Risk Aversion, Stochastic Dominance, and Rules of Thumb: Concept and Application

Vivek H. Dehejia
Carleton University and CESifo

Email: vdehejia@ccs.carleton.ca

January 14, 2008

JEL classification code: D81

Consider the following generic and fairly narrowly defined choice problem. An individual must choose from amongst a discrete and finite set of lotteries. Suppose for concreteness that each lottery represents a monetary payoff, and that all the lotteries are constructed so as to be comparable. As a running example, the lotteries could represent incomes in different countries in given years, and comparability could be ensured, at least in principle, by converting to a common metric using inflation- and purchasing power parity-adjusted exchange rates. Each lottery is characterized by a corresponding distribution function, that is known with certainty. The uncertainty arises because, if the individual picks a particular distribution, he will receive a payoff that is a random draw from that distribution. How is he to choose amongst these lotteries? Assuming that his preferences are such that they admit of a Von Neumann-Morgenstern (VNM) expected utility representation greatly simplifies the problem. Now, the individual will pick the lottery that gives him the maximum level of expected utility.

If the individual is risk-neutral, so that his expected utility function is linear in the monetary payoff, the problem is not especially interesting. From conventional economic theory, we know that maximizing expected utility in this case will reduce to maximizing the expected payoff, given the linearity of the expectation operator. The individual will simply pick the lottery that has the highest corresponding expected value, assuming, as I shall do throughout, that this (as all other relevant moments) exists and is well-defined for all the lotteries. In our example, this would involve picking the country whose income distribution has the highest mean income, i.e., income per capita.

Things get more interesting if the individual is risk-averse. Now, there is no general rule for picking amongst these competing lotteries, and, in theory, the entire distribution function for each of the lotteries is relevant. With a finite, discrete set of lotteries, the problem can be solved, at least in a mechanistic way, by computing the realized level of expected utility, for a given utility function, for each of the distributions, and picking the distribution with the highest corresponding expected utility level. But this approach does not yield much, if any, insight. An alternative, indeed the "right" approach, would be to employ the concept of stochastic dominance, of various orders, as developed in the literature on the economics of uncertainty. This corresponds to the concept of revealed preference in standard utility theory. Stochastic dominance, roughly speaking, allows, for any pairwise comparison, to say which of two distributions is "better" than the other. First-order stochastic dominance (FSD) exists, intuitively, if one distribution lies everywhere "above" another one: in this case, it can be proved that anyone with VNM preferences will prefer the higher distribution. This is only a partial ordering, however, as FSD very often does not exist when comparing pairs of distributions. The next concept is that of second-order stochastic dominance (SSD): this exists, intuitively, if one distribution is less "risky" than another, and, in this case, it can be proved that anyone with VNM preferences that are risk-averse will prefer the less risky distribution. While this, too, is a partial ordering, it takes us quite a bit further towards the problem our individual is facing. (There is a dauntingly technical literature addressing higher orders of stochastic dominance, and corresponding restrictions on the utility function thereby implied in VNM and other frameworks, that I will not be exploring here.) Given a finite, discrete set of lotteries, there exists a finite, discrete set of pairwise comparisons amongst them: if one could find a particular distribution that SSD each of the others in a series of pairwise comparisons, one has found, pari passu, the expected utility maximizing choice.

The problem, as described above, seems dry and pedantic, and appears to have a narrow range, if any, of applications. The opposite is true. Consider, again, our running example, in which the lotteries are considered income distributions, and let us now make the problem a bit more precise as well as more concrete. Suppose, now, our individual is a putative migrant, considering which out of a group of n countries he should choose as his preferred location. For simplicity, assume that migration is costless. Thus, without loss of generality, assume that his current location is one of the n locations. The individual potentially cares about various characteristics of the locations. For simplicity, assume that he cares only about the income he may enjoy. Strictly, he should care about some measure of wealth, or lifetime or "permanent" income, if a forward-looking agent. Restrict the analysis to the flow of income, for simplicity. The individual does not know with certainty what level of income he will enjoy in each location. He knows the probability distributions over income levels in each of the n locations. Let us suppose, for simplicity, that he assumes that, were he to migrate to location i, the income he would thereafter enjoy is a random draw from the probability distribution over income corresponding to location i. This assumption is not necessarily innocuous: the putative migrant, if sophisticated enough, would presumably prefer to derive a conditional probability distribution, conditioning on relevant prior factors such as gender, education, work experience, etc. These considerations, too, are excluded for simplicity. Assume, further, that income is computed so as to ensure commensurability, i.e., each location's statistical agency uses an identical methodology, and that each location's statistics are expressed in a common metric, e.g., PPP-adjusted US dollars fixed to a base year.

With the problem thus defined, which location will he choose? As it happens, this is an exact instance of the generic type of problem that we have been discussing, and it is related to ongoing debates about the appropriate ways of comparing incomes internationally, far from an arcane and academic matter. As a pertinent example, in a recent review article in the influential policy journal, Foreign Affairs, Joseph Stiglitz (2005) comments on the fact that in recent times mean income has risen in the United States, while median income has actually fallen. He continues: "Consider the following thought experiment: If you could choose which country to live in but would be assigned an income randomly from within that country's income distribution, would you choose the country with the highest GDP per capita? No. More relevant to that decision is median income … As the income distribution becomes increasingly skewed, with an increasing share of the wealth and income in the hands of those at the top, the median falls further and further below the mean. That is why, even as per capita GDP has been increasing in the United States, U.S. median household income has actually been falling."

This rich quotation from Stiglitz contains, inter alia, an assertion about the evolution of the shape of the income distribution in the United States:

whether this is accurate or not is an empirical question, that I and Marcel Voia (in progress) take up elsewhere. It also contains a claim about what I would call a "rule of thumb" choice rule in the context of just exactly the type of generic problem we have been considering. In particular, Stiglitz seems to be suggesting that, when faced with a choice over a set of income distributions that are rightward skewed, a risk-averse individual would do better to pick the one with the highest median, not the highest mean. This would be, presumably, because the mean, in some sense, overestimates the "true" centre of the data, compared to the median. (The mean would, presumably, underestimate, in this sense, if the distribution were leftward skewed.) This would seem to accord with common sense. After all, a few very rich individuals raise US mean income, but do nothing to the median: so it is surely sensible to imagine that you are going to end up somewhere around the median, and discount the effect that Bill Gates and his ilk have upon the mean.

I will note, in passing, that there is also a more abstract, indeed philosophical, interpretation to the situation just described. One can imagine, not a putative migrant, but a rational agent, behind a Rawlsian veil of ignorance, deciding on what sort of society he would like to live in. The only modification is that there would be now, not a finite, but a (hopefully countably) infinite number of distributions to consider, and the choice of a particular distribution represents, not the choice of a location where one would like to migrate, but the shape of a just society to which one wishes to belong. John Rawls' celebrated investigation along these lines yielded the "maximin" rule: society should maximize the well-being of its least-well off member. One could imagine constructing a neo-Rawslian, call it Stiglitzian, political theory in which the chosen rule is to maximize the well-being of the median individual in society. I will not pursue these philosophical reflections further in this paper.

How does this common sense intuition, and the rule of thumb that it generates, i.e., "pick the highest median, not the highest mean", square with the theoretical concepts of stochastic dominance that we have just discussed? Strictly, there need be no correspondence between FSD, SSD, or other higher orders of stochastic dominance, and simple summary statistics such as the mean or median. Nor does expected utility theory provide any guidance, for there is no generally known class of sub-utility function for which median-maximizing behavior will typically be optimal. (The closest such result, that does not really help us, is in statistical decision theory, where minimizing a mean absolute deviation loss function yields the median as a solution.) My approach, therefore, will have to be heuristic. What I shall do is to consider a real world application, our running example of income distribution, with empirically constructed distribution functions, invoke the VNM framework, and check to see if there is a rough correspondence in practice, if not in theory, between the optimal solution for a risk-averse expected utility maximizer and the solution found by using a simple rule of thumb.

My application, therefore, is to the policy-relevant problem of international income comparisons, that flows naturally from the Stiglitz quotation. Here, I draw upon work in progress by myself and Marcel Voia, already mentioned. To make the location choice concrete, consider just two countries, the United States and Canada. This is a sensible comparison, as the various non-economic factors that might play a role in choosing between the US and Canada, such as geography, language, culture, history, social norms, etc., are about as close as they could be between any pair of developed countries. We have annual income data for the US and Canada for the years 1993 – 2001, but, when we eliminate years with missing data for one or the other country, we have comparable data for the following six years: 1993, 1994, 1996, 1997, 1998, and 1999. All data are converted into 1993 US dollars to allow cross section and time series comparisons. We can apply stochastic dominance tests to the six pairwise comparisons that we have, but, before doing so, it might be more intuitive to compare the solutions yielded by various rules of thumb to what would be optimal for an expected utility maximizer with a specific utility function. Consider the simplest case of risk aversion, logarithmic utility, which implies an Arrow-Pratt coefficient of relative risk aversion of unity. Table 1 summarizes the results obtained.

Along with the realized level of expected utility (for our individual with log utility), the table reports three summary statistics, mean, median, and mode, which could serve as potential rules of thumb. For each pairwise comparison, i.e., income in the US and Canada for a given year, we can see which has the higher level of expected utility, and then can check whether going by the rule of thumb, "pick the higher …" (mean, median, or mode) yields the "right" or "wrong" answer. The first thing to note is that in all of the comparable years, Canada yields higher expected utility than the US. For 1993, the mean and mode given the right answer, whereas the median, being identical, cannot choose. For 1994, the median and mode give the right answer, and the mean gives the wrong answer. This pattern is repeated in 1996 and 1997. In 1998, the mean, median, and mode all give the right answer, a pattern repeated in 1999.

What emerges from this application is at least qualified support for the Stiglitz intuition, and the "pick the higher median" rule of thumb. Only in 1993, with identical medians, the rule would not be able to choose, and then it could be supplemented with a sub-rule, "if medians are identical, pick the higher mean". With this modification, the median rule would choose perfectly, at least in this particular application.

For analytical completeness, we can conduct stochastic dominance tests on these distributions. The results are reported in Dehejia and Voia (in progress), and they confirm the intuition of the log utility special case: in each of the six comparable cases, the Canadian distribution SSD the US distribution. This makes more precise the appealing result noted above, that, at least in this particular application, a common sense rule of thumb, picking the higher median, yields the same result as a formal test for stochastic dominance, and hence is the "right" choice for a risk-averse expected utility maximizer to make.

One can imagine a number of other potential applications, which would be a useful way to test the merits of the various rules of thumb. Possibilities would include location choice (more generally), education choice, occupation choice, and comparison of stock portfolios, to name just a few. It remains for future work to construct tables analogous to Table 1, to test the various rules of thumb against the expected utility criterion and also to perform tests of stochastic dominance. The Stiglitz hypothesis, amongst others, remains an attractive possibility in the interim.

References

Dehejia, Vivek H., and Marcel Voia (in progress). "International Income Comparisons and Location Choice: Methodology, Analysis, and Implications", MS.

Stiglitz, Joseph E. (2005). "The Ethical Economist", in: Foreign Affairs, November/December.

Table 1

The data is conditional on a maximum income of 250000 Canadian dollars. 1993 is used as a base year.

USA data

year	# obs	Mean	Median	Mode	Expected Utility
1993	62721	21576	16690	1498.38	0.001069202
1994	62721	21744	15684	8689.29	0.000952034
1995	62721	21599	15731	6858.57	0.001238318
1996	65439	22332	15228	5372.53	0.000629032
1997	65438	22038	15223	7940.92	0.000490796
1998	65435	22023	15344	7978.89	0.000358032
1999	65435	22127	15644	7375.43	0.000391709
2000	x	x	X	X	x
2001	65445	20135	14094	11297.37	0.001292198

Note: x =missing data for the given year

Canadian Data

year	# obs	Mean	Median	Mode	Expected Utility
1993	29536	21583.84	16690	3117.14	0.001322103
1994	29362	21577.82	16267.4	9912.28	0.001149368
1995	x	x	X	X	x
1996	61064	21749.14	16303	8760.95	0.001598236
1997	61455	21957.46	16485	13464.01	0.000698163
1998	62140	22491.71	17125.5	9987.22	0.000529192
1999	58051	23049.02	17643.72	11497.9	0.000397064
2000	57380	23376.8	17770.38	14405.81	0.000435218

Note: x =missing data for the given year