Empirical Inputs for Normative Policy Decisions

Julian C. Jamison

World Bank *and* University of Exeter

First, thanks in advance to everyone for looking at this and for the discussion and feedback, which I am very much looking forward to. Although I have been thinking about these issues off and on for quite a few years, as you will see there is much that remains to be put in a concrete form. This topic is also probably unusual for EGAP, but I believe it is quite relevant for international governance and is susceptible to creative methodological approaches (including RCTs, but also preference elicitation surveys, theory, and administrative data) of the types many members know well. The good news is that even if you don't find it relevant, this will be a short pre-read!

I am hoping to jumpstart a research agenda dealing with aspects of "What is social welfare?" or "What should a well-meaning policy-maker try to maximize?" or "What should I put on the left-hand side of my regressions?" My sense is that this necessarily includes elements of philosophy and evolutionary psychology in addition to economics and political science. One answer would be that in a democracy, policy-makers should maximize whatever the population wants them to; this is unsatisfactory for a number of reasons, not least of which is that we have excellent evidence suggesting that people don't know (and can't easily predict) what they want.

These are normative questions at the end of the day, and there will be no single correct answer. Partly for that reason, I think social scientists have often shied away from tackling them head-on. The explicit mindset (at least within economics, which I know best) is generally: someone else decides the goal, and we can help them obtain it. At a practical level, of course, this often breaks down, and e.g. "free markets" becomes an implicitly recommended goal in and of itself. For that reason alone it might be better to be transparent about the underlying assumptions regarding where policy should or should not be headed, so that there is a healthy dialogue between technocratic advisors, the population at large, and policy-makers who have the final say in any given instance.

But even if we accept the 'pure' mindset of not entering the normative side of the policy debate at all, I believe social scientists have more tools at our disposal for informing that debate than we typically deploy. For instance, imagine that society needs to decide on a discount function for valuing potential climate change interventions. At the end of the day, that is a normative question without a right answer. But we can go further than simply taking the discount function as given (from somewhere) and seeing what it implies in terms of specific policies, without entering the debate as to what it should be – although we can fruitfully do that too (without claiming to have a monopoly on the answer). We can bring data to the question, regarding what individuals and groups do; survey responses; and so on. We can bring theory to the question: this type of discount function is incompatible with these principles, etc. Indeed in this setting all of this is already happening to some extent, which shows that the same inputs may well be equally or more useful even in more obviously normative contexts. If nothing else, in many cases it would constitute a useful step forward simply to shine some light on what are not uncommonly *ad hoc* (and/or implicit) policy decisions, e.g. choices that imply a particular discount function without even realizing it.

Naturally this is a broad (and probably ambitious) agenda. I'm happy to discuss anything related to it, but in what follows I'll spell out some initial thoughts and ideas in four somewhat more specific areas. Roughly speaking, they progress from least to most expansive in terms of scope and departure from tradition. Overall here are a few questions to keep in mind, where I would welcome feedback:

- Which of these general areas or specific sub-areas (or others) is most feasible to make convincing progress on?
- What methodological techniques or types of data would be most relevant / convincing in each case?
- Which areas are most likely to have policy impact, either conceptually (i.e. should have impact) and/or practically (i.e. will in fact have impact)?
- What are the most pertinent existing literatures for these types of topics?

1. Identifying mistakes

In neo-classical economics, individuals don't make (ex ante) mistakes. Revealed preference is a maintained and untestable assumption. Of course even the most die-hard economist would physically stop their toddler from walking into traffic – not (solely) for the parent's welfare, or for social welfare, but because they believe the child too will be overall better off this way. Similarly someone might stop a drunk friend from driving home, or a family member with dementia or psychosis from self-harm. One might even stop a friend from impulsive behavior not due to the possibility of physical harm but highly embarrassing or financially costly actions.

In each scenario we must have some model in mind for why the potential action contravenes the person's own interests; it's not random. Where does that model come from? Can we test it? What constellation of data would be most suggestive of a mistake?

> ➢ Asking in advance (or even costly self-commitment) not to be allowed to do it
> ➢ Consistently claiming regret afterward
> ➢ Neurobiological evidence of a depressed reward system
> ➢ 1000 similar individuals all choosing a different action under the same circumstances

Once we have a model or theory that appears to be fairly consistent with the intuitively clear cases, we can potentially extend it to less clear-cut cases (e.g. where someone is not obviously directly 'impaired' in their decision-making). Policy-makers should want to distinguish mistakes not only to think about interventions that will decrease their prevalence but also to determine what to legitimately 'count' in the social welfare function.

2. Intrapersonal tradeoffs

Humans care about relative social status (not just being better or richer than their neighbors, but also being thought of as a 'good' neighbor), and they also care about absolute status (e.g. having a large enough car to fit the whole family's camping gear). Which one matters more? Or rather, what weights should be placed on each? Humans care about money, and they care about health, but what weight should be placed on each? Is it better to have 20% higher income or a 10% longer lifespan?

Of course the answer may depend not only on personal preferences but also on personal circumstances (what is your current expected income and lifespan). Is it different for men and women, or across cultures? Here the idea is not that anyone is necessarily making a mistake, but that it is difficult to aggregate across seemingly incommensurate dimensions. Once again revealed preference would give us the answer if we believed it, but once again there is theory and evidence to suggest that it yields an incomplete picture at best. On the other hand, if we simply ask people theoretical survey questions about such tradeoffs, is that likely to give any better answer? Does it depend on how we ask? Can we compare indicators (e.g. subjective or objective well-being) for those experiencing different combinations across dimensions?

This has clear implications for policy. It is (relatively speaking) easy to figure out what a ministry of health should maximize, and what a ministry of education should maximize, and what a ministry of jobs should maximize, but what percentage of the central resources should each of them receive? What inputs should factor into that decision?

Introducing issues of 'internalities' and 'biases' (such as time-inconsistency) make the problem yet more complex. If different selves have different preferences, how do we do welfare analysis? Note that it may well be the case that regulations – e.g. forced retirement savings – actually increase overall freedom (libertarian paternalism without the paternalism…) if they sufficiently expand the choice set in the future. Not all of this is quantifiable, but some of it is.

3. Interpersonal tradeoffs

This topic is similar to the one above (especially from an aggregate social perspective, without reference to discrete identifiable individuals). For instance, global health typically uses a measure of burden of disease like DALYs (disability-adjusted life-years), which require disability weights for every affliction. It's not conceptually crucial whether we think of this as a comparison between a given individual either (say) going blind or becoming paraplegic; or whether we think of it as a comparison between the welfare of a blind individual and a paraplegic individual. But in both cases we need numbers to make policy, and the specifics can make a huge difference.

Historically the numbers have often come from expert opinion (panels of doctors), but there is so much richer data potentially available. Furthermore, theory can help inform the reconciliation of different responses from those who are already blind versus those who are not; and the distinction between chronically salient conditions (e.g. skin rashes, recurrent pain, depression) versus those where there is – at least seemingly – adaptation (e.g. blindness and indeed most other ailments). Can we improve on stated preferences here?

Of course all of this applies to deaths in addition to disabilities. What value should be placed on a death at age 1 (or age "12 hours before birth") versus age 20 versus age 80? What about across geographies? The value-of-statistical-life literature would suggest a much lower value in developing countries (where incomes are much lower), while a prioritarian philosophical approach might yield a higher value in developing countries (since those citizens have been treated inequitably by the random nature of their birth-place and are therefore owed more by society). Would veil-of-ignorance style experiments contribute to this debate? Or measures of the marginal value of consumption?

4. Unconditionally normative economics

Traditionally descriptive economics covers actual behavior (e.g. how will people react to new information?) and normative economics covers optimal behavior (e.g. how should people react, if they're trying to maximize their expected returns?). The latter almost always has a conditional element: **if** this is your utility function, what should you do? A truly normative approach would instead ask what your utility function should be.

As above, these are often questions that social scientists prefer to leave to philosophers, but I believe we have something to contribute (even without claiming to be able to determine the answer). A mother might think that her child would be 'better off' (with respect to something…) being patient, time-consistent, ambiguity-tolerant, pro-social, and conscientious. She would try to inculcate those qualities, and society might likewise try to 'build character' – both for the sake of the individuals involved but also in the latter case perhaps due to externalities.

Individuals also try to change their own preferences (e.g. learn to appreciate classical music, or fine wine), but what does it mean to prefer different preferences? How should one do welfare analysis if preferences change, and in particular if they change due to an individual's own choices and actions? Is it ever appropriate for policy-makers to try to influence preferences? What if voters request it? Is this different from public education? If not entirely, then we should be transparent that this is one of the objectives, and we should use our familiar tools to analyze it as we would anything else.