On the Impossibility of Mechansim Design.

Original Idea: Ali Adali, Selman Erol, Nizameddin Ordulu
This Handout: Kevin Hasker

January 27, 2016

1 Introduction

In this handout I will try to explain the two fundamental impossibility result in social choice and mechanism design. The first result—the Mueller-Satterthwaite theorem—seems on it's surface to be just another variation of Arrow's Impossibility theorem. It finds some interesting properties that social choice functions might want to satisfy and then proves that only dictatorial social choice functions can satisfy it.

But who cares? Just because a property seems interesting a-priori doesn't mean that it is actually important. For example, Arrow showed that no social choice rule can be binary. OK, we like binary choice rules but that's not that big of a loss. Or is it?

The thing about the properties analyzed in the Mueller-Satterthwaite theorem is that they are necessary if the social planner wants to implement the social choice rule. The Gibbard-Satterthwaite theorem shows this result.

The methodology used in this paper was originally developed by Ali Adali, Selman Erol, and Nizameddin Ordulu at a Turkish Math Camp. Semih Koray gave them the Mueller-Satterthwaite theorem and asked them to prove it. In a seminar Semih Koray told the faculty the basic methodology. Using that presentation as a basis I developed the following handout, where I try to simplify and clarify the proof as much as possible.

The model we will be working in has a finite set of alternatives (or outcomes or allocations), A, and each individual has a strict linear preference ordering θ_i over A.[1] We denote $\theta = \times_{i\in I}\theta_i$ as the preferences of society (note there are I people in this society) and Θ as the set of strict linear preference orderings over A.

Our object of analysis will be social choice functions:

Definition 1 *A* social choice function *is a function from* Θ^I *to* A, *or* $F : \Theta^I \to A$

Notice in this context the importance of a *function.* We are trying to analyze social decision making, not to find socially undesirable outcomes. If we weakened this function to a *correspondence* ($F(\theta) \subseteq A$) then what we really would be doing is ruling out possible outcomes, or listing what we *don't* want to happen. This is not what we are trying to do at this point, instead we are trying to state what occurs.

[1] "Strict" means that for every $a \in A$ and $b \in A\backslash a$ either a is strictly preferred to b or the reverse. "Linear" means that the preferences are anti-reflexive, transitive, and complete.

1.1 Mueller-Satterthwaite

There will only be two properties that we will assume for this theorem, the first is extremely desirable given that we properly define utility.

Definition 2 Weak Paretian *if at* θ *for all* $i \in I$ a *is strictly preferred to* b *by then* $F(\theta) \neq b$.

A proper definition of utility also takes into consideration the way I feel about you, and about your smoking cigarettes, whether you get enough exercise, whether I think you make too much money, whether I think you deserve to be happy. If this is the definition that we are using then the above assumption on some level seems like it is a *fundamental* to a desirable social choice function. How could we choose something when something else is strictly better for everyone? How could we say our choice reflects what is "good" for society in any sense?

An alternative assumption one can use is that $F(\cdot)$ is *onto*. This works, but it actually is the same as the above when we assume Maskin Monotonicity, I will prove this in a moment.

Definition 3 *The* lower contour set of a *according to* θ_i *is:*
$L(a, \theta_i) = \{b \in A | a \succeq_i b\}$
where $\succeq_i$ *is the preference ordering of* i.

Then the second property we will assume is Maskin Monotonicity.

Definition 4 F *is* (Maskin) Monotonic *if* $\forall \theta, \theta' \in L(A)^I$ *and* $\forall a \in A$

$$F(\theta) = a \text{ and } \forall i \in I \; L(a, \theta_i) \subseteq L(a, \theta'_i) \Rightarrow F(\theta') = a$$

Intuitively the argument if we were already choosing a and everyone decides that they like a more shouldn't we still choose a? Sounds nice, much more sensible than the "transitivity" and "binary" assumptions Arrow made in the first impossibility theorem. But it's impossible, unless you're willing to live with:

Definition 5 F *is* dictatorial *if there exists a unique* $i \in I$ *such that*

$$F(\theta) = \{a \in A | \forall b \in A, a \succeq_i b\}$$

But before we prove this result I just want to establish that:

Lemma 1 *Assume that* $F(\cdot)$ *is* onto *and* Maskin Monotonic*, then* $F(\cdot)$ *is* Weak Paretian.

Proof. *Assume not, or is a* θ *such or that* a *is strictly preferred to* b *by every person but* b *is chosen. Now consider* θ^{+a} *where we move* a *up to the top choice for every person. Or for every* $i \in I$ *and* $\forall c \in A, a \succeq_i c$*, but the preferences over all other alternatives is the same as in* θ_i*. Now* b*'s lower contour set has not changed from this transformation, thus* $F(\theta^{+a}) = b$.

Now by onto there exists a θ^* *such that* $F(\theta^*) = a$*, and at* θ^{+a} *the lower contour set of* a *is the largest for every* i*, thus* $F(\theta^{+a}) = a$*, and we have a contradiction.* ■

The original theorem assumed onto, but Weak Paretian is a more desirable social quality and thus we will consider this problem.

Theorem 1 *(Mueller-Satterthwaite) : If* A has more than three elements*, and* F *is* Weak Paretian *and* Maskin Monotonic *then* F *is dictatorial.*

The proof of this theorem, like all theorems, really only requires that you figure out the critical implication of the assumptions. The critical implication of Maskin Monotonicity is that there is some "worst" profile at which a is chosen, once you find one of these profiles if you increase a's desirability (ranking) for any person then a will still be chosen. To make this clearer let's represent a θ_i as a list, with the most desirable being listed (ranked) higher on the list and less desirable being listed (ranked) lower. We put the identity of the individual in question on the top of the list for clarity. For example, if $A = \{a, b, c\}$ then the six possibilities for θ_i are:

i	i	i	i	i	i
a	a	b	b	c	c
b	c	a	c	a	b
c	b	c	a	b	a

and a society if $I = \{1, 2\}$ might be something like:

1	2
a	a
b	c
c	b

and notice that at this profile $F(\theta) = a$ by the Weak Paretian property. Now the critical concept we need for the proof is:

Definition 6 θ^a *is a* critical profile of a *if* $F(\theta^a) = a$ *and if any one person decides* a *is worse then in the resulting preference profile* θ', $F(\theta') \neq a$.

You can imagine Θ^I as a finite subset of N^I (where N is the natural numbers). Now once we subject $\theta \in \Theta^I$ to the proper ordering we can have a sub-space where a is chosen, and another sub-space where b is chosen, and so on and so forth. In other words finding the critical profiles is the task of finding the boundaries of these sub-spaces. The key insight of these students was that once you find these boundaries you have characterized the social choice function on the entire space.

We will prove this theorem with a series of lemmas. The first two just establish some basic properties of critical profiles. We will then use one critical proof twice, examine the implications of these steps, and we will be done. Now to give an example of a (possible) critical profile say that:

$\theta^a =$

1	2
a	a
b	c
c	b

was a critical profile for a then at

$\theta' =$

1	2
b	a
a	c
c	b

,

we would know that a was not chosen. What would be chosen? $F(\theta') = b$, this is because if $F(\theta') = c$ then $F(\theta^a) = c$ because the lower contour set of c would not have gotten worse, or for all i $L(c, \theta'_i) \subseteq L(c, \theta^a)$ and by Maskin Monotonicity $F(\theta^a) = c$. We can write this more formally as:

Lemma 2 *Let θ and θ' be two profiles such that $F(\theta) \neq F(\theta')$, then there exists an i such that $L(F(\theta'), \theta_i) \not\supseteq L(F(\theta'), \theta_i')$.*

Proof. *Assume not, but then we know that for all i $L(F(\theta'), \theta_i) \supseteq L(F(\theta'), \theta_i')$, and thus by Maskin Monotonicity $F(\theta) = F(\theta')$.* ■

A rather obvious fact is that:

Lemma 3 *If θ^a is a critical profile of a and θ' is derived from θ^a by permuting the ranking of options each player ranks below a, then θ' is also a critical profile.*

Proof. *This is immediate from the definition of Masking Monotone, since the lower contour sets of a has not changed it must still be chosen.* ■

This allows us to "move around" options in the lower contour set. Thus we can get the critical options right below our a.

Now we get to the critical step. Could the following profile be critical?

$$\begin{array}{rcl} \theta^a & = & \begin{array}{c|c} 1 & 2 \\ \hline \text{b} & \text{c} \\ \text{a} & \text{a} \\ \text{c} & \text{b} \end{array} \\ F(\theta^a) & = & a \end{array}$$

by Lemma 2:

$$F(\theta') = F\left(\begin{array}{c|c} 1 & 2 \\ \hline \text{b} & \text{c} \\ \text{a} & \text{b} \\ \text{c} & \text{a} \end{array}\right) = b, F(\theta'') = F\left(\begin{array}{c|c} 1 & 2 \\ \hline \text{b} & \text{c} \\ \text{c} & \text{a} \\ \text{a} & \text{b} \end{array}\right) = c$$

but then we can create a third profile by increasing c in θ' and increasing b in θ'':

$$\theta^* = \begin{array}{c|c} 1 & 2 \\ \hline \text{b} & \text{c} \\ \text{c} & \text{b} \\ \text{a} & \text{a} \end{array}$$

we know that $F(\theta^*) = b$ because $F(\theta') = b$ and the lower contour set of b has only increased in size. We know that $F(\theta^*) = c$ because $F(\theta') = c$ and the lower contour set of c has only increased in size. This is a contradiction, more generally:

Lemma 4 *In θ^a there can not be two people who have different outcomes in the lower contour set of a.*

Proof. *Assume not, or that i prefers a to b and j prefers a to c. By Lemma 3 we can assume that i prefers a to b with no intervening outcomes, likewise c is just below a for j. Construct θ' by moving a below b in i's preferences, and θ'' by moving a below c in j's preferences; then $F(\theta') = b$ and $F(\theta'') = c$. Now construct θ^* by making both changes, have b just above a for i and c just above a for j. Then b's lower contour set has not changed between θ' and θ^*, or $F(\theta^*) = b$, but also c's lower contour set has not changed between θ'' and θ^* thus $F(\theta^*) = c$—a contradiction.* ■

Now what does this leave for us? Could the following be critical?

$$\theta^a = \begin{array}{c|c} 1 & 2 \\ \hline \text{a} & \text{c} \\ \text{c} & \text{a} \\ \text{b} & \text{b} \end{array}$$

No, because then we can increase c for 1 and b for 2 and have the same contradiction. Could the following be critical?

$$\theta^a = \begin{array}{c|c} 1 & 2 \\ \hline \text{c} & \text{c} \\ \text{a} & \text{a} \\ \text{b} & \text{b} \end{array}$$

No, because then a is not Pareto efficient. The only possibilities are:

$$\theta^a = \begin{array}{c|c} 1 & 2 \\ \hline \text{a} & \text{c} \\ \text{c} & \text{b} \\ \text{b} & \text{a} \end{array}, \tilde{\theta}^a = \begin{array}{c|c} 1 & 2 \\ \hline \text{b} & \text{a} \\ \text{c} & \text{b} \\ \text{a} & \text{c} \end{array}$$

and these are the *dictatorial* profiles for a. This can be formally stated as:

Corollary 1 *Since $F(\cdot)$ is Weak Paretian in θ^a one person must top rank a and everyone else bottom rank it.*

Proof. *If at least one person does not top rank a then by Weak Paretian whatever is better for all must be chosen. By Lemma 4 there can be at most one person who does this.* ■

Now all we have to do is go back and contemplate Lemma 3 again.

Corollary 2 *If i top ranks a in θ^a then i top ranks b in θ^b, or i is a dictator.*

Proof. *Assume not or that $j \neq i$ top ranks b in θ^b, then by Lemma 3 we can assume j top ranks b in θ^a. But then the lower contour set of b in θ^a is larger than it is in θ^b, or $F(\theta^a) = b$—a contradiction.* ■

Beautiful, ehh? And quite a bit simpler than the proof in the book. All we have to do is find the critical profile, understand a few of it's properties, and then see one complicated implication of it.

1.2 Gibbard-Satterthwaite.

Like I said the previous result is fundamental because of the importance of Maskin Monotonicity. Maskin Monotonicity is important because it's the inductive implication of *independent person by person monotonicity*.

Definition 7 *We say that a choice rule is* independent Person by Person Monotonic (IPM) *if* $\forall \{\theta_i, \theta'_i\} \in \Theta^2 \forall \theta_{-i} \in \Theta^{I-1}$*:*

$$F(\theta_i, \theta_{-i}) \in L(F(\theta'_i, \theta_{-i}), \theta'_i) \text{ and } F(\theta'_i, \theta_{-i}) \in L(F(\theta_i, \theta_{-i}), \theta_i)$$

Intuitively this means given $\{\theta'_i, \theta_{-i}\}$ the choice at $\{\theta_i, \theta_{-i}\}$ is worse—and vice-a-versa. This obviously must be satisfied in order to get people to reveal their type, formally:

Lemma 5 $F(\cdot)$ *is truthfully implementable in dominant strategies if and only if it is IPM*

Proof. Assume not, then clearly $\rho_i = \theta_i$ is better than $\rho_i = \theta_i'$ if i is of type θ_i', or $\rho_i = \theta_i'$ is better if i is of type θ_i. ■

And we get the big result when we consider the set of all strict preference orderings, $\bar{\Theta}$ over A.

Theorem 2 *(Gibbard-Satterthwaite) Suppose that* $|A| \geq 3$ *and* $F(\cdot)$ *is onto,* $\Theta_i = \bar{\Theta}$*, then* $F(\cdot)$ *is truthfully implementable in dominant strategies if and only if it is dictatorial.*

Proof. What we will essentially do is show that IPM implies Maskin Monotonicity and then refer to Mueller-Satterthwaite. So consider θ and θ' such that for all i $L_i(F(\theta), \theta_i) \subseteq L_i(F(\theta), \theta_i')$ we must show $F(\theta) = F(\theta')$.

What we will do is consider a path from θ to θ', $\theta^0 = \theta$, $\theta^I = \theta'$, and in θ^j for $j \in \{1, 2, 3, ...I-1\}$ if $i \leq j$ $\theta_i^j = \theta_i'$ and if $i > j$ $\theta_i^j = \theta_i$. We will prove that the social choice function is constant along this path, or the $F(\theta^j) = F(\theta)$ for all $j \in \{0, 1, 2, 3, ...I\}$. This will imply the social choice function is Maskin Monotone.

So first we must determine $F(\theta_1', \theta_{-1})$. By IPM we know that $F(\theta_1', \theta_{-1}) \in L(F(\theta), \theta_1)$. By assumption we know that for all i $L_i(F(\theta), \theta_i) \subseteq L_i(F(\theta), \theta_i')$ thus $F(\theta_1', \theta_{-1}) \in L(F(\theta), \theta_1) \subseteq L(F(\theta), \theta_1')$. Notice that $F(\theta)$ is the highest rank alternative in $L(F(\theta), \theta_1')$, and that by strict preference orderings everything else is strictly worse. Thus if $F(\theta_1', \theta_{-1}) \neq F(\theta)$ player i will report θ_1 instead of θ_1'. Thus we must have $F(\theta_1', \theta_{-1}) = F(\theta)$.

We can now go to the next step and it will be exactly the same as this, by induction we are done. ■