A limited "proof by example" of Arrow's Theorem

The following example illustrates the role of each of the various axioms in proving Arrow's Theorem. The example was taken from *Positive Political Theory, Vol. I* by David Austen-Smith and Jeff Banks (1999). For the example we will assume that there are three alternatives and two players, and that the players have strict preferences (which we've been assuming all along, but isn't actually necessary in order to prove Arrow's Theorem).

$N = \{1, 2\}$ and $X = \{x, y, z\}$. If individuals have strict preferences, there are 36 possible preference profiles. All can be summarized in the following 6×6 table:

	xyz	**xzy**	**yxz**	**yzx**	**zxy**	**zyx**
xyz						
xzy						
yxz						
yzx						
zxy						
zyx						

The rows represent the possible preferences Player 1 could have, and the columns represent the possible preferences Player 2 could have. An entry in the table represents the social ranking produced by our *social welfare function*, or *preference aggregation rule*, $\mathcal{F}$.

Recall, all that we assume about our aggregation rule is that it satisfies *Pareto, unrestricted domain, IIA, transitivity,* and *no dictator*.

<u>Unrestricted domain</u> tells us that we must be able to fill in every cell in our table with an ordering of x, y and z.

Step 1: Use <u>Pareto</u> to fill in as many cells as possible. Pareto tells us that if the two players feel unanimously about the ordering of two alternatives, then the social ordering equals their individual orderings. For example, in cell (1, 2) (i.e. Row 1, Column 2) Player 1 has preference ordering (xyz) and Player 2 has preference ordering (xzy). They feel unanimously that $x \succ y$ and that $x \succ z$. Therefore we put "xy, xz" in our cell.

	xyz	**xzy**	**yxz**	**yzx**	**zxy**	**zyx**
xyz	xyz	xy, xz	xz, yz	yz	xy	
xzy	xz, xy	xzy	xz		xy, zy	zy
yxz	xz, yz	xz	yxz	yz, yx		yx
yzx	yz		yz, yx	yzx	zx	yx, zx
zxy	xy	xy, zy		zx	zxy	zx, zy
zyx		zy	yx	yx, zx	zx, zy	zyx

Step 2: Make a tie-breaking assumption in ONE cell. Pareto only gets us so far. We will need to make an assumption about the outcome in *one* cell to finish filling in the table. Assume that $x \succ y$ in cell (1, 4). With only this assumption you'll see that we can fill in the rest of the table.

Note that the assumption we made doesn't actually matter—we could also assume that $y \succ x$ in cell (1, 4) and our result would still hold. What we're trying to show is that using only our assumptions of Pareto, U.D., IIA and transitivity we will always get a dictator. Assuming $x \succ y$ in cell (1, 4) will end up making Player 1 the dictator (since he prefers x to y). If we assumed that $y \succ x$ in cell (1, 4) then Player 2 would emerge as the dictator. Thus, this assumption can be made *without loss of generality*.

Step 3: Apply <u>IIA</u> to determine as many (x, y) rankings as possible. Using the rankings we generated through the Pareto assumption and our assumption that $x \succ y$ in cell (1, 4) we can start applying IIA. We'll begin with the (x, y) social rankings. We'll also simultaneously apply transitivity if we can; so if we have "xy, yz" in a cell we'll automatically write it as "xyz".

We apply IIA to the (x, y) rankings by looking at whether the players' individual (x, y) rankings are unchanged *across* two different cells. Then if we know the social (x, y) ranking in one cell, it must be the same in the other.

For example, we assumed that $x \succ y$ in cell (1, 4). In this cell, Player 1 has preference $x \succ_1 y$ and Player 2 has preference $y \succ_2 x$. Thus, *whenever* Player 1 prefers x to y and Player 2 prefers y to x, we must have the social ranking $x \succ y$. Therefore we can say that $x \succ y$ in cells (1, 3), (1, 6), (2, 3), (2, 4), (2, 6), (5, 3), (5, 4), and (5, 6), because in all of these cells Player 1 prefers x to y and Player 2 prefers y to x.

	xyz	**xzy**	**yxz**	**yzx**	**zxy**	**zyx**
xyz	xyz	xy, xz	xyz	xyz	xy	xy
xzy	xz, xy	xzy	xz, xy	xy	xy, zy	xy, zy
yxz	xz, yz	xz	yxz	yz, yx		yx
yzx	yz		yz, yx	yzx	zx	yx, zx
zxy	xy	xy, zy	xy	zxy	zxy	zxy
zyx		zy	yx	yx, zx	zx, zy	zyx

We're going to use this same logic to fill in the rest of the table.

Step 4: Apply <u>IIA</u> to determine as many (x, z) and (z, y) rankings as possible. We'll do this exactly the same way as in the previous step. In the previous step we used cell (1, 4) as our starting point. In this step we'll use cells (1, 4) and (5, 4) as our starting points.

In cell (1, 4) we know by transitivity that $x \succ z$. Thus, *whenever* Player 1 has preference $x \succ_1 z$ and Player 2 has $z \succ_2 x$, then socially $x \succ z$. This occurs in cells (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), and (3, 6).

Similarly, in cell (5, 4) we know by transitivity that $z \succ y$. Thus, *whenever* Player 1 has preference $z \succ_1 y$ and Player 2 has $y \succ_2 z$, then socially $z \succ y$. This occurs in cells (2, 1), (2, 3), (2, 4), (5, 1), (5, 3), (6, 1), (6, 3), and (6, 4).

	xyz	**xzy**	**yxz**	**yzx**	**zxy**	**zyx**
xyz	xyz	xy, xz	xyz	xyz	xy, xz	xy, xz
xzy	xzy	xzy	xzy	xzy	xzy	xzy
yxz	xz, yz	xz	yxz	yxz	xz	yxz
yzx	yz		yz, yx	yzx	zx	yx, zx
zxy	xy, zy	xy, zy	xy, zy	zxy	zxy	zxy
zyx	zy	zy	zyx	zyx	zx, zy	zyx

Step 5: Apply <u>IIA</u> to determine as many (y, z) and (z, x) rankings as possible. Again, we'll do this exactly the same way as in the previous step. In this step we'll use cells (3, 6) and (6, 3) as our starting points.

In cell (3, 6) we know by transitivity that $y \succ z$. Thus, whenever Player 1 has preference $y \succ_1 z$ and Player 2 has $z \succ_2 y$, then socially $y \succ z$. In what cells can we use this information?

Similarly, in cell (6, 3) we know by transitivity that $z \succ x$. Thus, whenever Player 1 has preference $z \succ_1 x$ and Player 2 has $x \succ_2 z$, then socially $z \succ x$. In what cells can we use this information?

	xyz	**xzy**	**yxz**	**yzx**	**zxy**	**zyx**
xyz	xyz	xyz	xyz	xyz	xyz	xyz
xzy	xzy	xzy	xzy	xzy	xzy	xzy
yxz	xz, yz	xz, yz	yxz	yxz	xz, yz	yxz
yzx	yzx	yzx	yzx	yzx	yzx	yzx
zxy	zxy	zxy	zxy	zxy	zxy	zxy
zyx	zy, zx	zy, zx	zyx	zyx	zx, zy	zyx

We're almost done; we've applied IIA to (x, y), (x, z), (z, y), (y, z) and (z, x) pairs. All that is left is to apply IIA to the remaining (y, x) pairs whose rankings are still unspecified.

Do you notice a pattern emerging in this table?

Step 6: Apply IIA to determine the remaining (y, x) rankings. In this step we'll use cell (4, 5) as our starting point. In this cell we know, by transitivity, that $y \succ x$. Thus, whenever Player 1 has preference $y \succ_1 x$ and Player 2 has $x \succ_2 y$, then socially $y \succ x$.

	xyz	**xzy**	**yxz**	**yzx**	**zxy**	**zyx**
xyz	xyz	xyz	xyz	xyz	xyz	xyz
xzy	xzy	xzy	xzy	xzy	xzy	xzy
yxz	yxz	yxz	yxz	yxz	yxz	yxz
yzx	yzx	yzx	yzx	yzx	yzx	yzx
zxy	zxy	zxy	zxy	zxy	zxy	zxy
zyx	zyx	zyx	zyx	zyx	zyx	zyx

Thus, we've shown that the consequence of assuming $x \succ y$ in cell (1, 4) and requiring our social welfare function $\mathcal{F}$ to satisfy transitivity, Pareto, IIA, and unrestricted domain is to make Player 1 a "dictator".

As mentioned earlier, if we had assumed that $y \succ x$ in cell (1, 4), then (by the symmetry of the example) Player 2 would have turned out to be the dictator.

Our only other choice was to make $x \sim y$ (i.e. to make them socially indifferent). However, this would have been impossible. To see this, note that we would have had $x \sim y \succ z$ in cell (1, 4). Now consider cell (2, 6), with profile ($x \succ_1 z \succ_1 y$), ($z \succ_2 y \succ_2 x$). By Pareto, cell (2, 6) has $z \succ y$. By IIA we get $x \succ z$ in cell (2, 6), since the players' (x, z) rankings are unchanged between cells (1, 4) and (2, 6). By transitivity we get $x \succ y$ in cell (2, 6). However, by IIA it must also be the case that $x \succ y$ in cell (1, 4), because the players' (x, y) preferences are unchanged between cells (1, 4) and (2, 6): *a contradiction*. Thus, in cell (1, 4), our $\mathcal{F}$ must decide strictly in favor of either Player 1 or 2, leading to the "dictatorship" that Arrow says must be true in general. □