's Theorem, a famous impossibility result in social choice theory. Originally stated in [1] by Kenneth J. Arrow in 1951, we have the unfortunate result that all "reasonable" voting rules, involving 3 or more alternatives, must be dictatorial. After our melancholic endeavor of proving Arrow's result, we will consider May's Theorem, which will provide slight relief, for the case of 2 alternatives. Finally, as a return to despair, we will consider the Gibbard-Satterthwaite Theorem, independently stated by Gibbard in [3] and Satterthwaite in [4].

We start with some definitions, following closely the exposition of [2].

Let $N = \{1, \ldots, n\}$ be a finite set of voters, and let A be a finite set of alternatives–or candidates. Consider the following definitions.

Definition 1 (Weak and Linear Orders)**.** *A binary relation on a finite set A is a weak order if it is both complete and transitive. A linear order is a weak order that is additionally antisymmetric. Denote the set of weak orders $\succeq$ on A by $\mathcal{R}(A)$ and the set of linear orders $\succeq$ on A by $\mathcal{L}(A)$. Note that $\succ$ denotes the strict part of $\succeq$.*

Weak orders are used to model preferences permitting ties, and linear orders are used to model strict preferences. The preference of $i \in N$ is denoted by $\succ_i$. We now define social welfare functions, the central object of Arrow's Theorem.

Definition 2 (Social Welfare Functions (SWFs))**.** *A social welfare function f is a map of the form $f : \mathcal{L}(A)^n \to \mathcal{R}(A)$.*

Think of Definition 2 as a type of voting rule: taking in a profile of voters' preferences $P = (\succeq_1, \ldots, \succeq_n)$, f returns aggregates the preferences into a weak order. Arrow suggested two natural axioms that SWFs should satisfy. We state them here.

Axiom 1 (Weakly Paretian)**.** *An SWF f is weakly Paretian if for all $x_1, x_2 \in A$, $x_1 \succ_f x_2$ whenever $x_1 \succ_i x_2$ for all $i \in N$.*

Axiom 2 (Independence of Irrelevant Alternatives (IIA))**.** *An SWF f is independent of irrelevant alternatives if for all $x_1, x_2 \in A$, the relative ranking of x_1 and x_2 by f depends only on the relative rankings of x_1 and x_2 provided by the individuals, and not on the individuals' rankings of some irrelevant alternative x_3.*

*Originally written to better understand basic ideas in social choice theory.

Intuitively, Axiom 1 specifies that if all voters rank one candidate over another, the SWF must reflect this preference. Axiom 2 ensures that SWF's ranking of two alternatives shouldn't depend on voters' preferences involving a third "irrelevant alternative." Axiom 2 can be thought of as loose guard against voters' having an incentive to strategize and misreport their true preferences. Informally, one shouldn't change whether they like apples or oranges more when additionally given the option of bananas.

Now, we define the central object of May's Theorem and the Gibbard-Satterthwaite Theorem: social choice functions.

Definition 3 (Social Choice Functions (SCFs))**.** *A social choice function f is a map of the form $f : \mathcal{L}(A)^n \to 2^A \backslash \emptyset$, where 2^A is the power set of A.*

Think of Definition 3 as a type of voting rule: taking in a profile of voters' preferences $P = (\succ_1, \ldots, \succ_n)$, f returns a set of "winners." If $|f(P)| = 1$, we say f is single-valued on P. In this case, we may use the semantics of $f : \mathcal{L}(A)^n \to A$. We say f is resolute if it is single-valued for all profiles.

The notion of dictatorial SWFs and SCFs proceeds as follows.

Definition 4 (Dictatorial SWFs and SCFs)**.** *An SWF f is dictatorial if there exists $i^* \in N$ such that for all $x_1, x_2 \in A$, $x_1 \succ_{i^*} x_2$ implies $x_1 \succ_f x_2$.*

A resolute SCF f is dictatorial if there exists $i^* \in N$ such that for all profiles P where i^* ranks x^* above all other alternatives, $f(P) = x^*$.

In both cases, we refer to i^* as the dictator under f.

We now state some natural axioms for SCFs, just as we did for SWFs.

Axiom 3 (Anonymity)**.** *An SCF f is anonymous if each pair of voters are interchangeable. That is, $f(P) = f(P^*)$ for profiles P and P^*, whenever P^* is obtained from P by swapping the ballots cast by two voters i and j.*

Axiom 4 (Neutrality)**.** *An SCF f is neutral if each pair of alternatives are interchangeable. That is, when P^* is obtained from P by swapping the positions of alternatives x_1 and x_2 in every ballot, $f(P^*)$ is obtained from P by a similar swap. Moreover, we say that f is imposed if there exists an unelectable candidate x; i.e. for no profile P does $x \in f(P)$.*

Axiom 5 (Monotonicity & Positive Responsiveness)**.** *An SCF f is monotone if for a preference P with $x \in f(P)$, and for P^* obtained from P by just having one voter rank x higher in their ballot, $x \in f(P^*)$. We say f over $|A| = 2$ is positive responsive if $x \in f(P)$ and for P^* obtained from P by just having one voter rank x higher in their ballot, $\{x\} = f(P^*)$.*

Intuitively, Axiom 3 specifies that the SCF treats all voters equally: a ballot cast by one voter yields the same preference as the same ballot cast by another voter. Note that nondictatoriality is a very weak form of anonymity. Axiom 4 ensures that permuting alternatives' identities on the ballots yields an analogous permutation in the results. Note that nonimposition is a very weak form of neutrality. Finally, Axiom 5 requires that alternatives are not negatively affected by voters ranking them higher.

Remark. *As we will explore soon, Axiom 5 is especially insightful when there are only two alternatives; positive responsiveness, in particular, helps us grapple with ties. Furthermore, a resolute SCF satisfies monotonicity if a winner x under profile P is still the winner when a voter increases the ranking of x in a profile P^*; that is, $f(P) = f(P^*)$.*

We now introduce two axioms about voter strategy.

Axiom 6 (Strategyproofness). *A resolute SCF f is strategyproof if whenever P^* is obtained from P by having voter i changing their preferences from $\succeq_i$ to $\succeq_i^*$, we have that $f(P) \succeq_i f(P^*)$.*

Axiom 7 (Down Monotonicity). *A resolute SCF f is down monotone if whenever P^* is obtained from P by having voter i changing their preferences from $\succeq_i$ to $\succeq_i^*$ by dropping some losing alternative $\ell \neq f(P)$, we have that $f(P) = f(P^*)$.*

Think of Axiom 6 as ensuring that if voter i has true preferences $\succeq_i$ but votes according to $\succeq_i^*$ in an attempt to strategize, they get no benefit under f; they consider $f(P)$ is at least as good as $f(P^*)$ under their original preferences. Axiom 7 is similar; if voter i strategizes by dropping a losing alternative, the outcome will remain the same. It is trivial to see that Strategyproofness implies Down Monotonicity.

Finally, we state an axiom about the notion of a Paretian SCF.

Axiom 8 (Paretian). *An SCF f is Paretian if $f(P)$ never contains a Pareto dominated alternative; for $x_1, x_2 \in A$, we say that x_1 dominates x_2 if every voter ranks x_1 over x_2.*

Axiom 8 requires that if all voters rank x_1 over x_2, x_2 shouldn't be a winner under f.

We will use Axioms 1 and 2 when it comes to Arrow's Theorem, Axioms 3, 4, and 5 with May's Theorem, and Axioms 4, 6, 7, and 8 for the Gibbard-Satterthwaite Theorem.

2 The Heart of Arrow's Theorem

We build up to Arrow's Theorem with some lemmas and definitions. Here, we follow the general argument of [2] with some reorganization and notational differences. For the remainder of this section, $|A| \geq 3$ unless otherwise stated.

Lemma 1 (Dictatorial SWF $\implies$ Weakly Paretian and IIA). *Any dictatorial SWF f is both weakly Paretian and IIA.*

Proof. Let f be a dictatorial SWF with dictator i^*. Consider arbitrary $x_1, x_2 \in A$.

If for all $i \in N$ $x_1 \succ_i x_2$, then it must be the case that $x_1 \succ_{i^*} x_2$, so $x_1 \succ_f x_2$. Therefore, f is weakly Paretian.

Since the ordering $\succ_f$ is equivalent to the ordering $\succ_{i^*}$, it is also immediate that the ranking of x_1 and x_2 under f is equivalent to that of that of i^* and doesn't depend on the preferences that i^* has on a third alternative. ∎

Definition 5 (Coalitions). *A subset $C \subseteq N$ is called a coalition. We say C is decisive over (x_1, x_2) if $x_1 \succ_f x_2$ whenever $x_1 \succ_i x_2$ for all $i \in C$. Additionally, we say C is weakly decisive over (x_1, x_2) if $x_1 \succ_f x_2$ whenever $x_1 \succ_i x_2$ for all $i \in C$ and $x_2 \succ_j x_1$ for all $j \notin C$.*

Lemma 2 (Field Expansion, Weakly Decisive $\implies$ Decisive). *A weakly Paretian and IIA SWF f, with a weakly decisive coalition C over (x_1, x_2), is a decisive for all alternatives.*

Proof. Consider mutually distinct $x_1, x_2, x_1', x_2' \in A$, and let C be weakly decisive over (x_1, x_2). We will show that C is decisive over the alternatives (x_1', x_2').

Let $x_1' \succ_i x_1 \succ_i x_2 \succ_i x_2'$ for all $i \in C$, and for all $j \notin C$, let $x_1' \succ_j x_1$, $x_2 \succ_j x_2'$, and $x_2 \succ_j x_1$, Since C is weakly decisive over (x_1, x_2), we must have $x_1 \succ_f x_2$. Since f is weakly Paretian, we also have $x_1' \succ_f x_1$ and $x_2 \succ_f x_2'$. By transitivity of $\succ_f$, we get $x_1' \succ_f x_2'$. So, C is decisive over (x_1', x_2').

Because the choice of (x_1', x_2') was arbitrary, C is decisive over all pairs of alternatives. ∎

Remark. *To be very explicit, the construction of preferences in Lemma 2 is to exploit two assumptions. Since* $x_1 \succ_i x_2$ *for* $i \in C$ *and* $x_2 \succ_j x_1$ *for* $j \notin C$*, we can use the weakly decisiveness of* C *to conclude* $x_1 \succ_f x_2$*. Since* $x_1' \succ_k x_1$ *and* $x_2 \succ_k x_2'$ *for all* $k \in N$*, we can use the weakly Paretian nature of* f *to conclude* $x_1' \succ_f x_1$ *and* $x_2 \succ_f x_2'$*.*

Importantly, in our proof, we also did not need to consider how voters outside C *rank* x_1' *versus* x_2'*. We obtained* $x_1' \succ_f x_2'$ *only from exploiting the weak decisiveness of* C*, the weakly Paretian SWF* f*, and the transitivity of* $\succ_f$*. Also, note that by IIA, we only needed to make sure that all* $i \in C$ *had* $x_1' \succ_i x_2'$*. If instead* $x_2 \succ_i x_1$*, a similar argument would hold.*

Lemma 3 (Splitting Coalitions)**.** *Let* $C \subseteq N$ *be a decisive coalition, with respect to some pair of alternatives. Additionally, let* $|C| \geq 2$*. Then, we can write* $C = C_1 \cup C_2$ *with* $C_1 \neq \emptyset$*,* $C_2 \neq \emptyset$*, and* $C_1 \cap C_2 = \emptyset$*, where either* C_1 *or* C_2 *is decisive over all pairs of alternatives.*

Proof. Recalling that $|A| \geq 3$, suppose $x_1 \succ_i x_2 \succ_i x_3$ for all $i \in C_1$, $x_2 \succ_j x_3 \succ_j x_1$ for all $j \in C_2$, and $x_3 \succ_k x_1 \succ_k x_2$ for all $k \notin C_1 \cup C_2$. Because C is decisive, $x_2 \succ_f x_3$. Then, either $x_1 \succ_f x_3$ or $x_3 \succeq_f x_1$.

Case 1 ($x_1 \succ_f x_3$): We see that the preferences in C_1 match those aggregated by f. Since f is IIA, whenever voters in C_1 rank x_1 above x_3, the SWF does the same. So, C_1 is weakly decisive over (x_1, x_3). But by Lemma 2, C_1 is decisive for all pairs of alternatives.

Case 2 ($x_3 \succeq_f x_1$): By transitivity, and $x_2 \succ_f x_3$, we have that $x_2 \succ_f x_1$. So, the preferences in C_2 match those aggregated by f. Since f is IIA, whenever voters in C_2 rank x_2 above x_1, the SWF does the same. So, C_2 is weakly decisive over (x_1, x_2). But by Lemma 2, C_2 is decisive for all pairs of alternatives.

We have shown the desired result. ■

Theorem 1 (Arrow, Weakly Paretian and IIA $\iff$ Dictatorial SWF)**.** *When* $|A| \geq 3$*, an SWF is weakly Paretian and IIA if and only if it is dictatorial.*

Proof. By Lemma 1, we need only show that an arbitrary weakly Paretian and IIA SWF f is dictatorial. Note that N is a decisive coalition since f is weakly Paretian. Then, we can apply Lemma 3 repeatedly to obtain smaller and smaller decisive coalitions. Once we obtain a singleton decisive coalition, we are done. The element of the singleton decisive coalition is the dictator under f. Note that the inductive argument is valid since N is finite. ■

3 May's Theorem

We state and prove May's Theorem, with relative ease compared to the previous section. We again follow [2].

Theorem 2 (May, Majority Rule is Best for 2 Alternatives)**.** *For two alternatives and an odd number of voters, majority rule is the unique resolute, anonymous, neutral, and monotone SCF.*

For two alternatives and any number of voters, it is the unique anonymous, neutral, and positively responsive SCF.

Proof. Let $A = \{x, y\}$. Trivially, majority rule satisfies all the above properties.

For uniqueness, with any other SCF, we choose a profile where x wins, but with fewer votes than y. Suppose we switch enough ballots to reverse the number of votes x and y each have. Monotonicity implies that x still wins; however, neutrality and anonymity implies that y wins. If x and y tie, meaning $\{x, y\} \in f(P)$, but x has fewer votes than y, positive responsiveness similarly contradicts neutrality and anonymity. ■

4 The Famous Gibbard-Satterthwaite Theorem

Desiring to end with pessimism, we now present the famous Gibbard-Satterthwaite Theorem. The machinery we have developed from Arrow's Theorem will provide intuition. Again, let $|A| \geq 3$ unless otherwise stated for this section. As in previous sections, we follow [2], and of course, we will now start with some definitions and lemmas.

Definition 6 (Blocking and Dictating Sets)**.** *Let f be a resolute SCF for $|A| \geq 3$ alternatives. Let $x_1, x_2 \in A$ be distinct and $X \subseteq N$ be a set of voters. We say that X can use x_1 to block x_2 and write $X^{x_1 \succ x_2}$ if for all profiles P where every voter in X ranks x_1 over x_2, then $f(P) \neq x_2$. We say that X is a dictating set if $X^{x_1 \succ x_2}$ holds for every pair of distinct alternatives x_1 and x_2.*

Lemma 4 (Push-Down)**.** *Let $x_1, x_2, y_1, \ldots, y_{|A|-2}$ be the $|A| \geq 3$ alternatives. Let f be resolute and down monotonic SCF for A, and let P be an arbitrary profile with $f(P) = x_1$. There exists a profile P^* with $f(P^*) = x_1$ such that*

1. *for each voter i with $x_1 \succ_i x_2$, $\succ_i^* = x_1 \succ x_2 \succ y_1 \succ \cdots \succ y_{|A|-2}$, and*
2. *for each voter i with $x_2 \succ_i x_1$, $\succ_i^* = x_2 \succ x_1 \succ y_1 \succ \cdots \succ y_{|A|-2}$.*

Proof. To obtain P^* from P have voters change their preferences by successively dropping $y_1, \ldots, y_{|A|-2}$ to the bottom of their ranking. By down monotonicity, $f(P^*) = f(P) = x_1$. ■

Lemma 5 (Blocking Condition)**.** *Let f be a resolute and down monotonic SCF. If there exists a profile P where $f(P) = x_1$ and for all $i \in X$, it is the case that $x_1 \succ_i x_2$, and for all $j \in N \backslash X$, $x_2 \succ_j x_1$, then it must be the case that X can use x_1 to block x_2; that is, $X^{x_1 \succ x_2}$.*

Proof. Suppose for contradiction, we have such a profile P, but it is not the case that $X^{x_1 \succ x_2}$. Consider P' where all voters $i \in X$ rank x_1 over x_2 but $f(P^*) = x_2$. Now, form P'' by having any P' voters $j \in N \backslash X$ that rank x_1 over x_2 drop x_1 below x_2. By down monotonicity, $f(P'') = x_2$.

Now, we use Lemma 4 to form P^* with $f(P^*) = x_1$ and P''^* with $f(P''^*) = x_2$. But note that $P^* = P''^*$, but f seemingly is not well-defined. This is the contradiction we seek. ■

Lemma 6 (Splitting Blocking Sets)**.** *Let f be a resolute Paretian and down monotonic SCF for $|A| \geq 3$ alternatives. Suppose $X^{x_1 \succ x_2}$, and $X = X_1 \cup X_2$ with $X_1 \cap X_2 = \emptyset$. Let x_3 be distinct from x_1 and x_2. Then, either $X_1^{x_1 \succ x_3}$ or $X_2^{x_3 \succ x_2}$.*

Proof. Consider a profile P where all voters $i \in X_1$ have $x_1 \succ x_2 \succ x_3 \succ \cdots$, all voters $j \in X_2$ have $x_3 \succ x_1 \succ x_2 \succ \cdots$, and all voters $k \in N \backslash X$ have $x_2 \succ x_3 \succ x_1 \succ \cdots$. Since f is Paretian, $f(P) \in \{x_1, x_2, x_3\}$. Since $X^{x_1 \succ x_2}$, $f(P) \neq x_2$. By Lemma 4, if $f(P) = x_1$, then $X_1^{x_1 \succ x_3}$ and if $f(P) = x_3$, then $X_2^{x_3 \succ x_2}$. ■

Lemma 7 (Blocking and Third Alternatives)**.** *Let f be a resolute Paretian and down monotonic SCF for $|A| \geq 3$ alternatives. Suppose $X^{x_1 \succ x_2}$ and let x_3 be distinct from x_1 and x_2. Then, $X^{x_1 \succ x_3}$ and $X^{x_3 \succ x_2}$.*

Proof. Note $X = X \cup \emptyset = \emptyset \cup X$ is a valid partition under Lemma 6. But also, since f is Paretian, $\emptyset$ cannot block any alternative with any other alternative. Then, applying Lemma 6 with $X_1 = X$ gives $X^{x_1 \succ x_3}$, and with $X_2 = X$ gives $X^{x_3 \succ x_2}$. ■

Lemma 8 (Field Expansion, Blocking $\implies$ Dictating)**.** *Let f be a resolute Paretian and down monotonic SCF for $|A| \geq 3$ alternatives. If $X^{x_1 \succ x_2}$, then X is a dictating set.*

Proof. Let x_1' and x_2' be distinct from each other in A. Supposing that $X^{x_1 \succ x_2}$, We show that $X^{x_1' \succ x_2'}$.

Case 1 ($x_1' = x_1$): Lemma 7 immediately gives $X^{x_1' \succ x_2'}$ for all $x_2' \neq x_1$ since x_2' is distinct from $x_1 = x_1'$ and x_2.

Case 2 ($x_1' \notin \{x_1, x_2\}$): Lemma 7 gives $X^{x_1' \succ x_2}$ since x_1' is distinct from x_1 and x_2. Now, we apply Lemma 7 again with the distinct alternative x_2' so $X^{x_1' \succ x_2'}$ for all $x_1' \neq x_2'$.

Case 3 ($x_1' = x_2$): Lemma 7 gives $X^{x_1 \succ x_3}$. We apply Lemma 7 again to get $X^{x_2 \succ x_3}$, and then one more time to get $X^{x_2 \succ x_2'}$ for all $x_2' \neq x_2$. Since $x_1' = x_2$, this gives us that $X^{x_1' \succ x_2'}$.

We've shown that assuming X can block x_2 with x_1, X can block arbitrary alternative x_2' with arbitrary alternative x_1'. So, X is a dictating set, as desired. ∎

Lemma 9 (Splitting Dictating Sets)**.** *Let f be a resolute Paretian and down monotonic SCF for $|A| \geq 3$ alternatives. If X is a dictating set with $X = X_1 \cup X_2$ and $X_1 \cap X_2 = \emptyset$, either X_1 or X_2 are also dictating sets.*

Proof. Let $x_1, x_2, x_2 \in A$ be distinct. By Lemma 6, since $X^{x_1 \succ x_2}$, either $X_1^{x_1 \succ x_3}$ or $X_2^{x_3 \succ x_2}$. But then Lemma 4 tells us that either X_1 or X_2 are dictating sets. ∎

Lemma 10 (Down Monotone $\implies$ Paretian)**.** *Let f be a resolute nomimposed down monotonic SCF. Then, f is Paretian.*

Proof. Suppose, for contradiction, that all assumptions are satisfied, yet f is not Paretian. Choose profile P such that every voter ranks x_2 above x_1 but $f(P) = x_1$. By nonimposition, P' exists with $f(P') = x_2$. If voters in P' rank x_1 over x_2, obtain P'' by having them drop x_1 right below x_2. Then, by down monotonicity $f(P'') = x_2$. Applying Lemma 4 on P to get P^* yields $f(P^*) = x_1$. Doing the same on P'' to get P''^* gives $f(P''^*) = x_2$. But, $P^* = P''^*$, a contradiction. ∎

Lemma 11 (Dictatorial SCF $\implies$ Resolute, Nonimposed, and Strategyproof)**.** *Any dictatorial SCF f is resolute, nonimposed, and strategyproof.*

Proof. Let f be a dictatorial SCF with dictator i^*. Consider arbitrary $x_1, x_2 \in A$.

By definition, a dictatorial SCF is resolute.

It is also obvious that i^* can make x_1 or x_2 win by ranking the desired winner over the other, so f is nonimposed.

Because by definition i^*'s preferences are exactly reflected in f, i^* has no incentive to strategize. ∎

Now, we state the Gibbard-Satterthwaite Theorem itself.

Theorem 3 (Gibbard-Satterthwaite, Resolute, Nonimposed and Strategyproof $\iff$ Dictatorial SCF)**.** *When $|A| \geq 3$, an SCF is resolute, nonimposed, and strategyproof if and only if it is dictatorial.*

Proof. By Lemma 11, we need only show that an arbitrary resolute, nonimposed, and strategyproof SCF f is dictatorial.

First, we note, as we did earlier, that it is trivial to see that strategyproofness implies down monotonicity. We also showed in Lemma 4 that down monotonicity implies Paretian. So, all strategyproof SCFs are Paretian, and we need only show that an arbitrary resolute, down monotonic, and nonimposed Paretian SCF f is dictatorial.

We will show that there exists singleton dictating set X, of which the element is the dictator. Since f is Paretian, N must be a dictating set, and we can inductively apply Lemma 9 to obtain the desired singleton; we never obtain the $\emptyset$ because it is not dictating since f is Paretian. We are done. ■

Looking back, we ask the reader to consider both the similarities and differences between the statements and proofs of both Arrow's Theorem and the Gibbard-Satterthwaite Theorem.

References

[1] Arrow, K.J.: Social Choice and Individual Values. Yale University Press (1951)

[2] Brandt, F., Conitzer, V., Endriss, U., Jérôme, L., Procaccia, A.D.: Handbook of Computational Social Choice. Cambridge University Press (2016), https://procaccia.info/wp-content/uploads/2020/03/comsoc.pdf

[3] Gibbard, A.: Manipulation of voting schemes: A general result. Econometrica **41**(4), 587–601 (1973), http://www.jstor.org/stable/1914083

[4] Satterthwaite, M.A.: Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. Journal of Economic Theory **10**(2), 187–217 (1975). https://doi.org/https://doi.org/10.1016/0022-0531(75)90050-2, https://www.sciencedirect.com/science/article/pii/0022053175900502