

HARRIS SCHOOL WORKING PAPER
SERIES 06.05D

CHAPTER 2: DUAL-UTILITIES *

Howard Margolis

**This is a draft chapter from Cognition and Extended Rational Choice
(forthcoming Routledge 2007).*

Chapter 2. Dual-Utilities

The NSNX dual-utility structure requires a substantive defense: why bother unless it performs *better* (as stressed at the end of Chapter 1) than a model that does not require that. But before taking that up, consider some concrete situations which have nothing to do with the NSNX proposal, but which exhibit just the structure that appears in NSNX.

Suppose a real person (call her Ellie) was asked to organize a reception for her school's major contributors. Among other things, she decided to spend \$500 on floral arrangements: which she knew that in fact was a rather modest amount for the occasion. What would Ellie say if we asked what the floral arrangements were worth to her, and claimed that Ellie would be irrational to allocate that \$500 unless the value to her (allowing of course for her altruistic interest in the organization) is at least \$500? Her reaction would surely be that the \$500 reflects her view of how some money entrusted to her to be spent for the organization could best be spent. It is what (she judges) the spending is worth to the organization, which will routinely be far more than she sees it as worth to herself, even allowing for her altruistic interest in the organization.

For Ellie and the flowers, however, Ellie is morally (and legally... she could be sent to jail!) bound to treat the \$500 as the organization's money, not her own, so the example doesn't show that a person must ordinarily have the divided preferences required by the dual-utilities model. That empirical claim is not seriously tested by the flowers example. Rather, the example only illustrates how the standard tradeoff argument might be misleading if preferences, in fact, have the dual character I've sketched. For NSNX, resources allocated to G-spending are used to obtain social benefits, and it would be odd to try to think of them as fungible with private value to the individual. If we put the question only in private terms, we shouldn't be surprised to get an answer that understates the actual willingness to spend resources dedicated to social spending.

But consider a country that wishes to impose restraints on overseas investment. Whether that is a good idea or not is beside the point here. In fact many governments have done it. So it can happen. And suppose that to allocate the quota of exportable capital, the country adopts a rule that other things equal, makes overseas investment by a firm easier to the extent that the investment would yield a higher rate of return than investment at home. Call this Rule A. Rule A seems reasonable, but the essential point

here is just that obviously a government might adopt such a rule.

Yet for political or equity reasons, the country also wishes to avoid allocating all the overseas investment it judges tolerable to one or a few firms that happen to face the most favorable opportunities at the moment. So a second principle is adopted (call it Rule B) that says that other things equal, the more a firm has already been allowed to invest overseas, the more difficult it ought to be for it to export even more capital.

And if you compare rules A and B here with NSNX efficiency and NSNX equity Rules 1 and 2 in Chapter 1, you will see that the parallel is exact. The G' and S' of the dual-utilities setup would here be reinterpreted as rate of return on overseas and domestic investment, respectively. The weighting function can take a mathematically identical form [$W = f(g/s)$, $W' > 0$], but now g and s are, respectively, a firm's overseas and domestic investment, rather than individual's social and private spending. The equilibrium condition is the same for both cases [$W = G'/S'$], and with trivial adjustments the whole discussion of the figures in Chapter 1 then carries over into this mainstream economics example.

For this case, although the mathematical setup comes out identical to that of the dual-utilities model, we are unlikely to think of the firm as having dual preferences. Each firm would ordinarily prefer to be free to invest wherever it wants. We would interpret each firm's choices as maximizing subject to a government-imposed constraint. So here the same formal structure proposed to account for individual choice has a perfectly conventional interpretation. No economist would see anything bizarre in the situation, or suppose that the way to deal with the problem would be to reformulate the framework in terms of some overall, ultimate firm maximand, within which the analogues of G' and S' (here, they would be rate of return on overseas versus domestic investment) were subsumed. Although an individual firm would see the situation as maximizing subject to a constraint, in its observed behavior the firm would choose as if it held dual preferences.

And now carry this one step further. Certainly situations could arise in which the dual structure is the actual preference structure of the firm. Suppose the firm wants to ward off pressure for government controls or unfavorable publicity. Then it could plausibly, that is, without condemning its management as unreasonable or irrational, indeed perhaps demonstrating the sophistication and prudence of management, choose to

adopt an investment policy governed by rules A and B. Now the dual-utilities balanced by the exact NSNX equilibrium condition captures the actual preference of the firm, though still there is no overall firm utility function, nor is any economic theorist likely to propose one. Analogously, a really zealous proponent of both evolutionary arguments and canonical notions of rational choice might even interpret the dual-utilities NSNX setup as a pragmatic outcome like this economic example, arising *as if* from a calculated effort to produce a species capable of large-scale, open-ended cooperation, given the constraints on what is feasible for that metaphorical rational actor, Ms. Evolution.

Or consider another, entirely different, example with the same result, but now it will unambiguously involve the unconstrained preferences of a single individual. Suppose Ellie has two children, one of whom could gain more from her financial help than the other. This child (G) might be handicapped, so he needs more help; he might be exceptionally talented, so he can make better use of whatever help he gets; or as a struggling artist he might have much less income than his otherwise equally deserving businessman brother (S). Thinking of G's greater gain from her help, Ellie might wish to use almost all her spare resources to help G. However, Ellie also feels she ought to treat both children equally. She might resolve her difficulty by adopting versions of NSNX rules 1 and 2 (or, equivalently, rules A and B of the investment example): (1') Other things equal, the larger the value to G of a marginal dollar compared to the value of that marginal dollar to S, the more likely the dollar should go to G. But (2') Other things equal, the more Ellie has already given more to G than to S, the less likely it is that the next dollar will be spent to further help G.

We would once again get the mathematical structure of the dual-utilities model. Now G' and S' are the marginal values of dollars to help son G and son S, respectively; g and s are the amounts she has spent on G and S; and the weighting function, W , tells Ellie how big the marginal value ratio (G'/S') has to be to warrant allocating still more to G over S. Once again, all, figures used in Chapter 1 to illustrate the NSNX equilibrium are readily reinterpreted to cover Ellie and her children.

And as with the firm's choices for the voluntary constraint version of the capital export example, there is no ambiguity about whether the structure represents the authentic preference structure of a single individual. Plainly it does. Some such structure

no doubt captures the essentials of choice for many parents, and occasionally for some firms as well.

So however odd the dual-utility structure might seem as a model of individual choice these examples show that formally equivalent analogues either easily could exist or in fact do exist in actual societies. They are not examples that look especially bizarre, or that look in any plausible sense irrational. When I label NSNX an extension of standard rational choice theory that is more than just arbitrarily claiming access to the honorific label "rational". Unless you want to claim that Ellie and the business firm in my examples are behaving irrationally, then NSNX *is* an extension of rational choice theory to cover a crucial aspect of choice in important empirical contexts.

And to conclude with a particular simple application before proceeding to a more abstract and technical discussion of general cases, consider the long debate over why (or whether!) it is rational to vote in an election with many millions of voters. The exclamation is warranted. For not only would a large proportion of the population have to be considered irrational (and if irrational in this, why suppose they are rational on other issues routinely analyzed with rational choice apparatus), but since a propensity to vote increases with increasing education, it would be the most sophisticated segment of the population which is most likely to choose irrationally. Might there not be some flaw in this assessment?

But the expected value to self-interest of one vote could not amount to as much as a penny when that vote is one among many millions. Noticing the low payoff to self-interest yields the notorious "paradox of rational voting." But would allowing for social (not just self-interested) motivation resolve the paradox? If Ellie were voting out of social concern, she might reasonably see a large social gain (easily \$1 billion or more) from the election of a particular presidential candidate. For \$1 billion is a small number compared to the trillion-dollar federal budget a president shapes, smaller still compared to the GNP his policies influence, and trivial compared to the costs should a president blunder into a nuclear war.

Hence if Ellie were to attach a numerical value to the advantage to the country of electing what she judges to be the better president, that number could easily be in the billions. Ellie's one vote would provide only a tiny increment (one part in tens of

millions) to the probability that Ellie's candidate actually wins. But since the social value is very large, the expected value of her vote can still be perfectly respectable (some tens of dollars) compared to the cost of voting.

So there seems to be no great puzzle about how to account for rational voting if Ellie were moved by social, not only by self-interested, motivation. Yet the puzzle returns when we carry the analysis one step farther.

Suppose it were possible to show that Ellie doesn't attach a value of anything like \$1 billion to the outcome? Then she could not be acting rationally (consistently) if she acted as if she did. There may be realistic levels of social valuation large enough to make a vote rational. But such values apparently aren't Ellie's values in the sense relevant for the standard sort of rational actor analysis. What counts for that purpose is the personal sacrifice Ellie would make to promote those values. If Ellie is guided by a single preference structure which incorporate her social as well as her private preferences (if she is guided by the single utility function of standard theory), what the presidential election is worth to Ellie is just what she would give up to secure the result she takes to be socially valuable.

So from that standard point of view, consider how much of a private reward Ellie would require to make her prefer an election outcome in which her presidential candidate loses. Before an election, we might ask Ellie how much money in her own pocket it would take to make her hope for the defeat of the candidate she favors. The dollar amounts you will find if you ask people about this will rarely exceed \$10,000. And that is far too small to make Ellie's vote rational when discounted by any reasonable estimate of the effect of her vote on the election (Margolis 1977).

So, on this argument, if Ellie is rational, and if she reports she would begin to hope her candidate would lose if some rich rascal were to give her \$10,000 in that event, then in terms of her preferences the election is worth no more than \$10,000. The point is not that Ellie is insincere if she says that the social value is vastly larger than \$10,000. But in terms of what that large social value is worth to Ellie -- in terms of what Ellie would personally give up to obtain that large gain for society -- the smaller number (\$10,000, not \$1 billion) is alone relevant. When that smaller number is discounted for the effect of one vote on the probable outcome, what is left of the value of Ellie's vote

could hardly be enough to explain her bothering to vote. It would be a small fraction of a penny. But we are not surprised to find Ellie voting. And if we ask her whether she thinks the chance that she will be killed in an auto accident driving to the polls is as large as the chance that her single vote changes the election outcome, she will almost certainly allow that must be true. I give a detailed discussion in Chapter 7 of SA&R. But we are still not surprised to find Ellie voting. Similarly, my colleague Gary Becker assures me that he is both rational and votes. And, alas, I feel the same myself. So are we all (probably you too) deceiving ourselves?

But this argument I've just given does not work if the way human beings actually behave fits the NSNX dual-utilities model. Rather, I want to claim that this Ellie argument is just another of the empirically somehow-crazy inferences that can be deduced from the standard model when that model is used in the context of social choice.

In the standard model, a person maximizes utility by choosing an allocation such that the utility of the last dollar allocated to social utility just equals that of the last dollar allocated to private utility. Hence asking Ellie how much money in her own pocket it would take to get her to prefer the election outcome she thinks socially perverse would give us logically certain knowledge about the actual value *to Ellie* of the social gain she expects from electing one candidate over another. And that should guide her choice if Ellie is rational and the standard model correctly captures human preferences.

But in the dual-utilities model that logic does not hold. Equilibrium is determined by a condition that, however much it may look like a usual first-order equilibrium condition for maximizing utility, in fact works on an entirely different principle. Its roots lie in competing Darwinian pressures favoring group versus self-interested behavior. And resources won to the social side of the individual balance are not spent to satisfy Ellie's private preferences any more than the money she spends on flowers in the very first example of this chapter. Ellie privately might allow she couldn't resist (or at least suspects she couldn't resist) hoping for what she takes to be the socially perverse outcome if it would put \$10,000 in her pocket. But she will not vote that way, since the voting comes from resources to satisfy her G-preferences, where \$10,000 to Ellie does not begin to affect her judgment of which candidate would be better. The G-preference are about what would (in Ellie's judgment) be good for society, not about what would be good for

Ellie. Ellie counts in that G-function, for she is part of society, but in the G-function she counts for no more than other people like herself.

So in the dual-utilities context, the marginal allocation is not one that balances marginal returns to some hypothesized grand preference function across self-interested and group-interested spending opportunities. Instead it is the allocation that brings the individual as close as feasible to a "fair-share" balance, feeling neither selfish nor exploited, between spending to satisfy private versus spending to satisfy social preferences. And resources allocated to G-spending of course are spent to maximize G-preferences.

So even if Ellie would truly give up no more than \$10,000 to guarantee her presidential choice, she will also rationally treat the social value of that choice to be very much more than she personally is willing to pay. There are simply two distinct questions: (1) about Ellie's social judgment, and (2) about how large a share of her personal resources she would spend to support that judgment in a context where she has the power, if she is willing to make a large enough sacrifice, personally to determine the social outcome.

Suppose (as is reasonable) that Ellie in fact would ordinarily devote only a small share of her resources to social G-preferences. Even so, the cost of voting will be very small as a fraction of her total G-spending. How much of her personal resources she would sacrifice is irrelevant to the voting context. No choice remotely on that scale is at stake. In terms of the seesaw diagrams of Chapter 1, any shift in S' or G' or W contingent on whether or how Ellie votes would be very small.

The two questions-about Ellie's judgment of the social value, and about how much Ellie would personally sacrifice to gain that value must be connected. But they are not the same question. To take the answer to the question about personal sacrifice as the answer to the question about Ellie's true assessment of the value of the election is not a necessary condition for rational choice. Rather, it's a mistake.

The NSNX rules require that as the price Ellie must pay for voting her social preference gets larger (hence the share of her resources required gets larger), her reluctance to sacrifice that much must grow, to the point at which she eventually takes the money offered for changing her vote. Somewhere between the amount she would pay out

of pure self-interest (the S-value of the election) and the vastly larger social value (the G-value), she will feel that she personally is being asked to do too much.

A side point not really essential for the discussion here, but foreshadowing the importance of cognitive issues later on, is to notice that it would be embarrassing for the dual-utilities model if in fact a voter who feels strongly about a presidential election would actually prefer \$10,000 in her own pocket to what she takes to be the clearly better social outcome. This could be tested only in a situation where Ellie believes the actual outcome somehow will really be determined by whatever she prefers. That would be hard to arrange! But some reasonable analogue might be constructed in an experiment of the sort we will be examining in later chapters. If we found support for Ellie really being unwilling to accept a large but far from devastating cost in that case, that would be an embarrassment for NSNX.

However, from the actual evidence we have about what happens when people face choices between large private sacrifices and social gains they judge very important (behavior in time of war, for example), it is apparent that sacrifice of self-interest in such contexts frequently -- even routinely, though not universally, is very much greater than we notice in everyday situations, and often far greater than the individuals themselves could anticipate.

In general, we make judgments by a process of anchor-and-adjust, starting from patterns of response that are already in our cognitive repertoire. But in the Ellie thought experiment, there is nothing familiar to anchor on which is anywhere near the hypothetical situation in which Ellie's vote decides a presidential election. So she can be expected to anchor unconsciously on something far removed (on a situation in which her choice does not decide for all of society), and the adjustment is then likely to fall far short. If we could actually put Ellie in the situation, she might be surprised at her willingness to spend for a social cause when her choice alone determines what happens.

More is at stake in this rationality of voting discussion than accounting for why a citizen might spend the modest resources (the bit of time) required to vote. How a person votes can't be plausibly assumed to be independent of why she is motivated to vote in the first place. So although the cost of voting is ordinarily very small for any individual, and far from enormous even for the whole electorate, the social consequences of how people

vote are enormous. And the same pattern of argument can be extended to social choice generally. Your G-preferences might include a preference for helping me, but your sense of what would be good for me might be very different from my sense of that. I might be mightily pleased with my recent behavior, and you might think what I have earned is a good kick in the ass. But my sense of what is good for all of us is going to be compatible with my sense of what would be good for me, and the same for you.

Return now to the main point of this chapter. So long as you suspect that NSNX *can't* be right, you will not be very open to the possibility that it might be right. Suppose it just cannot make sense that agents do not have *a* set of preferences but two irreducible sets, without any need for some overarching total preference which mediates whatever compromise might emerge: Then NSNX can't be right. The S-function is the usual self-interested utility function of standard theory, and the G-function is also familiar under the label of a Social Welfare Function, as in Harsanyi 1955. But Harsanyi's SWF (like Adam Smith's much earlier "higher self") invokes preferences for, so to speak, special occasions, when individuals rise above their everyday propensities. But the NSNX G-function is always in play. Choices may not show much effect from it in routine situations. But it is always there, while the Harsanyi or Adam Smith notion might be only what a person responds to when little or nothing is at stake for them personally. But since we can sometimes observe that choices that do not appear to be purely self-interested occur even at non-trivial private cost, surely we ought to be interested in how we might organize our understanding of what conditions evoke this higher self, and how it might interact with the self-interested economic man that we easily notice as we go through a day.

Nevertheless what might seem the right way to proceed with the NSNX equilibrium ($W = G'/S'$) is to set $U' = WS' - G'$, expecting to integrate that expression to find a NSNX total utility function. But there is no such function. Total utility so defined would be path-dependent, with no stable value, hence not a viable mathematical function. And on the Darwinian logic that underpins the model, there is no reason to suppose there any such function would exist. Logically, given the Darwinian foundation, this is not a problem at all. For all the actual work in a standard choice model is done by the equilibrium condition obtained by differentiating the utility function and setting it equal

to zero. So to pick an especially simple example where there is no ambiguity, if $U = 8x - x^2$, where x is the agent's location in some space, then (taking the derivative) agent will want to be where $0 = 8 - 2x$, hence at $x = 4$. When $x = 4$, utility = 16, and agent is in equilibrium. She cannot do better. Making x either larger or smaller will lower utility. That is what we want to know. Given her utility function and her menu of feasible choices, what would she do?

In contrast, for the dual utilities model the equilibrium condition appeals directly to the pair of Darwinian rules developed in Chapter 1 (or also to the exact analogues of those rules illustrated in this chapter with the capital export and Ellie's children examples.). Even if a sufficiently complicated mathematical function could be devised with the property that, when differentiated, it approximates $W = G'/S'$ as its first-order condition, it would only be a mathematical freak, having no useful role to play in the logic of the model. We might want to know the cardinal value of utility to the agent, and the cardinal value of utility to the group. And where that is manageable under standard theory, it would also pose no special problem for NSNX. For as already mentioned, the S-function and G-function are both familiar functions within economic theory, consistent with the classical axioms of rational choice. But no total utility function is required for that.

Someone well accustomed to working with the standard model will be tempted to suppose unless the theory is just wrong, there really ought to be a single maximand hidden somewhere but I have missed it. The dual-utilities argument will then look like a proof that $2 + 2 = 5$: you know there is something wrong with it, even if you haven't put your finger on just where the error lies. But if you think about the capital export example and the Ellie example you have a way to test that intuition (however clear it may seem) against actual cases which have a structure identical to NSNX. It is unlikely, I think, that you will end up feeling it would be sensible to demand some total utility function for either case.

I conclude on a polemical note. We now have half a century of post-Downs (1957) efforts to account for voting within the standard model of rational choice. The record is coming to look like one of the psychologists' *umwelt* experiments. Put a starving chicken, then a starving dog, on one side of a wide glass wall, with food visible on the

other side. Each animal starts out by trying to bang through the wall. Eventually the dog seems to catch on and, after some searching, finds its way around the wall. The unfortunate chicken starves to death, banging its head against the wall. To me it seems about time to consider alternatives to the standard model. But I observe that the chicken is still hard at work.