

PERSON-AFFECTING UTILITARIANISM AND POPULATION POLICY;
OR, SISSY JUPE'S THEORY OF SOCIAL CHOICE

"You don't know," said Sissy, half crying, "what a stupid girl I am. All through school hours I make mistakes. Mr. and Mrs. M'Choakumchild call me up, over and over again, regularly to make mistakes. I can't help them. They seem to come naturally.... [T]oday, for instance, Mr. M'Choakumchild was explaining to us about Natural Prosperity.... And he said, Now, this schoolroom is a Nation. And in this nation there are fifty millions of money. Isn't this a prosperous nation? Girl number twenty, isn't this a prosperous nation, and an't you in a thriving state?"

"What did you say?" asked Louisa.

"Miss Louisa, I said I didn't know. I thought I couldn't know whether it was a prosperous nation or not, and whether I was in a thriving state or not unless I knew who had got the money, and whether any of it was mine. But that had nothing to do with it. It was not in the figures at all," said Sissy, wiping her eyes.

"That was a great mistake of yours," observed Louisa.

"Yes Miss Louisa, I know it was now. Then Mr. M'Choakumchild said he would try me again. And he said, This schoolroom is an immense town, and there are a million of inhabitants, and only five and twenty are starved to death in the streets, in the course of a year. What is your remark on that proportion? And my remark was—for I couldn't think of a better one—that I thought it must be just as hard upon those who were starved, whether the others were a million, or a million million. And that was wrong too."

"Of course it was."

Then Mr. M'Choakumchild said he would try me once more... And I find (Mr. M'Choakumchild said) that in a given time a hundred thousand persons went to sea on long voyages, and only five hundred of them were drowned or burnt to death. What is the percentage? And I said Miss;" here Sissy fairly sobbed as confessing with extreme contrition to her greatest error; "I said it was nothing."

"Nothing, Sissy?"

"Nothing, Miss—to the relations and friends of the people who were killed. I shall never learn!"

Charles Dickens, *Hard Times*

The great insight of utilitarian ethical theories, the feature that makes them seem "so self-evidently correct" to so many of their supporters ([26], p. 84) is surely their insistence that morality is intimately associated with human welfare. As most non-utilitarian theorists would acknowledge, any plausible ethical theory must at least take into account as morally significant the influence of choices and actions on well-being ([23], p. 30). Sissy Jupe, however, cannot accept the classical utilitarian assumption that the misery of the few can be rendered less unfortunate by the prosperity of the many. This paper develops a utilitarian account of choice that explains why she is right. The theory presented also has important implications for population theory, and concerning the moral status of contingent future persons.

I. BRINGING HAPPINESS VS. RELIEVING SUFFERING

Classical utilitarians tell us that "actions are right in proportion as they tend to promote happiness, wrong as they tend to promote the reverse of happiness." Some add that "By happiness is intended pleasure and the absence of pain; by unhappiness, pain and the privation of pleasure" [15]. So far so good: Most utilitarians take this principle to pick out as clear and unambiguous a moral theory as one could want. But as it stands, the utilitarian directive is seriously ambiguous, for it tells us nothing about the relationship between pleasures and pains, or about how we are to reconcile the joint commands to 'promote happiness' and also to 'inhibit the reverse of happiness.' A Benthamic Calculus is no help in this regard, not simply because of its explicit hedonism (which, in any case, is rejected by many contemporary utilitarians), but because Bentham, like Mill and most contemporary utilitarians, fails to distinguish the positive injunction to "promote happiness" from the correlative negative injunction to "prevent unhappiness," and fails to examine the ways in which the two may conflict. Contemporary economic theories of utility are worse than useless in this regard, for they render it impossible to distinguish the two injunctions from one another ([30], pp. 274-5).

There are two important, and seldom noticed differences between these twin utilitarian commands. First, the positive utilitarian imperative to "maximize happiness" is insatiable, while the negative utilitarian command to "minimize misery" is satiable: no matter how much happiness we have, the positive principle tells us that more would always be better. But the negative principle ceases to generate any obligations once a determinate but demanding goal has been reached: if misery could be eliminated, no further obligation would be implied by the negative principle, even if it were possible to provide people (or non-human 'persons') with additional bliss. Second, the injunction to minimize misery is a *person-affecting* injunction: it implies an obligation to respond to the needs of people who already exist, but imposes no obligation to bring people in the world so that their miseries might be relieved. The injunction to maximize happiness, as commonly explained by utilitarians, does not restrict its scope to improving the lives of people who already exist, but goes further to recommend that we should bring

more people into existence whenever their existence would facilitate an increase in total or average utility. If happiness is good, and more of it is better, then the positive principle seems to tell us that it would be better to have more well-off people around so that their happiness could contribute to a larger utilitarian aggregate.

Classical utilitarians assume that these twin injunctions are complementary, and that pains and pleasures are commensurable so that they can balance one another out in a grand utilitarian aggregate. But it is far from obvious that pains and pleasures are commensurable in this way, and there is good reason to doubt that the twin utilitarian aims are even compatible—at least not without further explanation. We might well ask how pleasures and pains mix to produce an aggregate, and whether different ways of mixing them might be preferable to one another. Suppose Aristippus and Epicurus are offered a choice between two lives: L1 includes high-highs low-lows, while L2 is as calm and *ataraxic* as an Epicurean could wish. These two lives might well be equal on a utilitarian scale that weighs the aggregate happiness they contain, in one case because the highs are compensated by the lows, and in the other because there are no highs or lows. On the assumption that the pleasures of the former life would precisely outweigh the pains, would it be irrational for Aristippus to prefer L1 on grounds that it is more exciting? Would it be irrational for Epicurus prefer the L2 on grounds that it is less so?

'*Commensurabilism*' is the view that pleasures and pains can cancel one another out in the way that classical utilitarians usually assume. Suppose that Jeremy, a commensurabilist hedonic-state utilitarian, is offered the following proposal by his good friend:

Sadist I: Sam the Sadist offers to stimulate the pleasure centers of Jeremy's brain, if only he will allow Sam to inflict suffering that is no greater than the amount of compensatory pleasure offered in return. Like all sadists, Sam enjoys inflicting suffering. He removes dental fillings and applies drills and sharp dental probes directly to the nerves in his victim's teeth. But Sam is a more reasonable sadist than most: he is willing to "pay" for his enjoyment, by replacing the fillings and providing compensatory hedonic benefits. Thus he claims that he leaves his victims no worse off than they would otherwise have been.

Ignore, for the moment, the interesting question whether Sam's sadistic preferences should be taken into account from the moral point of view, for we have other concerns. The question on which I would like to focus attention is this: Should a commensurabilist hedonist like Jeremy be indifferent between acceptance and rejection of this offer? Most of us, I believe, would reject it, even if trustworthy Sam promised us that the compensatory pleasure would be more than enough to compensate us, so that we would actually be better off as a result.

Let us set aside another potentially confusing issue that might cloud our examination of this example and others: It is worth noting that if we adopt revealed preference theory [RPT] and suppose that happiness is exclusively a function of preferences and their satisfaction, then the example will seem nonsensical. Those who accept RPT would argue that the pleasure Sam offers to Jeremy is equal to the pain he anticipates inflicting *only if* Jeremy is indifferent: Jeremy's indifference determines the quantity of pleasure that Sam will need to provide. This introduces another classical problem in the theory of choice, for it is in Jeremy's interest to overstate the disvalue of the pain Sam proposes to inflict, since he can then bargain for a larger hedonic benefit. There is no easy way to insure that Jeremy will accurately reveal the amount of pleasure that would compensate him, for even if he feels indifferent when offered 100 units of compensatory pleasure, he has good reason to demand 200. Since the defense of RPT is couched in psychological behaviorism, this is a problem. For the assumption that Jeremy has a reservation value beneath which he will not go presupposes that the value Jeremy places on 100 units of compensatory pleasure is sometimes independent of the value he will reveal in his choices.

Behaviorism is no longer a defensible psychological theory, even though its vestigial organs live on in the economic theory of choice. Those who find the implicit behaviorism of RPT implausible must find some other way to determine when pleasures and pains balance out, and this is a problem for any theory that regards preference satisfaction as relevant from the moral point of view (that is to say, for all plausible moral theories). But even if this problem can be solved, we may still be uncertain what to say about whether Jeremy should be indifferent between acceptance or rejection of Sam's offer, or whether Aristippus and Epicurus' decisive but divergent preferences are irrational. For these reasons, RPT cannot be the basis for a *normative* theory of utilitarian choice even if it were the best theory of preference to underlie an economic theory of consumer choice. We might agree that individuals are typically the best judges of their own interests, and that therefore we should not interfere with the choices of others. But this attitude toward others' preferences is consistent with acknowledgment that pleasures and pains may not cancel one another out like equivalencies in the numerator and denominator of a fraction.

In the case of *Sadist I* it seems unlikely that either choice is rationally required or forbidden, and we may acknowledge Jeremy's right to make his own choice. But the problem of comparing pains and pleasures (or, in a non-hedonic case, disappointments and satisfactions) becomes more poignant when Sam's activities involve more than one person:

Sadist II: Before meeting Sam, Alph is already miserable and Beth already is happy. Sam satisfies his sadistic impulses by torturing Miserable Alph. But although he is a sadist, Sam believes that he is obliged to leave the world no worse than he found it, and accepts a commensurabilist utilitarian theory that prescribes when

one situation is better or worse than another. He therefore offsets Alph's pain by providing compensatory benefits for Happy Beth.

As Sam is hauled off by the police, should good utilitarians join him in his indignant protest that he is innocent, since he has left the world no worse than he found it? Derek Parfit argues that they should not. He recommends the following

Compensation principle: One person's burdens cannot be compensated by benefits provided for someone else.¹

If suffering is bad, and if outcomes that include uncompensated suffering are *ceteris paribus* worse than those that do not, then Sam has failed to leave the world "no worse off than he found it." Parfit claims that the *Compensation principle* is clearly true, but if so, it is notoriously difficult for most utilitarian theories to explain why. Indeed, the objection that 'we cannot rightly pay for the bliss of the fortunate many by inflicting misery and suffering on the unfortunate few' is often given as ground for rejecting utilitarian ideals altogether. According to Norman Daniels, it is utilitarianism's rejection of the *Compensation principle* that founds the common claim that utilitarian theories inappropriately ignore the boundaries between persons ([4], p.95). Kantians might explain the *Compensation principle* with the "dogmatic assertion" that individuals should not be used as mere means in the pursuit of utilitarian maxima.² At the least, any theory that would require us to reject the *Compensation principle* would come in conflict with a widely shared and deep seated (if not ineradicable) moral judgment or "intuition."

Other ground could be given in favor of this principle, but here I will recommend it only for its attractiveness, and will consider the relationship between this principle and Utilitarian theories of choice. I will argue that utilitarians can agree with Parfit's *Compensation principle* if they accept that burdens and benefits, misery and happiness, are not two sides of the same coin, but that they identify two different evaluative currencies altogether. For example, one might maintain one's commitment to utilitarianism while adding that one is never justified in ignoring the injunction to minimize misery even when one has an opportunity to increase the bliss of those who are not miserable. Perhaps it is a view like this one that underlies Sissy Jupe's judgment that it is just as hard on those who starve, whether the others are "a million or a million million." If what we care about is the minimization of misery and deprivation and if we also accept the *Compensation principle*, then we will not care that the *proportion* of miserable persons to well-off persons is small, we will instead focus our attention on the sum total of suffering. Before developing this into a more substantial utilitarian theory for ordering outcomes, we will consider the relative weight of the twin utilitarian imperatives in another context.

II. POPULATION CHOICES

The ambiguity of the twin utilitarian injunctions is especially confusing when we apply them to decisions about population policy. Some of these confusions generate great paradoxes, as Henry Sidgwick glimmeringly suspected and as Derek Parfit has dazzlingly shown. These paradoxes have in turn driven some theorists toward surprisingly baroque and sometimes brilliantly subtle revisions of the utilitarian view.³ The examples that follow reproduce the problem utilitarians face in Sadist II in the context of population choices. I consider both total and average utilitarian views: total utilitarians recommend that we should strive to maximize the total amount of utility that everyone enjoys, while average utilitarianism tells us that we should maximize the value we get when we divide the grand total by the number of people who enjoy it.

Maximizing Total Utility: Total utilitarianism [TU] does not take account of the distribution of well being and misery, only its sum total. Under the commensurabilist assumption, TU instructs that we should *ceteris paribus* be indifferent between the following alternatives:

Total 1: The next generation will contain 10 billion blissful people and no miserable people.

Total 2: The next generation will contain 100 billion blissful people and X miserable people, where X is chosen so that the total disutility of X miserable people is equivalent to the additional utility provided by the 90 billion blissful additions. That is, where $U(*)$ is a utility function, choose X such that:

$$U(90 \text{ Billion Blissful People}) + U(X \text{ Miserable People}) = 0.$$

Total utilitarians claim that Total 1 and Total 2 are equally choiceworthy, since they both contain the same utilitarian total of well-being. But those who accept Parfit's Compensation principle will join Sissy Jupe in the judgment that it is just as hard on the miserable no matter how numerous are the blissful. Since there are suffering people in Total 2, and since their suffering is uncompensated, the *Compensation principle* would lead one to conclude that Total 1 is better than Total 2. After all, the 90 billion additional blissful people are not worse off in Total 1 than they would have been in Total 2, for they would never exist at all in Total 1.

Maximizing Average Utility: Average utilitarians [AU] tell us to maximize average utility. Again, this injunction implies nothing about the distribution of well-being and misery within the population: if the average utility in two outcomes is equivalent, we should be indifferent between one in which everyone's well-being is the equal, and another in which one person is super-extravagantly blissfully happy, while everyone else is miserable. In population choices, AU recommends indifference between the following alternatives:

Average 1: The next generation contains 9 Billion Blissful People each of whom enjoys a happy life at a Utility Level of B.

Average 2: The next generation contains 9 Billion Blissful people, X BlissPlus People, and 9 Billion Miserable People, where BlissPlus people are even better off (much better off) than blissful people, and where X is chosen so that:

$$\frac{U(X \text{ BlissPlus People}) + U(9 \text{ Billion Miserable People})}{(9 \text{ Billion} + X)} = B$$

The average utility level is the same in *Average 1* and *Average 2*. But those who accept Parfit's Compensation Principle will again note that *Average 2* contains uncompensated suffering and a lot of miserable people, while *Average 1* does not. Acceptance of the compensation principle would lead one to conclude that *Average 1* is better than *Average 2*.

Some utilitarians will bite the bullet, and will reject the Compensation principle [13]. Those who are more firmly committed to the Compensation principle than to utilitarianism may be led to conclude that utilitarianism is an unacceptable theory [4]. I will urge that the problem is not with utilitarianism but with the commensurabilist assumption that well-being and misery can balance one another out like the two sides of a see-saw. In section III I develop a non-commensurabilist utilitarian theory for comparing outcomes. This theory maintains the tight utilitarian association between morality and welfare, but without violating the Compensation principle. Section VI will incorporate this theory into a broader account of obligation.

III. EVALUATING OUTCOMES: A NON-COMMENSURABILIST THEORY OF UTILITARIAN 'BETTERNESS'

Theories are *utilitarian* if they rank outcomes in terms of their relative 'betterness' and if they further hold that the 'betterness' relation takes account only of the well- or ill-being of the persons who will exist if one outcome comes about rather than another. Theories are *non-commensurabilist* if they reject the claim that misery and well-being are commensurable qualities that balance one another out in a grand utilitarian aggregate. Only non-commensurabilist theories are consistent with the Compensation principle. Is it possible to develop a utilitarian theory for ordering outcomes that does not violate the Compensation principle? I will show that it is. But the Compensation principle limits the structure of such a theory. The requirement that one person's burdens cannot be compensated by benefits to another person implies that the obligation to minimize misery is lexically prior to the correlative obligation to maximize well-being.

It is worthwhile to re-emphasize that terms like 'happiness,' 'well-being,' 'misery,' and 'ill-being' should be understood to refer to whatever makes life go

well or ill, and should be interpreted as evaluative criteria that apply to complete lives, and not necessarily to momentary mental states like pleasure, satisfaction, pain, or frustration. Momentary hedonic states like pleasure and satisfaction may indeed contribute to making lives go well, just as pain and frustration may contribute to making some lives go badly. But for the purposes of this theory, we need not take a stand on whether hedonic states are the only determinants of well-being and misery. For now I will simply describe a system for ordering outcomes. A full defense of this system, including comparison with plausible alternatives, must wait for another time.

Utilitarians can accept the compensation principle if they formulate the negative utilitarian imperative as a first principle for the ordering of prospective outcomes:

Negative Ordering Principle [NOP]: Outcomes are worse if they contain more misery, less bad if they contain less.

Fred Feldman plausibly holds that misery and unhappiness have a different moral status when they are freely chosen, compensated within a life, or deserved as punishment [8]. On this view, NOP should take account only of ill-being that is undeserved, uncompensated, and unchosen. I raise this issue only to set it aside, for acceptance or rejection of Feldman's view will not affect the argument made here. In ordering outcomes according to their relative betterness, the first step is to use NOP to divide outcomes into indifference classes—classes of outcomes that contain the same amount of misery or ill-being. NOP orders these indifference classes according to the aggregate misery they contain. For those indifference classes that contain more than one member, positive ordering principles may be used to further order outcomes within these sets. As a first approximation, we might adopt the following:

Positive Ordering Principle [POP]: Outcomes are better if they contain more happiness, less good if they contain less.

To avoid conflict with the Compensation principle, POP may appropriately be applied only for further refinement of the ordering given by NOP. On this view, NOP is thus lexically prior to POP. We must distinguish different interpretations of POP that apply to its application in different contexts. In particular, different interpretations of POP are associated with its implications for the ordering of outcomes that contain different numbers of persons. Many different number choices will be settled by NOP, but some will remain, since NOP gives no ground for ordering outcomes that contain the same amount of misery, regardless of the number of well-off persons they contain. Before considering the difficult issues raised by those different number choices that remain after application of NOP, we can see that POP has the following implication for the ordering of outcomes that contain the same number of persons:

Positive Ordering Principle for Same Number Choices [POP-SN]:

If the same number of people would exist in either of two outcomes, O1 and O2, O1 is better than O2 if those who would exist in O1 are better off than those who would exist in O2.

POP-SN tells us that we should make those who exist as well off as possible. As stated here, POP-SN is more or less identical to Derek Parfit's "Same Number Quality Claim," which he calls Principle Q ([21], p.360). Like Parfit's *Principle Q*, POP-SN is subject to a variety of interpretations. A weak paretian interpretation of POP-SN requires that at least one person is better off in O1 than in O2 while no one is worse off in O1 than in O2. A strong paretian interpretation requires that everyone must be better off in O1 and O2. Non-paretian interpretations might require that there be less aggregate misery, or more aggregate happiness, or some combination of these requirements. While I cannot sort these interpretations out here, any of them is consistent with the view I develop in later sections of this paper.

How should comparisons be made when outcomes contain different numbers of persons? Four alternative views have been defended concerning the purported value of adding additional well-off people to the world: Richard Hare argues that it is better if more well-off people exist [10]. Christoph Fehige argues that it not better if more people exist, regardless of how well-off they would be [7]. In a similar vein, Jan Narveson claims that the existence of additional well-off people is a matter of moral indifference ([17], [18]). And Thomas Hurka argues that additional people make outcomes better when total population is low, but that the value of additional people diminishes as population levels rise [12]. These views correspond to four different refinements on the POP that might be used to order outcomes in different number choices:

Positive Ordering Principles for Different Number Choices [POP-DN]:

POP-DN1: Outcomes are better when they contain more well-off persons [Hare].

POP-DN2: Outcomes that contain more people are *ceteris paribus* no better than those that contain fewer, no matter how well-off the additional people would be [Fehige].

POP-DN3: The value of additional persons is a decreasing function of population size [Hurka].

POP-DN4: The existence of additional well-off persons makes outcomes neither better nor worse [Narveson].

All of these, with the possible exception of POP-DN3, accept that it is worse when more miserable people exist. While they are not all inconsistent with one another, they clearly have quite different implications concerning the moral status of future persons, and there is little agreement about which we should accept. Arguments that have been provided in favor of one of these principles over others usually appeal to our intuitions, but intuitions provide a less reliable when they are applied to novel domains, and population theory is surely a novel decision domain.

But even if agreement cannot be reached on the choice among these alternatives, it is important to note that different number choices that involve more or less ill-being have already been settled by the application of NOP, and the set of possible outcomes already has a high degree of order, given by NOP and POP-SN, before we even come to these divisive issues. In most practical circumstances, our ability to describe or predict likely outcomes is limited, and our descriptions will be coarse grained enough that most different number choices will be definitively settled by NOP. The negative principle is far less controversial than the positive principle in its implications for most different number choices,⁴ since its application does not raise metaphysical questions about the relative value of persons and the well-or-ill-being that they enjoy. Although the theory of obligation developed later in section VI is consistent with any of the positive ordering principles for different number choices identified above, it is worthwhile to take a moment briefly to sort out some of the arguments that motivate the different views, and my own reasons for preferring POP-DN4 over the alternative views. Section IV will discuss some of the issues that divide advocates of the different positive principles listed above.

IV. THE VALUE OF WELL-BEING AND THE VALUE OF WELL-OFF PERSONS

The different principles deserve, and have received, more extensive consideration than they can be given here, and the issues that separate their advocates are among the most divisive issues in normative population theory. If one accepts the *Compensation principle*, however, the set of cases over which these divisions range is rather small and the set of prospective outcomes already has a high degree of order before we even come to these issues.

Richard Hare's argument for the view that the existence of additional well-off people makes outcomes better is based on his view that to be moral is to have analogous preferences for analogous situations. Since some real life people want to have been born, Hare's view implies that morality requires us to prefer the existence of any other people who would, if they existed, want to have been born.⁵

Some people are persuaded by this argument. But one wonders what it is that people think they are preferring when they view themselves as having a preference for having been born. Such a preference might have its source in vague fears about the emptiness of nonexistence, or nebulous worries about *what it would have been like for them* had they never been born. Since there is *nothing* that it would have

been like *for them* if they had never been born, such sources would render this purported preference poorly formed or poorly articulated. In order to have a preference among options, one must have at least some idea what those options would involve, and when alternatives are vaguely or nebulously understood, people may be mistaken in their beliefs about what they prefer. Partha Dasgupta writes "The impossibility of imagining our own nonexistence gives spurious credence to the view that nonexistence must be a long dismal night from which we must try to rescue people" ([6], p.114). If one suspects that people's purported 'preference for having been born' might derive from such a spurious source, one may reasonably reject Hare's argument. The Epicureans believed that the fear of death would be extinguished once people had a proper understanding of the state of nonexistence. Since nonexistence before birth is usually far less frightening to people than post-natal nonexistence (death), it is perhaps even more plausible to think that a proper understanding of nonexistence would extinguish or render superfluous Hare's purported preference for having been born.⁶

Christoph Fehige's argument for the claim that existence is never preferable to nonexistence is presented in one of the most interesting papers to have appeared on this subject since the publication of Parfit's path-breaking book. The claim is founded on an antecedent theory of preference, *Anti-Frustrationism*, for which Fehige offers a detailed independent argument. According to an anti-frustrationist, "we don't do any good by creating satisfied extra preferences. What counts about preferences is not that they have a satisfied existence, but that they don't have a frustrated existence" ([7], p. 497). On this view, satisfaction of a preference at best brings us to the level of well-being we would have enjoyed if we had not had the preference at all. Since non-existent people have no preferences to be frustrated, this view implies that they enjoy the same level of well-being that they would have enjoyed if they had existed and had all of their desires fully satisfied. Thus Fehige concludes that people who might have existed but who never do exist are always at least as well off as actual persons who exist and who are sure to enjoy less than complete preference satisfaction. It follows from this theory of preference that nonexistence is always at least as good as existence. This theory of preference has ancient roots in the works of Epicurus and his followers, but also perhaps in Sophocles, who writes "Not to be born is, past all prizing, best; but when a man hath seen the light, this is next best by far, that with all speed he should go thither, whence he hath come."⁷ Unlike Sophocles, Fehige does not claim that people are better off if they die sooner; for many of us this would frustrate our desires to go on living. But like Sophocles, he claims that those possible people who are never born are better off than we less fortunate actual persons, since the terminally unborn will never suffer frustration of their desires. Those who find this view absurd on its face should surely examine Fehige's brilliant contemporary defense of it. While Fehige begins by defending this theory of preference directly, he later demonstrates that his view implies plausible solutions for many of the most perplexing paradoxes of population theory. Even without the direct argument Fehige offers, the tidiness of

his solution to Parfit's *Mere-Addition Paradox* should lead population theorists to give his argument the most respectful consideration.

Although I cannot discuss all of the features of Fehige's detailed and careful defense of this theory of preference, I believe that his view may reasonably be rejected. In particular, anti-frustrationism may pass too quickly over the very plausible possibility that some preferences might be enjoyable and desirable in themselves, and that some pleasures may constitute positive benefits rather than simply the removal of painful longings. Sometimes people are better off if they have preferences to be satisfied than they would have been had they not had those preferences. Perhaps some desires are like afflictions which we seek to free ourselves, either by extinguishing or satisfying them. But surely there are other desires and preferences possession of which makes our lives better. Why else would people take pains to develop expensive and refined tastes if their possession didn't make life richer by virtue of the rare but profound satisfaction enjoyed when these tastes are satisfied? Fehige is perhaps too quick to conclude that there is no independent value for sweet longings and cultivated preferences, and that no desires are desirable for their own sake.⁸

It may help to compare *anti-frustrationism* with an ancient precursor: the Epicureans recommended that we should endeavor to extinguish sexual desire altogether so that its frustration would not trouble our happiness. But they also allowed that those who find such desires difficult to extinguish might permissibly satisfy them so to be rid of them. Later Epicureans (Lucretius) added that in such circumstances these desires should be satisfied without any passionate feeling or attachment, since turbulent passions would undermine the calm and unperturbed happiness of the Epicurean ideal. Most find this argument unconvincing, and for good reason: sometimes turbulent passions, loves, and attachments make our lives richer, even when they disturb the placidity of our happiness and even though they render us susceptible to loss and pain. If some satisfactions provide positive benefits that are worth pursuing for their own sake, these benefits cannot be explained by an account of happiness as absence of pain and painful desires. But Fehige's anti-frustrationism is not quite the same as the Epicurean view: in particular, Fehige does not characterize unsatisfied desires as *painful*, and can explain the benefit of acquiring new preferences in terms of the preference that a persons may have to do just that. Anti-frustrationism allows that it may be reasonable to acquire preferences as a way of avoiding the frustration of meta-preferences. But like Epicureanism, anti-frustrationism offers a negative account of well-being as the absence of frustrated or unsatisfied preferences or desires, and we may have reason to want an account of desire that better captures the positive role that desires, satisfactions, and passions play in our lives. While this brief argument cannot do justice to the subtlety of Fehige's position, we may still have good reason to want an account of preference that recognizes this positive role.

Thomas Hurka famously argues that additional persons have positive value, but that the value of additional persons diminishes as total population size increases [12]. Hurka brilliantly shows that this hypothesis about the value of persons collects

many widely shared considered judgments about extinction and the value of individuals. But like Hare, Hurka completely separates the value of additional persons from the value that their lives have for them. For Hurka, the value of a person is decisively dependent on population size, and when population size is low, Hurka implies that it might even be appropriate to bring miserable persons into existence for the purpose of perpetuating the species, and that sometimes it would be better to increase population than to increase the welfare of existing people ([12], p.504; [11] p.142). This would violate the *Compensation principle*, since these people would suffer personal misery in the service of an impersonal population total. For these reasons, and once again acknowledging that my brief discussion here can not do justice to Hurka's subtle argument, I believe that one may reasonably reject this account of the value of persons in the context of population choices.

These questions raise issues at the foundation of ethics, concerning the value of well-being and the value of persons: Is well-being is valuable because it is *good for people*, or are people valuable because their lives contain well-being? The problem does not arise in the case of ill-being or misery, because the negative utilitarian injunction places no independent value on the existence of any evaluative unit, independently of the persons who might possess it. No one would suggest that the injunction to minimize misery might impose an obligation to bring people into existence that their misery might be relieved. Because I find it implausible that the value of well-being could be independent of its value to those persons who enjoy it, and because acts that fail to bring well-off persons into existence are worse for no one, I am inclined to accept Narveson's view that it is neither better nor worse to bring additional well-off persons into existence.

Those who are antecedently convinced by the argument for one of the alternative views discussed in this section may well be unconvinced by the brief responses given to those arguments in the discussion above. But as I have argued, the set of options over which this disagreement ranges is small, and agreement on these issues may not be necessary for agreement about what we should do. As noted earlier, I believe that the theory of obligation I develop in what follows could be accepted by advocates of any of these views. But one important reason theorists have given for opting for an impersonal theories like those recommended by Hare and Hurka is that they believe that only impersonal theories can give an acceptable account of population choice. If successful, the theory developed in section V and VI undermines that reason.

V. TWO DEONTIC ASYMMETRIES

The examples considered in sections I and II motivate the suggestion that misery and well-being may not balance one another out in the way that classical utilitarians typically assume. Sections III and IV take steps toward development of a non-commensurabilist theory of utilitarian betterness. In this section I move from the evaluation of outcomes to the evaluation of actions in terms of their

permissibility, impermissibility, and obligatoriness. In particular, I consider a set of asymmetries in our commonsense understanding of our obligations concerning the value of promoting well-being versus inhibiting misery. In section VI, I show how the theory of betterness developed in III and IV can be incorporated into a plausible theory of obligation.

It is sometimes thought that there is an asymmetric relation between the normative status of acts that prevent harm, and those that provide extra benefits for persons who are already well off and who are not experiencing harms. This gives us the

First Asymmetry Thesis: We have an obligation to prevent harms, at least when we can do so without excessive risk or cost, but no obligation to provide additional benefits for persons who are well off and who are not miserable or deprived or experiencing harms. It may be a good thing to benefit those who are adequately well-off, but it is not morally obligatory.

Suppose we have the opportunity to provide an additional benefit to Keith, who is already quite well off and who is living a happy and fulfilling life. But we could instead use our resources to improve the lives of others who would be doomed to lives squalor and misery in the absence of our aid. Those who accept the *First Asymmetry Thesis* might acknowledge that it would be nice to provide Keith with an additional benefit, but would not claim that failure to do so would be morally reprehensible. Failure to provide a benefit to someone who is well off might show that one lacks certain virtues of beneficence or generosity, but would not be a serious moral wrong. But failure to prevent harm and misery, when one could have done so without risk or excessive effort on ones own part may sometimes be a serious moral wrong.

In a slightly more fanciful context, suppose that Well-off Keith is exceptionally efficient at turning resources into well-being: if we used our resources to feed the hungry or house the homeless, we would decrease their misery by a given amount, but if we instead gave them to Keith, his welfare level, already quite high, would go through the roof.⁹ Almost everyone would agree that our obligation to aid the miserable should take precedence, even if Keith's prospective utility gain would be larger than theirs. It would be nice to send Keith's welfare level through the roof, but it is not morally obligatory, and it would be wrong to do so at the expense of the hungry or the homeless.

A second deontic asymmetry arises in the case of "genesis choices,"—choices that will result in the existence of persons who would not exist otherwise:

Second Asymmetry Thesis: We have an obligation not to cause people to exist if those who would exist will be miserable and deprived, but we have no obligation to cause people to exist who

would be happy ([7], section 7; [14], p.100; [17]; [18]; [21], p. 391).

Those who accept the *second asymmetry thesis* believe that people are not doing something wrong when they choose not to have a child who would be happy. But if one knows that any child one has will be miserable, then it would be wrong to have children. In general, those who accept the *second asymmetry thesis* hold that when our choices might cause the existence of well-off persons who would not have existed otherwise, the happiness those people would enjoy does not by itself constitute a reason for bringing them into existence. But when our choices might cause the existence of miserable people, the misery they would suffer does constitute a reason for not bringing them into existence.

It is plausible to think that this asymmetry has its source in the difference between misery and well-being, and in the differences between the positive and negative utilitarian commands: when one fails to bring into existence a person who would be well-off, the person who would have existed is not well-off or badly-off as a result, since nonexistent persons cannot be well-off or badly off. But if one brings into existence a miserable person, that person's misery is a consequence of one's choice. It seems likely that the folk theory of obligation that underlies these asymmetries has roots that stretch to the same source that makes the *Compensation principle* so plausible. The unhappiness of the miserable cannot be compensated by benefits to the happy, no matter how efficient the happy may be at converting benefits into welfare.

VI. AN IMPURE CONSEQUENTIALIST THEORY OF OBLIGATION

Pure consequentialists hold that there is no sharp distinction between the theory of betterness and a theory of obligation. They believe that our obligation is always to do what would be best, and actions are better or worse depending only on the betterness or worseness of the outcomes at which they aim. Pure consequentialist theories have come under criticism on ground that they conflict with commonsense moral judgments: such theories accord no weight to promises, integrity, or moral desert. Some have argued that such theories demand too much of us, and that they eliminate entirely some commonsense normative categories like supererogation—actions that are good, but which exceed the requirements of obligation. I hope that the *impure* theory I develop here will avoid such objections.

Some terminological issues must be settled before we will have an adequate interpretation of the theory I will propose: To say that an agent A has an obligation to do X is to say that A's failure to do X would reveal that A is a *schmuck*. For the moment, I rely on the reader's intuitive understanding of what it means to be a *schmuck*. Unless a contravening obligation or prerogative justifies an agent in overriding or rebutting a *prima facie* obligation, A has a duty to do what obligation assigning principles prescribe. On this understanding, the function of a theory of

obligation is to distinguish the *schmucks* from people who are minimally morally decent. Most people are imperfect, and sometimes fail to do what they have obligations to do: that is, most of behave like *schmucks* in at least some of our actions and choices. But there are some people whose general propensity to do what they are morally obliged not to do justifies the judgment that they are *schmucks tout court*. I believe that this account of obligation captures well both our common usage, and the usage of obligation-talk in most discussions of moral theory. In effect, this account makes the theory of obligation depend on the theory of what it is to be a *schmuck*—a theory of vice.

To say that it would be good for A to do X even though X is not obligatory is to say that A has a moral reason to do X. But while A is no saint if she fails to do what would be good but is not obligatory, she is not thereby a *schmuck* either. Thus one function of a theory of beneficence is to separate those who are only minimally decent from those who are saints and heroes. Again, I believe that this captures commonsense and philosophical notions of supererogation. (See [28]) On this account, the theory of moral goodness and supererogation depends on a theory of virtue.

I would like to suggest a theory of obligation and supererogation that is consistent with the Compensation principle and with the non-commensurabilist theory of utilitarian betterness developed in sections III and IV. If it is acceptable, this theory would also explain the two deontic asymmetries described in the previous section. As in previous sections, it should be understood that the unit of comparison is complete life expectations and not momentary hedonic states. The theory is not a pure consequentialist theory, since it allows that people do not always have an obligation to do what will result in the best outcome, and since it leaves room for actions that are supererogatory. I will call this theory the

Impure Consequentialist Theory of Obligation [ICTO]:

1. *Negative Principle of Obligation* [NPO]: Actions that reduce (or minimize) misery are prima facie obligatory.
2. *Positive Principle of Beneficence* [PPB]: Actions are good if they increase well-being. Actions are better or less-good depending on their propensity to promote well-being: better when they do so more effectively, less good when they do so less effectively.

In the non-commensurabilist theory for the ordering of outcomes, the negative utilitarian principle was given priority over the positive principle. This was necessary to render the theory consistent with the *Compensation principle*. In ICTO, the negative principle is understood as obligation-generating, while the positive principle is not. In this case, the difference in treatment of positive and negative consequences gains support from four different sources: First, the ordering of our

obligations described by ICTO is consistent with the non-commensurabilist theory of betterness developed in section III. Second, it is consistent with and supportive of the commonsense judgments of those who accept the first deontic asymmetry discussed in section V. In general, most of the actions we classify as obligatory or supererogatory fit ICTO fairly well. Third, this account of the separation between obligation and supererogation reflects the difference in urgency between acts that are necessary to prevent the perpetuation of harmful circumstances, on the one hand, and acts that would provide additional benefits for persons who are already adequately well-off on the other hand [24]. Failure to respond to others' urgent needs, when one could easily do so, is a grave moral wrong. There is no similar moral urgency concerning actions that provide extra benefits to those who are already well enough off.

Finally, this account of the ordering of our obligations and supererogations explains the intuitions of those who believe that there is an obligation not to have children who would suffer and who would be badly off, but who do not believe that there is an obligation to have children who would be well off. ICTO is *consistent with* the view of those who believe that there is a good moral reason to bring more well-off persons into existence, but it is also consistent with the view of those like Narveson and myself, who hold that such actions are, from the moral point of view, a matter of indifference.¹⁰

Disagreement about what would be better need not yield disagreement about what we have obligations to do. But there is good reason to want one's theory of obligation to capture the moral difference between outcomes that involve the existence of additional well-off people, and outcomes that involve the existence of additional people who are badly off, for in the context of genesis choices, the value of well-being is formally different from the disvalue of misery. When we choose not to bring into existence a person who would be happy, there is no one who is worse off. But when we fail to meet the obligation assigned by NPO by bringing a miserable person into existence, there exists a badly-off person who would not have been badly off but for our moral mistake. It is a virtue of ICTO that it captures this difference and its moral significance.

Different substantive moral theories will allow different circumstances to defeat or override the obligation assigned by NPO. Some hold that such obligations can only be overridden by contravening *prima facie* obligations of equal or greater moral 'weight.' Others hold that there are *agent centered prerogatives* that justify an agent in giving more weight to her own interests than to the interests of others. Since ICTO is offered not as a complete moral theory, but as an account of the way in which consequences and well-being should be taken into account in the context of moral choice, we need not try to stipulate here all of the circumstances that might justify or excuse an agent in failing to fulfill the *prima facie* obligations assigned by NPO. Utilitarian theories are often criticized on grounds that they demand too much. While ICTO is less demanding than other utilitarian theories, it is still quite demanding, and probably implies that most of us fail to fulfill our obligations much of the time. The degree to which ICTO is demanding is partly a function of our

circumstances (what misery are we positioned to respond to?) and partly a function of the circumstances under which our obligations may be defeated.

VII. IS ICTO A PERSON-AFFECTING THEORY? AND DOES IT MATTER?

It is sometimes claimed that our theories should be *person-affecting* in their treatment of well-and-ill-being, but it is not always clear what this requirement is supposed to mean. Sometimes this claim is associated with the dictum that it is good to make people happy, but that we should be indifferent about making happy people [18]. Sometimes it is associated with the claim that the part of morality that is concerned with human well-being should be explained entirely in terms of what would be good or bad for those people who will be better or worse off ([21], p. 39; [9], p. 66), and at other times with the claim that all and only those people who are or will be actual fall within the scope of morality ([1], p. 70; [11]). It is clear not only that these proposals are different from one another, but that some are not consistent with others. This ambiguity makes it difficult to say what philosophers mean when they claim that moral theories should be (or should not be) person affecting.

On some accounts, ICTO is not a person-affecting theory, since it incorporates obligations to some persons who will never be actual: it implies a *prima facie* obligation not to bring people into existence if they would be miserable. There is no person to whom this obligation is owed unless it is violated. On the other hand, ICTO captures what many have wanted in a person-affecting theory: The value of well-being is a function of its value for persons, not vice versa, and the disvalue of misery is also a function of its value for those who suffer it [1]. ICTO is also consistent with Narveson's insistence that it is good to make people happy, but that we may be indifferent about making happy people [18]. I believe that this justifies regarding ICTO as a person-affecting theory. But surely its plausibility as a theory of choice does not depend on the aptness of this ambiguous label.

VIII. SOME COMMON OBJECTIONS

Because ICTO is closely related to negative utilitarianism [NU], it is subject to several of the obvious objections that have been levelled against NU. Most theorists seem to have regarded these objections as decisive, and to my knowledge, no one has seriously considered or defended negative utilitarianism since Karl Popper briefly proposed that it would "add clarity to ethics" if we would "formulate our demands negatively, i.e. if we demand the elimination of suffering rather than the promotion of happiness."¹¹ Indeed, some theorists have not made the distinction between negative utilitarianism as articulated in NPO and a quite different view that has been defended by Narveson and others: Narveson claims that we have obligations not to harm, but no obligation to give aid.¹² Obviously Narveson's

theory of obligation is different from ICTO, which assigns a positive obligation to aid others whenever doing so would minimize misery. In this respect, ICTO is similar to commonsense reflective morality, since most people would agree that there is a duty to aid others in at least some circumstances—that is, there are circumstances in which failure to aid others would clearly reveal to the world that one is a *schmuck*. Here I consider three common objections to views like ICTO that incorporate the more plausible negative utilitarianism represented by NPO and its role in ICTO.

(i) *The Marginalist Objection*: Some economists and philosophers regard lexical orderings as intrinsically irrational. They would conclude that the non-commensurabilist theory of ‘betterness’ developed in section III is unacceptable, since the value of relieving suffering is taken to be prior to the value of providing additional happiness to the well-off. This first objection can be dismissed as a marginalist prejudice. In order to make the theory of consumer choice work properly, economists need to assume that people’s indifference curves are smooth and twice-differentiable, but the fact that commensurable values are easier to manipulate mathematically is no reason to think that *value* is structured in this way. A theory that forbids lexical ordering would also force us to accept that Seneca’s unwillingness to sell his honor for no-matter-how-much monetary profit is also an irrational (since lexical) ordering.

(ii) *The Anti-Natalist Objection*: Several theorists have argued that negative utilitarian views imply that we have an obligation never to have children. Hermann Vetter argues that this is an implication of any view that accepts Narveson’s claim that we (i) have a duty not to conceive a child if we could reliably foresee that it will be miserable, but (ii) we have no duty to conceive a child even if we could be sure that it will be very happy. But because we don’t know the future, having children always involves imposing on them some risk that they will not be happy. Vetter writes: “It will be seen immediately that the act “do not produce the child” dominates the act “produce the child” because it has equally good consequences as the other act in one case [if the child is or would have been happy], and better consequences in the other [if the child is or would have been unhappy]. So it is to be preferred to the other act as long as we cannot exclude with certainty the possibility that the child will be more or less unhappy; and we never can.”¹³

If ICTO implied that people always have an obligation not to conceive children (and by implication, that people who have children are all *schmucks!*), it might indeed be an unacceptable theory. Fortunately, the Vetter dominance argument fails, because it fails to take into account all of the morally relevant considerations at stake in the decision to bring a child into existence. Our prospective children may contribute to making our lives better, and to making the lives of others better as well. Thus *failure* to conceive a child will put at risk the welfare of all those who might have been better off (or less badly off) if one’s child had existed. However, ICTO may imply that parents have certain obligations with respect to their children: since the choice to have a child is justified only if our decision to conceive them leaves the world less badly off than it would otherwise have been, we have good

reason to teach our children not only that they should flourish and find happy lives for themselves, but also that they have obligations to those who are badly off. If we have good reason to believe that our children will have happy lives, and that they will make the world a better place in some respects, then it is permissible to have them.

Far from implying that it is always wrong to conceive a child, ICTO implies that people sometimes have a *prima facie* obligation to do so if the existence of their child would make the world a less miserable place overall. The happiness that the children might enjoy is not, on this view, an independent reason for conceiving them. But their ability to contribute to the lives of others is an independent reason of this sort. But even in this case, the obligation to have a child would still be a *prima facie* obligation, and a full moral theory would need to include an account of the circumstances under which this obligation might be defeated.¹⁴ Acceptance of ICTO will not turn us into organic baby-production machines.

(iii) *The Apocalypse Objection*: But what if one could painlessly stop the entire human endeavor all at once? Some have argued that the best way to minimize misery would be painlessly to kill off all sentient beings, perhaps by blowing the planet to smithereens with bombs. Since NU and ICTO imply that we have an obligation to minimize misery, it has been argued that they must also imply that we have an obligation to destroy the planet and all conscious beings on it. No theory with this implication can be acceptable ([27], pp. 542-3; [11], p.60).

Those who take ICTO or something like it to be a complete moral theory could respond by emphasizing the extent to which the case for the continuation of our species can rest on the utilities of people who are or will be actual even if humankind dies out.¹⁵ One might then focus on the frustrated hopes and plans, and disappointed expectations that would result. Some have argued fairly persuasively that it is far from clear that the continuation of our species would be justified if not for these contingencies ([1], p.71). But there is another way around this objection. While ICTO does provide an excellent account of the way in which welfarist considerations should be taken into account in moral decision making, one might also urge that welfarist considerations are not the only morally relevant considerations. Other considerations might come into play in two ways: First, we might have a contravening *prima facie* obligation not to kill other people. Such an obligation might then defeat the *prima facie* obligation to alleviate suffering and would provide the basis for a plausible argument against the annihilation of our world. Since most people accept that we *do* have a *prima facie* obligation not to kill other people, this might be sufficient.

But second, we might accept that ICTO describes one species of moral value among others, and that the continuation of the species has only person-affecting value of a sort that is not incorporated in ICTO, but which is not ruled out either. I have offered no argument for the claim that only person-affecting values are significant from the moral point of view, and cannot imagine what such an argument would look like. Few would deny that person-affecting welfarist considerations represent one important kind of moral value and have a special moral status, but

some people hold that there are other types of value as well. If the continued existence of persons has value in itself, this value must be explained and justified in non-person affecting terms. For example, environmental ethicists have sometimes argued that the value of animal species and ecosystems is a non-person-affecting value. Some readers may agree that there are other species of value, and that their pursuit can sometimes override or rebut our *prima facie* obligation to relieve suffering and to aid those who are badly off. If the perpetuation of the species has value in itself, as Hurka, for example, claims that it does, then this value would need to be taken into account before we could have a complete understanding of our obligations. ICTO is offered as a theory of the way in which person-affecting welfarist values are relevant to moral decision making, but there may be no reason to believe that all relevant values can be summed up in a simple utilitarian rule.¹⁶

X. CONCLUSION

If accepted, ICTO would explain why Jeremy would not be irrational to reject Sam's offer in *Sadist 1*, and why Sam makes the outcome worse by torturing Alpha and compensating Beth in *Sadist 2*. ICTO would explain why *Total 1* and *Average 1* are respectively better than *Total 2* and *Average 2* and would account for the two deontic asymmetries described in section V. It would explain why Sissy Jupe is right to think that the suffering of some cannot be compensated, in the grand utilitarian equation, by the well-being of others. ICTO would have practical implications for social choice theory and for theories of poverty and development: it implies that we should regard poverty as getting worse when there are *more miserable impoverished people*, and should regard development efforts as successful when they reduce total destitution, not when they increase GNP or decrease the proportion of population that lives in destitution. It would thus sanction a shift from relative or proportionate measures of progress toward absolute measures.

The theory developed here provides a way to order prospective outcomes without running into many of the paradoxes and confusions that plague alternative theories. ICTO has generally plausible implications for population choices—at least, its implications seem much less *implausible* than those of alternative views. It captures the most compelling considered convictions that lie behind the claim that welfare-theory should be person-affecting. It implies that well-being is valuable because it is good for people, denies that people are good because they can serve as loci on which well-being may be heaped, and incorporates the claim that it is good to make people happy, but not necessarily good to make happy people. Finally, ICTO forbids the purchase of bliss for some at the price of misery for others. As far as I know, it is the only utilitarian theory that possesses these virtues. Who could ask for anything more?¹⁷

NOTES

¹ See Parfit ([21, p. 337], and Daniels ([40], p. 170). Parfit adds one qualification to the *Compensation Principle*: Our burdens can sometimes be compensated by benefits to those we love, but they cannot be compensated by benefits to total strangers.

² Some utilitarians never mention Kant without using the word 'dogmatic.' But it is clear that one may be a nondogmatic Kantian (see the work of Barbara Herman, Christine Korsgaard, or Thomas Hill) as that one may dogmatically propound the utilitarian view. Bentham and James Mill are safe, dead examples of dogmatic utilitarians. Indeed, J.S. Mill passed the same judgment on Bentham. Of living utilitarians, we need say nothing if it is not good -- perhaps the dogmatic variety became extinct in the mid-Nineteenth century.

³ See especially [6], [8], [12], [17], [19], [26] and [30].

⁴ There is a set of different number choices concerning which the Negative principle is highly controversial. These will be discussed in section VIII.

⁵ See Fehige ([7], p. 439) for a brief examination of Hare's view. The concise presentation of Hare's view given here is adapted from Fehige, and from Hare ([10], pp. 70-73).

⁶ In act III Scene ii, Hamlet whines "Who would fardels bear, To grunt and sweat under a weary life, But that the dread of something after death, The undiscovered country from whose bourn No traveler returns, puzzles the will And makes us rather bear those ills we have Than to fly to others that we know not of?" If we interpret Hamlet's fear as fear of nonexistence rather than fear of unknown prospects for life after death (admittedly a textual stretch), then Hamlet's fear of death might be mistaken for a preference for life (or for existence). Such misconceptions about the evil of nonexistence are still widely shared. Is it implausible to suppose that these misconceptions might be the source of Hare's speculation that people can rationally prefer existence to nonexistence? Hare considers and rejects the possibility that our preference for having been born might be inextricably associated with fear of death ([10], p. 70-71). It seems unlikely that Hare's argument or this one can settle the question once and for all.

⁷ Sophocles, *Oedipus at Colonus*, at 1225 in the text. Cited in Dasgupta ([5], p. 116). For similar sentiments see Job 3:3 and Ecclesiastes 4:2-3.

⁸ Actually Fehige discusses this possible objection. He writes: "All this is to say that nobody wants new and satisfied preferences, or that it would be irrational, or morally irrelevant to want them. Mary's wish to have additional wishes conveys instrumental value on their existence--the value of avoiding to frustrate another (*viz.*, the highest order) wish. (. . .) The argument here is not a crusade against wanting more. All it denies is that even if people *don't* want to want more, wanting more (and getting it) would benefit them." ([7], p. 434). Once again, my discussion here cannot settle the question. I can only point toward the sources of disagreement. Those who find Fehige's independent argument for *anti-frustrationism* convincing may be unimpressed by my objections here.

⁹ Perhaps Keith is a "utility monster" like those described by Nozick ([20, p. 41]. The example here was inspired by Allen Buchanan, who suggested during my oral prelim exam that Keith Lehrer just gets more enjoyment out of things than most other people do. Anyone unconvinced by the *First Asymmetry Thesis* might consider sending contributions to Keith Lehrer at the University of Arizona. I would suggest that such contributions should go to OXFAM instead.

¹⁰ Not that parents should be or will be indifferent about bringing more well-off children into the world, but their decision to do so, on Narveson's view, is not better because of the happiness that these children will enjoy.

¹¹ Popper, [22], vol. I, chap. 5, n. 6(2). Also vol. i, chap. 9, n. 2. Quoted in Smart ([27], p. 542).

¹² Narveson, [16]. See Glover ([9], pp. 66-69) for criticism of this view.

¹³ Vetter ([29], p. 301). Heyd ([11], p. 60). Interestingly, Vetter does not regard this as an objection. He argues that we really do have an obligation not to have children, and that it would be a good thing if the human race could quietly go extinct.

¹⁴ There is much more to be said in response to Vetter's objection. Having children changes parents too, sometimes making them more responsible, kinder, and more understanding people than they would have been if they had not had children. If one has reason to think that parenthood would change one in these ways, then the additional good one might do as a result might constitute an additional reason for conceiving a child, not decisive in itself perhaps, but not irrelevant either. The general point is that

Vetter's game theoretic analysis of the problem leaves out consideration of the interests of persons other than the child herself.

¹⁵ Bennett [1] emphasizes these costs on pp. 70-71. My presentation of the case owes much to his earlier presentation of it.

¹⁶ It has also been suggested that this kind of negative utilitarian view implies Parfit's Repugnant Conclusion. In Wolf [30] I argue that some negative utilitarian views do not imply a repugnant version of the repugnant conclusion.

¹⁷ I received helpful comments on this project from Christoph Fehige, Frederick Ferre, Nick Fotion, Claudia Mills, Michael Tooley, and colloquium participants at the University of Colorado. I would like to thank them, and to absolve them from responsibility for any faults—or at least from any of the paper's faults!

BIBLIOGRAPHY

1. Bennett, J.: 1978, "On Maximizing Happiness." In R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations*, 61-73. Philadelphia: Temple University Press.
2. Broome, J.: 1992, *Counting the Cost of Global Warming*, Cambridge: White Horse Press.
3. Broome, J.: 1994, "The Value of a Person", *The Proceedings of the Aristotelian Society*, supp. vol. LXVIII.
4. Daniels, N.: 1988, *Am I My Parents' Keeper? An Essay on Justice Between the Young and the Old*, New York: Oxford University Press.
5. Dasgupta, P.: 1974, "On Optimum Population Size." In *Economic Theory and Planning*, Ashok Mitra, ed. Calcutta: Oxford University Press.
6. Dasgupta, P.: 1994, "Savings and Fertility: Ethical Issues", *Philosophy and Public Affairs*, 23:99-127.
7. Fehige, C.: 1997, "A Pareto Principle for Possible People." In Christoph Fehige and Ulla Wessels (eds.), *Preferences*, de Gruyter: Berlin/New York.
8. Feldman, F.: 1996, "Justice, Desert, and the Repugnant Conclusion." Forthcoming.
9. Glover, J.: 1977, *Causing Death and Saving Lives*, NY: Penguin Books.
10. Hare, R.: 1993, *Essays on Bioethics*, New York: Oxford University Press.
11. Heyd, D.: 1992, *Genethics: Moral Issues in the Creation of People*, Berkeley: University of California Press.
12. Hurka, T.: 1983, "Value and Population Size," *Ethics*, 93:496-507.
13. Leslie, J.: 1989, "The Need to Generate Happy People," *Philosophia* 19:29-33
14. McMahan, J.: 1981, "Problems of Population Theory", *Ethics* 92:96-127.
15. Mill, J. S.: 1979/1861, *Utilitarianism*, Indianapolis: Hackett Publishing Co.
16. Narveson, J.: 1967, *Morality and Utility*, Baltimore MD: Johns Hopkins Press.
17. Narveson, J.: 1967b, "Utilitarianism and New Generations," *Mind* 76:62-72.
18. Narveson, J.: 1973, "Moral Problems of Population," *The Monist* 57:62-86.
19. Ng, Y-K.: 1989, "What Should We Do About Future Generations? The Impossibility of Parfit's Theory X," *Economics and Philosophy*, 5:235-253.
20. Nozick, R.: 1974, *Anarchy, State and Utopia*, New York: Basic Books.
21. Parfit, D.: 1982, *Reasons and Persons*, Oxford: Clarendon Press.
22. Popper, K.: 1962, *The Open Society and its Enemies*, Princeton N.J.: Princeton University Press.
23. Rawls, J.: 1971, *A Theory of Justice*, Boston: Harvard University Press.
24. Scanlon, T.: 1975, "Preference and Urgency," *The Journal of Philosophy*, 72(19):655-669.
25. Sidgwick, H.: 1981/1907, *The Methods of Ethics*, Indianapolis: Hackett Publishing Co.
26. Singer, P.: 1976, "A Utilitarian Population Principle", in Michael Bayles, ed. *Ethics and Population*, Cambridge: Schenkman Publishing Co.
27. Smart, R. N.: 1958, "Negative Utilitarianism," *Mind*, 67:542-3.
28. Urmson, J. O.: 1958, "Saints and Heroes." In A. I. Melden, ed., *Essays in Moral Philosophy*, Seattle: University of Washington Press, 198-216.
29. Vetter, H.: 1971, "Utilitarianism and New Generations," *Mind* 80:301-2.

30. Wolf, C.: 1996, "Social Choice and Normative Population Theory: A Person-Affecting Solution to Parfit's Mere Addition Paradox," *Philosophical Studies*.81(2-3):263-282.