A fair pivotal mechanism for nonpecuniary public goods

Marcus Pivato, Trent University

November 4, 2011

Abstract

The Clarke pivotal mechanism is inappropriate for nonpecuniary public goods, because the assumption of quasilinear utility is invalid, and because the mechanism gives disproportionate influence to wealthier voters. But by introducing a 'stochastic' Clarke tax, we can convert any separable utility function into a quasilinear one. Also, by stratifying a large population by wealth, and applying different 'weights' to the votes from different wealth-strata, we can ensure that the mechanism is *fair* in the sense that the voters in different strata all have equal influence (on average) over the outcome. These weights can be fine-tuned to their optimal values over time, by using the rich dataset generated by a series of large-population referenda. The result is a fair, strategy-proof implementation of weighted utilitarian social choice over nonpecuniary public goods.

Keywords. pivotal mechanism; strategy-proof implementation; nonpecuniary public good; utilitarian; inequality

1 Introduction

Let $\mathcal{A}$ be a menu of social alternatives, which involve the provision of pure public goods (i.e. nonrivalrous and nonexcludable). Let $\mathcal{I}$ be a set of voters. For each $i \in \mathcal{I}$, let u_i be the cardinal utility function of voter i over the alternatives in $\mathcal{A}$, and suppose that i's joint utility over $\mathcal{A}$ and money is *quasilinear*. Thus, if alternative a is chosen and voter i pays a tax t_i, then her utility will be $u_i(a) - c_i\, t_i$, where c_i is the (constant) marginal utility of money for voter i. The social planner wishes to find the element of $\mathcal{A}$ which maximizes aggregate utility, but does not know the true values of the utility functions u_i.

One solution to this problem is the Clarke (1971) *pivotal mechanism.* Each voter i announces a monetary value or *bid* $v_i(a)$ for each alternative a in $\mathcal{A}$ (so that $v_i(a) - v_i(b)$ measures how much i prefers a over b). The social planner then chooses the alternative with the highest aggregate bid, and levies a 'Clarke tax' against any 'pivotal' voters. The Clarke tax is structured such that it is a dominant strategy[1] for voter i to set

[1]That is: a strategy which is utility-maximizing for i, regardless of the actions of the other players.

$v_i(a) = u_i(a)/c_i$ for each a in $\mathcal{A}$.[2] If every voter deploys her dominant strategy, then the mechanism selects the alternative a in $\mathcal{A}$ which maximizes the weighted utilitarian sum

$$\sum_{i \in \mathcal{I}} \frac{u_i(a)}{c_i}. \tag{1}$$

In other words, the pivotal mechanism is a *strategy-proof* implementation of the social choice rule defined by maximizing (1): no voter ever has any incentive to strategically misrepresent her utility function.[3]

This mechanism is ideal in a *purely pecuniary* decision problem, where the voters have purely financial interests about the public good, so that 'utility' can be measured in dollars. For example, the voters might be the stakeholders (i.e. employees, customers, shareholders, and creditors) of various firms in an industrial district, and the elements of $\mathcal{A}$ might be various proposals to build or improve public roads through that district. In this case, $u_i(a)$ is the extra revenue which stakeholder i expects to personally receive (minus the extra taxes she expects to pay) if proposal a is implemented. Thus, $c_i = 1$ for all i in $\mathcal{I}$, and the mechanism selects the outcome which maximizes the aggregate financial gain for the voters.

However, many public goods are *purely nonpecuniary*: they affect the subjective well-being of the voters, rather than their income. For example, suppose a municipal government must decide how to divide a fixed[4] budget between sanitation and waste disposal, public cultural events and festivals, the preservation of historic edifices, and the construction and maintenance of public buildings, monuments, plazas, parks, playgrounds and recreational facilities. Or suppose a federal government must divide a fixed budget between national parks and wilderness reserves, public radio and television stations, academic research, and public health. Of course, some voters may have some pecuniary interest in some of these decisions. But for most voters, such public goods are relevant mainly in how they affect quality of life.[5] Many other political issues have little pecuniary relevance; they are mainly about the conflict between the values or ethical sensibilities of different voters.

In all these examples, it seems desirable to choose the social alternative which will maximize aggregate utility. But for nonpecuniary decisions, the pivotal mechanism has two obvious problems. First, the assumption of quasilinear utility is not realistic; it is more realistic to suppose the marginal utility of money is *declining* for each voter (e.g. due to satiation). This leads to the second problem: the Clarke mechanism

[2] See e.g. Proposition 23.C.4 of Mas-Colell et al. (1995) or Lemma 8.1 of Moulin (1988).

[3] The terms 'strategy-proof', 'dominant-strategy incentive-compatible', 'dominant-strategy truth-revealing' and 'demand-revealing' are all used interchangeably in the literature.

[4] The budget must be fixed, because otherwise the decision would also have a pecuniary component.

[5] Of course, any potential change to a voter's quality of life can be given a pecuniary value in terms of her 'willingness to pay' (WTP). But her WTP for the potential change depends on the price and marginal utility of all other goods in her current consumption bundle. In particular, it depends on her current level of wealth. Thus, WTP is fundamentally different from the 'purely pecuniary' effects in the previous paragraph. The meaning of a sum of pecuniary effects is clear: it is just the net impact on aggregate income, measured in dollars. The meaning of a sum of nonpecuniary WTPs is not clear.

seems to be inequitable. The political 'influence' of voter i on the weighted utilitarian sum (1) is proportional to $1/c_i$, and which is (*ceteris paribus*) proportional to her level of wealth. In other words, rich voters generally have more influence than poor voters.

For example, in 2007, 10% of Americans amassed nearly 50% of all income earned in the United States, after having averaged over 45% during the previous decade (Atkinson et al., 2011, Table 1). Thus, if people's bids in the pivotal mechanism are roughly proportional to their income (which seems plausible), then this 10% alone could effectively control the outcome. The pivotal mechanism would devolve into a plutocracy. This is not only unjust; it also undermines the democratic legitimacy of the mechanism, and makes it unlikely that it will ever be adopted by any democratic society.[6]

The first problem is relatively easy to resolve: instead of a dollars, we must levy the Clarke tax in some units which are guaranteed to be linear in cardinal utility. If we assume that each voter has a von Neumann-Morgenstern utility function, this can be done using a suitably constructed lottery, as we explain below.

The second problem is more difficult. Although it seems intuitively obvious that rich voters have more influence over the pivotal mechanism than poor voters, it is difficult to make this intuition precise without making strong (and questionable) assumptions about interpersonal comparisons of cardinal utility. And even if such interpersonal comparisons were meaningful in theory (so that we could precisely quantify the 'unfairness' of the mechanism), it is not clear how this unfairness could be rectified in practice. There is no known way to obtain from each voter i the true value of c_i on some interpersonal cardinal utility scale. At best, we might be able to estimate the ratio c'_i between voter i's marginal utility for money, and the 'intensity' of her political preferences. If $c'_i < c'_j$, then voter i effectively exerts more influence over the pivotal mechanism than voter j (i.e. she will generally bid larger sums of money). But we might have $c'_i < c'_j$ for many reasons; it may be partly because i is richer than j, it may be partly because i is less materialistic or has less expensive tastes than j, and it may be partly because i honestly has stronger political preferences than j. It seems impossible to disentangle these effects.

Indeed, it is necessary to examine carefully what we mean by 'unfair'. It does not seem unfair if voter i exerts more influence than j because i has very strong preferences about public policy, while j is politically apathetic. Nor does it seem unfair if i exerts more influence because she is simply less materialistic than j, and is willing to endure a greater sacrifice of material consumption to achieve her political goals. It only seems unfair if i has more political influence *simply because* she is richer than j, and for no other reason. Thus, our goal should be to isolate this last effect.

To do this, we will stratify the population of voters according to their level of wealth, and examine the statistical distribution of voting behaviour within each wealth

[6]If pivotal voting is proportional to *wealth*, rather than *income*, then the inequality becomes even more extreme. According to Saez and Kopczuk (2004), the wealthiest 1% of Americans alone control more than 20% of all wealth in the U.S. According to some other estimates, the wealthiest 5% of Americans possessed at least 62% of all wealth in the U.S. in 2007, whereas the poorest 80% collectively owned less than 15%; see (Wolff, 2010, Table 2) and (Allegretto, 2011, p.5).

stratum. If the statistical distribution of voting behaviour is the same in Stratum A as it is in Stratum B, then voters in Stratum A exert, *on average*, the same political influence as voters in Stratum B. (Of course, individual voters in Stratum A may exert greater or lesser influence than the Stratum A average, due to factors such as more intense political preferences, or less expensive material tastes.) If we implement some 'wealth-adjusted' version of the pivotal mechanism, such that voters of all wealth strata exert the same influence, on average, then we can say that this mechanism is 'fair' in the sense that it does not give more power to rich voters than poor voters.

To make this intuition precise, we must make several assumptions:

1. The population $\mathcal{I}$ of voters is large enough that we can stratify voters according to wealth, and still have enough voters in each stratum to obtain good statistics.

2. We are not facing a single referendum, but a series of many referenda on different issues. Thus, the statistics acquired from earlier referenda can be used to 'tune' the parameters of the mechanism for later referenda.

3. Voters' political preference intensities are statistically independent of their wealth stratum, and the statistical distribution of preference intensities is unchanging over time. Thus, any statistical difference we observe between the average voting intensity of different wealth strata is evidence of 'unfairness'.

This paper is organized as follows. Section 2 introduces the *nonpecuniary pivotal mechanism*, and shows that, under certain plausible assumptions, it is not only strategy-proof, but converges rapidly to a mechanism which is 'fair' in the sense that all wealth strata have roughly the same influence. Section 3 discusses an application to taxation and redistribution. Appendix A contains all proofs. Appendix B is an alphabetized index of notation.

2 The nonpecuniary pivotal mechanism

Suppose $\mathcal{I} = \mathcal{I}_1 \sqcup \mathcal{I}_2 \sqcup \cdots \sqcup \mathcal{I}_N$, where, for each n in $[1 \ldots N]$, all voters in stratum $\mathcal{I}_n$ have roughly the same net wealth.[7] (For example, we might set $N := 100$, and define $\mathcal{I}_n$ to be the nth percentile-interval of the wealth distribution.) For all n in $[1 \ldots N]$, let $\varphi_n > 0$ be a positive 'fee'. (Heuristically, these fees should be chosen so that the average marginal utility of φ_n dollars for voters in stratum $\mathcal{I}_n$ is about the same as the average marginal utility of φ_m dollars for voters in stratum $\mathcal{I}_m$, in a

[7]The precise definition of a voter's 'net wealth' is complicated. For simplicity, we could define it to be the average, over all members of the voter's household (including dependents), of that household member's total financial assets and marketable physical assets, minus liabilities, plus the net present value of that member's projected lifetime earnings from labour income (where we set this value to zero for dependents, and otherwise extrapolate future earnings based on, say, the past 12 months of labour income). The information necessary to compute each voter's net wealth is already available to most governments; it is the information which is used to assess of income tax, capital gains tax, property tax, etc.

sense which will be made precise by condition (F2) below.) We refer to the N-tuple $\boldsymbol{\varphi} := (\varphi_1, \varphi_2, \ldots, \varphi_N)$ as the *fee schedule*.

Imagine a series of referenda, occurring at times $t = 1, 2, 3, \ldots$. Let $\mathcal{A}_t$ be the menu of social alternatives for the referendum occurring at time t. We assume that each voter i in $\mathcal{I}$ is an expected-utility maximizer (i.e. satisfies the von Neumann-Morgenstern axioms). Let $u_i^t : \mathcal{A}_t \longrightarrow \mathbb{R}_+$ be voter i's vNM utility function over $\mathcal{A}_t$, and let $u_i^\$: \mathbb{R} \longrightarrow \mathbb{R}$ be her (possibly nonlinear) vNM utility function for net wealth.[8] Assume without loss of generality that $\min_{a \in \mathcal{A}_t} u_i^t(a) = 0$ (add a constant to u_i^t if necessary, to achieve this). We now suppose that i's joint vNM utility over $\mathcal{A}_t$ and wealth is *separable*; that is, if alternative a is chosen and voter i is left with a net wealth of w dollars, then her utility will be $u_i^t(a) + u_i^\$(w)$. We now come to the first component of our mechanism.

(P1) For all n in $[1 \ldots N]$, the stratum $\mathcal{I}_n$ is randomly split into two equal-sized subgroups, $\mathcal{I}_n^+$ and $\mathcal{I}_n^-$. (Each voter knows her subgroup assignment). Let φ_n^+ be slightly larger than φ_n, and let φ_n^- be slightly smaller than φ_n. (For example, we might set $\varphi_n^+ := 1.001\,\varphi_n$ and $\varphi_n^- := 0.999\,\varphi_n$.)

(P2) For all i in $\mathcal{I}$, and each a in $\mathcal{A}_t$, voter i declares a *value* $v_i^t(a)$ in $[0, 1]$ for alternative a.[9] We require that $\min_{a \in \mathcal{A}_t} v_i^t(a) = 0$.

(P3) Given the data $\mathbf{v} := (v_i^t)_{i \in \mathcal{I}}$, society chooses the alternative a^* in $\mathcal{A}_t$ which maximizes the sum $V(a) := \sum_{i \in \mathcal{I}} v_i^t(a)$.

(P4) Voter i is *pivotal* if there is some other b in $\mathcal{A}_t$ with $V(a^*) - V(b) \leq v_i^t(a^*) - v_i^t(b)$. In this case, define $p_i^t(\mathbf{v}) := \sum_{j \in \mathcal{I} \setminus \{i\}} [v_j^t(b) - v_j^t(a^*)]$.

Note that $0 \leq p_i^t(\mathbf{v}) \leq v_i^t(a^*) - v_i^t(b) \leq 1$.

(P5) For all n in $[1 \ldots N]$, any pivotal voter i in subgroup $\mathcal{I}_n^\pm$ now faces a gamble: with probability $p_i^t(\mathbf{v})$, she pays a fee of $\varphi_n^\pm$ dollars, while with probability $1 - p_i^t(\mathbf{v})$, she pays nothing. We refer to this gamble as a *stochastic Clarke tax*.

To understand this mechanism, let i be in $\mathcal{I}_n^+$, and suppose voter i's net wealth at time t is w_i^t. If $c_i^t := u_i^\$(w_i^t) - u_i^\$(w_i^t - \varphi_n^+)$, then the expected utility cost imposed upon i by the stochastic Clarke tax is $c_i^t \cdot p_i^t(\mathbf{v})$, which is a *linear* function of $p_i^t(\mathbf{v})$. Suppose $u_i^t(a) \leq c_i^t$ for all a in $\mathcal{A}_t$. Then, by a simple modification of the standard analysis of the pivotal mechanism, it is easy to show that voter i's dominant strategy is to set $v_i^t(a) = u_i^t(a)/c_i^t$ for all a in $\mathcal{A}_t$. If all voters deploy their dominant strategies, then the alternative chosen in (P3) will be the alternative in $\mathcal{A}_t$ which maximizes the weighted utilitarian sum in Eq.(1).

[8] Here, $\mathbb{R}$ is the set of real numbers, and $\mathbb{R}_+$ is the set of nonnegative real numbers.

[9] If the mechanism is working properly, then the function v_i^t should be a scalar multiple of u_i^t.

However, if $u_i^t(a) > c_i^t$ for some a in $\mathcal{A}_t$, then voter i's dominant strategy is to set $v_i^t(a) = 1$; in this case, we say i *hits the ceiling*. If enough voters hit the ceiling, then the outcome of (P3) may no longer maximize the utilitarian sum (1).

For all i in $\mathcal{I}$, let $V_i^t := \max_{a\in\mathcal{A}_t} v_i^t(a)$; then V_i^t measures the 'influence' of voter i over the outcome of referendum t. We define

$$I \;:=\; |\mathcal{I}| \quad \text{and} \quad \overline{V}^t \;:=\; \frac{1}{I}\sum_{i\in\mathcal{I}} V_i. \tag{2}$$

Thus, $\overline{V}^t$ measures the per capita average influence of any voter during referendum t. For all n in $[1\ldots N]$, we also define

$$I_n \;:=\; |\mathcal{I}_n| \quad \text{and} \quad \overline{V}_n^t \;:=\; \frac{1}{I_n}\sum_{i\in\mathcal{I}_n} V_i^t. \tag{3}$$

Thus, $\overline{V}_n^t$ measures the per capita average influence of a voter in stratum n on the outcome of referendum t. We say that the fee schedule $\boldsymbol{\varphi}$ was *perfectly fair* in referendum t if:

(F1) $V_i^t < 1$ for all voters i in $\mathcal{I}$; and

(F2) $\overline{V}_n^t = \overline{V}^t$ for all n in $[1\ldots N]$.

Condition (F1) says that no voter hit the ceiling; this ensures that every voter's dominant strategy was a scalar multiple of her true utility function. Condition (F2) means that each wealth stratum had, on average, the same influence over the referendum as every other wealth stratum.

Unfortunately, it will not generally be possible to guarantee that $\boldsymbol{\varphi}$ is perfectly fair. Instead, let $\epsilon > 0$ be some small but positive 'error tolerance'. We say that the fee schedule $\boldsymbol{\varphi}$ was *ϵ-fair* in referendum t if

($\mathbf{F1}_\epsilon$) $\#\{i \in \mathcal{I};\;\; V_i^t = 1\} < \epsilon \cdot I$.

($\mathbf{F2}_\epsilon$) $1 - \epsilon < |\overline{V}_n^t/\overline{V}^t| < 1 + \epsilon$ for all n in $[1\ldots N]$.

Condition (F1_ϵ) says that *almost* nobody hit the ceiling, and (F2_ϵ) says all strata had *almost* the same influence. Unfortunately, we cannot even know whether a fee schedule is ϵ-fair until after the referendum has occurred. However, assuming that the statistical distribution of votes is roughly the same from one referendum to the next, we can compute in advance the probability that $\boldsymbol{\varphi}$ will be ϵ-fair in a referendum. Let $0 < p < 1$ and let $\epsilon > 0$. Given a particular statistical distribution of voter behaviour, we say that the fee schedule $\boldsymbol{\varphi}$ is *(p, ϵ)-fair* if it has a probability of at least p to be ϵ-fair in a referendum where the behaviour of the voters is randomly generated according to this distribution.

Our goal now is to design a (p, ϵ)-fair fee schedule, for the empirically observed distribution of voter behaviour. This is the purpose of the second component of our mechanism: to use historical data to 'tune' the fee schedule $\boldsymbol{\varphi}$ so that it will converge to (p, ϵ)-fairness over time. Let $\boldsymbol{\varphi}^t = (\varphi_1^t, \varphi_2^t, \ldots, \varphi_n^t)$ be the fee schedule at time t. Fix a constant $\lambda > 1$. Construct $\boldsymbol{\varphi}^{t+1}$ as follows:

(R1) Let $E_t := \#\{i \in \mathcal{I};\ \ V_i^t = 1\}/I$. If $E_t \geq \epsilon$, then for all n in $[1 \ldots N]$, set $\varphi'_n := \lambda \cdot (E_t/\epsilon) \cdot \varphi_n^t$. Otherwise, if $E_t < \epsilon$, then set $\varphi'_n := \varphi_n^t$ for all n in $[1 \ldots N]$.

(R2) For all n in $[1 \ldots N]$, set $\varphi_n^{t+1} := (\overline{V}_n^t / \overline{V}^t)^{s_n} \cdot \varphi'_n$, where

$$s_n \ := \ \frac{\log(\overline{V}_n^{t,+}) - \log(\overline{V}_n^{t,-})}{\log(\varphi_n^{t,+}) - \log(\varphi_n^{t,-})}, \quad \text{with } \overline{V}_n^{t,+} \ := \ \frac{1}{|\mathcal{I}_n^+|} \sum_{i \in \mathcal{I}_n^+} V_i^t \ \text{ and } \ \overline{V}_n^{t,-} \ := \ \frac{1}{|\mathcal{I}_n^-|} \sum_{i \in \mathcal{I}_n^-} V_i^t.$$

Rule (R1) says that, if too many voters hit the ceiling, then all fees in the schedule should be adjusted upwards in proportion to the number of voters who hit the ceiling. Rule (R2) says we should then further adjust the fee of stratum n up (respectively, down) if the average influence of that stratum was higher (respectively, lower) than the population average. (Heuristically, s_n estimates the per capita average elasticity of disutility with respect to the fee φ_n^t for stratum n.)

We refer to the sequence of referenda described by rules (P1)-(P5) and (R1)-(R2) as the *nonpecuniary pivotal mechanism*. We shall now see that, for any $\epsilon > 0$ and $0 < p < 1$, if the strata are large enough and the statistical distribution of voter preferences satisfies certain regularity conditions, then this mechanism will rapidly converge to a (p, ϵ)-fair fee schedule.

Formally, let $\mathbb{N} := \{0, 1, 2, 3, \ldots\}$. For all i in $\mathcal{I}$ and all t in $\mathbb{N}$, let $U_i^t := \max_{a \in \mathcal{A}_t} u_i^t(a)$, where $\mathcal{A}_t$ and u_t^i are as defined prior to (P1). Thus U_i^t measures the 'intensity' of voter i's preferences on referendum t. Here is our first assumption:

(U) For all t in $\mathbb{N}$, there is a probability distribution μ_t on $\mathbb{R}_+$ such that U_i^t is a μ_t-random variable, for all i in $\mathcal{I}$. Furthermore, $\{U_i^t;\ \ i \in \mathcal{I} \text{ and } t \in \mathbb{N}\}$ is a set of independent random variables.

Assumption (U) says all strata have the same statistical distribution of political preference intensities on any particular referendum,[10] and there is no correlation of preference intensities between different referenda or between different voters.

For all i in $\mathcal{I}$, and all $\varphi > 0$, let $C_i^t(\varphi) := u_i^\$(w_i^t) - u_i^\$(w_i^t - \varphi)$ be the 'cost' (in utility) of a fee of size φ for voter i at time t. In particular, if voter i is in stratum $\mathcal{I}_n$, and deploys her dominant strategy for the mechanism (P1)-(P5), then $v_i^t(a) = \max\{1,\ u_i^t(a)/C_i^t(\varphi_i^t\}$ for every alternative a in $\mathcal{A}_t$. Thus,

$$\text{for all } n \text{ in } [1 \ldots N] \text{ and all } i \text{ in } \mathcal{I}_n, \qquad V_i^t \ = \ \min\left\{1,\ \frac{U_i^t}{C_i^t(\varphi_n^t)}\right\}. \qquad (4)$$

Let $\mathcal{C}$ be the space of all nondecreasing functions from $\mathbb{R}_+$ to itself. Here is our second assumption:

[10] Note that we do *not* assume that all strata have the same distribution of political *preferences*, but only the same distribution of preference *intensities*. In general, different strata will have different preference distributions. (Indeed if all strata had the same preference distribution, then 'fairness' would be unnecessary: we could simply allow one stratum to entirely control the mechanism, and obtain the same outcome).

(C) For all n in $[1 \ldots N]$, there is a probability distribution ρ_n on $\mathcal{C}$, such that:

> **(C1)** For every t in $\mathbb{N}$, the set $\{C_i^t\}_{i \in \mathcal{I}_n}$ is a set of independent, ρ_n-random elements of $\mathcal{C}$.
>
> **(C2)** For every t in $\mathbb{N}$, and every i in $\mathcal{I}_n$, the random variables U_i^t and C_i^t are independent.
>
> **(C3)** For any $\epsilon > 0$, there is some constant $\overline{\varphi}_n^\epsilon > 0$ with the following property. For all t in $\mathbb{N}$, if U_t is a μ_t-random variable and C_n is an independent, ρ_n-random function, then $\text{Prob}\left[U_t \geq C_n(\overline{\varphi}_n^\epsilon)\right] < \epsilon$.
>
> **(C4)** There is a decreasing, continuously twice-differentiable function $V_n : \mathbb{R}_+ \longrightarrow [0, 1]$ such that $V(0) = 1$ and $\lim_{\varphi \to \infty} V(\varphi) = 0$, and such that for any $\varphi \geq 0$ and any t in $\mathbb{N}$, $V_n(\varphi)$ is the expected value of the random variable $\min\{1,\ U_t/C_n(\varphi)\}$, where U_t and C_n are as in (C3).

In words: $V_n(\varphi)$ the *expected influence* which a random voter in stratum n would have on the outcome of referendum t, if $\varphi_n^t = \varphi$. Assumption (C4) says that this function is well-behaved, and the same for all referenda.[11] Assumption (C3) says that it is highly improbable that a voter's political preference intensity will be huge, when measured in monetary terms. Assumptions (C1) and (C2) say there is no correlation between voters, or across time periods.

For example, suppose $\epsilon = 0.01$ in (C3); then $\overline{\varphi}_n^\epsilon$ is the minimum fee required such that less than 1% of the voters in stratum $\mathcal{I}_n$ would be willing to pay more than $\overline{\varphi}_n^\epsilon$ dollars to change the outcome in a typical referendum. For a typical middle-class stratum, we would expect $\overline{\varphi}_n^{0.01}$ to be perhaps a few thousand dollars.

Our first result says that, if the set $\mathcal{I}$ of voters is large enough, and we divide it into N equal-sized subgroups $\mathcal{I}_1, \ldots, \mathcal{I}_N$, then there exists a (p, ϵ)-fair fee schedule.

Proposition 1 *Assume* (U) *and* (C), *and let* $0 < V^* < 1$ *be any constant.*

> **(a)** *For all n in $[1 \ldots N]$, there exists a unique φ_n^* in $\mathbb{R}_+$ such that $V_n(\varphi_n^*) = V^*$.*

Now let $0 < \epsilon, p < 1$, and suppose that

$$\text{(5.1)} \quad I \;\geq\; \frac{8\sqrt{N^3+1}}{\epsilon\, V^* \sqrt{1-p}}, \qquad \textit{and} \qquad \text{(5.2)} \quad I_1 = I_2 = \cdots = I_N = \frac{I}{N}. \qquad (5)$$

> **(b)** *There is a constant $K > 0$ such that, for any t in $\mathbb{N}$, if $|\varphi_n^t - \varphi_n^*| < K\,\epsilon$ for all n in $[1 \ldots N]$, then $\boldsymbol{\varphi}^t$ will satisfy* $(\text{F}2_\epsilon)$ *with probability p or higher.*[12]
>
> **(c)** *If V^* is close enough to zero, then $\boldsymbol{\varphi}^t$ will also satisfy condition* $(\text{F}1_\epsilon)$ *with probability p or higher.*

[11] Note that we do *not* assume that the individual response functions $\{U_i^t/C_i^t\}_{i \in \mathcal{I}_n}$ are well-behaved (or equivalently, that the cost functions $\{C_i^t\}_{i \in \mathcal{I}_n}$ are well-behaved). In principle, these functions could be nondifferentiable, or even discontinuous. We only require their *average* to be well-behaved.

[12] K is proportional to the slopes of the functions $V_1, \ldots, V_N$ near the values $\varphi_1^*, \ldots, \varphi_N^*$.

For example, if $N = 10$, $\epsilon = 0.01$, $p = 0.99$, and $V^* = 0.5$, then it suffices for $I \geq 507,000$ to satisfy inequality (5.1); this is the population of a medium-sized city. If we make V^* small enough, and define $\boldsymbol{\varphi}^* := (\varphi_1^*, \ldots, \varphi_N^*)$ as in Proposition 1(a), then Proposition 1(b,c) guarantees that the fee schedule $\boldsymbol{\varphi}^*$ will be $(0.99, 0.01)$-fair; thus, there would be no need for the calibration rules (R1) and (R2). But to know what value of V^* is 'small enough', and to compute the corresponding values of $\varphi_1^*, \ldots, \varphi_N^*$, we must know the exact structure of the probability distributions $\{\mu_t\}_{t=1}^{\infty}$ and $\rho_1, \ldots, \rho_N$ in assumptions (U) and (C). Since we don't know their exact structure, the calibration rules (R1) and (R2) are still necessary.

The effect of rules (R1) and (R2) can be heuristically understood as follows. Iterating rule (R2) effectively causes the values of $(\varphi_1^t, \ldots, \varphi_N^t)$ to converge to the values $(\varphi_1^*, \ldots, \varphi_N^*)$ described in Proposition 1(a) (for some unspecified value of V^*). Thus, after enough iterations of (R2), the conditions of Proposition 1(b) are satisfied, so that $\boldsymbol{\varphi}^t$ satisfies (F2$_\epsilon$) with probability p. Meanwhile, iterating rule (R1) effectively decreases the value of V^* which is being targeted (by uniformly increasing all of $\varphi_1^t, \ldots, \varphi_N^t$). Thus, after enough iterations of (R1), the hypothesis of Proposition 1(c) is also satisfied, so that $\boldsymbol{\varphi}^t$ also satisfies (F1$_\epsilon$) with probability p. At this point, $\boldsymbol{\varphi}^t$ is (p, ϵ)-fair. The next two propositions provide a more precise description of this calibration process. First, we need one more technicality. For any $\epsilon > 0$, define

$$L(\epsilon) \quad := \quad \frac{\max\{\log(\overline{\varphi}_n^\epsilon / \varphi_n^0)\}_{n=1}^N}{\log(\lambda)}, \tag{6}$$

where $(\varphi_1^0, \ldots, \varphi_N^0)$ is the initial fee schedule at time 0, and $\overline{\varphi}_1^\epsilon, \ldots, \overline{\varphi}_N^\epsilon$ are as in assumption (C3), and where λ is as in rule (R1). The behaviour of the function L depends on the shape of the distributions $\{\mu_t\}_{t=1}^{\infty}$ and $\rho_1, \ldots, \rho_N$ in assumptions (U) and (C). For our purposes, the important thing is that typically, $L(\epsilon) \to \infty$ relatively slowly as $\epsilon \searrow 0$. For example, under reasonable hypotheses, we have $L(\epsilon) = \mathcal{O}(\log(1/\epsilon))$ as $\epsilon \searrow 0$.[13] Furthermore, $L(\epsilon)$ will be small if our initial guess φ_n^0 was not too far from $\overline{\varphi}_n^\epsilon$. For example, suppose $\lambda = 1.26 \approx \sqrt[3]{2}$; then we will have $L(\epsilon) \leq 6$ as long as $\varphi_n^0 \geq \overline{\varphi}_n^\epsilon / 4$ for all $n \in [1 \ldots N]$.

Proposition 2 *Let $0 < \epsilon, p < 1$, and suppose $I > 1/\epsilon\sqrt{1-p}$. If only the calibration rule* (R1) *is applied during each referendum, then there will almost surely come a time T_p^ϵ such that, for all $t > T_p^\epsilon$, condition* (F1$_\epsilon$) *will be satisfied with probability p or higher. The expected value of the random variable T_p^ϵ is at most*

$$\frac{1}{1-p}\, L\left(\epsilon - \frac{1}{I\sqrt{1-p}}\right). \tag{7}$$

[13]That is: there is some constant $k > 0$ such that $0 \leq L(\epsilon) < k \log(1/\epsilon)$ for all sufficiently small $\epsilon > 0$. For example, suppose that, for all t in $\mathbb{N}$, the μ_t-random variable U_t has mean $\overline{U}$ and variance σ_1^2, and that, for any n in $[1 \ldots N]$ and $\varphi > 0$, the (independent) ρ_n-random variable $C_n(\varphi)$ has mean $\overline{C}_n(\varphi)$ and variance σ_2^2. Let $\sigma^2 := \sigma_1^2 + \sigma_2^2$. Then Chebyshev's inequality yields $\overline{\varphi}_n^\epsilon \leq \overline{C}_n^{-1}\left(\overline{U} + \sigma/\sqrt{\epsilon}\right)$. Thus, if there is some $s \in \mathbb{R}$ such that $\overline{C}_n(\varphi) = \mathcal{O}(\varphi^s)$ as $\varphi \to \infty$, then $\overline{\varphi}_n^\epsilon = \mathcal{O}(\epsilon^{-1/2s})$ as $\epsilon \searrow 0$. Thus, $L(\epsilon) = \mathcal{O}(\log(1/\epsilon))$ as $\epsilon \searrow 0$.

For example, let $\epsilon := 0.01$, and suppose we want to ensure that condition ($F1_\epsilon$) is violated in less than 4% of all referenda. If $I \geq 10\,000$ and $L(0.0095) \leq 6$, then 150 iterations of rule (R1) will usually suffice to reach this goal. (To see this, set $p := 0.96$ in Proposition 2.)

If $p \approx 1$, and $t > T_p^\epsilon$, then Proposition 2 says that condition ($F1_\epsilon$) will be satisfied with very high probability, so that rule (R1) will almost never be invoked after time T_p^ϵ. Thus, after time T_p^ϵ, we can focus on the dynamics of rule (R2) only. We will now show that (R2) causes the fee schedule $\boldsymbol{\varphi}^t$ to converge to the fee schedule $\boldsymbol{\varphi}^*$ described in Proposition 1(b).

Consider the random variable $V_n^t := \min\{1,\ U_t/C_n(\varphi_n^t)\}$ (where U_t and C_n are as in (C3)). Since $0 \leq V_n^t \leq 1$, the variance of V_n^t is less than 1. If every voter deploys her dominant strategy, then Eq.(4) and assumptions (U), (C1) and (C2) imply that the random variables $\{V_i^t\}_{i\in\mathcal{I}_n}$ are independent and identically distributed to V_n^t. Assumption (C4) says $V_n(\varphi_n^t)$ is the expected value of V_n^t, while definition (3) says $\overline{V}_n^t$ is the average of $\{V_i^t\}_{i\in\mathcal{I}_n}$. Thus, the Central Limit Theorem says $\overline{V}_n^t = V_n(\varphi_n^t) + \gamma_n^t$, where γ_n^t is some random variable with mean zero, variance less than $1/I_n$, and an 'almost Gaussian' distribution.

In practice, I_n will be very large, so that $|\gamma_n^t|$ will be extremely small, with very high probability. For example, if $N = 10$ and each $\mathcal{I}_n$ represents one decile-interval of the wealth distribution of a polity with 10 million voters, then $I_n = 10^6$. Then we will have $|\gamma_n^t| < 0.004$, with probability greater than 99.99%. Thus, $\overline{V}_n^t \approx V_n^t(\varphi_n^t)$. For simplicity, in the next proposition we will assume that this approximation is exact.

Proposition 3 *Suppose that:*

(S1) $\overline{V}_n^t = V_n(\varphi_n^t)$ *all t in $\mathbb{N}$ and all n in $[1 \ldots N]$.*

(S2) *There is some V^* such that $\overline{V}^t = V^*$ for all t in $\mathbb{N}$.*

Suppose that only rule (R2) *is applied during each referendum.*

- **(a)** *Let $(\varphi_1^*, \ldots, \varphi_N^*)$ be as defined in Proposition 1*(a). *For any $\delta > 0$, there exists $T_0(\delta) > 0$ such that $|\varphi_n^t - \varphi_n^*| < \delta$ for all $t \geq T_0(\delta)$ and all n in $[1 \ldots N]$. Furthermore, $T_0(\delta) = \mathcal{O}\left(\sqrt{\log(1/\delta)}\right)$.*

- **(b)** *For any $\epsilon > 0$, there exists $T_1(\epsilon) > 0$ such that* ($F2_\epsilon$) *is satisfied for all $t \geq T_1(\epsilon)$. Furthermore, $T_1(\epsilon) = \mathcal{O}\left(\sqrt{\log(1/\epsilon)}\right)$.*

- **(c)** *If the functions $V_1, \ldots, V_N$ in* (C4) *are isolestic*[14]*, then $T_0(\delta) = T_1(\epsilon) = 1$.*

Of course, Proposition 3 does not exactly describe the behaviour of rule (R2), because assumptions (S1) and (S2) are both approximations. But by setting $\delta := K\,\epsilon^2$ and combining Propositions 1(b) and 3(a), we obtain the following heuristic statement:

[14]That is: if $V_1, \ldots, V_N$ have constant elasticity —e.g. $V_n(\varphi) = (1 + \varphi/c_n)^{s_n}$ for some $s_n < 0$ and $c_n > 0$.

Suppose that $\overline{V}_n^t \approx V_n(\varphi_n^t)$ *all* t *in* $\mathbb{N}$ *and all* n *in* $[1 \ldots N]$*, and there is some* V^* *such that* $\overline{V}^t \approx V^*$ *for all* t *in* $\mathbb{N}$*. If* $I_1, \ldots, I_N$ *and* I *satisfy the conditions* (5)*, and only rule* (R2) *is applied during each referendum, then for any* $\epsilon > 0$*, there exists* $T_0(\epsilon) > 0$ *such that* $\boldsymbol{\varphi}^t$ *will satisfy* (F2$_\epsilon$) *with probability* p *or higher. Furthermore,* $T_0(\epsilon) = \mathcal{O}\left(\sqrt{\log(1/\epsilon)}\right)$.

The convergence described in Propositions 2 and 3 is extremely fast. Obviously, these results are idealizations. First of all, both rules (R1) and (R2) will be applied during each iteration, complicating the analysis (although (R1) will be invoked less and less often). Second, assumptions (S1) and (S2) are both approximations. Nevertheless, Propositions 2 and 3 strongly suggest that, under assumptions (U) and (C), the nonpecuniary pivotal mechanism will converge rapidly to a (p, ϵ)-fair fee distribution.

3 Taxation and redistribution

So far, we have assumed the government has a fixed, exogenous revenue stream; the nonpecuniary pivotal mechanism is used to decide how this revenue should be allocated towards various nonpecuniary public goods. But there is also the question of how to structure the system of taxes which generate this revenue stream in the first place. More generally, there is the question of how to structure a system of transfers (i.e. taxes and benefits), not only to generate revenue, but also to redistribute wealth for the sake of altruism or social justice. For simplicity, we will refer to such a system as a *tax schedule* (even though the effective 'tax' on some voters may be negative).

Assuming a fixed revenue target of R dollars per year, the choice of tax schedule is a 'zero-sum' policy problem: every tax schedule has the same net pecuniary impact on society (namely, it extracts R dollars), but different schedules will have different pecuniary costs/benefits for different people. The classic pivotal mechanism is indecisive on such zero-sum questions: every tax schedule will get the same level of aggregate support from voters (namely $-R$), and none will appear to be 'socially preferred' to any other.

A tax schedule affects each voter in two ways. First, it has a private, purely pecuniary impact on the voter: the net financial cost/benefit which she expects to receive from the schedule, both due to direct monetary transfers and due to the indirect economic distortions generated by the tax. But the tax schedule can also generates nonpecuniary public goods, by achieving certain goals of altruism and social justice, and perhaps through the indirect social benefits of greater material equality.[15] We want to use the nonpecuniary pivotal mechanism from §2 to decide the structure of this nonpecuniary public good. The problem is that each voter's private, pecuniary interest makes it difficult to obtain accurate information about her nonpecuniary preferences over the social justice aspect of the tax schedule.

[15] For example, Wilkinson and Pickett (2010) have argued that there is robust empirical evidence that lower material inequality in a society is correlated with greater levels of trust and civic engagement, and lower levels of violent crime, emotional stress, mental illness, and physical disease.

However, at a purely pecuniary level, a tax schedule has two parts, which we will call *personal* and *impersonal.* The *personal* part consists of taxes or benefits targeted at specific, identifiable voters. This includes income tax, dividend tax, capital gains tax, residential property tax, welfare payments, unemployment insurance, state pensions, and vouchers (e.g. to purchase food or education). The *impersonal* part of the schedule consists of taxes or benefits which cannot be tied to a particular voter. This includes corporate income tax, commercial property tax, value-added taxes, excise taxes, import tariffs, and subsidized or state-provided food, housing, education and health care. It also includes any economic distortions caused by any of these taxes and subsidies (e.g. distortions in the labour market due to the income tax schedule).

We will now discuss a way to use the nonpecuniary pivotal mechanism to optimally determine the *personal* part of the tax schedule. First, divide the households into K equally sized, randomly chosen groups —call them $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_K$. These groups are *not* the 'wealth strata' considered in §2; each group should be statistically representative of the entire population. For example, we might divide households into twelve groups, depending on the birth-month of the eldest member of the household. Or we might divide them into ten groups, depending upon the last digit in the social insurance number of the eldest household member. Each of the groups $\mathcal{G}_1, \ldots, \mathcal{G}_K$ must provide exactly $1/K$ of the government's revenue target; however, different groups might provide this revenue through *different* personal tax schedules, as we now describe.

The voters are also divided into K equally sized 'juries' —call them $\mathcal{J}_1, \ldots, \mathcal{J}_K$. The members of jury $\mathcal{J}_k$ are randomly chosen from $\mathcal{I} \setminus \mathcal{G}_k$. Thus, no member of group $\mathcal{G}_k$ can be part of the household of any member of jury $\mathcal{J}_k$, but other than this restriction, group membership and jury membership are statistically independent, and each jury is a statistically representative sample of the whole population. The personal tax schedule for group $\mathcal{G}_k$ will be decided by jury $\mathcal{J}_k$, using the nonpecuniary pivotal mechanism. We impose the following requirements:

- K is reasonably large (e.g $K \geq 10$) so that members of $\mathcal{G}_k$ comprise only a small fraction $(1/K)$ of the friends and family of a typical voter in $\mathcal{J}_k$, and also so that any taxation-induced economic distortions in group $\mathcal{G}_k$ have little direct pecuniary impact on voters in $\mathcal{J}_k$.

- The tax schedules for all K groups are decided simultaneously. Thus, a voter in $\mathcal{J}_k$ has no incentive to manipulate the tax schedule of $\mathcal{G}_k$, in the hope of setting a precedent or receiving some sort of *quid pro quo* for her own group.

- No juror knows which jury she belongs to until after the referendum is over. Thus, a voter in $\mathcal{J}_k$ cannot strategically vote for a tax schedule which helps her friends or hurt her enemies in $\mathcal{G}_k$ (because she doesn't know she is in $\mathcal{J}_k$).

If these conditions are satisfied, then a juror in this mechanism has little or no private pecuniary interest in the personal tax schedule she is voting on; for her, this is an almost purely nonpecuniary public good (involving social justice within an as-yet unspecified group $\mathcal{G}_k$ which is disjoint from her own household). Thus, her dominant strategy is to reveal her true utility function with respect to this public good.[16]

Since the juries $\mathcal{J}_1, \ldots, \mathcal{J}_K$ are all large, statistically representative samples from the same population, we expect that the K personal tax schedules they select will all be virtually identical, *ex post*. The result is a strategy-proof implementation of weighted utilitarian social choice over the personal tax schedules, decided purely on nonpecuniary social justice grounds, without interference from private pecuniary incentives.

This mechanism cannot be applied to the *impersonal* tax schedule, because we cannot isolate the people who will be affected by an impersonal tax from the jurors who will vote on it. A partial solution is to replace impersonal taxes/subsidies with roughly equivalent personal taxes/subsidies whenever possible. For example, taxes on corporate profits could be replaced with (personal) dividend taxes of equal revenue yield. Instead of applying a value-added tax to goods and services, we could provide an income tax deduction for any income which is saved or invested (with an equal-sized tax on any cash which is withdrawn from said savings or investments). Thus, a taxpayer would pay a lower tax on any income she saves or invests, and a higher tax on any income which she spends on consumption; this would act like a 'personal value-added tax'. Instead of subsidizing or publicly providing food, housing, education, and healthcare, we could provide vouchers to citizens to purchase these services privately.

The remaining instruments in the impersonal tax schedule cannot be decided with the pivotal mechanism. These include Pigouvian instruments (to internalize externalities), countercyclical instruments (for macroeconomic stabilization), 'seed money' or tax breaks directed at nascent industries (for industrial policy), and perhaps other instruments designed to correct market distortions introduced by the personal tax schedule. These instruments are best designed by technocrats, not by referenda.

Conclusion

The classic pivotal mechanism is a strategy-proof implementation of weighted utilitarian social choice amongst pecuniary public goods. We have modified this mechanism to obtain a fair, strategy-proof implementation of weighted utilitarian social choice amongst nonpecuniary public goods. But in reality, no public good is purely pecuniary or purely non-pecuniary. The examples in Section 1 all roughly approximate one extreme or the other. But many public goods are not even 'approximately' pure: they generate a substantial amount of both pecuniary and nonpecuniary costs/benefits, for a substantial proportion of voters. These 'hybrid' public goods include: law enforcement, urban zoning laws, most roads and public transportation systems, public education, regulations regarding goods and services, and of course, the composition of the government itself. Neither the classic pivotal mechanism nor the nonpecuniary

[16] Of course, many voters in a particular economic class may vote for tax schedules which favour this class, either because most of their friends and family come from the same class, or out of some general sense of 'class solidarity'. But they cannot expect any *personal* pecuniary gain from this vote.

pivotal mechanism seems appropriate for these questions.

There remain four other unresolved problems. First: all revenue from the Clarke tax must be destroyed, or the mechanism is not strategy-proof. Thus, the pivotal mechanism is not efficient. This problem has been extensively studied, and several more or less satisfactory solutions have been proposed for the classic pivotal mechanism. For example, Green et al. (1976) and Green and Laffont (1979) showed that, under reasonable assumptions, the *per capita* inefficiency introduced by the Clarke tax goes to zero like $1/\sqrt{I}$ as $I \to \infty$ (where I is the number of voters). Gary-Bobo and Jaaidane (2000) and Faltings (2004) suggest that the pivotal mechanism could be applied to randomly selected, statistically representative jury, with Clarke tax revenues being redistributed to non-jurors. (For example, in the multi-jury system of Section 3, any Clarke tax revenues from $\mathcal{J}_k$ could be redistributed to $\mathcal{I} \setminus \mathcal{J}_k$.) Bailey (1997) and Cavallo (2006) suggest strategy-proof tax refund schemes which are revenue-neutral on average. For example, in the large-population, multi-referendum scenario considered in this paper, the simplest solution is perhaps to collect all Clarke tax revenue in a fund, and pay every voter $\overline{t}$ dollars out of this fund every year, where $\overline{t}$ is the annual per capita average Clarke tax over the previous ten or twenty years.

Second, the nonpecuniary pivotal mechanism is even more informationally intensive than the classic pivotal mechanism (especially the version proposed in Section 3). This creates some technological challenges, especially since all votes must remain confidential, so that voters cannot be bribed or intimidated, or coordinate their actions in voting blocs. In particular, the identities of pivotal voters must remain secret, so that they cannot be retroactively rewarded or punished by someone seeking to manipulate the outcome. It is not clear that this is feasible.

Third, we have assumed that each voter's joint utility function over wealth and nonpecuniary public goods is *separable*. But this is false; a large gain or loss of wealth will generally change a voter's preferences over nonpecuniary public goods. However, for relatively small variations of wealth (such as those implied by the fee schedule $\boldsymbol{\varphi}^t$), separability may be an adequate approximation. A more fundamental problem is that the use of a stochastic Clarke tax assumes that the voters are von Neumann-Morgenstern expected utility maximizers —an assumption which is empirically false (Kahneman and Tversky, 2000). In reality, many voters may fail to reliably identify their dominant (i.e. truth-revealing) strategy, due to cognitive distortions. Thus, v_i^t might *not* be a scalar multiple of u_i^t, for many i in $\mathcal{I}$. However, hopefully it will generally be a good enough approximation that step (P3) of the mechanism will still maximize the weighted utilitarian sum (1).

Fourth, we have assumed the nonpecuniary pivotal mechanism operates with a budget of *fixed* size (either when deciding nonpecuniary public goods in Section 2, or when deciding personal tax schedules in Section 3). This budget size must be fixed in advance, because otherwise these decisions would involve an inextricable pecuniary component. But how should society determine the size of this budget? This is a one-dimensional policy problem, over which most voters presumably have single-peaked preferences; thus, the decision could be made by simple majority vote, which would be a strategy-proof implementation of the choice of the median voter. But it is not clear that the median is the welfare-maximizing choice; it only aggregates voters' *ordinal* preferences, not their cardinal utilities. The optimal size of the government remains an open problem.

Appendix A: Proofs

The next lemma is used in the proofs of Propositions 1 and 2.

Lemma A.1 *Let $r := \sqrt{1/(1-p)}$, and suppose $I > r/\epsilon$. Let $\delta := \epsilon - r/I$ (so $0 < \delta < \epsilon$). Define $\overline{\varphi}_1^\delta, \ldots, \overline{\varphi}_N^\delta$ as in assumption* (C3). *If $\varphi_n^t \geq \overline{\varphi}_n^\delta$ for all $n \in [1 \ldots N]$, then the fee schedule $\boldsymbol{\varphi}^t$ will satisfy* (F1$_\epsilon$) *with probability greater than p.*

Proof. Define E_t as in rule (R1); we must show that $\mathrm{Prob}\,[E_t \geq \epsilon] < 1 - p$. For all $n \in [1 \ldots N]$, let U_t and C_n be as defined in (C3), and let $p_n := \mathrm{Prob}\,\big[U_t \geq C_n(\overline{\varphi}_n^\delta)\big]$. Then $p_n < \delta$, by assumption (C3). For all $n \in [1 \ldots N]$, we have

$$\frac{1}{I_n}\#\big\{i \in \mathcal{I}_n\ ;\ U_i^t \geq C_i^t(\overline{\varphi}_n^\delta)\big\} \underset{(*)}{=} \gamma_n + p_n \underset{\text{(C3)}}{<} \gamma_n + \delta. \tag{A1}$$

Here, γ_n is some random variable with mean zero and variance less than $1/I_n$, and $(*)$ is because assumptions (U), (C1) and (C2) together imply that we are averaging a set of I_n independent random variables with mean p_n and variance less than 1. Now,

$$\big\{i \in \mathcal{I}\ ;\ V_i^t = 1\big\} \underset{(\diamond)}{=} \bigcup_{n=1}^{N} \big\{i \in \mathcal{I}_n\ ;\ U_i^t \geq C_i^t(\varphi_n^t)\big\} \underset{(\dagger)}{\subseteq} \bigcup_{n=1}^{N} \big\{i \in \mathcal{I}_n\ ;\ U_i^t \geq C_i^t(\overline{\varphi}_n^\delta)\big\}. \tag{A2}$$

Here, $(\diamond)$ is by Eq.(4), $(\dagger)$ is because $\varphi_n^t \geq \overline{\varphi}_n^\delta$ for all $n \in [1 \ldots N]$, and C_i^t is nondecreasing, for all $i \in \mathcal{I}$. Thus,

$$I\,E_t \;=\; \#\big\{i \in \mathcal{I}\ ;\ V_i^t = 1\big\} \underset{(\dagger)}{\leq} \sum_{n=1}^{N}\#\big\{i \in \mathcal{I}_n\ ;\ U_i^t \geq C_i^t(\overline{\varphi}_n^\delta)\big\} \underset{(*)}{<} \sum_{n=1}^{N} I_n\,(\delta{+}\gamma_n).$$

where $(\dagger)$ is by formula (A2), while $(*)$ is by inequality (A1). It follows that

$$E_t \;<\; \frac{1}{I}\sum_{n=1}^{N} I_n\,(\delta+\gamma_n) \;=\; \frac{1}{I}\left(\sum_{n=1}^{N} I_n\right)\delta + \frac{1}{I}\sum_{n=1}^{N} I_n\,\gamma_n \underset{(\dagger)}{=} \delta+\gamma, \tag{A3}$$

where γ is some random variable with mean zero and variance less than $1/I$, and where $(\dagger)$ is because $(I_1\,\gamma_1), \ldots, (I_N\,\gamma_N)$ are independent random variables (by (U), (C1) and (C2)) with mean zero and variances less than $I_1, \ldots, I_N$ respectively (so their sum $I\,\gamma$ has mean zero and variance less than $I_1 + \cdots + I_N = I$). Thus,

$$\mathrm{Prob}\,[E_t \geq \epsilon] \underset{(*)}{<} \mathrm{Prob}\,[\delta+\gamma \geq \epsilon] \underset{(\diamond)}{=} \mathrm{Prob}\,[\gamma \geq r/I] \underset{(\dagger)}{\leq} \frac{1}{r^2} \underset{(\ddagger)}{=} 1-p,$$

as desired. Here, $(*)$ is by inequality (A3), $(\diamond)$ is by the definition of δ, $(\dagger)$ is by Chebyshev's inequality (because γ has mean 0 and variance $1/I$), and $(\ddagger)$ is by the definition of r. □

Proof of Proposition 1. (a) For all $n \in [1 \ldots N]$, the Intermediate Value Theorem yields a unique φ_n^* such that $V_n(\varphi_n^*) = V^*$, because V_n is continuous and decreasing, by assumption (C4).

(b) **Claim 1:** *There exists a constant $k > 0$ such that, for all $n \in [1 \ldots N]$ and any small enough $\epsilon > 0$, if $|\varphi - \varphi_n^*| < k\,\epsilon$, then $V_n(\varphi)/V^* \in (1-\epsilon, 1+\epsilon)$.*

Proof. For all $n \in [1 \ldots N]$ we have $V_n(\varphi_n^*)/V^* = 1$, and the function V_n is continuously differentiable by (C4). Thus, Taylor's theorem says there is some $k_n > 0$ and $\overline{\epsilon}_n > 0$ such that, for all $\epsilon \in (0, \overline{\epsilon}_n)$, if $|\varphi - \varphi_n^*| < k_n \epsilon$, then $V_n(\varphi)/V^* \in (1-\epsilon, 1+\epsilon)$. Now let $k := \min\{k_1, \ldots, k_N\}$. ◇ `Claim 1`

Fix $\delta \in (0,1)$, and suppose $|\varphi_n^t - \varphi_n^*| < k\,\delta/2$ for all $n \in [1 \ldots N]$.

Claim 2: (i) *For all $n \in [1 \ldots N]$,* $\text{Prob}\left[\overline{V}_n^t/V^* \not\in (1-\delta, 1+\delta)\right] < \dfrac{4}{(\delta\, V^*\, I_n)^2}$.

(ii) $\text{Prob}\left[\overline{V}^t/V^* \not\in (1-\delta, 1+\delta)\right] < \dfrac{4}{(\delta\, V^*\, I)^2}$.

Proof. Consider the random variable $V_n^t := \min\{1,\ U_t/C_n(\varphi_n^t)\}$ (where U_t and C_n are as in (C3)). Since $0 \leq V_n^t \leq 1$, the variance of V_n^t is less than 1. If every voter deploys her dominant strategy, then Eq.(4) and assumptions (U), (C1) and (C2) imply that the random variables $\{V_i^t\}_{i \in \mathcal{I}_n}$ are independent and identically distributed to V_n^t. For all $n \in [1 \ldots N]$, assumption (C4) says that the expected value of V_n^t is $V_n(\varphi_n^t)$, while Claim 1 says $V_n(\varphi_n^t) = h_n\, V^*$ for some $h_n \in (1 - \frac{\delta}{2}, 1 + \frac{\delta}{2})$. Meanwhile, definition (3) says that $\overline{V}_n^t$ is the average of the I_n i.i.d. random variables $\{V_i^t\}_{i \in \mathcal{I}_n}$. Thus,

$$\overline{V}_n^t \quad = \quad V_n(\varphi_n^t) + \gamma_n^t \quad = \quad h_n\, V^* + \gamma_n^t, \tag{A4}$$

where γ_n^t is some random variable with mean zero and variance less than $1/I_n$. Thus,

$$\begin{aligned} \overline{V}^t \ \underset{(*)}{=}\ \frac{1}{I}\sum_{i \in \mathcal{I}} v_i^t \ &=\ \frac{1}{I}\sum_{n=1}^{N}\sum_{n \in \mathcal{I}_n} v_i^t \ \underset{(\dagger)}{=}\ \frac{1}{I}\sum_{n=1}^{N} I_n \overline{V}_n^t \ \underset{(\diamond)}{=}\ \frac{1}{I}\sum_{n=1}^{N} I_n (h_n\, V^* + \gamma_n^t) \\ &=\ \frac{V^*}{I}\sum_{n=1}^{N} I_n\, h_n \ +\ \frac{1}{I}\sum_{n=1}^{N} I_n \gamma_n^t \quad = \quad V^*\, h + \gamma^t, \end{aligned} \tag{A5}$$

where $h \in (1 - \frac{\delta}{2}, 1 + \frac{\delta}{2})$, and where γ^t is some random variable with mean zero and variance less than $1/I$. Here, $(*)$ is by Eq.(2), $(\dagger)$ is by Eq.(3), and $(\diamond)$ is by Eq.(A4).

Now, for all $n \in [1 \ldots N]$, we have

$$\begin{aligned} \left(\overline{V}_n^t/V^* \geq 1 + \delta\right) \ &\underset{(\diamond)}{\iff}\ \left(h_n + \gamma_n^t/V^* \geq 1 + \delta\right) \iff \left(\gamma_n^t/V^* \geq 1 + \delta - h_n\right) \\ &\underset{(*)}{\implies}\ \left(\gamma_n^t/V^* > \tfrac{\delta}{2}\right) \iff \left(\gamma_n^t \geq \delta\, V^*/2\right), \end{aligned}$$

where $(\diamond)$ is by Eq.(A4), and $(*)$ is because $1+\delta-h_n > \frac{\delta}{2}$ because $h_n < 1+\frac{\delta}{2}$. A very similar argument shows that

$$\left(\overline{V}^t_n/V^* \leq 1-\delta\right) \implies \left(\gamma^t_n \leq -\delta\, V^*/2\right), \quad \text{and thus,}$$

$$\text{Prob}\left[\overline{V}^t_n/V^* \not\in (1-\delta, 1+\delta)\right] \leq \text{Prob}\left[|\gamma^t_n| \geq \delta\, V^*/2\right] \underset{(*)}{<} \frac{4}{(\delta\, V^*\, I_n)^2},$$

where $(*)$ is by Chebyshev's inequality. This proves part (i) of the claim.

In a similar way, using Eq.(A5) and the fact that $1-\frac{\delta}{2} < h < 1+\frac{\delta}{2}$ we can show that

$$\left(\overline{V}^t/V^* \not\in (1-\delta, 1+\delta)\right) \implies \left(|\gamma^t| \geq \delta\, V^*/2\right), \quad \text{and thus,}$$

$$\text{Prob}\left[\overline{V}^t/V^* \not\in (1-\delta, 1+\delta)\right] \leq \text{Prob}\left[\gamma^t \geq \delta\, V^*/2\right] \underset{(*)}{<} \frac{4}{(\delta\, V^*\, I)^2},$$

where $(*)$ is by Chebyshev's inequality. This proves part (ii). $\diamond$ `Claim 2`

Now let $\delta := \epsilon/4$.

Claim 3: *If* $\overline{V}^t_n/V^* \in (1-\delta, 1+\delta)$ *and* $\overline{V}^t/V^* \in (1-\delta, 1+\delta)$, *then* $\overline{V}^t_n/\overline{V}^t \in (1-\epsilon, 1+\epsilon)$.

Proof. Since $\overline{V}^t_n/V^* > 1-\delta$ and $\overline{V}^t/V^* < 1+\delta$, we have

$$\frac{\overline{V}^t_n}{\overline{V}^t} > \frac{(1-\delta)V^*}{(1+\delta)V^*} = \frac{1+\delta-2\delta}{1+\delta} = 1-\frac{2\delta}{1+\delta} \underset{(*)}{>} 1-2\delta \underset{(\dagger)}{=} 1-\frac{\epsilon}{2} > 1-\epsilon.$$

Here $(*)$ is because $1+\delta > 1$, so that $\frac{2\delta}{1+\delta} < 2\delta$. Meanwhile, $(\dagger)$ is because $\delta = \epsilon/4$.

Meanwhile, since $\overline{V}^t_n/V^* < 1+\delta$ and $\overline{V}^t/V^* > 1-\delta$, we have

$$\frac{\overline{V}^t_n}{\overline{V}^t} < \frac{(1+\delta)V^*}{(1-\delta)V^*} = \frac{1-\delta+2\delta}{1-\delta} = 1+\frac{2\delta}{1-\delta} \underset{(*)}{<} 1+4\delta \underset{(\dagger)}{=} 1+\epsilon,$$

where $(\dagger)$ is because $\delta = \epsilon/4$, and $(*)$ is because $\delta < 1/4$, so that $1-\delta > 3/4 > 1/2$, so that $\frac{2\delta}{1-\delta} < 4\delta$. $\diamond$ `Claim 3`

Thus,

$$\text{Prob}\left[\boldsymbol{\varphi}^t \text{ violates } (\text{F2}_\epsilon)\right] = \text{Prob}\left[\exists\, n \in [1 \ldots N] \text{ with } \frac{\overline{V}^t_n}{\overline{V}^t} \not\in (1-\epsilon, 1+\epsilon)\right]$$

$$\underset{(\dagger)}{\leq} \text{Prob}\left[\frac{\overline{V}^t}{V^*} \not\in (1-\delta, 1+\delta) \text{ or } \exists\, n \in [1 \ldots N] \text{ with } \frac{\overline{V}^t_n}{V^*} \not\in (1-\delta, 1+\delta)\right]$$

$$\leq \text{Prob}\left[\frac{\overline{V}^t}{V^*} \not\in (1-\delta, 1+\delta)\right] + \sum_{n=1}^{N} \text{Prob}\left[\frac{\overline{V}^t_n}{V^*} \not\in (1-\delta, 1+\delta)\right]$$

$$\underset{(\diamond)}{\leq} \quad \frac{4}{(\delta\, V^*\, I)^2} \; + \; \sum_{n=1}^{N} \frac{4}{(\delta\, V^*\, I_n)^2} \quad \underset{(\ddagger)}{=} \quad \frac{4}{(\delta\, V^*\, I)^2} \; + \; N \cdot \frac{4}{(\delta\, V^*\, I/N)^2}$$

$$= \quad \frac{4\,(N^3+1)}{(\delta\, V^*\, I)^2} \quad \underset{(*)}{<} \quad 1-p$$

where (†) is by Claim 3, ($\diamond$) is by Claim 2, (‡) is by equation (5.2), and ($*$) is by inequality (5.1) (with $\delta = \epsilon/4$).

Thus, Prob$[\boldsymbol{\varphi}^t$ satisfies $(\mathrm{F2}_\epsilon)] > p$, as desired. Set $K := k/8$ to complete the proof.

(c) Let $r := \sqrt{1/(1-p)}$; then inequality (5.1) implies that $I > r/\epsilon$. Let $\delta := \epsilon - r/I$, and define $\overline{\varphi}_1^\delta, \ldots, \overline{\varphi}_N^\delta$ as in assumption (C3). Now let $V^* < \min\{V_1(\overline{\varphi}_1^\delta), \ldots, V_N(\overline{\varphi}_N^\delta)\}$. Thus, if we select $\varphi_1^*, \ldots, \varphi_N^*$ as in part (a), then $\varphi_n^* > \overline{\varphi}_n^\delta$ for all $n \in [1 \ldots N]$, because the functions $V_1, \ldots, V_N$ are decreasing, by assumption (C4). Thus, if $\boldsymbol{\varphi}^t$ is close enough to $\boldsymbol{\varphi}^*$, then $\varphi_n^t \geq \overline{\varphi}_n^\delta$ for all $n \in [1 \ldots N]$. Then Lemma A.1 implies that the fee schedule $\boldsymbol{\varphi}^t$ will satisfy $(\mathrm{F1}_\epsilon)$ with probability greater than p. □

Proof of Proposition 2. Recall that $E_t := \#\{i \in \mathcal{I};\;\; V_i^t = 1\}/|\mathcal{I}|$, and $\boldsymbol{\varphi}^t := (\varphi_1^t, \ldots, \varphi_N^t)$. We will analyze the convergence of a simplified version of rule (R1):

(R1*) If $E_t \geq \epsilon$, then set $\boldsymbol{\varphi}^{t+1} := \lambda \cdot \boldsymbol{\varphi}^t$. Otherwise, if $E_t < \epsilon$, then set $\boldsymbol{\varphi}^{t+1} := \boldsymbol{\varphi}^t$.

Clearly, the convergence of (R1*) will be slightly slower than (R1) (because it multiplies $\boldsymbol{\varphi}^t$ by a slightly smaller factor, so $\boldsymbol{\varphi}^t$ grows more slowly as $t \to \infty$). Thus, it suffices to establish the desired conclusion for (R1*).

For any $T \in \mathbb{N}$, let $S_T := \#\{t \in [1 \ldots T];\;\; E_t \geq \epsilon\}$. Note that S_T is a random variable, because $\{E_1, \ldots, E_T\}$ are random variables (because V_i^t is a random variable for every $i \in \mathcal{I}$ and $t \in \mathbb{N}$). Rule (R1*) implies that $\boldsymbol{\varphi}^T = \lambda^{S_T}\, \boldsymbol{\varphi}^0$. Let $S^* := \min\{s \in \mathbb{N};\;\; \lambda^s\, \boldsymbol{\varphi}^0$ satisfies $(\mathrm{F1}_\epsilon)$ with probability greater than $p\}$; we must determine how quickly S_T reaches S^* as $T \to \infty$. Let $T^*_{\epsilon,p} := \min\{t \in \mathbb{N};\;\; S_T \geq S^*\}$. Note that $T^*_{\epsilon,p}$ is a random variable, because $\{S_t\}_{t=1}^\infty$ are random variables.

Claim 1: *$T^*_{\epsilon,p}$ is almost surely finite, and* $\mathbb{E}[T^*_{\epsilon,p}] \leq S^*/(1-p)$.

Proof. Let $\{B_t\}_{t=1}^\infty$ be a Bernoulli process (i.e. a sequence of independent, identically distributed, $\{0,1\}$-valued random variables) with Prob$[B_t = 1] = 1-p$ for all $t \in \mathbb{N}$. For any $t \in \mathbb{N}$, $n \in [1 \ldots N]$, and $i \in \mathcal{I}_n$, equation (4) implies that $V_i^t = 1$ if and only if $U_i^t \geq C_i^t(\varphi_n^t)$. Thus,

$$E_t \quad = \quad \frac{1}{I} \sum_{n=1}^{N} \#\big\{i \in \mathcal{I}_n \;;\; U_i^t \geq C_i^t(\varphi_n^t)\big\}. \tag{A6}$$

For all $t \in \mathbb{N}$, define

$$E'_t \quad := \quad \frac{1}{I} \sum_{n=1}^{N} \#\big\{i \in \mathcal{I}_n \;;\; U_i^t \geq C_i^t(\lambda^{S^*-1}\, \varphi_n^0)\big\}. \tag{A7}$$

Then $\{E'_1, E'_2, E'_3, \ldots\}$ are independent random variables, by assumptions (U), (C1), and (C2). Furthermore, for all $t \in \mathbb{N}$, we have

$$\text{Prob}\left[E'_t \geq \epsilon\right] \;=\; \text{Prob}\left[\lambda^{S^*-1}\,\boldsymbol{\varphi}^0 \text{ violates } (\text{F1}_\epsilon)\right] \;\underset{(*)}{\geq}\; 1-p \;=\; \text{Prob}\left[B_t = 1\right], \tag{A8}$$

where $(*)$ is by the definition of S^*. Thus, for all $T \in \mathbb{N}$, if we define the random variables

$$S'_T \quad := \quad \#\{t \in [1\ldots T]\;;\; E'_t \geq \epsilon\} \quad \text{and} \quad S''_T \quad := \quad \sum_{t=1}^{T} B_t,$$

then inequality (A8) implies that S'_T stochastically dominates S''_T. Thus, if we define the random hitting times $T' := \min\{t \in \mathbb{N};\ S'_t \geq S^*\}$ and $T'' := \min\{t \in \mathbb{N};\ S''_t \geq S^*\}$, then T'' stochastically dominates T'. Thus, $\mathbb{E}[T''] \geq \mathbb{E}[T']$. But T'' is a Pascal (or 'negative binomial') random variable of type $(S^*, 1-p)$, which is almost-surely finite. Thus, T' is also almost-surely finite. Furthermore, $\mathbb{E}[T''] = S^*/(1-p)$. Thus, $\mathbb{E}[T'] \leq S^*/(1-p)$. Now, if $t < T^*_{\epsilon,p}$, then

$$\begin{aligned} E_t \;&\underset{(*)}{=}\; \frac{1}{I}\sum_{n=1}^{N} \#\big\{i \in \mathcal{I}_n\;;\; U_i^t \geq C_i^t(\varphi_n^t)\big\} \\ &\underset{(\dagger)}{\geq}\; \frac{1}{I}\sum_{n=1}^{N} \#\big\{i \in \mathcal{I}_n\;;\; U_i^t \geq C_i^t(\lambda^{S^*-1}\,\varphi_n^0)\big\} \;\underset{(\diamond)}{=}\; E'_t, \end{aligned} \tag{A9}$$

where $(*)$ is by Eq.(A6) and $(\diamond)$ is by Eq.(A7), and where $(\dagger)$ is because

$$\Big(t < T^*_{\epsilon,p}\Big) \iff \Big(S_t \leq S^* - 1\Big) \Longrightarrow \Big(\forall\, n \in [1\ldots N],\ \ \varphi_n^t = \lambda^{S_t}\varphi_n^0 \leq \lambda^{S^*-1}\,\varphi_n^0\Big)$$
$$\underset{(\ddagger)}{\Longrightarrow} \Big(\text{for all } n \in [1\ldots N] \text{ and all } i \in \mathcal{I}_n,\ \ C_i^t(\varphi_n^t) \leq C_i^t(\lambda^{S^*-1}\,\varphi_n^0)\Big).$$

(Here $(\ddagger)$ is because the function C_i^t is nondecreasing, for every $i \in \mathcal{I}$ and $t \in \mathbb{N}$.) Now, inequality (A9) implies that $S_t \geq S'_t$ for all $t \in \big[0\ldots T^*_{\epsilon,p}\big)$. Thus, $T^*_{\epsilon,p} \leq T'$. Thus, $T^*_{\epsilon,p}$ is almost surely finite, and $\mathbb{E}[T^*_{\epsilon,p}] \leq \mathbb{E}[T'] \leq S^*/(1-p)$. ◇ Claim 1

Now, let $\delta := \epsilon - 1/I\sqrt{1-p}$, and let $\overline{\varphi}_1^\delta, \ldots, \overline{\varphi}_N^\delta$ be as in assumption (C3). Let $\overline{S} := \min\{s \in \mathbb{N};\ \ \lambda^s\,\varphi_n^0 \geq \overline{\varphi}_n^\delta \text{ for all } n \in [1\ldots N]\}$. Thus, for all $t \in \mathbb{N}$, if $S_t \geq \overline{S}$, then Lemma A.1 implies that $\boldsymbol{\varphi}^t$ will satisfy (F1_ϵ) with probability greater than p. Thus, $S^* \leq \overline{S}$. But it is easy to verify that $\overline{S} = L(\delta)$, where $L(\delta)$ is defined by expression (6). This, together with Claim 1, implies that $T^*_{\epsilon,p}$ is almost-surely finite, and $\mathbb{E}[T^*_{\epsilon,p}] < L(\delta)/(1-p)$.

Finally, observe that, since rule (R1) increases the fee schedule $\boldsymbol{\varphi}^t$ faster than (R1*), we must have $T_p^\epsilon \leq T^*_{\epsilon,p}$. This completes the proof. □

Proof of Proposition 3. (a) For all $n \in [1 \ldots N]$, and all $\lambda \in \mathbb{R}$, define $L_n(\lambda) := \log(V_n(e^\lambda))$. The function $L_n : \mathbb{R} \longrightarrow \mathbb{R}$ is finite, decreasing, and continuously twice-differentiable everywhere on $\mathbb{R}$, by assumption (C4). For all $n \in [1 \ldots N]$ and $t \in \mathbb{N}$, if $\lambda_n^t := \log(\varphi_n^t)$, and s_n is defined as in (R2), then it is easy to see that

$$\log(\overline{V}_n^t) \underset{\text{(S1)}}{=\!=} \log(V_n(\varphi_n^t)) = L_n(\lambda_n^t) \quad \text{and} \quad s_n \approx L_n'(\lambda_n^t). \tag{A10}$$

Let $L^* := \log(V^*)$ and $\lambda_n^{t+1} := \log(\varphi_n^{t+1})$. Then taking the logarithm of both sides in rule (R2), and substituting the identities in (A10) and (S2), we get

$$\lambda_n^{t+1} = \lambda_n^t + s_n \cdot \left(L_n(\lambda_n^t) - L^*\right) \approx \lambda_n^t + L_n'(\lambda_n^t) \cdot \left(L_n(\lambda_n^t) - L^*\right), \tag{A11}$$

Formula (A11) is the *Newton-Raphson method*; when iterated, the sequence of values $\{\lambda_n^t\}_{t=1}^\infty$ converges rapidly to the (unique) value λ_n^* such that $L_n(\lambda_n^*) = L^*$. Indeed, under hypothesis (C4), there is some $k > 1$ such that

$$|\lambda_n^t - \lambda_n^*| = \mathcal{O}(k^{-t^2}) \qquad \text{as } t \to \infty. \tag{A12}$$

But $\varphi_n^t = \exp(\lambda_n^t)$ and $\varphi_n^* = \exp(\lambda_n^*)$, and the exponential function is continuously differentiable. Thus, Taylor's theorem and Eq.(A12) imply that $|\varphi_n^t - \varphi_n^*| = \mathcal{O}(k^{-t^2})$ as $t \to \infty$. In other words, there is some constant $B_0 > 0$ such that $|\varphi_n^t - \varphi_n^*| < B_0\, k^{-t^2}$ for all sufficiently large $t \in \mathbb{N}$. For any $\delta > 0$, let

$$T_0(\delta) := \sqrt{\frac{\log(B_0) - \log(\delta)}{\log(k)}}.$$

If $t \geq T_0(\delta)$, then $B_0\, k^{-t^2} \leq \delta$, so that $|\varphi_n^t - \varphi_n^*| < \delta$ for all $t \geq T_0(\delta)$. Finally, observe that $T_0(\delta) = \mathcal{O}\left(\sqrt{\log(1/\delta)}\right)$.

(b) For all $n \in [1 \ldots N]$, recall that $V^* = V_n(\varphi_n^*)$ and $\varphi_n^t = \exp(\lambda_n^t)$ and $\varphi_n^* = \exp(\lambda_n^*)$. Thus,

$$\left|\log\left(\frac{V_n(\varphi_n^t)}{V^*}\right) - 0\right| = \left|\log\left(\frac{V_n(\varphi_n^t)}{V_n(\varphi_n^*)}\right)\right| = \left|\log\left[V_n(\exp[\lambda_n^t])\right] - \log\left[V_n(\exp[\lambda_n^*])\right]\right|$$
$$= \left|L_n(\lambda_n^t) - L_n(\lambda_n^*)\right| \underset{(*)}{=\!=} \mathcal{O}(k^{-t^2}) \quad \text{as } t \to \infty.$$

Here, $(*)$ is by Taylor's theorem and Eq.(A12), because the function L_n is continuously differentiable by (C4). Now, $\exp(0) = 1$, and the exponential function is also continuously differentiable near zero, so a second application of Taylor's theorem implies that

$$\left|\frac{V_n(\varphi_n^t)}{V^*} - 1\right| = \left|\exp\left[\log\left(\frac{V_n(\varphi_n^t)}{V^*}\right)\right] - \exp(0)\right| = \mathcal{O}(k^{-t^2}) \quad \text{as } t \to \infty.$$

Thus, there is some $B_1 > 0$ such that $\left|\frac{V_n(\varphi_n^t)}{V^*} - 1\right| < B_1\, k^{-t^2}$, for all large enough $t > 0$. For any $\epsilon > 0$, define

$$T_1(\epsilon) := \sqrt{\frac{\log(B_1) - \log(\epsilon)}{\log(k)}}.$$

If $t \geq T(\epsilon)$, then $B_1 \, k^{-t^2} \leq \epsilon$, so that $\left|\frac{V_n(\varphi_n^t)}{V^*} - 1\right| < \epsilon$. By assumptions (S1) and (S2), this yields $\left|\frac{\overline{V}_n^t}{\overline{V}^t} - 1\right| < \epsilon$, for all $n \in [1 \ldots N]$; this is equivalent to (F2$_\epsilon$). Finally, observe that $T_1(\epsilon) = \mathcal{O}\left(\sqrt{\log(1/\epsilon)}\right)$.

(c) If V_n is isolestic, then L_n will be a linear function with slope s_n; in this case, the Newton-Raphson formula (A11) will cause $\lambda_n^{t+1} = \lambda_n^*$ for any choice of initial condition λ_n^t. That is: (R2) will converge in a single iteration to the optimal value of φ_n. □

Remark. Newton-Raphson (which is the content of rule (R2)) is only one of many numerical methods for finding the root of a function, and it is the reason we require the regularity hypothesis (C4). If we wish to weaken hypothesis (C4), and assume only that the functions V_n are continuous and increasing, then we can use the *bisection method* instead; this will require a suitably modified form of rule (R2). The price we pay is slightly slower convergence in Proposition 3: we would have $T(\epsilon) = \mathcal{O}(\log(1/\epsilon))$.

Appendix B: Notational Index

Symbol	First use	Brief Description
$\mathcal{A}_t$	↑ (P1)	the menu of alternatives during referendum t.
a, b	§1	generic elements of $\mathcal{A}_t$.
c_i	§1 only	(constant) marginal utility of money for voter i.
c_i^t	↓ (P5)	the disutility of the fee φ_n^t for voter i at time t.
$C_i^t(\varphi)$	↑ Eq.(4)	the disutility of a fee of size φ for voter i at time t.
$\mathcal{C}$	↑ (C)	the space of nondecreasing functions from $\mathbb{R}$ to $\mathbb{R}$.
ϵ	(F1$_\epsilon$),(F2$_\epsilon$)	a small 'error tolerance', used in definition of 'ϵ-fair' .
E_t	(R1)	$\#\{i \in \mathcal{I};\ V_i^t = 1\}/I$.
φ_n^t	↑ (P1)	the stochastic Clarke tax 'fee' for stratum n in referendum t.
$\varphi_n^{t,-}$	(P1)	slightly smaller than φ_n^t.
$\varphi_n^{t,+}$	(P1)	slightly larger than φ_n^t.
$\boldsymbol{\varphi}^t$	↑ (P1)	$\boldsymbol{\varphi}^t = (\varphi_1^t, \varphi_2^t, \ldots, \varphi_N^t)$ is the fee schedule for referendum t.
$\overline{\varphi}_n^\epsilon$	(C3)	a fee large enough that $\text{Prob}\left[U_t \geq C_n(\overline{\varphi}_n^\epsilon)\right] < \epsilon$.
γ	Eq.(A3)	a random variable, mean 0, variance less than $1/I$.
γ^t	Eq.(A5)	a random variable, mean 0, variance less than $1/I$.
γ_n^t	↑ Prop.3	a random variable, mean 0, variance less than $1/I_n$.
$\mathcal{G}_k$	§3	a group for personal tax schedule assignment.
$\mathcal{I}$	§1	the set of voters.
i, j	§1	generic voters in $\mathcal{I}$.
$\mathcal{I}_n$	↑ (P1)	the nth 'wealth stratum' of voters in $\mathcal{I}$.
$\mathcal{I}_n^+$	(P1)	a random selection of half the voters in $\mathcal{I}_n$.
$\mathcal{I}_n^-$	(P1)	the other half of the voters in $\mathcal{I}_n$.
I	Eq.(2)	the cardinality of $\mathcal{I}$.
I_n	Eq.(3)	the cardinality of $\mathcal{I}_n$
$\mathcal{J}_k$	§3	the jury deciding personal tax schedule for group $\mathcal{G}_k$.

Symbol	First use	Brief Description
$L(\epsilon)$	Eq.(6)	$\max\{\log(\overline{\varphi}_n^\epsilon/\varphi_n^0)\}_{n=1}^N/\log(\lambda)$.
μ_t	(U)	probability distribution of the random variables $\{U_i^t\}_{i\in\mathcal{I}}$.
n	$\uparrow$ (P1)	an element of $[1\ldots N]$, indexing a wealth stratum.
N	$\uparrow$ (P1)	the number of wealth strata (typically $N=10$).
$\mathbb{N}$		$\{0,1,2,3,\ldots\}$; the set of natural numbers.
p	$\downarrow$ (F2$_\epsilon$)	the probability that the fee schedule $\boldsymbol{\varphi}$ will be ϵ-fair.
$p_i^t(\mathbf{v})$	(P4)	$\sum_{j\in\mathcal{I}\setminus\{i\}}[v_j^t(b)-v_j^t(a^*)]$; probability of stochastic Clarke tax.
$\mathbb{R}$		the set of real numbers.
$\mathbb{R}_+$		the set of nonnegative real numbers.
ρ_n	(C)	probability distribution of random variable C_i^t, for all $i\in\mathcal{I}$ and $t\in\mathbb{N}$.
t	$\uparrow$ (P1)	time (indexes the sequence of referenda).
T_p^ϵ	Prop.2	time after which (F1$_\epsilon$) is satisfied with probability p.
$T_0(\delta)$	Prop.3(a)	time after which $\lvert\varphi_n^t-\varphi_n^*\rvert<\delta$ for all $n\in[1\ldots N]$.
$T_1(\epsilon)$	Prop.3(b)	time after which (F2$_\epsilon$) is satisfied.
u_i^t	$\uparrow$ (P1)	voter i's vNM cardinal utility function over $\mathcal{A}_t$.
$u_i^\$$	$\uparrow$ (P1)	voter i's vNM cardinal utility function for money.
U_i^t	$\uparrow$ (U)	$\max_{a\in\mathcal{A}_t} u_i^t(a)$, the 'intensity' of i's preferences.
$v_i^t(a)$	(P2)	voter i's declared value for alternative a in $\mathcal{A}_t$.
$\mathbf{v}$	(P3)	$\mathbf{v}:=(v_i^t)_{i\in\mathcal{I}}$.
$V(a)$	(P3)	$V(a):=\sum_{i\in\mathcal{I}}v_i^t(a)$.
V_i^t	$\uparrow$ Eq.(2)	$\max_{a\in\mathcal{A}_t} v_i^t(a)$; voter i's 'influence' on referendum t.
$\overline{V}^t$	Eq.(2)	$\frac{1}{I}\sum_{i\in\mathcal{I}}V_i$.
$\overline{V}_n^t$	Eq.(3)	$\frac{1}{I_n}\sum_{i\in\mathcal{I}_n}V_i^t$.
$\overline{V}_n^{t,+}$	(R2)	$\frac{1}{\lvert\mathcal{I}_n^+\rvert}\sum_{i\in\mathcal{I}_n^+}V_i^t$.
$\overline{V}_n^{t,-}$	(R2)	$\frac{1}{\lvert\mathcal{I}_n^-\rvert}\sum_{i\in\mathcal{I}_n^-}V_i^t$.
$V_n(\varphi)$	(C4)	expected influence of a stratum n voter, given fee φ.
V^*	Prop.1	a target value for $V_1(\varphi_1^*),\ldots,V_N(\varphi_N^*)$ and/or $\overline{V}^t$.
w_i^t	$\downarrow$ (P5)	voter i's wealth level prior to referendum t.

References

Allegretto, S., March 2011. The state of working america's wealth. Tech. Rep. 292, Economic Policy Institute briefing paper.

Atkinson, A., Saez, E., Piketty, T., 2011. Top incomes in the long run of history. Journal of Economic Literature 49 (1), 3–71.

Bailey, M. J., April 1997. The demand revealing process: To distribute the surplus. Public Choice 91 (2), 107–26.

Cavallo, R., 2006. Optimal decision-making with minimal waste: Strategyproof redistribution of VCG payments. In: Proc. of the 5th Int. Joint Conf. on Autonomous Agents and Multi Agent Systems (AAMAS'06). Hakodate, Japan, pp. 882–889.

Clarke, E., Fall 1971. Multipart pricing of public goods. Public Choice 11, 17–33.

Faltings, B., 2004. A budget-balanced, incentive-compatible scheme for social choice. In: Conference on Agent-mediated E-commerce (AMEC'04). pp. 30–43.

Gary-Bobo, R. J., Jaaidane, T., May 2000. Polling mechanisms and the demand revelation problem. Journal of Public Economics 76 (2), 203–238.

Green, J., Kohlberg, E., Laffont, J., 1976. Partial equilibrium approach to the free rider problem. Journal of Public Economics 6, 375–394.

Green, J., Laffont, J., 1979. Incentives in Public Decision Making. North-Holland, Amsterdam.

Kahneman, D., Tversky, A. (Eds.), 2000. Choices, Values, and Frames. Cambridge UP, Cambridge, UK.

Mas-Colell, A., Whinston, M. D., Green, J. R., 1995. Microeconomic Theory. Oxford UP, Oxford, UK.

Moulin, H., 1988. Axioms of cooperative decision making. Cambridge University Press, Cambridge.

Saez, E., Kopczuk, W., June 2004. Top wealth shares in the united states, 1916-2000: Evidence from estate tax returns. National Tax Journal 57 (2), 445–487.

Wilkinson, R., Pickett, K., 2010. The Spirit Level. Penguin.

Wolff, E. N., March 2010. Recent trends in household wealth in the united states: Rising debt and the middle-class squeeze. Tech. Rep. 589, The Levy Economics Institute of Bard College., Annandale-on-Hudson, NY.