Summer August 26, 2021

Marquis de Condorcet and von Neumann: Voting Paradox, Economic Growth and Linear Programming

Lester G Telser

8/26/21

Marquis de Condorcet and von Neumann: Voting Paradox, Economic Growth and Linear Programming

L . G . Telser

1 Introduction

In common with many great discoveries von Neumann' s minimax theorem (1928) had unintended consequences. It was von Neumann's original intention to describe the best conservative way to play 2-person zero-sum games. It turned out that this theorem is the mathematical foundation of linear programming. George Dantzig's invention of the simplex method for solving linear programming problems is arguably a major source of economic growth not only in the last half of the 20th century but also to the present. (See Forward in Dantzig and Thapa, 1997).

The expected value of payoffs is used not because it implies transitivity but because it underlies the minimax theorem. The minimax theorem needs mixed strategies. A bivariate game could not have a saddlevalue were it confined to pure strategies. Whether payoffs are measured in money or by an increasing function of money called utility, the validity of the minimax theorem depends on a well chosen, convex combination of payoffs. The weights of the convex combination are interpreted as probabilities. Only on this interpretation does the convex combination become the expected value of the payoffs. Of course, the expected value of the payoffs does not describe the actual payoff in any single play of the game. That the expected value of payoffs is transitive is not why it is used.

This interpretation is implicit in von Neumann and Morgenstern (1947, sec. 3.5.2). Also worth reading is their deep analysis of how to handle comparisons among different goals.(1947, Appendix, especially A.3.4).

Long before von Neumann and Morgenstern, Marquis de Condorcet studied voting mathematically. He discovered the intransitivity of choices among three alternatives given by majority rule. He states, "Thus the form of the assemblies which decide men's lot is much less important for their happiness than the enlightenment of those who sit in these assemblies and the progress of reason will contribute more to the happiness of peoples than the form of political constitutions." Condorcet, 1785, p. 57.

2 Transitivity and the v. Neumann & Morgenstern Utility Indicator

Transitivity seems to have acquired sterling credentials owing to its place in the v. Neumann - Morgenstern cardinal utility function. My brief description hardly does justice to their subtle reasons. They introduce their utility function to meet objections that a mixed strategy creates risky monetary payoffs. Other things equal, it appears that more money is always better than less. However, a mixed strategy is not simple. A player may win more money with a lower probability or win less money with a higher probability. Which the player would prefer is not obvious because it depends on two things, the amount of money and the probability of getting it. If the probabilities were the same, then presumably everyone would prefer more to less money. If the money were the same, then presumably everyone would prefer the higher to the lower probability of getting the money. The pair, money and probability, is a vector with two coordinates. To derive transitivity from a vector one must transform it to a scalar. Von Neumann & Morgenstern did this in several steps. First, they replace money by a utility indicator that moves pari passu with the amount of money. Now a player decides on the basis of the utility of the amount of money, U(A), not on the amount of money, A. Since the utility indicator is a scalar, A > B > C implies U(A) > U(B)> U(C). So far there is an implication of transitivity. The choice between A and B is clear if A>B and the probability of getting A is not less than the probability of getting B. However, the choice is between the joint commodity, an amount of money and the probability of getting it, the choice is not between amounts of money ignoring the probabilities of getting the amounts of money. Von Neumann and Morgenstern reduce the vector to a number by assuming the choice is made on the basis of *expected* utility, not on the basis of *realized* utility. Expected utility is an average of the realized utilities of the various outcomes weighted by the probabilities of these outcomes.

The solution is not straightforward for this joint product, an amount of money and a probability. Assume the vector (A, p) is preferred to the vector, (B, q), written $(A,p)\succ(B,q)$. The joint product (A, p) is preferred to the joint product (B, q), if p and q are probabilities, so they are positive numbers between 0 and 1. Also assume that $(B, q)\succ(C, r)$, where r is a probability. If there were transitivity, then

(1) $(A,p) \succ (B,q)$ and $(B,q) \succ (C,r)$ would imply $(A,p) \succ (C,r)$.

However, the conclusion in (1) does not follow from anything assumed so far. Assume $A > B > C$. Also assume the bigger is the prize, the smaller the probability of getting it so that $p < q < r$. Finally, assume these probabilities sum to 1, $p+q+r=1$. All these assumptions do not suffice to establish transitivity in (1). Transitivity depends on the shape of the utility indicator. Take the simplest case, linear utility and assume that

(2) $p A > q B$, $p < q$, & $q B > r C$, $q < r$.

Now (2) implies transitivity because

(3) $p A > q B > r C$,

It must not escape attention that to get (3) the model is enriched by the assumption not only that the size of the prize varies inversely with the probability of getting it but also that the inequalities in (2) hold. The latter is an assumption about the world, not about individual preferences.

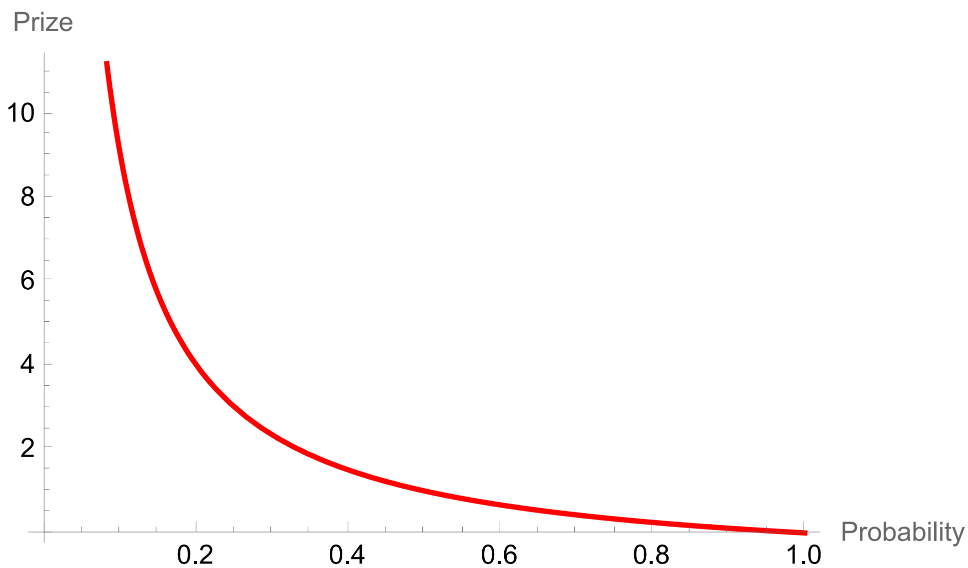
Utility becomes a random variable because the prize is a random variable. Von Neumann and Morgenstern replace utility as a random variable by a scalar, *expected* utility. It is a weighted average of the prizes with weights of the utility of each possible prize equal to the probability of winning it.

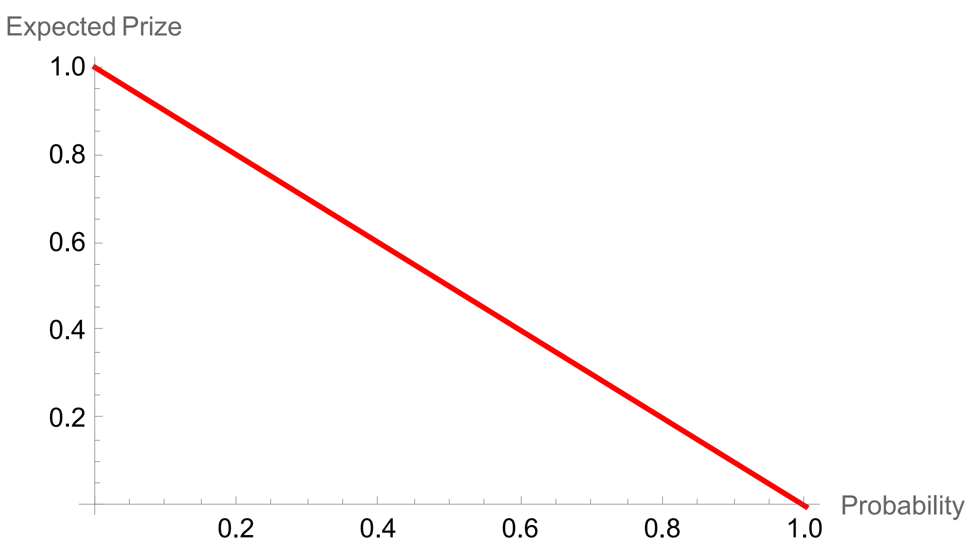
An instructive example shows how the inverse relation between the probability of winning a prize and the size of the prize is pertinent. Let p denote the probability of winning the prize w. Let r denote the return and let

(4) $w = \frac{1}{p} - 1$

(5) $r = \boxed{\begin{array}{c|c} 0 & \text{probability} = 1 - p \\ \hline w & \text{probability} = p \end{array}}$

Therefore, E(r) = 0 (1-p) + w p = 1 - p. The expected value of the return is 1- p .





Here transitivity and certainty is worth zero.

Because this example is unfamiliar, it needs some explanation. If X is a random variable so that $\Pr\{X \le x\} = \int dF(x)$, then the expected value of X is given by E(X)=$\int x\, dF(x)$. The formula

w= (1/p) -1 seems to make the random variable depend on the probability. However, this is consistent with standard probability theory. In standard form, let f(w) be the probability density function of the random variable w. Hence E(w) =$\int w\, f(w)\, dw$. Equation (4) yields

p=1/(1+w) so let f(w)=1/(1+w) be the pdf of w. Now

$\int f(w)\, dw = \int \frac{1}{1+w}\, dw = \log(1+w)$. For a proper pdf, confine the random variable w to the finite range from 0 to an arbitrary, finite upper bound V. Given this finite upper bound,

(5) E(w) = $\int \frac{w}{1+w}\, dw = w - \log(1+w)$.

Dividing by the term log (V+1) to have a proper pdf,

(7) E(w)=(w - log(1+ w) $|_0^V) / \log(V+1) = (V / \log(V+1) - 1$.

Since V/log (1+V) is an increasing function of V, the expected value is unbounded

3 Linear Programming in the Theory of Demand

According to the minimax theorem, an mXn matrix A has a saddle value given by the nonnegative m-tuple x_o and the nonnegative n-tuple y_o so that for all nonnegative x and y,

(1) $x\, \mathrm{A} y_o \leq x_o\, \mathrm{A} y_o \leq x_o\, \mathrm{A} y$.

In the application to games the coordinates of (x, y) are interpreted as probabilities. The vectors in the pair (x_o, y_o) are the saddle points and $x_o\, \mathrm{A} y_o$ is the saddlevalue of A.

In 1932 von Neumann gave a lecture to the mathematics department, Princeton University, on economic growth in equilibrium. His interpretation of the saddlevalue and saddle points is entirely different than his 1928 article on game theory that interprets x and y as probabilities. In the application to economic growth, the nonnegative coordinates of the x-vector describe n inputs to m production processes. It is the physical combination of the inputs required to make the various outputs used as inputs in the economy. An even more surprising change is the interpretation of the coordinates of the y-vector. No longer probabilities, now they are the relative prices of inputs and outputs in the economy. Almost 15 years later George Dantzig invented the simplex method that brought to life practical application of this result. It became known as linear programming. Dantzig vividly describes von Neumann's major role in this invention (Dantzig and Thapa, 1997).

A novice learning the standard economic theory of demand is told that a consumer selects that bundle of goods which maximizes utility subject to a budget constraint. More thoughtful people, including D. H. Robertson, wonder what is utility. They are told it is a rubbery thing. What it is actually does not matter as long as it stretches more, the more commodities it takes into account. After this assurance comes a long story about real income and how price changes affect the quantity bought. A student who emerges from learning all this and who gets a job in the marketing department of some business has few tools to handle actual demand problems.

A critic may claim that my use of linear programming to describe what is best to buy tacitly allows transitivity in through the back door. This deserves a closer look. Assume a buyer decides what to buy from a pertinent class of various commodities on the basis of two considerations, the requirements that m different commodities can satisfy and their prices. The pertinent class includes those commodities that are related by the requirements they can satisfy. Think of specific means of transportation such as automobiles of various types.

The mXn matrix $A = [a_{i\,j}]$ describes a collection of m related commodities in terms of the n requirements they satisfy. Thus $a_{i\,j}$ is the amount of requirement j in a unit of commodity i. This does not assume each commodity can satisfy all n requirements. It does assume each commodity can satisfy at least one. Let the coordinates of the n-vector r denote the minimal amounts of the desired requirements. This allows choice of several different commodities to satisfy the n requirements. The coordinates of the m-vector p are the unit prices of the m commodities. The coordinates of the m-vector y are the nonnegative quantities of the m commodities. A buyer seeks a combination of the m commodities that satisfies the n requirements at the least total cost. The problem is
(1) Primal: Min p y with respect to $y \geqq 0$ subject to $A\,y \geqq r$.
Because the objective is the total cost in monetary terms, a scalar, it may seem the smaller is the outlay, the better so the objective is

transitive. This is wrong. The best bundle is not necessarily the cheapest. The best bundle is the one that satisfies all requirements at the least cost. A commodity in the best bundle is not necessarily the cheapest. The most expensive commodity can be in the solution. As long as there is more than one requirement, the linear programming algorithm shows how to find the solution. The solution of the dual problem finds shadow prices for the requirements.

A serious complication faces both the linear programming algorithm and the standard demand theory. It concerns social aspects of commodities, studied more than a century ago by Thorstein Veblen (1899). The desirability of many commodities depends on the prestige they confer and on society's opinions about them. Money itself is a leading example of a commodity whose public aspects are paramount. Money is not only a numeraire for making comparisons but it also derives its value from its acceptability as a means of payment. The inequalities should incorporate the social and public aspects of commodities and describe the users of the commodities, who they are, how many there are and their relations to the individual buyer.

Although the algorithm ignores the public and social aspects of commodities, still it furnishes useful information about buyers' choices. A solution has at most as many commodities as requirements. Because commodities are usually more numerous than the requirements they can satisfy, a solution excludes many commodities. A fall in prices of unbought commodities need not induce a buyer to replace those actually bought. Similarly, a rise in the prices of the purchased commodities need not cause removing them from the best bundle. Also, the model applies to an individual buyer. Different buyers can have different requirements. Even if all buyers face the same prices, their purchases can differ. This complicates analysis of the total effects of price changes.

The companion to the primal problem is the dual problem. The dual problem seeks the shadow prices of the n requirements that *maximizes* the value of the requirements. The coordinates of the n-vector x are the shadow prices of the n requirements. They must be nonnegative. The shadow prices impart a value to the requirements derived from commodity prices according to the nature and amount of requirements that the commodities can satisfy.

(2) Dual: Max x r with respect to $x \geqq 0$ subject to $x A \leqq p$.

Only those commodities are bought that yield as much value from the requirements that they satisfy at their unit prices. A commodity that fails this test is not bought.

From $A y \geqq r$, it follows that $x A y \geqq x r$. From $x A \leqq p$, It follows that $x A y \leqq p y$. Hence $p y \geqq x A y \geqq x r$. Let x^o denote a solution of the dual and y^o a solution of the primal.

(3) $p y \geqq p y^o = x^o A y^o = x^o r \gneqq x r$.

Expression (3) says the least a buyer pays equals the most a seller gets.

4 Money as a Numeraire

Beginning with Bernoulli's proposed solution of the St Petersburg paradox that supposes a dollar is worth more to the poor than to the rich, the fashion in economic theory has become utility, not money, in the theory of demand. By the end of the 19th century it was a widely accepted axiom of economics that ordinal utility would suffice for a theory of demand for private goods. A public good is not so simple. If the benefit of a public good is widespread equally among the individuals in the community, then payment for it should vary inversely with the wealth of the beneficiary. In the private sector only a monopolist would maximize his net return in like manner. A monopolist could charge the rich more than the poor one for the same good. It is only the presence of rivals who can use the same technology that prevents exploitation by means of such price discrimination. The supplier of a public good, the state, is the epitome of monopoly. Even if the public somehow decides what and how much public goods to have, transitivity remains a red herring. It is so not only owing to Condercet's analysis of the deficiencies of majority rule but also, and more importantly, because a public good must also satisfy many different requirements. Hence many different public goods compete for different purposes by satisfying different requirements. Monetary measures of the costs and benefits still command best way th handle these jobs.

It may seem preposterous to apply linear programming algorithms to public goods. A few examples can dispel doubts. The defense budget contains many cases where linear programming is actually used to choose among alternatives. Planes for the Air Force, ships for the Navy, radar systems, communication gear, armored vehicles are only a few examples where linear programming is used. Some inequalities are in use for public goods that are motivated by considerations unusual in the private sector. These include effects on employment, carbon emissions, diversity of employees, regional equity and so forth.

Even to mention public goods is an invitation to controversy. It clarifies the issues by remarking how capital goods come into existence. A capital good is durable. Those who finance it do so in the belief that in due course they will get enough revenue to pay the total cost of the capital good including interest. A capital good must be productive enough over its life including the costs of maintenance and repair to make this happen. Whether a capital good is the property of an individual, a private company or any other institution in the economy, the arithmetic of amortization is the same. Every accountant knows the formula. It is in every pocket calculator. Outlays on capital goods that fail this test will not be embraced by the private sector. Which organizations will finance such goods and on what terms enters another terrain.

5. Goals of Game Theory

There are two main candidates, not necessarily rivals, to be the goal of game theory; the first is normative, the second is positive. The normative goal describes how to play a game well. The positive goal is scientific. It aims to make a model that predicts how people play. A thought experiment can explain this, the 2-person game, 'Rock, Scissors, Paper'. Each player has three alternatives. Simultaneously, each may show a fist for rock, two fingers for scissors or a spread out hand for paper. If they show the same thing, then nobody wins or loses and the payoff is 0. Rock beats scissors, scissors beats paper and paper beats rock. The payoff matrix of this game for the row player is

(1)

row/column	Rock	Scissors	Paper
Rock	0	1	−1
Scissors	−1	0	1
Paper	1	−1	0

If the row player picks row i with probability x_i and the column player picks column j with probability y_j, then the row player's expected return is

(2) $E(\text{row}) = x_1(y_2 - y_3) + x_2(y_1 - y_3) + x_3(y_1 - y_2)$.

Therefore, the row player's expected return is zero no matter what row he chooses if the column player is equally likely to choose each column. However, if the column player somehow knows which row his adversary will choose, then he would not make his selections at random if he wants to win. There is also an advantage of delegating the selection to a device capable of random selections. It provides security since neither player knows what the robot will choose. A disadvantage is that the column player who uses a random device thereby delegates playing the game to a robot. Indeed, if the two play often enough, then their average return would be close to zero. Using a mixed strategy to play the game introduces insurable risk plus boredom. Those who wish to play this game for pleasure, entertainment or excitement would not delegate play to a random device. They might believe they could guess their adversary's pick. They might even derive pleasure from playing against a slot machine.

This game is not a game of chance. Chance enters only by virtue of the minimax that requires a mixed strategy. The two players are not obliged to employ the mixed strategy that yields the minimax. They follow the dictates of this mixed strategy if and only if they seek a conservative mode of play on the hypothesis they choose to play. A conservative person can have the certainty of neither gain nor loss by not playing.

The 3 choice version of this class of games is exceedingly simple. It may give the impression that the minimax theorem is trivial. It is far from being trivial. Even the 4 choice version is decidedly more challenging than the 3 choice version. The payoff matrix for the row player in the 4 choice version is T4 as follows.

(3) $\quad T4 = \begin{array}{|c|c|c|c|} \hline 0 & 1 & -1 & 1 \\ \hline -1 & 0 & 1 & -1 \\ \hline 1 & -1 & 0 & 1 \\ \hline -1 & 1 & -1 & 0 \\ \hline \end{array}$

The expected return to the row player if both players use a mixed strategy is given by
(4)
E(row)=
$x_1(y_2 - y_3 + y_4) + x_2(-y_1 + y_3 - y_4) + x_3(y_1 - y_2 + y_4) + x_4(-y_1 + y_2 - y_3)$
The saddle points for the saddlevalue are equal and set $x_4 = y_4 = 0$. Consequently, the saddlevalue = 0.

The saddlevalue = 0 for all games in this class. However, there is a surprising difference between the games with an even number and an odd number of choices. Let m denote the number of choices. If m is even, then $x_m = y_m = 0$ and the remaining values of the x's and y's =1/(m-1). The minimax strategy in effect converts an even m game into an odd m game. For odd m, all choices have the same chance of being chosen, 1/m.

Repeated play of this game produces a sequence of independent random variables. With 3 alternatives in the thought experiment, the variance is 2/3. With m alternatives, the variance is (1-1/m). The density function is 1/3 only for m = 3 as shown in Figure 1. Compare it with Figure 2. The bigger is the number of alternatives, the smaller the probability of 0. In the limit, the probability of a tie is 0 and the variance is 1.

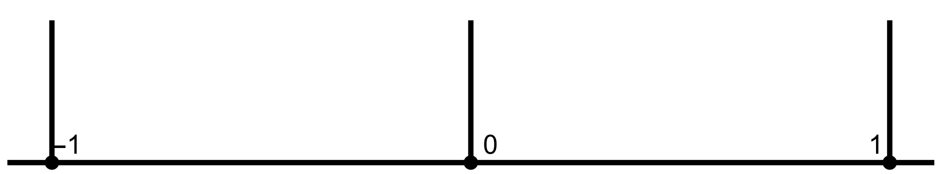


Figure 1 m = 3

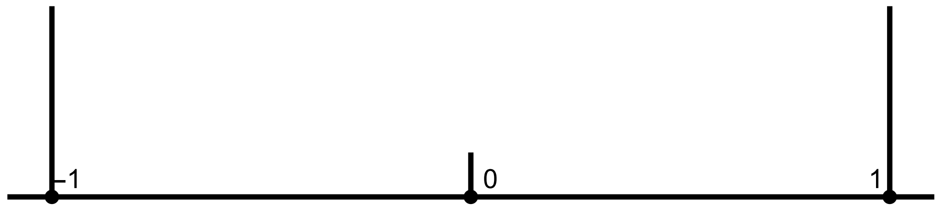


Figure 2, m = 10

This argument concludes that game theory can give poor descriptions of how people play games. A main contribution of game theory is the discovery of the minimax theorem, the mathematical foundation of linear programming. Granting this, and without the simplex algorithm invented by Dantzig, the minimax theorem would be a theoretical curiousum.

References

Condorcet, Marquis de. 1785. The Academician under the Old Regime. In Baker, Keith Michael. 1976. Condorcet: Selected Writings. Indianapolis: Bobbs-Merrill.

Dantzig, George B. and Mukund N. Thapa. 1997. Linear Programming. New York: Springer Verlag.

von Neumann, John. [1928] 1959. On the Theory of Games of Strategy. trans. Sonya Bargmann in Contributions to the Theory of Games. IV. Princeton: Princeton University Press.

______. [1932] 1945-6. A Model of General Economic Equilibrium. Review of Economic Studies. 13:1-9.

von Neumann, John and Oskar Morgenstern. 1947. The Theory of Games and Economic Behavior. 2d ed. Princeton University Press.

Robertson, D. H. 1952. Utility and All That. London: Allen & Unwin.

Telser, L. G. 2011. A Realistic Model of Demand. Part I. Unpublished. Lester G. Telser, Selected Works on BEPress.

Veblen, Thorstein. 1899. The Theory of the Leisure Class. New York: Macmillan.

PDF