Is *homo economicus* self-interested? That question might seem as silly as asking whether the Pope is Catholic. Most people, including most economists, would be inclined to say, “Of course.” Indeed the first sentence in the Wikipedia entry for “*homo economicus*” states that it “is the concept in many economic theories of humans as rational and narrowly self-interested actors.” Economists typically assume that people’s behavior responds to financial incentives -- which certainly looks like material self-interest. Entrepreneurs aim to maximize their profits. Consumers strive to acquire the largest bundle of commodities they can afford. 

Nor is this view new. As Adam Smith wrote in Book I, Chapter 5 of *The Wealth of Nations*,

> But man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only. He will be more likely to prevail if he can interest their self-love in his favour. . . We address ourselves, not to their humanity, but to their self-love, and never talk to them of our own necessities, but of their advantages.

Though Smith here allows that people are benevolent, he doesn’t think that much will come of it. What motivates people is their “self-love” or concern for their own advantage. So it looks as if interpreting *homo economicus* as fixated on his or her own narrow material self-interest is amply justified.

But the entrepreneur who crushes a union may give away half her income to charity or to a political candidate devoted to an ideology whose implementation might ultimately be harmful to her earnings. The consumer who shops at Walmart rather than the corner grocery to save some money may use that money to pay for his children’s ballet classes. The same individual who takes one job rather than another because the salary is larger, may turn the whole pay check over to family members. Conversely, someone may choose a life of financial self-sacrifice
in pursuit of the self-interested goal of securing the admiration of others. The connection between so-called “economic rationality” and self-interest is not so obvious as it may appear. Addressing the question of whether economic rationality implies self interest requires clarifying both what mainstream economists claim about human nature and what constitutes self-interest. Self-interest apparently enters economics in four main ways:

1. Via the theory of rationality, which maintains that each individual is motivated only by his or her own preferences.
2. Via the theory of consumer choice, which maintains that consumers care only about the size of their own bundle of commodities.
3. Via the theory of the firm, which maintains that those engaged in production care only about their own financial returns and leisure.
4. Via welfare economics, which identifies preference satisfaction with individual welfare. If welfare is preference satisfaction and, as asserted by the theory of rationality, people are motivated only by their own preferences, then individuals are motivated only by their own welfare.

Section 1 begins with the theory of rational choice, since it shapes all the other parts of economics I shall discuss. Section 2 turns to the task of clarifying what counts as self-interest. Section 3 examines whether rationality implies self-interest. Section 4 discusses theories of consumer choice and of the firm and their implications concerning self-interest. Section 5, the longest section in the paper, considers whether welfare economics assumes that individuals are self-interested. Section 6 offers conclusions concerning whether (and for what purposes) mainstream economics does or should treat individuals as self-interested.

1 Rationality
Mainstream economics is built around a theory of rational choice, which includes a theory of rational preference and a theory of rational belief. In many applications, it is reasonable to
idealize and to suppose that the agents have perfect knowledge of all relevant facts. In that
case nothing more need be said about belief. Several axioms specify what it means for
preferences to be rational. I shall consider only the simplest and thinnest account. For
convenience, these axioms are usually formulated with respect to a relation R of weak
preference, such as $xR_ay$ – where an individual $A$ weakly prefers alternative $x$ to $y$ if and only if
$A$ prefers $x$ to $y$ or is indifferent between $x$ and $y$. Since preference and weak preference are
interdefinable, it makes no substantive difference whether the axioms are formulated in terms of
preference or weak preference.

The standard axioms governing preferences are the following three:

$$\forall x \; xRx$$  reflexivity

$$\forall x,y \; xRy \text{ or } yRx$$  completeness

$$\forall x,y,z \; \text{if } xRy \text{ and } yRz, \text{ then } xRz$$  transitivity

These axioms impose a formal structure on preferences without specifying what preferences
are. In contexts of risk and uncertainty, most mainstream economists embrace a richer account
of rationality, which I shall not discuss in this essay. The relata of the preference relation are
sometimes limited to immediate objects of choice (as in revealed preference theory), but in
predicting choices, economists in fact consider how choices bear on states of affairs that
individuals value that are not themselves immediate objects of choice. For example, although
those who buy stocks cannot choose whether they increase or decrease in value, they prefer
that the stocks they own increase in value, and it is because of this preference that the belief
that a stock will increase in value leads the agent to purchase it.

Although textbooks sometimes list only the three axioms stated above, in fact
economists rely on additional assumptions concerning preferences and their relations to choices.
First, they assume that preferences are reasonably stable. Although the preferences of rational
individuals may change, they do not flip-flop from moment to moment. Otherwise inferences
about preferences drawn from past choices would have no predictive value concerning future choices. Second, preferences between two alternatives do not depend on the existence of other alternatives. Consider the following apocryphal anecdote (from *Wikipedia*) concerning Sidney Morgenbesser:

After finishing dinner, Sidney Morgenbesser decides to order dessert. The waitress tells him he has two choices: apple pie and blueberry pie. Sidney orders the apple pie. After a few minutes the waitress returns and says that they also have cherry pie at which point Morgenbesser says "In that case I'll have the blueberry pie."

Morgenbesser's preferences are not incomplete. Nor are they intransitive, because transitivity says nothing about how preferences over one set of alternatives \{apple pie, blueberry pie\} should be related to preferences over a different set of alternatives \{apple pie, blueberry pie, cherry pie\}. If rational preferences are to imply consistent choices, then they must satisfy an additional condition:

(Context independence) The preferences of a rational agent among alternatives $x$ and $y$ does not depend on what other alternatives are available.

But to require context independence as a condition of rationality may seem absurd. Rational individuals with conventional tastes concerning wines prefer red to white when the main course is beef and white to red when the main course is flounder. To avoid condemning such preferences as irrational, economists idealize further and suppose that the alternatives among which individuals are choosing are complete states of the world. Preferences among these should be context independent.

One more axiom is needed. Like stability and context independence, economists rarely formulate it explicitly, but it is crucial. This axiom links preference to choice, and one way to express it is as follows:

(Choice determination) Among the alternatives that they correctly believe to be feasible, agents choose one at the top of their preference ranking.
Completeness and transitivity imply that individuals rank alternative actions. Some actions are not feasible, and it is impossible for agents to choose them. As soon as one relaxes the assumption of perfect knowledge, one must, however, countenance the possibility that individuals will fail to choose their most preferred alternative, because they mistakenly believe that it is not feasible. Among the remaining actions – those that are both feasible and are believed to be feasible – rational agents will choose the action that is at the top of their ranking or, if there is a tie for first place, one of the actions that is at the top of their ranking.

These axioms concerning preferences and the relations between preference and choice (given constraints and beliefs) constrain the interpretation of “preference.” In particular, they imply:

1. Preferences are not just matters of taste. They are cognitively demanding.

2. Preferences are comparative. To say that an agent prefers some alternative is elliptical. To prefer some alternatives is to prefer it to other alternatives.

3. Preferences are “total” comparisons. Since (given constraints and beliefs) preferences determine choices, no factor influences choices except via influencing preferences. In forming their preference rankings, individuals take into account everything that they take to be relevant to choice.

4. Preferences are subjective states that combine with beliefs to determine choices.

The interpretation of preferences that best fits both the axioms and the practice of economics takes preferences to be total subjective comparative evaluations. Although economists have defended other interpretations of preferences when reflecting on what they do, in their practice they take preferences to be total subjective comparative evaluations.¹

2 Self-Interest

When is an action self-interested? This might be a question about how the expected benefits to

¹ The account of preferences sketched here follows Hausman (2012).
the agent compare to the expected benefits of alternatives or about how the realized benefits compare to what the benefits of alternatives would have been. And one should distinguish the answers the agent gives to these questions from the answers that third parties might give. To maintain that an agent A who prefers x to y is self-interested or that an agent A who chooses x rather than y is acting in his or her own interest can thus mean four different things:

<table>
<thead>
<tr>
<th></th>
<th>ex ante</th>
<th>ex post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjective</td>
<td>1. <em>Subjective ex ante self-interest.</em> The expected benefit to A of x as appraised by A is larger than that of y.</td>
<td>2. <em>Subjective ex post self-interest.</em> The actual benefit to A of x as appraised by A is larger than what A believes the benefit of y would have been.</td>
</tr>
<tr>
<td>Objective</td>
<td>3. <em>Objective ex ante self-interest.</em> The objective expected benefit to A of x is larger than that of y.</td>
<td>4. <em>Objective ex post self-interest.</em> The actual benefit to A of x is larger than the objective expectation of what the benefit to A of y would have been.</td>
</tr>
</tbody>
</table>

Let us call the senses of self-interest in the top row, “subjective” self-interest, and the senses in the bottom row “objective” self-interest. When individuals have perfect knowledge, these four senses coincide. When the subjective probabilities of individuals coincide with objective probabilities, subjective and objective self-interest coincide, but there will still be a distinction between those actions that are expected to promote an individual's interests and those which in fact do so. Since individuals do not have perfect knowledge and their expectations differ, actions that are self-interested in one of these four senses need not be self-interested in any of the other senses.

One might reasonably complain that none of these four senses fully captures what it means to say that someone is self-interested, because none of them mentions *motivation*. The fact that A’s preference ranking of the alternatives happens to coincide with a ranking that would be described as self-interested in any of these four ways does not make someone self-interested, because A might not have been *aiming* to benefit himself or herself. A may expect
that x will make a greater contribution to A’s well-being than y, but prefer x to y for some other
reason. A’s preference for x over y might have been just the same if the actual or expected
benefits of y to A had been larger than the benefits of x.

On a view of preferences like the one to which mainstream economists subscribe, in
which preferences encompass all the considerations relevant to choice, aiming to benefit
oneself is identical to preferring whatever one judges to have a larger expected benefit for
oneself. On the view of deliberation and choice sketched in section 1, individuals are self-
interested only if they are self-interested in at least one of the four senses defined above. But
this fact points to an inadequacy in that view of deliberation and choice, not to the spuriousness
of the concern about motivation. There is a real difference between agents whose preferences
just happen to match their rankings of alternatives in terms of expected benefit and agents who
hold the preferences that they do because that preference ranking coincides with their rankings
of alternatives in terms of expected benefit to themselves. To distinguish these two possibilities
requires a richer structure of the relations between values and choices than is standard – one in
which, in addition to preferences among alternatives, individuals place values on attributes or
characteristics of alternatives that influence their preferences. The self-interested individual has
preferences of one of the four kinds because of the value they place on a particular attribute of
alternatives: namely its contribution to the agent’s own welfare.

I shall, however, set aside this complication, which does not affect the conclusions I shall
draw.

3 Rationality and self-interest
Since the choice determination axiom maintains that (given beliefs and constraints) A’s choices
are determined exclusively by A’s preferences, many have supposed that the theory of
rationality implies self-interest. But it is a mistake to jump so quickly to this conclusion. Self-
interest in any of the senses above concerns the content (and causes) of an individual’s
preferences, not whether the preferences of agents guide their choices. Agents who prefer to sacrifice their interests to benefit or to harm others are altruistic or malevolent rather than self-interested, even though their actions are no less determined by their preferences than the actions of those who are self-interested. Since preferences include every consideration that agents take to be relevant to their decisions, maintaining that preferences determine choices does not imply that individuals are self-interested.

It thus seems obvious that rationality does not imply self-interest. The fact that A ranks $x$ above $y$ by itself says nothing about A’s (or any body else’s) expectations concerning the effects of $x$ and $y$ on A’s life or about the actual effects of $x$ and $y$ on A. But if preference satisfaction constitutes welfare (as welfare economists apparently maintain), then the fact that A prefers $x$ to $y$ makes it the case that $x$ is better for A than $y$! Given a preference-satisfaction theory of well-being, rationality implies subjective self-interest, and there is no such thing as objective self-interest that is distinct from subjective self-interest. If, ex post, I prefer $x$ to $y$, and this subjective fact makes it the case that $x$ better for me than $y$, then it is impossible that $y$ is in fact better for me than $x$ is.

But individuals may (of course) falsely believe that $x$ is better for them than $y$, and individuals are perfectly capable of distinguishing between what they believe to be better, all things considered, and thus more choiceworthy, and what they believe to be better for themselves. In both cases, A’s preferring $x$ to $y$ does not imply that $x$ is better for A than $y$. So the purported connection between rationality and self-interest via the putative rationality of acting on complete and transitive preferences collapses. More will be said about the connections between preference satisfaction and welfare in section 5.

4 Consumer choice and production
Since the theory of rationality purports to be a purely formal theory that does not specify the content of preferences, it is unsurprising that it says nothing about whether individuals are self-
interested. The theory of consumer choice and theories of the firm, in contrast, make assertions concerning what people prefer. So perhaps it is here that self-interest enters economics.

The theory of consumer choice takes the objects of preference to be bundles of commodities and services. If one takes this literally, it implies that people care about nothing except commodities and services, but clearly the theory is not intended as a general account of human nature. Rather it supposes that this characterization of preferences is a reasonable first approximation when one is attempting to predict or explain choices at the supermarket or mall. Does this approximation commit economists to the view that consumers are in those contexts (at least as a good first approximation) self-interested?

It appears that the answer is yes, because the theory assumes that each individual cares only about his or her own bundle of commodities, unlike those who are altruistic or malevolent and who consequently care about the consumption of others. But the question is more complicated than it might appear, because it is unclear what counts as a “commodity” or “service” that belongs to agent A’s consumption bundle. If A contributes to a charity vaccinating children in Somalia, does that mean that the doses of vaccine and the actions of the aid organization are commodities and services in A’s consumption bundle, alongside of the new shoes that A also purchases? One could specify that something counts as a commodity or service only if it promotes A’s welfare or interest. But that would trivialize the question of whether the theory of consumer choice implies that individuals are self-interested, and for the purposes of explaining and predicting A’s behavior, there is no reason to do this. The theory of consumer choice allows anything that has a price to count as a commodity or service. Moreover, although economists assume that most of these things will be used or consumed by individuals and expected by individuals to benefit themselves, nothing in the theory of consumer choice implies that the things that people buy will be used by their purchasers, expected to benefit their purchasers, or in fact be beneficial to their purchasers.

The theory of the firm relies heavily on two motivational assumptions: that individuals
respond to financial incentives and that otherwise individuals prefer leisure. Since most people

take leisure to contribute to their well-being, and it appears that only I can consume my own

leisure, it might seem obvious that the desire for leisure is self-interested. But leisure is often

instrumental to other ends, and these are not necessarily directed toward one's own welfare.

Individuals may seek leisure in order to promote some cause that matters to them or to please

or to assist someone they care about. The other motivational assumption, that financial

incentives determine individual choices, seems to imply that individuals are narrowly materially

self-interested, and it does rule out some non-self-interested motives. For example, it rules out

the possibility that moral commitments or altruism leads people to accept lower wages from

public-interest law firms than from firms representing corporate interests or that moral

commitments lead them to accept lower returns from a “socially responsible” investment fund.

Financial incentives do not explain such choices. Standard theories of production imply that

such motivations are unimportant, just as the theory of consumer choice abstracts from the

altruistic motives that lead individuals to pay more to patronize a local merchant or the

malevolent motives that lead consumers to avoid stores owned by people whose skin color or

religion is not to their liking.

Standard theories of consumer choice and of production thus rule out some non-self-

interested behavior and at the same time, they are largely compatible with a view of individuals

as entirely self-interested. Furthermore, in ruling out some sorts of non-self-interested

motivation, standard mainstream positive economics may encourage the view that individuals

are self-interested. It may also be that many economists believe that people are almost entirely

self-interested.

But self-interest is not built into positive economics. Neither consumer choice theory nor

the theory of production state or imply that individuals are always self-interested. Just as the

theory of consumer choice is silent about the motives that drive consumer choices, so the

theory of the firm is silent about why individuals want more money or more leisure. Indeed, the
simplest versions of the theory of the firm assume that the firm is structured so as to maximize profits, which is unlikely to be consistent with entirely self-interested motivation on the part of the employees. One can say the same thing about classical political economy. J. S. Mill, for example, regarded individuals as seeking wealth, as seeking to reproduce, and as diverted from either of these ends by a desire for “the present enjoyment of costly indulgences” (1843, Bk. VI, chap. 9, sec. 3). Although reproduction might be regarded as self-interested, the purposes behind the accumulation of wealth and costly indulgences need not be self-interested.

5 Welfare economics

Welfare economics rests on the view that an agent A is better off with \( x \) than \( y \) if and only if A prefers \( x \) to \( y \). Since the relation “is at least as well-off with \( x \) as with \( y \)” must be reflexive and transitive, this view of welfare requires that weak preferences also be reflexive and transitive. Context independence also seems to be required, but not completeness: there seems to be no reason that how well someone’s life is going in one state of affairs must always be comparable to how well it is going in another state of affairs. From a subjective \textit{ex ante} perspective, incompleteness of judgments concerning well-being ubiquitous. If asked whether a new bridge or change in the tax code will make me better off, the honest answer is often, “I haven’t the faintest idea.”

Even if economists suppose that it is generally harmless to ignore the incompleteness of (intrapersonal) welfare comparisons, the dependence of preference on belief falsifies the identification of welfare and preference satisfaction. Jack may believe that the expected benefit to him of a new road is greater than the expected benefit to him of a new park or airport because he does not anticipate the noise and congestion that the road causes. After the road is built, he may believe that it in fact has provided him with more or fewer benefits than the airport would have, because of other false beliefs. Even if, as many economists believe, individuals are the ultimate authorities on what they value, they are not the ultimate authorities on the facts.
When self-interested individuals are mistaken about the consequences or properties of alternatives, satisfying their preferences need not make them better off.

This obvious conclusion does not show that welfare economics does not depict individuals as self-interested, but it does decisively refute three of the four interpretations of the claim that individuals are self-interested. Owing to false beliefs, the realized outcomes of their actions may not promote their well-being whether it is appraised objectively or subjectively, and their actions will not be objectively expected to promote their well-being. But false beliefs do not break the link between preferences and subjective expectation of benefit.

But welfare economics is of course concerned with promoting welfare, not expectations of welfare, no matter how misconceived. It accordingly identifies what is in fact good for individuals with what they prefer. Why, despite the obvious problems caused by false beliefs, do welfare economists persist with this identification?

A couple of equivocations may help explain why welfare economists identify welfare with preference satisfaction. Preferences in economists are typically represented by utility functions. It is provable that preferences that satisfy the axioms discussed above and some other technical conditions can be represented by ordinal utilities – that is by numbers that are arbitrary apart from representing the preference ranking. For a finite set of alternatives, this representation theorem is trivial. One can imagine Jack listing all alternatives on a very large sheet of lined paper. If Jack prefers apple to blueberry pie, then he enters “apple pie” on a higher row than where he enters “blueberry pie.” If he is indifferent between blueberry and cherry pie, then he enters “cherry pie” on the same row where he entered “blueberry pie.” Because Jack’s preferences are complete, every alternative goes somewhere, and because his preferences are transitive every alternative has only one location. Jack can then assign numbers to rows, with higher numbers attached to higher rows. Since utilities in this way indicate preferences, economists can replace talk of preferences with talk of utilities. Doing so is in principle harmless, but it makes possible an equivocation. Since the classical utilitarians took “utility” to refer to
pleasure and welfare, they would have interpreted the greater utility of apple pie for Jack as the
greater pleasure and welfare Jack experiences. In that way the purely formal representation of
preference by utility functions slides into the substantive (and false) claim that preference
satisfaction and welfare coincide.

Speaking of preference “satisfaction” creates a second opportunity for equivocation. If
Ann prefers a year of ballet lessons for her son to a less expensive spa membership for herself,
then the lessons satisfy her preferences better than the spa. If she feels more satisfied, then
she is better off. So preferring what looks like a benefit to her son to a benefit to herself is
actually a self-interested pursuit of greater satisfaction. Stated this baldly the equivocation is
not hard to see. A state of the world satisfies my preferences insofar as it conforms to them.
“Satisfaction” in this sense has no implications concerning feelings.

False belief is not the only reason to question the welfare economist’s identification of
welfare with preference satisfaction. A further problem concerns the connection between
preference and self-interest. As discussed in section three, self-interest concerns the content of
preferences, not whether preferences determine choices. Is there any reason why an agent’s
preference ranking should coincide with his or her ranking of alternatives in terms of expected
welfare? One possibility is that the satisfaction of well-informed preferences constitutes an
individual’s welfare. But if one thinks about it, this claim is bizarre. Take some state of affairs
that seems to be irrelevant to my welfare and to Woody Allen’s welfare, such as the number of
provinces in Uzbekistan. If I prefer there to be 12 provinces rather than any other number, and
he prefers there to be 11 rather than any other number, then the actual state of affairs makes
me somewhat better off than I might have been, and it makes Woody Allen somewhat worse off.
How is it possible for the satisfaction of my preference that Uzbekistan have 12 provinces to
enhance my welfare?

I can think of three answers, but none of them vindicates the view that preference
satisfaction constitutes welfare. First, finding out that the number of provinces in Uzbekistan is
as I preferred may give me pleasure. But this answer assumes that welfare is pleasure, not preference satisfaction. According to the preference satisfaction view, the fact that there are 12 provinces in Uzbekistan makes me better off whether or not I ever learn this fact. Second, I might know that there being 12 provinces in Uzbekistan will benefit me in some other way, such as enabling me to win a prize on a quiz show. But then it is not the fact that my preference is satisfied that benefits me, but something else. Third, I might have some worthwhile project that I am pursuing that would be successful if there were 12 provinces in Uzbekistan rather than some other number. But in that case what makes me better off is the successful pursuit of some worthwhile project, not preference satisfaction.

Preference satisfaction does not constitute welfare even when preferences do not depend on false beliefs. If people are self-interested, it is not because merely possessing a preference ranking and acting on it makes them self-interested. But preference satisfaction and welfare might nevertheless often coincide if individuals seek their own good and know what choices are good for themselves. There are many contexts in which it is plausible to suppose that people are predominantly self-interested. If, in addition, their preferences are not cognitively distorted and that their knowledge of what is good for them is at least as good as the knowledge of third parties, then their preferences reflect their judgments of what is in their own interests, and insofar as their judgments are correct, their preferences are a good guide to their welfare.

According to this evidential view, Ann’s preference for ballet classes for her son over a spa membership for herself is a guide to what is good for Ann if she seeks to benefit herself and knows what is in her benefit. Her judgment that the ballet classes are better for her thus depends on some other notion of “better for” than “is preferred to”, since the preference depends on the prior judgment. And if welfare economists are going to rely on the generalization that people’s preferences are a good guide to what is good for them, economists must also have some view of welfare that is prior to preference satisfaction. I suspect that for many economists this tacit theory of welfare is hedonistic, but it need not be; and in fact economists
do not need any philosophical theory of welfare at all. Various generalizations about the sorts of things that are typically beneficial to people would suffice.

Is it plausible to maintain that preferences such as Ann’s are self-interested – that is, that Ann’s preference for ballet classes for her son over a spa membership for herself depends on her judgment that the ballet classes are better for herself than the spa membership? After all, the ballet classes are for her son, not for her. One can describe some circumstances in which the ballet classes for her son are better for Ann. Her son might otherwise make her life miserable pester her for the classes or chastising her for her selfishness in joining the spa instead. Or the ballet classes might be an investment in a career for her son that will provide returns to Ann in the form of prestige for the mother of such an accomplished dancer or a more financially secure retirement. Or it might be that her son’s joy in ballet directly gives Ann pleasure. The fact that the ballet classes are for her son, while the spa is for her does not rule out self-interest.

On the other hand, this recounting of the many circumstances in which the lessons are better for Ann falls far short of proving that her choice of the ballet lessons is self-interested (let alone that all of her choices are self-interested). Since people do so many things that have no obvious connection to their interests or that are apparently detrimental to their interests, how can welfare economists justify identifying welfare with the satisfaction of well-informed preferences?

There are three ways to proceed. First, economists might take a broad permissive view of the interest or welfare of the agent and indeed of the agent himself or herself. Thus, economists often take individuals as representatives of households and their self-interest as the interest of the household. The ballet lessons might then count as self-interested in virtue of promoting the interest of the household whether or not they promote Ann’s interests. Welfare economists can then maintain that if self-interest is understood broadly, then in fact people are overwhelmingly self-interested. But I do not think that welfare economists have established this
generalization, and I am genuinely in doubt about whether it is true. Welfare economists are often persuaded by the fact that individuals in modern societies respond to financial incentives – and indeed they do. But as already pointed out, the fact that individuals prefer more money to less says nothing about what they plan to do with the money. Furthermore, maintaining that individuals pursue the interests of their households rather than specifically their own interests in effect surrenders the claim that individuals are self-interested.

Second, welfare economists might commit themselves to a hedonistic theory of motivation: the anticipation of pleasurable or painful sensations explains why individuals prefer one thing to another.\(^2\) They might claim (implausibly in my view) that both the apparent altruist who spends weeks in terrible conditions tending to the victims of the ghastly civil wars in the Congo and the wine connoisseur who spends weeks touring the vineyards of Bordeaux are after pleasures. These pleasures are different, but desirable mental states are both the ultimate aim of actions and the constituents of welfare. Call this the “warm-glow” theory of motivation.

To maintain that all human actions aim at pleasurable mental states is a sweeping empirical generalization, which economists have done little to confirm. As a claim about the conscious goals of actions, the warm-glow theory is obviously false. I do not explicitly consider how much pleasure I will feel if I assign a B or a C to an exam I am grading, nor do I make my decision on the basis of such a conscious anticipation. Perhaps I am different from most people in this regard (as in many other regards), but I still count as a counterexample. The only plausible version of the warm glow theory maintains that regardless of people’s self-awareness or self-conceptions, the only ultimate things they care about are their own pleasures and pains, which constitute their welfare and explain their actions. A plausible account of desire formation in terms of reinforcement lends some apparent support to this view, but preferences as total comparative evaluations are rarely the direct results of reinforcements by pleasures or pain, and the extent to which the pleasures and pains accompanying outcomes determine desires is

\(^2\) For example, see Dolan 2008, p. 70 and Dolan and Kahneman 2008, p. 217.
controversial. Furthermore, the fact that a pleasurable experience of a state of affairs caused a desire for that state of affairs does not by itself show that the desire for that state of affairs is in fact the desire for a pleasurable experience.

A third approach is different. Why not jettison any concern with welfare and focus normative economics on satisfying preferences? That immediately liberates economists from taking any view of welfare or of the extent to which individuals are self-interested. For example, Robert Sugden writes,

> He never considers the possibility that welfare economics might not be about the assessment of individual well-being from the perspective of a philanthropist, but instead about the non-judgmental aggregation of individuals' preferences, interpreted as their separate subjective evaluations.  

As an account of welfare economics, this view is indefensible, because it either denies that normative economics should be "about the assessment of individual well-being" or it implicitly (and mistakenly) identifies aggregating preferences with promoting welfare. What justifies normative economics as envisioned by Sugden? If one is benevolent and wants people to live well, then one has a good moral reason to support policies that promote their welfare – provided that the policies respect other moral commitments and constraints. But what reason does Ann have to help satisfy Albert's preferences when Ann does not believe that doing so would Albert better off or bring about some other valuable state of affairs? Why should Ann care about whether Albert's preferences are satisfied or frustrated, regardless of the content of the preferences or the effects of their satisfaction on Albert's mental states?

I can think of two answers, but neither of them is very good, and neither commits normative economics to the view that individuals are self-interested. First, one might argue that a standing commitment, other things being equal, to help satisfy the preferences of others is a way of showing respect for them. This claim is questionable and the reason it might give me to

---

3 (2013, p. 98).
help satisfy the preferences of others is weak. Sometimes I show respect for others by attempting to persuade them to pursue different aims. The other reason one might give why social policies should aim (other things being equal) to satisfy individual preferences, regardless of their content, relies on a questionable interpretation of the notion of popular sovereignty. One might argue that since the people collectively are ultimately sovereign, policies must respect their preferences. I think this argument is a mistake. Popular sovereignty requires that one respect the will of the populace concerning how decisions should be taken. It implies rule by plebiscite only if that is what people collectively prefer. If most people determine that they would like social policies to respect preferences, regardless of their content, then popular sovereignty would support Sugden’s proposal. But there is no reason to believe that most people support policy-making that respects preferences, regardless of their content. And in any case, conforming to popular sovereignty does not imply promoting self-interest. Mainstream normative economics largely stands or falls with the generalization that people’s preferences are a good guide to their interests.

6 Conclusions
Positive economics contains some approximations that restrict the range of non-self-interested motivations individuals can have as well as, somewhat surprisingly, some approximations that actually rule out self-interest. Normative economics, in contrast, becomes arbitrary and unmotivated unless it attributes largely self-interested preferences to people. In order for policies that satisfy Albert’s preferences to promote his welfare, it must be the case that that policies that aim to satisfy Albert’s preferences are, on average, better at actually promoting Albert’s welfare than alternative policies. When four conditions are met, there is a good case for basing policy on satisfying preferences:

1. Individuals are good judges of what promotes their well-being.
2. Individuals are better informed concerning the relevant facts than the policy maker is.
3. Individual preferences and choices are not badly distorted by cognitive flaws such as framing effects.

4. The preferences of individuals are largely directed toward promoting their own welfare (or the welfare of their household or those with whom they closely identify).

When the third and fourth conditions are satisfied, people will prefer alternatives because they believe they are better for themselves or those they identify with. When in addition the first two conditions are satisfied, then their preferences will be a better guide to their welfare than the judgment of the policy maker. In many domains it may be reasonable to assume, as a first approximation that these conditions are on average satisfied and that there is good reason to base policy on people’s preferences.

So just as the Pope is really Catholic, so economists – at least insofar as they are concerned with welfare – are committed to the view that there is an important domain of behavior in which individuals are self-interested – at least in a permissive sense of self-interest. But they need not be the reductionists they are often taken to be, who regard people as knowing the price of everything and the value of nothing. Although largely compatible with a view of individuals as self-interested, their positive theories are not committed to self-interest at all, and their view of welfare economics requires only that there be a significant social domain in which a very permissive sense of self-interest predominates. That is not such an unreasonable view.

References


