Aggregating Subjective Well-Being for Marginal Policy Adjustments*

Daniel J. Benjamin
Cornell University and NBER

Ori Heffetz
Cornell University

Miles S. Kimball
University of Michigan and NBER

Nichole Szembrot
Cornell University

First Draft: December 30, 2012
This Draft: December 30, 2012

Abstract

We propose a theoretically-grounded approach for aggregating estimates of the effects of policy on SWB for different groups into a policy change. The approach—the "as-if voting mechanism"—has two central features: (1) the scales for SWB are treated as ordinal and non-comparable across groups, and (2) any group's gains from deviating from truthfully reporting its SWB levels on the survey that is used for estimating policy effects are second order. The resulting policy recommendation is a marginal adjustment of policy from the status quo. When used for multiple policies, the as-if voting mechanism incorporates information on intensity of preference—despite treating the SWB scales as ordinal—by implicitly giving each individual a budget constraint for votes and allowing her to allocate the votes across policies. Among its desirable features, the mechanism identifies and implements policy changes that are relatively unimportant to any one individual but that would benefit many.

JEL Classification: D69, H0, I38

Keywords: subjective well-being, happiness, life satisfaction, aggregation, policy

* We thank Gregory Besharov, Steve Coate, Ted O'Donoghue, Marc Fleurbaey, Alex Rees-Jones, Alois Stutzer, and the theory seminar audience at Michigan for valuable early conversations and comments. For financial support we are grateful to Cornell's Institute for Social Science, S. C. Johnson Graduate School of Management, and NIH/NIA grants R01-AG040787 to the University of Michigan and T32-AG00186 to the NBER.
E-mail: db468@cornell.edu, oh33@cornell.edu, mkimball@umich.edu, nes66@cornell.edu.

Proposals for the government to track "national well-being" through subjective well-being (SWB) surveys have been gaining momentum in Britain, France, the United States, and other countries. We consider how SWB data might be used for policy purposes. Two key questions are:

1. How should an individual's responses to different survey questions, which elicit different aspects of well-being, be aggregated into a proxy for the individual's utility?
2. Once the effect of a policy on each individual's utility proxy has been estimated, how should these effects be aggregated across individuals to guide policy?

Our previous paper, "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference" (Benjamin, Heffetz, Kimball, and Szembrot, 2012) addressed the first question. This paper addresses the second.

Several considerations motivate the large-scale joint enterprise of collecting SWB data and using it in a theoretically-grounded way to inform policy (see Layard, 2005; Diener, 2006, signed by 50 researchers; Stiglitz, Sen, and Fitoussi, 2009). First, people care deeply about many aspects of their lives that can be affected by policy but are not well measured by GDP and other traditional economic indicators. Second, in contrast with holding referenda to determine policies, estimating the effect of a policy on SWB sidesteps individuals' lack of information and misconceptions about policy, and remains practical even with a large number of policy questions. Indeed, SWB survey responses may ideally serve as all-purpose poll data for answering unanticipated policy-effect questions as they arise.

As the setup for our analysis, we imagine that there is a government agency that conducts a national SWB survey, and we take as given that the agency has some credible method for estimating the effects of policies on a utility proxy (which may be a linear combination of responses to different SWB questions, as described in our earlier paper). We assume that the agency generates such estimates for different groups of citizens. These estimates are then used to recover each group's implied preferences over policies.

The mechanism we propose for aggregating these estimates has two central features: (1) the scales of SWB measures are treated as ordinal and non-comparable across groups;[1] and (2) the mechanism is non-manipulable: every (individual and) group has an incentive to respond truthfully to the SWB survey. Ordinality is important not only because of well-known objections to assuming interpersonal comparability of utility, but also because different people might use SWB response scales differently. Non-manipulability is important because once SWB survey responses are explicitly used to guide policy, individuals may misreport their true SWB levels if they can benefit from doing so (Frey and Stutzer, 2012). For example, if SWB survey responses were treated as cardinal and added up to get a "social welfare function"—as is done implicitly in empirical work that estimates effects of policy on SWB averaged across all individuals—citizens could increase their impact on policy by reporting exaggerated feelings (and might be organized to do so by interest groups).

The mechanism we propose is a *local* mechanism. It uses only the orientation of indifference surfaces in the neighborhood of a status-quo policy vector. The resulting policy recommendation is a marginal adjustment of policy. Two attractions of a local approach to social choice are that local policy effects can be estimated empirically more credibly than global effects, and that designing policy *de novo* is usually less practical than adjusting policy (much as Feldstein, 1976, argued that a theory of tax reform may be more useful than a theory of optimal taxation). Moreover, we believe that merely adjusting policy is especially attractive in the particular context of SWB data. At least initially, using these data in a limited way (as a supplement to more familiar inputs to policy-making) is prudent because the enterprise is still exploratory and untested. In addition, although our formal description of the mechanism is static—a point to which we return below—the notion of marginal policy adjustments fits with the dynamic spirit of national-SWB proposals: monitoring national well-being on an ongoing basis, and using the data to frequently evaluate and incrementally modify policy.

---

[1] In order to estimate policy effects on a group, the agency will have to pool SWB data across individuals in the group. Doing so is consistent with treating SWB data as merely ordinal for each individual—and accurately recovering the group's preferences—as long as the individuals comprising the group locally have the same implied preferences over policies. Notice that for the effect of a single policy, either everyone's SWB will increase or everyone's SWB will decrease, and hence the average effect will have the correct sign. And when the agency estimates the effects of two policies, since everyone has the same marginal rate of substitution, the distribution of the SWB effect of one policy will be a positive constant times the distribution of the SWB effect of the other policy. Hence the ratio of the average effects will correctly recover the shared marginal rate of substitution (even if individuals use the SWB response scales differently, as long as the estimation is local).

Our approach is not restricted to the use of SWB data, but it relies on having some way of estimating the effects of policy on an empirical proxy for ordinal utility. While combinations of SWB measures provide immediate candidates for such a proxy, we emphasize the need for work to deal with a variety of issues that bedevil the measurement of SWB (see, e.g., Adler, 2012), which we do not address.

I. Formal Description of the As-If Voting Mechanism

Suppose that there are $P \geq 1$ policies. Each element of $\boldsymbol{p} \equiv (p_1, \dots, p_P)$ represents the level of one policy and is a real number (e.g., the federal tax on distilled spirits) that the government can adjust independently of other policies. Let $\boldsymbol{p}_0$ denote the status-quo policy vector. There are $G > 1$ groups in the population. For each group g, let ϕ_g denote its fraction of the population, and let $u_g(\boldsymbol{p})$ denote the ordinal utility that is assumed identical for each individual in the group. We assume that each $u_g(\cdot)$ is strictly quasi-concave and continuously differentiable.

A government agency conducts a national SWB survey and, based on the survey responses, constructs a utility proxy for each individual. The agency then estimates $\frac{\partial r_g}{\partial p_j}$, the average effect of each policy j on the reported utility proxies for the individuals in group g, which yields a measure of the group's preferences over policy based on the SWB reports (as discussed in footnote 1). We assume that the $\frac{\partial r_g}{\partial p_j}$'s are observed but the $\frac{\partial u_g}{\partial p_j}$'s are not; these may differ if groups misrepresent their SWB responses. The agency's aggregation mechanism is its algorithm for mapping the $\frac{\partial r_g}{\partial p_j}$'s into a policy change $\Delta\boldsymbol{p}$.

The as-if voting mechanism that we will propose has exogenous parameters: $m_1, m_2, \dots, m_P$, and ε. Note that the P policies are measured in different units, e.g., tax rate vs. dollars of spending. The m_j's make them comparable by pinning down a distance metric: $m_j > 0$ is the amount of change in policy j (e.g., dollars tax per proof gallon) corresponding to "1 policy-unit." The parameter $\varepsilon > 0$ is the "step size" of the mechanism. It is the Euclidean distance, in terms of policy units, of the *maximum* policy change that the mechanism can prescribe (which is only achieved if, locally, all groups have the same preferences). The key features of the mechanism do not depend on the choice of these parameters, but we discuss in section III below how they would matter in practice.

To describe the mechanism, additional notation will be useful. For each group g, let $\boldsymbol{\nabla u_g} \equiv \left(\frac{\partial u_g}{\partial p_1}, \dots, \frac{\partial u_g}{\partial p_P}\right)$ denote the utility gradient at $\boldsymbol{p}_0$, and let $\boldsymbol{\nabla r_g} \equiv \left(\frac{\partial r_g}{\partial p_1}, \dots, \frac{\partial r_g}{\partial p_P}\right)$ denote the reported well-being gradient at $\boldsymbol{p}_0$. Define the matrix $\boldsymbol{M} \equiv \mathrm{diag}(m_1, m_2, , \dots, m_P)$, and note that any policy change $\Delta\boldsymbol{p}$ in natural units is the change $\boldsymbol{M}^{-1}\Delta\boldsymbol{p}$ in policy units. If $\|\boldsymbol{M\nabla r_g}\| > 0$, define $\widetilde{\boldsymbol{\nabla r_g}} \equiv \varepsilon \frac{\boldsymbol{M\nabla r_g}}{\|\boldsymbol{M\nabla r_g}\|}$, which is the reported well-being gradient in policy units, normalized to have length ε. If $\|\boldsymbol{M\nabla r_g}\| = 0$, define $\widetilde{\boldsymbol{\nabla r_g}} \equiv \boldsymbol{0}$. We refer to $\widetilde{\boldsymbol{\nabla r_g}}$ as group g's "ε-vote."

Example: Imagine two groups, Young and Old, and two policies, the federal tax on distilled spirits and spending on national parks. Suppose the relevant government agency estimates that a \$1 increase in the tax per proof gallon on distilled spirits increases the utility proxy of the Old by 3 units and decreases the utility proxy of the Young by 3 units. The agency also estimates that a \$100-million increase in spending on national parks decreases the Old's utility proxy by 2 units and increases the Young's by 3 units. With these data, $\boldsymbol{\nabla r}_{\text{Old}} = \left(\frac{3 \text{ utils}}{\$1 \text{ per proof gallon}}, \frac{-2 \text{ utils}}{\$100 \text{ mil.}}\right)$ and $\boldsymbol{\nabla r}_{\text{Young}} = \left(\frac{-3 \text{ utils}}{\$1 \text{ per proof gallon}}, \frac{3 \text{ utils}}{\$100 \text{ mil.}}\right)$.

In order to apply the mechanism, the agency must pin down what amount of change in each policy is equal to 1 policy unit, which, together with ε, determines for each policy the maximum change in that policy that could be generated by the mechanism. Suppose that one policy unit corresponds to a \$1.25 per proof gallon change in the tax on distilled spirits, or a \$125 million change in federal funding for the National Park Service: $m_1 = \$1.25$ per proof gallon and $m_2 = \$125$ million.

The ε-votes are

$$\widetilde{\boldsymbol{\nabla r}_{\text{Old}}} = \varepsilon\left(\frac{\$1.25\left(\frac{3\text{ utils}}{\$1}\right)}{\sqrt{\left[\$1.25\left(\frac{3\text{ utils}}{\$1}\right)\right]^2+\left[\$125\text{mil.}\left(\frac{-2\text{ utils}}{\$100\text{mil.}}\right)\right]^2}}, \frac{\$125\text{mil.}\left(\frac{-2\text{ utils}}{\$100\text{mil.}}\right)}{\sqrt{\left[\$1.25\left(\frac{3\text{ utils}}{\$1}\right)\right]^2+\left[\$125\text{mil.}\left(\frac{-2\text{ utils}}{\$100\text{mil.}}\right)\right]^2}}\right) = (0.83\varepsilon, -0.55\varepsilon) \text{ and}$$

$$\widetilde{\boldsymbol{\nabla r}_{\text{Young}}} = \varepsilon\left(\frac{\$1.25\left(\frac{-3\text{ utils}}{\$1}\right)}{\sqrt{\left[\$1.25\left(\frac{-3\text{ utils}}{\$1}\right)\right]^2+\left[\$125\text{mil.}\left(\frac{3\text{ utils}}{\$100\text{mil.}}\right)\right]^2}}, \frac{\$125\text{mil.}\left(\frac{3\text{ utils}}{\$100\text{mil.}}\right)}{\sqrt{\left[\$1.25\left(\frac{-3\text{ utils}}{\$1}\right)\right]^2+\left[\$125\text{mil.}\left(\frac{3\text{ utils}}{\$100\text{mil.}}\right)\right]^2}}\right) = (-0.71\varepsilon, 0.71\varepsilon).$$

As illustrated in Figure 1, each is a vector of length ε that points in the direction of maximal increase in the utility proxy for the group.

The aggregation mechanism that we call "as-if voting" can be described as:

(1) $$\boldsymbol{M}^{-1}\Delta\boldsymbol{p} = \sum_{g=1}^{G} \phi_g \, \widetilde{\boldsymbol{\nabla r_g}} \, .$$

In words, the policy change (in policy units) is the vector sum of the ε-votes, with each group's vote weighted by its group size.

<u>Example (continued):</u> Suppose that the population shares are $\phi_{\text{Old}} = \frac{2}{5}$ and $\phi_{\text{Young}} = \frac{3}{5}$. The mechanism prescribes a policy change (in policy units) of $\frac{2}{5}(0.83\varepsilon, -0.55\varepsilon) + \frac{3}{5}(-0.71\varepsilon, 0.71\varepsilon) = (-0.09\varepsilon, 0.22\varepsilon)$. Figure 2 illustrates this vector addition, with the resultant policy change from $\boldsymbol{p}_0$ to $\boldsymbol{p}_1$ shown by the solid vector.

Suppose ε, the maximum possible policy change, is 1 policy unit. Then the mechanism prescribes $(-0.09, 0.22)$. Converting back to natural units: reduce the tax by $0.09 \times \$1.25$ per proof gallon $= \$0.11$ per proof gallon, and increase spending on national parks by $0.22 \times \$125$ million $= \$25.3$ million.

For ε small, the mechanism is non-manipulable up to a first-order approximation: each group g has an incentive to ensure that its ε-vote is calculated using its true preferences, $\boldsymbol{\nabla r_g} = \boldsymbol{\nabla u_g}$. To see why, note that $\boldsymbol{M\nabla u_g}$ is the true gradient of well-being in policy units. Thus, up to first order, $\phi_g \varepsilon \frac{\boldsymbol{M\nabla u_g}}{\|\boldsymbol{M\nabla u_g}\|}$ is group g's most-preferred policy change of size $\phi_g \varepsilon$ starting from *any* initial policy vector that is within an ε-ball around $\boldsymbol{M}^{-1}\boldsymbol{p}_0$. Next, note that regardless of other groups' ε-votes, the policy vector to which group g's vote is added, namely $\sum_{g' \neq g} \phi_{g'} \widetilde{\boldsymbol{\nabla r_{g'}}}$, is necessarily less than ε away from $\boldsymbol{M}^{-1}\boldsymbol{p}_0$. Therefore, for each group g, it is (up to a first-order approximation) a dominant strategy to ensure that its ε-vote is $\varepsilon \frac{\boldsymbol{M\nabla u_g}}{\|\boldsymbol{M\nabla u_g}\|}$. A group can ensure that the government agency calculates its ε-vote using its true preferences by accurately reporting (up to a monotonic transformation) its SWB levels, which the agency uses to calculate the impact of policy on the group's SWB. No group can benefit by exaggerating swings in SWB because its impact is constrained to have size $\phi_g \varepsilon$.

It is easiest to understand why we call the mechanism "as-if voting" in the scalar case where there is a single policy dimension. Normalizing m to 1, equation (1) specializes to:

$$\Delta p = \varepsilon \sum_{g=1}^{G} \phi_g \operatorname{sign}\left(\frac{dr_g}{dp}\right). \tag{2}$$

In this case, the mechanism simply counts the fraction of citizens in favor of increasing p and the fraction in favor of decreasing p and changes p by that vote margin times ε. (Indifferent citizens abstain.) Unlike in an actual referendum, citizens do not need to have correct beliefs about how the policy will affect them; the agency adjusts policy based on its calculation of how each citizen *would* vote under perfect information, based on its estimate of how the citizen will be affected. In this one-dimensional case, the mechanism is exactly incentive compatible for a range of small enough ε's.

In the multiple-policy case, the mechanism in equation (1) can be interpreted "as if" each citizen is optimally allocating votes for or against each of the *P* policies, subject to the constraint that the sum across policies of the squared magnitude of votes cannot exceed ε. Formally, each group g chooses $\Delta \boldsymbol{p_g}$ to maximize $u_g(\boldsymbol{p_0} + \Delta \boldsymbol{p_g})$ subject to the constraint $(\boldsymbol{M}^{-1}\Delta \boldsymbol{p_g})'(\boldsymbol{M}^{-1}\Delta \boldsymbol{p_g}) \leq \varepsilon^2$. With Lagrange multiplier λ, the first-order condition is $\boldsymbol{\nabla u_g} - \boldsymbol{2}\lambda(\boldsymbol{M}^{-1})'\boldsymbol{M}^{-1}\Delta \boldsymbol{p_g} = 0$. This equation implies that the optimal $\boldsymbol{M}^{-1}\Delta \boldsymbol{p_g}$ is proportional to $\boldsymbol{M\nabla u_g}$, and the constraint implies that it has length ε. Therefore, $\boldsymbol{M}^{-1}\Delta \boldsymbol{p_g} = \varepsilon \frac{\boldsymbol{M\nabla u_g}}{\|\boldsymbol{M\nabla u_g}\|}$. (As above, this solution is approximately equal to the solution to the group's maximization of $u_g(\boldsymbol{p_0} + \sum_{g' \neq g} \Delta \boldsymbol{p_{g'}} + \Delta \boldsymbol{p_g})$, for ε sufficiently small.) The mechanism sidesteps the Gibbard-Satterthwaite impossibility theorem, which rules out incentive-compatibility when voting on three or more possible outcomes, by restricting attention to first-order incentive-compatibility for local policy adjustments.

This voting interpretation helps clarify two important features of the aggregation mechanism. First, when *P* > 1, the mechanism incorporates information on a group's intensity of preference. It does so—despite treating utility as ordinal and giving each citizen equal weight averaged across the policy dimensions—by weighting a group's preference *relatively* more strongly along the policy dimensions where a marginal policy adjustment would have a relatively bigger effect on utility. An implication is that, ideally, the mechanism would be used with as many policy dimensions included as possible, thereby allowing as much intensity information as possible to be accounted for.

Second, because the voting constraint is quadratic, each group will allocate at least a small amount of its voting mass to every policy issue. The mechanism will therefore enact policy changes that are relatively unimportant to any one individual but that would benefit many. It can even disfavor policies, such as restrictions on trade, that benefit a few individuals a great deal but hurt many individuals a little bit.

II. Properties of the As-If Voting Mechanism

We now review five properties of the mechanism and contrast it with alternative approaches to aggregation. The first and second properties, already emphasized above, are using only ordinal utility information and being non-manipulable. As far as we are aware, all alternative ideas for how to use SWB data for policy are manipulable. For example, a common practice among empirical SWB researchers is to measure the effect of a policy in dollars by estimating the amount of additional income that an individual would need to receive in order for her SWB to remain unchanged. Any aggregation rule that used such estimates would give survey respondents an incentive to understate the effect of income on their SWB, in order to magnify the dollar value that a policymaker imputes to a change in their SWB.

The third property is a local version of the Paretian principle: if some direction of policy change would make all groups better off, then the mechanism will implement change in that direction. Formally, for any policy-change vector $\boldsymbol{\delta}$ such that $\boldsymbol{\delta}' \cdot \left(\boldsymbol{M}\boldsymbol{\nabla}\boldsymbol{u}_{\boldsymbol{g}}\right) > 0$ for all g, it follows that $\boldsymbol{\delta}' \cdot \left(\sum_{g=1}^{G} \phi_g \, \varepsilon \frac{\boldsymbol{M}\boldsymbol{\nabla}\boldsymbol{u}_{\boldsymbol{g}}}{\|\boldsymbol{M}\boldsymbol{\nabla}\boldsymbol{u}_{\boldsymbol{g}}\|}\right) > 0$ and hence $\boldsymbol{\delta}' \cdot (\boldsymbol{M}^{-1}\boldsymbol{\nabla}\boldsymbol{p}) > 0$. Since groups will as-if vote on every issue, this Paretian property implies that—even if citizens disagree about issues that are important to them—the mechanism will find and implement areas of agreement on other issues.

Fourth, the mechanism satisfies anonymity: permuting the identities of SWB survey respondents does not affect the policy change vector generated by the mechanism. Because each individual contributes equally to the resultant policy change, we call the version of the mechanism described above "one person, one as-if vote." That feature could be relaxed if it were desired to give the preferences of some individuals, such as poorer people, more weight in the aggregation rule. As long as the weights are not a function of the SWB survey responses, the mechanism remains non-manipulable.

How does the mechanism escape from the dictatorial conclusion of Arrow's impossibility theorem? Strictly speaking, the mechanism falls outside the scope of standard social-choice theorems because it does not generate a social preference ordering. However, the mechanism could be interpreted as generating a *local* social preference ordering, where the policy change vector is the gradient of the local social indifference curve. Arrow's theorem extends to this case of local social preferences (Inada, 1964). However, the conclusion of Arrow's theorem does not hold for the as-if voting mechanism because the mechanism violates the theorem's independence of irrelevant alternatives (IIA) assumption. Figure 3 illustrates an example. Applied to individuals' preferences only for policy dimension 1 (Figure 3b), the mechanism prescribes an increase from $\boldsymbol{p_0}$ to $\boldsymbol{p_1}$. Incorporating preferences over dimension 2 (Figure 3a), however, the mechanism prescribes a *decrease* in policy 1. IIA requires that the social preference between $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$ does not depend on preferences over dimension 2, but here, the social indifference curves switch direction. In our view, the violation of IIA is not objectionable because it results from the mechanism taking into account more information about intensity of preference as the number of policy dimensions increases.

The fifth property satisfied by the mechanism is that it only uses information about preferences local to the status quo. Doing so contrasts with other approaches to social aggregation. For example, consider Fleurbaey, Schokkaert, and Decancq's (2009) income-equivalence approach in the case of two goods, health and income. Fleurbaey et al. propose to use (possibly non-linear) regressions of SWB on health and income to estimate individuals' indifference curves. Figure 4 illustrates two individuals' indifference curves that go through their current consumption bundles. Fleurbaey et al. then propose to fix some "reference expansion path"—in their example, perfect health—

and calculate the dollar value of an individual's consumption bundle as the amount of income she would need in order to be indifferent to a consumption bundle on the reference expansion path. These money-metric utilities can then be aggregated using standard methods in public economics. Relative to the income-equivalence approach, the as-if voting mechanism has the advantage (in addition to non-manipulability) that it does not rely on estimating preferences over counterfactual consumption bundles that may be far from the status quo.

III. Discussion

We now return to the choice of the exogenous distance-metric parameters, $(m_1, m_2,, \ldots, m_P)$. Note that all else equal, the mechanism with a larger m_j will generate a policy change that is larger in magnitude in dimension j (and smaller in the other dimensions). The mechanism is non-manipulable regardless of these parameters—as long as citizens believe that the government will actually carry out the policy change dictated by the mechanism. Therefore, in principle, it is clear how to determine the m_j's: in order not to jeopardize non-manipulability, they should be set so as to accurately reflect the political will to move one policy relative to another.

In practice, however, the degrees of freedom afforded in the choice of the m_j's may be the mechanism's biggest limitation. While we have emphasized that the mechanism avoids manipulation by citizens, the government's ability to manipulate the mechanism's output by choosing the m_j's may warrant great concern (Frey and Stutzer, 2012). To minimize this risk, it is crucial that the agency that determines the m_j's be independent and non-political.

While we only analyze the static performance of the mechanism, in practice it would be applied iteratively, so understanding its dynamic properties would be a key extension. When applied dynamically, the formal approximate non-manipulability result will break down because each group could use its current SWB reports to influence the dynamic policy adjustment path. As with the second-order manipulation incentive in the static context, however, we suspect that this additional incentive to manipulate would have a small effect in practice because a great deal of information about other groups' preferences is needed in order to successfully implement an advantageous manipulation.

Assuming everyone truthfully responds to the SWB survey, the mechanism is equivalent to a hill-climbing algorithm if all individuals have the same preferences. If the shared preferences are convex and have a global optimum, then the mechanism should typically converge to it as long as the step size is sufficiently small. (If the shared preferences are non-convex, then the mechanism could converge toward a local optimum that is not a global optimum.) If individuals differ in their preferences, we conjecture that the iterated version of the mechanism might in some circumstances have a limit cycle and not converge at all. We view non-convergence of the iterated version of our mechanism as an undesirable property similar in spirit to intransitivity of a global preference ordering; indeed, the name "as-if voting" serves as a reminder that the mechanism does not entirely escape the weaknesses of actual voting as a social choice mechanism (such as intransitivity) that we seem willing to tolerate in democracies. However, loops on those policy issues where citizens are nearly evenly divided would be small, and would not prevent the mechanism from implementing substantial changes in other dimensions in which there is widespread agreement.

Three other possible extensions of our approach deserve mention. First, our analysis does not allow for discrete policy choices. Indeed, it hinges on taking a local approximation of preferences and ensuring that the policy adjustment is sufficiently small. When policy choices are discrete at a point in time, it may sometimes be possible to convexify policy either by the fraction of time at each discrete policy level, or by the probability of each discrete policy level.

Second, we have assumed that the policy effects $\boldsymbol{\nabla r_g}$ are known, but in reality they would be estimated with uncertainty. An extension to this case would be valuable.

Third, we have assumed that policies can be adjusted independently. When policies are instead linearly dependent—for example, if lower taxes necessitate lower spending so that there is effectively only one tax-and-spend dimension—the matrix $\boldsymbol{M}$ would be positive definite instead of diagonal. After diagonalizing it by an orthonormal transformation, the analysis can proceed as above, with the same calculations and geometric intuitions.

We view this paper primarily as proposing a new approach to aggregating SWB that merits further development. Nonetheless, there are several immediately actionable implications for empirical researchers that would facilitate exploring realistic applications of the mechanism. First, researchers should report at what quantile of SWB respondents the estimated effect of a policy is zero; the mechanism as applied only to that policy dimension would dictate adjustment proportional to the implied vote margin. Second, when studying the effect of more than one policy, researchers should report the marginal rate of substitution between policies within each sociodemographic group. Third, researchers should report results broken down by a standard set of sociodemographic cells (e.g., age, sex, income, marital status, and race). That way, even if different papers estimate the effects of different policies, these effects could be combined to calculate an ε-vote for each group, with groups defined by the sociodemographic cells. Fourth, in order to facilitate such comparisons across studies that use related but different SWB-based utility proxies, it would be helpful to report effect sizes in comparable units (e.g., the cross-sectional standard deviation of the proxy).

References

Adler, Matthew D. 2012. "Happiness Surveys and Public Policy: What's the Use?" University of Pennsylvania Law School Research Paper 12–36, http://ssrn.com/abstract=2076539.

Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Nichole Szembrot. 2012. "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference." NBER Working Paper No. 18374.

Diener, Ed. 2006. "Guidelines for National Indicators of Subjective Well-Being and Ill-Being." *Applied Research in Quality of Life*, 1: 151–157.

Feldstein, Martin. 1976. "On the theory of tax reform." *Journal of Public Economics*, 6, 77-104.

Fleurbaey, Marc, Erik Schokkaert, and Koen Decancq. 2009. "What good is happiness?" CORE Discussion Paper 17, March.

Frey, Bruno, and Alois Stutzer. 2012. "The use of happiness research for public policy." *Social Choice and Welfare*, 38(4): 659–674.

Inada, Ken-ichi. 1964. "On the Economic Welfare Function." *Econometrica*, 32(3): 316–338.

Kahneman, Daniel, and Alan B. Krueger. 2006. "Developments in the Measurement of Subjective Well-Being." *Journal of Economic Perspectives*, 20(1): 3–24.

Layard, Richard. 2005. *Happiness: Lessons from a New Science*. The Penguin Press.

Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. www.stiglitz-sen-fitoussi.fr

Figure 1

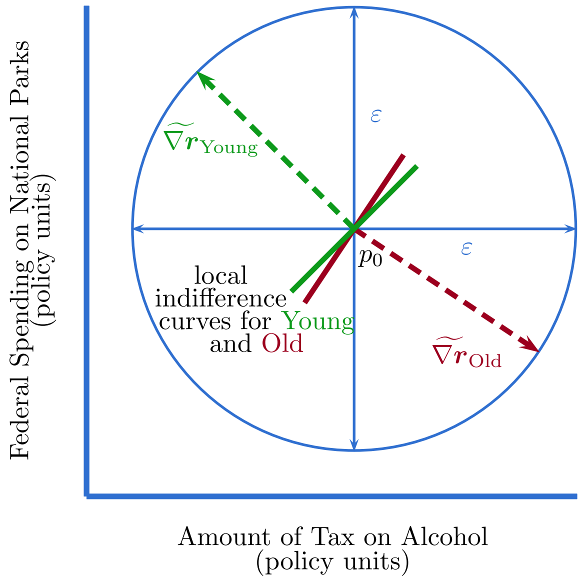

Figure 2

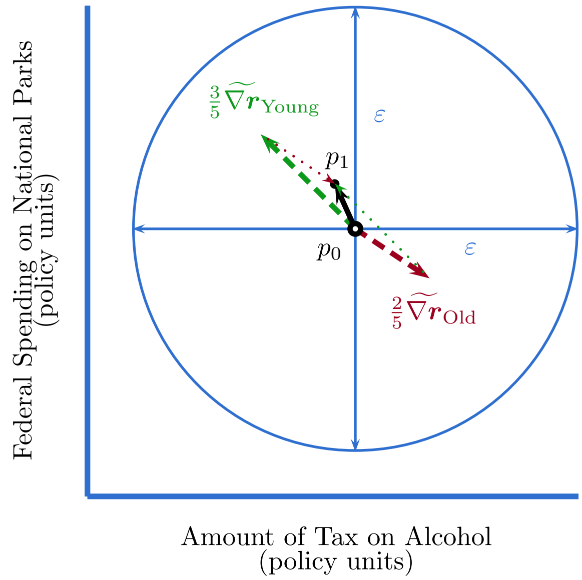

Figure 3a: Policy Change Generated by the Mechanism in 2 Dimensions

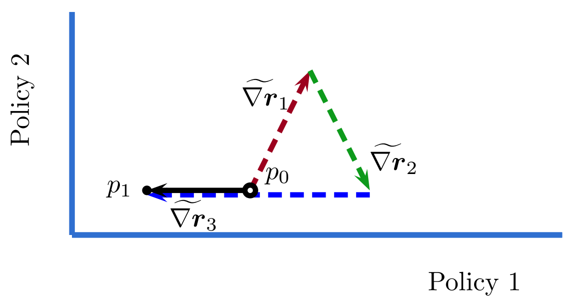

Figure 3b: Policy Change Generated by the Mechanism in 1 Dimension

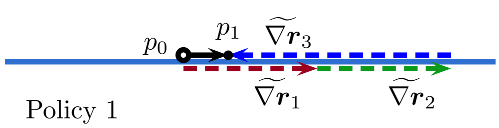

Figure 4